

UNIVERSITY of PETRA

English Department

## **Assessment of Google's Translation of Legal Texts**

**تقويم ترجمات جوجل للنصوص القانونية**

A thesis submitted in partial fulfillment of the requirements for the degree of  
Master of Arts in Translation

By

Basel Abd-Alkareem Hijazi

Supervised by

Professor Abdul Baki As-Safi

Amman, Jordan

June 2013

# Assessment of Google's Translation of Legal Texts

Submitted by

Basel Abd-ALKareem Hijazi

In partial fulfillment of the requirements for the degree of  
Master of Arts in Translation

The thesis is approved by:

Committee Chair -----

Examining Committee

1. -----
2. -----
3. -----
4. -----

*Dedication*

*To my mother and father who have inspired me throughout  
my personal and professional life.*

## **Acknowledgements**

I would first like to express my deep gratitude to my supervisor Professor Abdelbaki As-Safi whose guidance, shared knowledge, mental and psychological support made this work possible. I consider myself tremendously lucky to have been supervised by Professor As-Safi who has created a research environment that is thrilling, productive, light-hearted, motivating and cooperative. Professor As-Safi's energetic commitment to research and mentorship has rubbed off on me. He has given me far more time and guidance than I could have expected or asked for.

I am particularly indebted to the committee members for the feedback they provide and their patience in thoroughly reading through the thesis.

It is also my duty to record my thankfulness to the University of Petra and in particular Dr. Nehal Ameera who has been abundantly helpful starting from the day I joined the university up to the end of the course of this thesis.

I take this opportunity to extend my appreciation to my parents for their generous financial and psychological support. I should not forget my wife for her moral support throughout my MA studies. I express my gratitude to my two kids, Karim and Ahmad, who have been very patient and understanding during the period of my study.

Finally, I would like to express my deepest gratitude to the Royal Jordanian Air Force for granting me the permission to further my studies, and for assisting me in different aspects throughout the duration of my MA Program.

## **Abstract**

In recent years, research as well as software applications have been concentrating on Machine Translation (MT). This is due to many factors, the most important of which is the increasing need to create online communication between different parts of the world and between people speaking different languages. Translation between two languages, which are distant and have different structures, e.g. Arabic and English, poses a challenge for linguists who aim at developing a system which could contribute towards more accurate Machine Translation.

The current thesis conducts an experiment on *Google Translate*, which is considered one of the most popular ‘web-based’ free tools. An experiment is carried out on fourteen English-authentic articles extracted from six different legal contracts. These articles are fed into the system which in return produces Arabic outputs as translations of the articles.

An assessment is performed to analyze the performance of the tool in handling the legal text on two main levels, the lexical and the syntactic respectively. An error analysis is also furnished on both levels by categorizing the errors under the two levels. Each level includes three subcategories of problems that recur in the assessment. The lexical level breaks into polysemy and homonymy, legal doublets, and legal adverbs. On the other hand, the syntactic level includes problems of morphological parsing, concord and modality.

The overall assessment shows that the system is not feasible in the field of legal translation as such practice is characterized by paramount precision that *Google Translate* fails to achieve. However, the system can usually furnish a gist of the input which can help end users to figure out the subject matter of the source text.

## الملخص

تركزت الدراسات و تطبيقات برمجة الحاسوب في السنوات الأخيرة على الترجمة الآلية. و يعود ذلك لأسباب عدة من أهمها الحاجة المتزايدة للتواصل عبر الانترنت بين الدول المختلفة في العالم و كذلك لتواصل الشعوب التي تتحدث لغات مختلفة. و تشكل الترجمة بين لغتين متباينتين في التركيب تحدياً للغويين الذين يسعون إلى تطوير نظام قادر على الإسهام في ترجمة آلية أدق.

تجري هذه الرسالة تجربة على مترجم جوجل الذي يعد واحداً من أشهر الأدوات المجانية المستخدمة على الشبكة العنكبوتية. و تشمل هذه التجربة أربع عشرة فقرة قانونية مستخرجة من ستة عقود قانونية مختلفة حيث يتم إدخال هذه الفقرات في النظام الذي بدوره يقوم بتزويد ترجمة تلك الفقرات.

و يقيم أداء النظام عن طريق تحليل تعامل الأداة مع النصوص القانونية من خلال مستويين أساسيين و هما المستوى المعجمي والمستوى النحوي. و يضم كل مستوى ثلاثة مستويات فرعية تحتوي كل منها على المشاكل التي تكررت في هذا التقويم. و ينقسم المستوى المعجمي إلى الوحدات المعجمية المحولة و المتعددة المعنى و العبارات المزدوجة العبارات و الظرف القانوني, و من جهة أخرى يحتوي المستوى النحوي على الأعراب الصرفي و التوافق بأنواعه و أوجه الفعل.

يظهر التقويم الكلي أن النظام قاصر في مجال الترجمة القانونية و التي تتسم بدقة متناهية يصعب على مترجم جوجل تحقيقها. و مع ذلك يستطيع النظام تزويد المستخدم بجوهر المدخل الذي يساعد في كشف معنى موضوع النص الأصلي.

## List of Abbreviations

3GS	third singular
AI	artificial intelligence
CAT	computer aided/assisted translation
MT	Machine Translation
N	noun
NLP	natural language processing
PAST-PART	past participle
PL	plural
PRES-PART	present participle
SG	singular
SL	source language
ST	source text
SMT	Statistical Machine Translation
TL	target language
TT	target text
V	verb

# *Table of Contents*

Dedication.....	III
Acknowledgments.....	IV
Abstract.....	V
List of Abbreviations.....	WIII
Table of Contents.....	VIII
<b>1. Chapter One: Introduction.....</b>	<b>1</b>
1.1. Preliminaries .....	1
1.2. Problem.....	3
1.3. Hypotheses.....	3
1.4. Objectives.....	3
1.5. Significance.....	4
1.6. Methodology .....	4
1.7. Scope .....	5
1.8. Review of Literature.....	6
<b>2. Chapter Two: Machine Translation (General Facts) .....</b>	<b>9</b>
2.1. Preliminaries .....	9
2.2. Types of Machine Translation.....	9
2.3. Types of Analysis.....	12
2.4. Machine Translation History.....	16
2.5. Difficulties Facing Traditional Approaches to MT.....	23
2.6. Statistical Approach to MT.....	26
2.7. Probability Theory.....	27
2.8. Alignment.....	30

2.9.	<i>Google Translate (GT)</i> .....	35
2.9.1.	Google’s Approach.....	36
<b>3.</b>	<b>Chapter Three: Legal Discourse Characteristics .....</b>	<b>38</b>
3.1.	Preliminaries.....	38
3.2.	Characteristics of Legal Language.....	39
3.3.	Types of Legal Writing .....	43
3.4.	Contracts as a Sub-Legal Text.....	44
3.5.	The Difficulty of Translating Legal Texts.....	45
3.6.	Translating Legal Texts using Machine Translation.....	48
<b>4.</b>	<b>Chapter Four: Evaluation of Machine Translation Output .....</b>	<b>49</b>
4.1.	Introduction.....	49
4.2.	Types of MT Evaluation.....	50
4.3.	Human vs. Automatic Evaluation of MT.....	50
4.4.	Methodology and Corpus.....	56
4.5.	Qualitative Assessment.....	58
4.5.1.	Lexical Level Assessment.....	59
4.5.1.1	Polysemy and Homonymy.....	59
4.5.1.2	Couplings.....	67
4.5.1.3	Legal Adverb (here/there+ preposition).....	69
4.5.2.	The Syntactic Level Assessment.....	72
4.5.2.1.	Morphological Parsing.....	72
4.5.2.2.	Concord.....	74
4.5.2.3.	Modality: shall/may.....	77

<b>Conclusions and Recommendations .....</b>	<b>83</b>
<b>Bibliography.....</b>	<b>87</b>
<b>Appendices.....</b>	<b>I</b>
Appendix 1 (the test set, <i>Google Translate</i> output, and proposed translation)	II
Appendix 2 Buckwalter Transliteration Chart.....	XI

# Chapter One

## Introduction

### 1.1. Preliminaries

The thought of using the computer to translate or assist translating human languages is almost as old as the computer itself. Machine Translation (MT) is an area of information technology and applied linguistics dealing with the translation of human languages. Computer technology has been applied to technical translation in order to improve one or both of the following factors (1) *Speed*: Translation by or with the aid of machines can be faster than manual translation, and (2) *Cost*: Computer aids in translation can reduce the cost per word of a translation. In addition, the use of MT can result in improvements in quality, particularly in the use of consistent terminology within a text or for a particular kind of client.

MT is one of the applications of Natural Language Processing (NLP). Also, it is called Automatic Translation which entails the process that utilizes computer software to translate texts from one natural language to another. It therefore follows that MT is not simply substituting words for other words, but like human translation it involves the application of complex linguistic rules especially in morphology, syntax and semantics which means that the computer could be used to translate from the source language (SL) into the target language (TL). It could translate an entire document automatically and then presents it to a human. MT should not be confused with Computer Aided Translation (CAT) which covers a wide range of tools such as spelling and grammar checkers, terminology managers, unilingual and bilingual electronic dictionaries, etc. The translation tool under

assessment in the current research, *Google Translate*, is a web-based MT system which is able to produce output of texts unaided.

However, it turns out that really “good” MT is so hard to attain. The task has fatigued the best computing resources of every generation attempting it. Nevertheless, MT is going stronger than ever, fired up by the globalization of the Net. Today, all over the world, software designers, programmers, hardware engineers, neural-network experts, AI specialists, linguists, and cognitive scientists are enlisted in the effort to teach computers how to port words and ideas from language to language (Abu-Al-Sha’r & Zughoul, 2009).

Arabic is one of the languages which has been tackled since the early days of MT; however, few systems have dealt with the Arabic language due to its syntactic characteristics which are different from Latin characteristics. The Arabic language, which is the mother tongue of more than 300 million people, presents significant challenges to many NLP applications since Arabic is a nonconfigurational language.

Legal texts have been chosen in this research due to their “conventional” nature of the structure formats in use and the formulaic nature of the kind of texture employed in the field of translation (Hatim, 1997:38). In addition, legal discourse is characterized by a ‘close-knit texture’ (ibid: 18). Due to the characteristics the legal discourse has, the approach that should be followed to render a legal text is ‘literal’.

Because of globalization, developments in information technology and the international giant strides in communication, translating legal texts is of paramount importance. As many businesses are investing in the Arab world whose economies have proven to be robust in comparison with the west economies, the necessity for translating different legal documents (contracts, certificates, and agreements) has surfaced.

## **1.2. Problem**

It is widely known that a lot of people depend on MT to translate different kinds of documents for the two aforementioned reasons. Unfortunately, they are unaware of the grave consequences resulted from erroneous outcomes especially when precision in such documents is a priority.

Although *Google Translate*, the system under assessment, has won an international competition for English-Arabic and English-Chinese Machine Translation, translated texts can often include apparently nonsensical and obvious errors. This is due to the fact that *Google Translate* adopts a Statistical Machine Translation approach (SMT) rather than a dictionary/grammar approach.

## **1.3. Hypotheses**

The thesis will adopt the following hypotheses:

- a. It is hypothesized that MT and *Google Translate* in particular has not so far produced accurate translation even if the text type translated is formulaic such as legal language. The inaccuracies should be verified and classified according to the level they pertain to, i.e. lexis or syntax.
- b. It is also hypothesized that *Google Translate* is able to provide a gist of the input rather than a functional equivalent of the legal source text.

## **1.4. Objectives**

The current thesis aims at assessing the performance of *Google Translate* in handling legal discourse, particularly contracts. The research touches upon the main errors made by the system by categorizing such errors under two main categories: lexical and syntactic.

The researcher aims at constructing a legal test set (articles extracted from different contracts). This test set can be used in further research for the purpose of comparing the performance of other MT systems.

### **1.5. Significance**

The main significance of the current research is the error analysis conducted by the researcher. The errors analyzed are systematic, and they recur in different test examples. Google as a corporation is a very huge entity which allocates a lot of resources (time, money and effort) for research and development. Such analysis is of paramount importance for people working on developing *Google Translate* in particular and any other MT systems in general.

In addition, the researcher furnishes parallel corpora (translated articles, couplets and adverbs) which could benefit existing and future systems designed to handle legal language, and more specifically the systems that adopt statistical approaches.

This study serves as a warning to lawyers, law librarians and laypeople who are after translating legal documents for official use. The research hints that outcomes could only serve as a gist for the original.

To the best of the researcher's knowledge, legal language has not been assessed in the field of MT and *Google Translate* in particular. The originality of the method, test set, results and conclusions are of good value to MT developers, translators and researchers in the field of computational linguistics.

### **1.6. Methodology**

In the current research, six different contracts regulating different matters are chosen. Fourteen articles are extracted based on certain criteria. Articles (the input)

are fed individually into *Google Translate*. The tool performs the translation to produce outputs. The outputs are scrutinized against a proposed translation furnished by the researcher in order to locate salient lexical and syntactic errors. The errors that recur in the study are categorized and discussed in chapter four.

The thesis comprises four chapters. Chapter One is an introduction which consists of the problem, the hypotheses, the significance of the study, the scope of the study, among other introductory items. Chapter Two furnishes historical overview of MT tracing its revolution, success and failure. It discusses the different approaches adopted in MT ranging from direct approaches to statistical ones. It also touches upon the difficulties that MT faces. Finally, an overview of *Google Translate* is furnished. Chapter Three is devoted to legal register, i.e. the salient characteristics of this language variety which distinguishes it from other varieties. Difficulties in translating legal documents are discussed as well as the role of MT in the process of translating legal documents. Chapter Four examines the translations produced by *Google Translate*. The chapter is of an evaluative nature, i.e. legal texts will be assessed as objectively as possible based on solid criteria. The thesis ends with conclusions and recommendations based on the assessment and error analysis.

## **1.7. Scope**

The research is unidirectional, i.e. from English into Arabic. Authentic English legal texts are fed into *Google Translate*. An assessment is conducted on the Arabic output compared with the English input. Further research could be conducted on Arabic-English translation. Moreover, the research is limited to legal texts. Further studies could be conducted on other types of texts.

The assessment in the current research does not include all the errors and problems exhibited in the outcome. It rather focuses on salient lexical errors which are: polysemy/homonymy, legal doublets, and legal adverbs. The next level of assessment is the syntactic wherein morphological parsing, concord and legal modality are the focal point of the analysis. It is important to notice that other problems are worth noting such as (passive structures, tenses, word order, etc.). These problems could be dealt with in further research as it is beyond the scope of the current thesis.

## **1.8. Review of Literature**

Arabic is one of the languages which has been tackled since the early days of MT; however, few systems have dealt with the Arabic language due to its syntactic characteristics which are different from Latin characteristics (Mokhtar et al., 2000). Ibrahim (1991) has discussed the problem of the English-Arabic translation of the embedded idioms and proverb expressions in English sentences. Rafea et al., (1992) try to do something different. They have developed an English-Arabic MT system, which translates a sentence from the domain of the political news of the Middle East. Maalej (1994), on the other hand, discusses the MT of English nominal compounds into Arabic. It has been motivated by their frequent occurrence in referring and naming in all text-types. El-Desouki et al., (1996) discuss the necessity of modular programming for English-Arabic MT. A translation of an English subset of a knowledge base, written in KROL (Shaan, 2000), to the corresponding Arabic phrases is described in El-Saka et al., (1999). Mokhtar et al., (2000) developed an English-Arabic MT system, which is applied to abstracts from the field of Artificial Intelligence. Al-Najjar (2004), a renowned researcher in the field of MT, designs a lexicon specified for MT from English into

Arabic. In an earlier study, he discusses the relation between linguistics and MT. (Al-Najjar 1999).

From an evaluative perspective, Al-Salman (2004) attempts to carry a comparative study to assess the effectiveness of three different MT programs. Al-Wasiti (2005) stresses that the quality of machine translation does not reach the same level of human translation. He demonstrates his conclusion by translating three text types from English into Arabic using *Al-Wafi*. Abdo (2007) tackles the rendition of pronouns in MT. His analysis shows that the system under investigation erroneously renders the pronouns resulting in ill-formed structures. Abdul-Hameed (2008) attempts on devising a framework for evaluation of MT. He conducts an assessment on three different tools: *Al-Nakel*, *Al-Arabi 2.00*, *Golden Al-Wafi 1.00* and *Al-Mutarjim Al-Araby 3.00*. The findings of his evaluation show that the systems mentioned above produce average or below average quality. Analogues to Abdul-Hammed, Abu-Al-Sha'r & Zughoul (2009) evaluate the translations of six different online services in which *Google Translate* is among them. They reveal that the services produce texts that are incomprehensible. However, *Google Translate* produces better quality outputs when it comes to translating English into Arabic.

Al-Dabbagh (2013) has conducted an assessment on *Google Translate* by choosing four different text types, namely the journalistic, the economic, the scientific and the technical, two of which are extracted from web pages and the other two from books. She has found that the system produces Arabic texts that abound with lexical, grammatical and textual flaws. The analysis indicates that the errors recur regardless of the text type, text length, text difficulty and input mode. In another study, Al-Dabbagh (2010) carries out a questionnaire which investigates how the readers rate the quality of translated texts by *Google Translate*. Her

findings are that the system fails in providing its users with a general idea about the translated texts.

Most findings in the above studies confirm the fact that MT still needs more time and effort to be excreted by researchers and MT developers in order to produce acceptable translations. Such findings are concluded on the basis of rigorous research and assessment regardless of the text type being chosen. The current research conclusions coincide with the above findings since *Google Translate* fails to handle legal discourse.

## **Chapter Two**

### **Machine Translation: General Facts**

#### **2.1. Preliminaries**

Machine Translation is generally perceived as the application of computer and language science. Hutchins (1994:13) defines Machine Translation as “the computerized systems responsible for the production of translation with or without human assistance”. However, he distinguishes between computer-based translation tools which support translators such as dictionaries, remote terminologies, etc. and MT which is the automation of the full translation process.

Nirenburg (1987:2) defines the task of MT very simply by stating that “the computer must be able to obtain as input a text in one language SL and to produce as output a text in another language TL, so that the meaning of the TT (target text) is the same as that of the ST (source text). From Nirenburg’s definition, we can observe that MT definition is not very much different from (human) translation and therefore, this could pave the way to anticipate how difficult and complex the process is.

#### **2.2. Types of Machine Translation**

Machine Translation can be divided into two main systems: bilingual systems and multilingual systems. The former systems are designed for two particular languages. These systems could be uni-directional i.e. operate in one direction, for example, from Arabic into English, or they could be bi-directional in which they operate both ways. On the other hand, multilingual systems are designed for more than one pair of languages. These systems provide translations of one language to any one or more languages within the same system.

### **2.2.1. The Direct Approach**

Another classification of MT is by the approach or strategy the system adopts. There are generally three basic approaches in MT: the first approach is the “direct approach” which is dedicated to one particular pair of languages. As its name suggests, the translation is direct from the ST text into the TT. Little analysis is carried out on such systems. To Tucker (1978), the direct translation strategy passes each sentence of the text to be translated through a series of principal stages in which the output of each stage is the input to the next. These strategies draw heavily on well developed dictionaries, morphological analysis, and text processing software. Linguistic theories and parsing principles are of little importance due to the fact that surface structures are dealt with using this method while deep structure analysis is abandoned. An example of direct approach is the Georgetown system.

### **2.2.2. The Interlingual Approach**

The second approach is the interlingual approach. Unlike the direct approach, the interlingual approach goes through two phases of translation. The first is from SL into interlingua (IL) and from IL into TL. Mel’chuck (quoted in James 1980), in a discussion of translation theory, defines interlingua as “a system which encompasses, as is desirable for translation, the analysis characteristic of the SL and the synthesis characteristic of the TL text”. The SL analysis is SL-oriented and it is not related to the TL by any means. In addition, programs designed for TL synthesis are TL oriented and have nothing to do with the structure of the SL. One distinctive feature of such systems is economy of effort in construction compared with the direct approach systems. Tucker (1978: 25) states that:

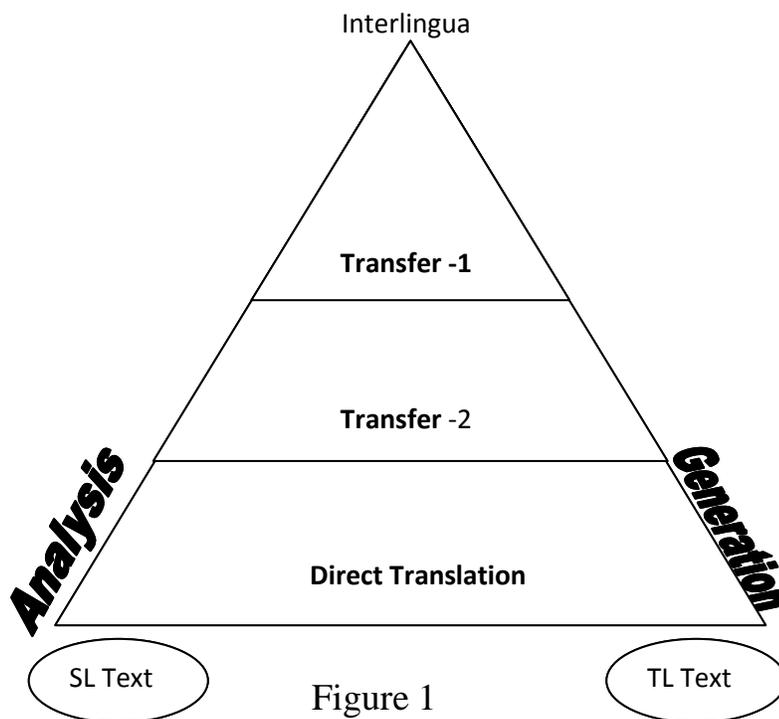
a genuine interlingua must be able to express the meaning of the text to be translated. This type of research can proceed despite the debate as to whether it is possible to capture and formalize the human encyclopedic knowledge that is a necessary part of language understanding. This is because one can work with

subworlds and sublanguages in the hope of producing a translation scheme which can later extend beyond the experimental stage.

In the current research, the sublanguage being under investigation is of legal nature. In this respect, Somers (2003:6) believes that “MT could work if the input text was somehow restricted”. Therefore, this research accords with Tucker, Somers, among others that MT systems should limit their domain to a particular field or area of discourse. The TAUM-METEO system is an example of an operational transfer-based MT in a limited subject matter domain which is weather reports. This system was a big success in the history of MT for the scarce reliance on post-editing.

### **2.2.3. The Transfer Approach**

The third approach is the transfer approach which is carried out in three major stages. The first stage converts the STs into SL-oriented representations. In the second stage, a transfer is carried at both levels (lexical and structural) into equivalent TL- oriented representations. The final stage is the generation of the TT. Three dictionaries are necessary in this approach: an SL dictionary, a bilingual dictionary and a TL dictionary. One main difference between the interlingua and the transfer approach is that the latter tackles the inherent ambiguities in SL and TL, while in the interlingual approach, ambiguities in each and every bit in the ST should be dealt with in order to have a proper TT. Tucker (1978) has illustrated the process of the transfer approach in a model which explicates the transfer stages in the following figure:



The model represents the stages in which the transfer could take place. There are two levels of transfer to show that MT developers are to decide which method to adopt. Transfer 2 represents the least complex module compared with Transfer 1 as little analysis is carried out, i.e. transfer is performed on purely syntactic surface structure markers. The researcher has modified Tucker's triangle by adding the direct approach at the bottom of the triangle and the Interlingual approach at the top to illustrate the three different traditional approaches to MT. To simplify the above diagram, the Interlingual approach tries to eliminate the transfer task, and the direct approach tries to do without analysis and generation.

### 2.3. Types of Analysis

Since computers can recognize only symbols which are part of its own character set: letters, punctuation marks, spaces, numbers, and various other symbols, analysis should be carried out on different levels of linguistic description.

Butler (1985) believes that analysis should be carried out on the level of graphology, phonology (the sound system), lexis (vocabulary), syntax (the combination of words in grammatical constructions), and semantics (meaning).

### **2.3.1. Graphological Analysis**

A graphological analysis is considered one of the simplest matters in the analytical process. The computer can easily count individual letters and words appearing in a text and generate word-lists, indexes, concordances, or statistical information. Therefore, most computational work on texts has relied on analysis at this level.

### **2.3.2. Lexical Analysis**

After generating a word-list, index, or concordance, lexical analysis takes place. As words have inflected forms, the computer has to decide the lemma for the given word, i.e. its origin. For example, the verb ‘to walk’ may appear in different forms, ‘walk’, ‘walking’, ‘walks’, etc. the base form of the word ‘walk’ is called lemma, the word one looks up from the dictionary. The activity of grouping the different lexemes of one lemma is called lemmatization in the science of computational linguistics. Lemmatization is therefore the algorithmic process of determining the lemma of a given word which is considered to be one of the most challenging tasks in MT.

### **2.3.3. Syntactic Analysis**

Syntactic analysis is sought next in MT which requires the recognition of clause and phrase boundaries, and the classification of clauses (main or subordinate), phrases (nominal, verbal, prepositional, etc), words (as nouns, verbs, adjectives, or other parts of speech). Explicitly, such analysis is of an intricate nature which needs an automatic parsing system which is not 100% accurate. Nirenburg (1987) illustrates the importance of the knowledge of the syntactic structure by furnishing the following example:

(1) *The coach lost a set*

Without syntax, a translator cannot decide whether *coach* is a noun or verb, *lost* is a verb or adjective, *set* a noun, verb or an adjective. This knowledge is found in grammar books of English in which the human translator should be aware of before indulging himself in translation. In MT, there is a special processing unit called ‘syntactic parser’ which applies this knowledge to the input text and produces its syntactic structure. However, syntactic analysis might not be enough in some cases:

(2) *Old men and women*

In (2), the modifier “old” could modify the noun “men”, or it could modify both nouns “men and women”. In such instances, this issue constitutes a problem to human translators, let alone MT systems. More comprehensive, such ambiguities lead to a big number of parses.

#### **2.3.4. Semantic Analysis**

Analogous to syntactic analysis, semantic analysis is considered a challenging task to MT developers. This is due to the fact that “meaning” is an abstract concept compared with the linguistic form. There are different approaches to semantic analysis. One approach is to examine the collocations of the SL and TL. Collocation is simply how words associate with other words; some items strongly collocate with each other, for example (*rancid* with *butter*, *addled* with *egg*). Other collocations have a wide range of items in which they could collocate with. *Bad* or *good* as adjectives could associate with a large number of words. The latter type is of a problematic area in MT. The second approach is to create a list of words and analyze it automatically by means of a thesaurus held with the computer system to isolate predominant themes in the text. Butler (1978) summarizes the complexity of the semantic analysis by stating that:

fully automatic translation of high quality accepting any kind of text as input, is an ideal which will not be achieved in the immediate future. Progress towards this ideal is limited by the extent of our knowledge about human languages. In particular, the semantic analysis of language, which is arguably the most important area for Machine Translation, is still at rudimentary stage, even for English. Because of these limitations, many experts in Machine Translation have decided to focus on much more modest aims (ibid: 32).

The present achievements in MT coincide with Butler's hypothesis specifically in the area of the quality of the outcome. However, more research is directed towards translating from one language to many and vice versa with lower ambitious outcomes; *Google Translate* is an example of a service which can translate between 58 languages instantly.

## **2.4. Machine Translation History**

After discussing certain principles and problems facing MT, a brief history is furnished in the coming section. Although MT was a dream in the 17<sup>th</sup> century, it did not come to light until late 20<sup>th</sup> century. Hutchins (1994) is one of the renowned researchers who wrote in the history of MT among others such as Tucker and Nirenburg (1984), and As-Safi (2004).

Hutchins (1994) classifies the history into four main periods:

1. Precursors and Pioneers, 1933-1956
2. The decade of high expectations and disillusion 1956-1966
3. The quiet decade 1967-1976
4. MT research in the 1980s

### **2.4.1. Precursors and Pioneers, 1933-1956**

The first MT work was made in solid proposals by George Artsouni and Petr Smirnov-Troyanskii. Their work was simply a storage device on a paper tape

which could be used to find the equivalent of any word in another language. Troyanskii's proposals, on the other hand, were more solid in which he predicted three stages of mechanical translation. To As-Safi (2004), Astrouni's work is of paramount importance as it consists of the key principles which recent MT systems heavily depend on.

A British cryptographer Warren Weaver and Andrew D. Booth discussed the possibility of using computers in translation. Booth collaborated with Richard H. Richens of Cambridge in the field of MT. In 1949, Weaver issued a memorandum which highlighted the idea of MT and suggested methods and talked about prospects.

A few years later, the USA enrolled in the research through its reputable universities (Washington University in Seattle, University of California at Los Angeles, and Massachusetts Institute of Technology (MIT) among others). The first MT researcher was appointed at the MIT in 1951 and a year later it hosted a conference on MT, attended by 18 individuals interested in the field. At that time, it was obvious that a full MT system is far-fetched without rigorous research. In addition, it was expected that human intervention is indispensable by preparing texts to be translated (pre-edit) or revising output (post-edit). Georgetown University witnessed the first MT system in January 1954 which was the outcome of the collaboration of Leon Dostert with IBM. The system could translate 49 Russian sentences into English using 250 words and 6 grammar rules. Although simple and scientifically insignificant, their work was of great importance in stimulating more funds in MT research.

#### **2.4.2. The Decade of High Expectations and Disillusions 1956-1966**

It should be highlighted that computers at that time were inadequate for accomplishing the task of MT. Therefore, researchers focused on technical aspects by improving hardware components of the computer while other researchers grouped their efforts in developing suitable language processing tools. Other researchers kept on studying the theoretical issues especially in Europe and the Soviet Union. Most US research was for Russian-English translation due to political and military reasons.

Erwin Reifler (University of Washington) followed the word for word approach by constructing a large bilingual dictionary where it was not only used to look up equivalents, but also for solving grammatical problems without using syntactic analysis. Gilbert King at the IBM Corporation led to more practical outcome in 1958. He developed a system which was installed for the US Air Force which produced translation, usually crude and barely intelligible, for many years until replaced by Systran in 1970.

RAND adopted an empirical approach through the years 1950-1960. It neglected any linguistic theory and emphasized the statistical analysis approach instead. A computer program was written using bilingual glossaries and simple grammar rules. The outcome was looked at by post editors who in return indicated errors. Consequently, the glossaries and the grammar rules were revised, and the text would be translated again until satisfactory results were achieved. It is worth noting that at RAND, David Hays developed the first parser based on dependency grammar.

A similar approach was adopted at Georgetown University which was carried by Leon Dostert. In fact, The Georgetown group was the largest in the US

with different points of view by the different groups. There were mainly four groups; the first group adopted Ariadne Lukjanow's method which in return exhibited excellent results. The second group had Paul Garvin who utilized the syntactic analysis method and prepared the linguistic basis for the demonstration system in 1954. The third group developed the 'sentence by sentence' method by Anthony Brown who followed the empirical cyclical method. The fourth group which was headed by Michael Zarechnak used the 'general analysis' method which was adopted and named Georgetown Automatic Translation (GAT). The main feature of Zarechnak approach was the three levels of analysis: morphological, syntagmatic, i.e. agreement of nouns and adjectives, government of verbs, modification of adjectives, etc. and syntactic. GAT was implemented on the SERNA system which was mainly designed by Peter Toma who later on designed Systran.

At the MIT, research was going on hand in hand with other researches in the US. It was initiated by Bar-Hillel in 1951. Victor Yngve took over from 1953 until 1965, the year which witnessed the end of the project. Yngve believed in syntax being the core of MT as he was influenced by Chomsky and the transformational theory. In fact, Chomsky took part in the project for two full years. However, the syntactic transfer approach started to show its deficiencies. The project came across the semantic barrier and Yngve acknowledged in the mid 1960s that for MT to succeed, it should 'understand' what translating is.

In the University of Texas, the Linguistic Research Center (LRC) was established by Winfried Lehmann in 1958. Not different from MIT, they followed the syntactic transfer approach. Their collaboration resulted in developing the METAL system.

It was surprisingly evident that no American group followed the interlingual approach. However, it was adopted at Cambridge Language Research in England as well as at Milan by Silvio Ceccato. Not very much different from the intense research in the US, research was vigorous in the Soviet Union where both empirical and theoretical approaches were followed. In mid 1960's different groups from different parts of the world were established including most European countries, Japan, Mexico, and China.

Explicitly, MT research showed high expectations in the 1950's. It was predicted that many inventions were imminent and the world would witness a fully automated systems. Unfortunately this was not the case. Bar-Hillel (1960) criticized the assumption that the goal is to create a fully automated system which can produce excellent quality translations. He stated that MT should adopt less ambitious goals. Large grants from government, military and private sources had been awarded. In the USA, \$12 to \$20M was spent on MT research. As a result, the famous committee ALPAC (Automatic Language Processing Advisory Committee) was formed by the government sponsors of MT in the US to figure out if its money had been well spent.

In its negative report in 1966, it stated that MT was slower, less accurate, and twice as expensive as human translation and there is no immediate or predictable prospect of useful MT. The report's effect was immense on MT with very damaging consequences; it brought research in the US to a grinding halt for more than a decade and MT was recognized as a complete failure. Somers (2003) states that the conclusions of the ALPAC report should not have been surprising. The translation process by computers was hampered by primitive technology and the under-estimation of the difficulty by the researchers who are mostly mathematicians and electrical engineers, rather than linguists.

Indeed, theoretical (formal) linguistics was in its “infancy” at that time: Chomsky’s revolutionary ideas were only just gaining widespread acceptance (ibid: 5). The two reasons behind the failure of MT at that time shed light on the necessity of the inclusion of linguists in the team of MT projects as long as computational linguistics is considered a branch of linguistics. Therefore, linguists, programmers and mathematicians among others should work hand in hand in order to achieve desirable results in the field of MT. Bennett (2003: 157) highlights the importance of linguistics in MT by stating that “Linguistics has not solved the problems of MT, but it can help the researcher to reach solutions, by offering a range of observations, techniques and theories that may be adopted and extended within the MT enterprise.”

### **2.4.3. The Quiet Decade**

As its name suggests, one can expect that researchers ceased to work on MT. This was true in the US as activities were not significant. However, the situation was different in Canada and Europe. In Canada, the bilingual policy led to the launch of an important research group at the University of Montreal. In Europe great demand on translating scientific, legal, and administrative documents was present then.

Research began in the 1970 at Montreal, Canada where the *TAUM* project team was able to design a system called *METEO* which is considered one of the most important achievements in the history of MT. In this period, research focused on Interlingua approach in which two main groups were actively working on this project at that time. The first group established by Bernard Vauquois at Grenoble University developed an Interlingua system called *CETA* for translating Russian mathematics and physics texts into French. The second group was in the University

of Texas which was mentioned earlier with its *METAL* system. However, both *CETA* and *METAL* faced many problems due to the rigidity of the analysis and inefficient parsers.

#### **2.4.4. MT Research in the 1980s**

Since the failure of the Interlingua approach exhibited by *CETA* and *METAL*, researchers started to look for an alternative. Consequently, the Grenoble group known as *GETA* developed their renowned *Ariane* system which is considered the most advanced linguistic-based transfer system. A similar project was carried by Makoto Nagao who developed *the Mu* system at the University of Kyoto in Japan. His system was operational for use by the Japanese Information Center for Science and Technology. *SUSY* system was the product of research at Saarbrücken. What makes *SUSY* different is its in-depth treatment of inflected languages such as Russian and German among other languages which were under investigation.

The most prominent project of that period is the *Eurotra* project of the European Community. The project was well known as the largest and most ambitious in the world. The goal of the project was to build a multilingual transfer system for translation among the Community languages.

MT research was not exclusive to North America, Europe and Japan. Research was vigorous in Korea, Taiwan and China. The Soviet Union showed an increased interest in MT after the quiet decade and the *ALPAC* report. One significant feature of the period was the urgency on research on NLP within the mainframe of artificial intelligence (AI). Hutchins (1994) justifies the need of AI in research on MT by stating that “since translation is concerned primarily with conveying content, any MT system must be capable of 'understanding' the meaning

of texts by reference to extra-linguistic knowledge.” The most important project which draws on AI and NLP was carried at Carnegie-Mellon University in Pittsburgh which examines the construction of knowledge-based MT systems.

Many researchers, however, believe that AI is not applicable in MT. Their prophecy lies behind the fact that understanding the text requires some linguistic analysis especially if the domain is not of a technical or scientific nature. The semantic general lexicon, theme-rhyme structures, nominalization, tenses, etc. play a vital role in determining the meaning of a text. In addition, experience with AI approaches to parsing and disambiguation, which seek models with deep understanding, has shown that hand-coded syntactic constraints and preference rules are time consuming to build, do not scale up well and are fragile in the face of the extensive use of metaphor in language. (ibid: 1994)

In the 1990's up to the moment, vigorous research has been conducted on MT and Computer-Aided (or -Assisted) Translation which is widely known as CAT. CAT tools would range from multilingual word processing, spell checkers, synonym list (thesauri), etc. to machine aided translation systems which could do a draft to be post-edited by the translator. Such tools are used by professional translators and lay people alike, the latter group for the purpose of translating emails and World Wide Web pages. MT researchers, on the other hand, have continued to set themselves high ambitions. The main difference between MT and CAT is that the former is designed to replace the translator while the latter is there to support and aid him or her through automating terminology lookup activities and re-using previously translated texts.

## **2.5. Difficulties Facing Traditional Approaches in MT**

MT is considered one of the most important applications in NLP. Unfortunately, it is a tough predicament. The important question remains: Why is MT difficult? Arnold (2003) answers this question by attributing the difficulties to two factors: the first is on the nature of the translation while the second part pertains to the ability of the computers. He believes that translation is not only difficult for computers, but is also difficult for humans.

Arnold (2003:119) believes that “[T]ranslating is a many-faceted skill that goes well beyond mere competence in two languages”. In particular, the task of the translator is to take a text in one language SL and produce a text in another language TL which is in some sense equivalent. In addition to the slippery notion of ‘equivalence’ which is difficult to pin down, and can vary greatly depending on the kind of text involved, this is not the sole objective of the translator whose work should be clear, unambiguous, interesting, persuasive, elegant, poetic, gripping, etc., depending on the text being translated. While this could be achievable, it is somehow unfair, especially when such process is to be automated. One should not neglect the fact that translation is a creative task where translators are expected to act as cultural mediators who convey to the readers of the TL what may be obvious to the readers of the SL. The previous facts about the process of translation amount up to the arduous task of the translator and machine alike.

Computers, on the other hand, are just devices that follow rule, albeit with considerable speed and precision. Rule following can rarely produce a creative work especially the kind of creativity required for the process of translation. Arnold (2003:121) summarizes the limitations of computers in the field of MT, namely the inability of computers to:

1. Perform vaguely specified tasks.
2. Learn things (as opposed to being told them).
3. Perform common-sense reasoning.
4. Deal with some problems where there are a large number of potential solutions.

The necessity of finding precisely formulated rules is a cornerstone in computer programming. Much of the difficulty in MT arises from the ability to find straight forward formulations. Learning is also a problematic area from a computational perspective; “what a computer needs to know, it must be told, in the form of explicit rules, written by humans” (ibid: 121). Coding Common-sense reasoning involves huge amount of data about the world to be computed which is a strenuous task. The fourth problem is connected with time required to achieve the task depending on the probabilities such task has. This issue is not significant nowadays as computers have improved their hardware and speed tremendously.

One can find different inexpensive or free services that call themselves ‘translation programs’ which produce low-quality translations. This reality contradicts the goal of many NLP researches who aim at producing error-free texts that read fluently in the target language (TL). To be more systematic, we should examine the different approaches mentioned above elucidating each challenge MT researchers are faced with.

The first and simplest approach as we mentioned above is the direct approach which translates ‘word-for-word’. Linguists as well as translators can easily predict the challenge. There is no one-to-one correspondence between words in different languages. ‘Cake’ in English is ‘قالب حلوى’ in Arabic. This example shows that the equivalence is one-to-two. Another problem is the lexical ambiguity. ‘Suit’ could have different senses depending on its context where it

could be translated as a ‘lawsuit’ or ‘set of garments’. In order to render the right sense, one needs to examine the context larger than individual words.

Another challenge for the direct approach is that languages have different word orders. English favors the SVO sentential word order while Arabic prefers the VSO. A simple word-for-word translation will get the word order in the TL wrong.

The latter problem is addressed by the transfer approach which draws on parsing the ST, and then it transfers the parse tree of ST into a syntactic tree in the TT using proper rules. Finally it generates the translation from this syntactic tree. Still, we are faced with syntactic ambiguity. The example of ‘اطال النوم’ could be translated as ‘he prolonged sleep’ instead of ‘he slept for a long time’. It is evident here that syntactically correct translation often has inappropriate semantics. Lederer (2003) contributed to the grammatical ambiguity by giving the example of the sentence “gas pump prices rose last time oil stocks fell.” Each word in this sentence can be either a noun or a verb. She adds that “one can imagine the number of instructions which must be given to the machine so that it can disambiguate each word and place it in its correct grammatical category.”

At last, the Interlingua approach (which does not rely on literal translations) has the advantage of translating into a large number of languages. However, the disadvantage of such approach is the complex process of designing efficient and comprehensive knowledge representation formalisms. This is due to the amount of ambiguity that has to be resolved to translate from a natural language to a knowledge representation language (Manning & Shutze, 2002).

The need for a new approach has surfaced after many years of research and development on the previous approaches without achieving the ultimate goal of

MT. Statistical methods reemerge<sup>55d</sup> as a result. It was first suggested by Warren Weaver in 1949. However, efforts in that direction were soon abandoned due to the limitation of computational capabilities. The recent availability of bilingual, machine-readable texts has stimulated interest in methods for extracting linguistically valuable information from such texts. The initial proposal of using statistical approach in MT was by Brown et al., (1993) who suggested a statistical approach to MT from English into French.

## **2.6. Statistical Approach to MT**

Due to the above mentioned difficulties, linguists started to think of empiricist approaches rather than rationalist ones. Most linguistics and especially natural language processing were completely dominated by a rationalist approach. An empiricist approach to NLP suggests that “we can learn the complicated and extensive structure of language by specifying an appropriate general language model, and then inducing the values of parameters by applying statistical, pattern recognition, and machine learning methods to large amount of language use” (Manning & Shutze, 2002:5). In this respect, Arnold (2003:138) sees that “translation is a very complex process involving many factors, the appeal of some kind of statistical methodology should be clear.”

Unlike traditional approaches which draw on observing a large amount of language use situated within its context in the world, statistical NLP uses texts, and regards the textual context. A body of text is called corpus (a Latin word which means body) and corpora for a multiple collections of texts.

Other researchers distinguish between statistical and traditional approaches by referring to rule-based MT systems versus data-driven systems (Bennet & Gerber, 2003). The former systems are created manually by linguists who write

lexicon, grammar and translation rules while the latter systems learn patterns automatically by the computer from texts. Some of the data-driven systems learn wholly automatically, and other systems are hybrids, adopting some mixture of manual development with automatic learning.

Statistical NLP approach to syntactic and semantic analysis is different as well. It concentrates on the lower levels of grammatical processing. Semantic analysis resides in the distribution of contexts over which words and utterances are used. Statistical NLP approach seeks to solve syntactic and semantic disambiguation by automatically learning lexical and structural preferences from corpora. Manning and Shutze (2002:19) believe that

statistical models offer a good solution to the ambiguity problem: statistical models are robust, generalize well, and behave gracefully in the presence of errors and new data. Thus, Statistical NLP methods have led the way in providing successful disambiguation in large scale systems using naturally occurring texts.

Therefore, linguists started to see a prospect in statistical methods which provide practical solutions to real problems. This advancement is simply the reason that prevented traditional methods of achieving any major success.

## **2.7. Probability Theory**

The Statistical approach is based on the probabilistic theory which deals with predicting how likely it is that something will happen. The current research will not delve into such theory in details as it believes that it is related to another discipline (mathematics). Brown et al., (1993) explicates the statistical approach and the probability theory work together in MT designed to translate French into English. Since the current research is on a different language pair i.e. English and Arabic, a modification to their explication is sought for (the example is on an attempt of the statistical MT to translate an Arabic sentence into English): When

presented with an Arabic sentence  $a$ , we imagine that the writer had in mind an English sentence  $e$ , but that  $e$  was somehow distorted in translation so that it came out as  $a$ . The task of MT is just to produce  $e$  when presented with  $a$ . Translation is seen as an instance of transmission down a *noisy channel* (like a telephone line), and there should be a way to extract the original input (the English sentence  $e$ ). It is obvious then that  $a$  is likely to occur depending on which sentence the writer had in mind. The probability theory is in need at this stage as we want the sentences that give  $a$  the highest probability among the English sentences. In examining the following example, 1(b) has a higher probability and it is of course the right choice.

- 4 a. كم الساعة  
b. *what time is it?*  
c. *what hour is it?*

In this respect, Manning and Shutze (2002:486) states that the SMT system in reference to the above explanation should consist of:

1. Language model: the language model assigns a probability to  $a$  for each sentence of English.
2. Translation model: which assigns a probability to every sentence in English.
3. Decoder: in this step, the best candidate for  $e$  is sought for according to the above model.
4. Translation probabilities: such probabilities are estimated using EM (Expectation Maximization) algorithm. The basic idea of the mentioned algorithm is that it solves the credit assignment problem. i.e., if a word in the source is strongly aligned with a word in the target, then it is not available anymore to be aligned with other words in the target.

Despite the fact that the statistical approach is intuitively appealing, it is still prone to significant problems. Arnold (2003) attributes the deficiencies of SMT to two reasons. The first pertains to a different version of the problem of description while the second relates to the quality of the available statistical models. He states that “[T]he statistical version of the problem of description is the problem of sparse data” (ibid: 140). In reference to example 4, in order to know that *b* is more probable than *c*, huge amounts of text must be analyzed to discover that *time* appears more often than *hour* in this context. The problem is even worse when examining the aligned parallel corpora (collections of texts in two languages which are supposed to be translations of each other). Arnold (2003:139) adds “whatever the probability of seeing an expression on its own, the probability of seeing it as the translation of some other expression must generally be lower”.

In her official webpage, Suliaiti, an independent researcher and an Arabic corpus consultant, lists the different Arabic corpora available. There are about 20 corpora listed, five of which are for the purpose of MT and NLP. Other corpora, which are classified on source, medium, size, purpose and material, are used for different purposes such as lexicography, pedagogy, speech recognition, etc. The list shows the scarcity of Arabic corpora in general and parallel corpora in particular as more research should be targeted at such essential field which aids many practical and applied linguistics applications.

The second hurdle to SMT Arnold presents is the statistical models (mentioned above by Manning and Shutze) utilized in the system. The standard example of a monolingual statistical model is bigram model which is used as well in speech recognition. Such a model is based on the assumption that the probability of any given word sequence can be figured out by the joint probability of each word occurring by giving the preceding word. Arnold illustrates the process as follows:

The probability associated with '*The cat died*' is the joint probability of '*The*' occurring as the first word in a sentence, of '*cat*' occurring given that the preceding word was '*The*', and the probability of '*died*' occurring given that the preceding word was '*cat*'. The basic data for such a model is thus observations about the relative frequency of various pairs of words (bigrams), (ibid: 139).

From the above illustration, one can easily conclude that such models could be applied to certain languages which syntax and word order play an important role in its syntax while other languages such as Arabic are freer in word order and such models are doomed to failure.

In the translation model, a generalization might assume that the probability of *a* appearing as the translation of word *e* depends on the predecessors of *a* and *e*. It is obvious however that there are shortcomings to the bigram model as the probability declines sharply in an example such as *The big fat black cat* compared to *the big cat* where *the* exerts the same kind of effect while the probability of the sequence *big fat and black* is less.

## **2.8. Alignment**

In addition to drawing on the probability theory, statistical MT developers consider alignment as the backbone of their domain. Alignment can be divided according to the size of the unit of language. It starts with the text as the largest unit until it goes down to the word.

Although it is not part of the translation process *per se*, text alignment is used to create lexical resources such as bilingual dictionaries and parallel grammars, which then improve the quality of MT. It is evident that the availability of parallel texts or bitexts which include the same content in several languages is of paramount importance. Such kind of texts can be found in parliamentary proceedings and official documents in countries with multiple official languages. Another source could be found in magazines and newspapers published in more

than one language. It is axiomatic to conclude that text alignment is an obligatory first step for making use of multilingual corpora.

After selecting the appropriate text alignment method, the task gets more intricate by knowing which paragraphs or sentences correspond to their counterparts in the other language (sentence and paragraph alignment). The next problem lies in learning which word tends to be translated by which other word. Manning and Shutze (2002:469) clarify the difficulty by stating that “translators do not always translate one sentence in the input into one sentence in the output.... Human translators change and rearrange material so the output text will flow well in the TL, even when they are translating material from quite technical domains”.

The example below (ibid: 469) shows extracts from a document in English and its French version. The middle and right columns show the French and English versions with arrows connecting parts that can be viewed as translations of each other. The italicized text in the left-handed column is a literal translation of the French text. The two sentences show a variation in word order and content. Moreover, large pieces of content can disappear. For example, the final English words *achieved above-average growth rates*. These words were implied in the French version.

*with regard to  
the mineral water  
and the lemonades  
soft drinks*

Quant aux a  
Les eaux  
Minerals et eux  
Les limonades

According to

{Our survey,} 1988

*they encounter  
still more  
users*

Elles rencontrent  
Toujours plus  
D'adeptes.

{Sales} of

*indeed  
our servey*

En effet  
{ Notre sondage }

{Mineral water and soft  
drinks} were

*makes stand out  
the sales*

Fait ressortir  
{ Des ventes }

{Much higher}

*clearly superior*

{ Nettement  
Superieures }

{Than in 1978,}

Reflecting

*to those in 1987*

{ A celles de 1978, }

{The growing popularity}  
of these products

*for cola-based  
drinks*

Pour { les boisson a  
base de cola }

{Cola drinks}  
manufacturers

*especially*

{ Notamment }

{In particular}  
Achieved above  
average growth rates

Figure 2

An Arabic-English alignment is even worse. Lee (2004) presents a novel technique to enhance the quality of SMT between English and Arabic through morphological analysis. This analysis is done for the reason that the two mentioned languages have different morphological structures. She states that “so called a word (separated by a white space) in Arabic often corresponds to more than one independent word in English, posing a technical problem to the source channel models. In the English-Arabic sentence alignment shown in Figure 3, Arabic word *AlAHmr* (written in Buckwalter transliteration.<sup>1</sup>) is aligned to two English words ‘the red’, and *llmEARdp* to three English words ‘of the opposition’. The research opts for transliteration in the following figure to show how alignment is executed on morphemes in Arabic. Figure 3 shows the two aligned sentences in English and Arabic before the morphological processing.

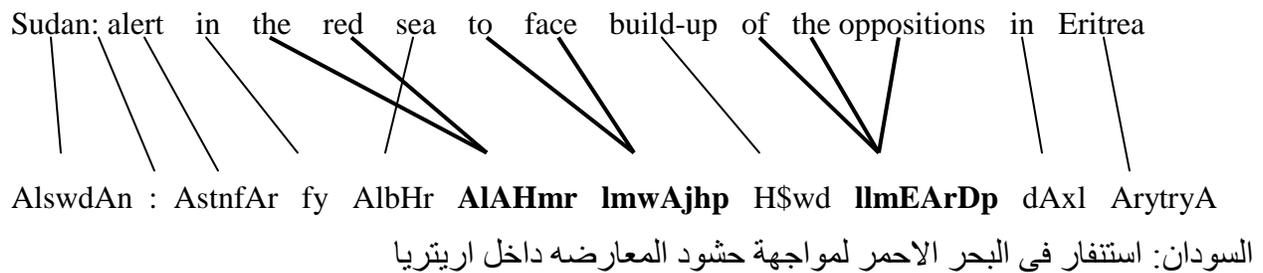


Figure 3

Technically, there are many methods to align sentences from parallel corpora. Each method has its distinctive features. The following are the different alignment methods suggested by Manning and Shutze (2002):

**2.8.1.1. Length-Based methods:** As its name suggests, the method aligns sentences depending on the length of the sentence in the parallel corpora. The salient feature of such method is its efficiency in rapid alignment of large

<sup>1</sup> A table of bukwalter transliteration is furnished in Appendix 2.

quantities of text. Gale and Church, Brown, Wu among others approached the above-mentioned method.

**2.8.1.2.** Offset alignment by signal processing technique: These techniques do not opt for aligning a sentence but rather position offsets in the two parallel texts. Church, Fung and McKeown are the main researchers who adopt such technique.

**2.8.1.3.** Lexical methods of sentence alignment: This method is much different from the previous two as it is more robust since such methods use lexical information to guide the alignment process. Kay and Roscheisen, Chen and Haruno and Yamazaki among others who opt for the lexical methods.

It should be noted that there are many methods to align sentences; however, it chiefly depends on the type of the parallel corpora being in use and the two pairs being under translation.

The last step in the alignment process is targeted at the smallest meaningful unit of language, i.e. word alignment. One of the most important uses of word alignment is the creation of bilingual dictionaries and terminology database. This is done mainly by following two steps. The first is to extend the text alignment into a word alignment. The second step is to use a criterion to select aligned pairs taking into consideration that there is enough evidence to include them in the bilingual dictionary.

It is worth noting that the statistical approach is not only beneficial to MT, but it could also benefit lexicographers in constructing bilingual dictionaries since it depends on matching words from authentic contexts and their translations. One of the drawbacks of bilingual dictionaries is the inclusion of phrasal verbs. Such downside could be avoided using statistical methods.

The current research has chosen *Google Translate* as the MT service to be assessed. It is worthwhile discovering some of the main features of this tool.

## 2.9. Google Translate.

*Google Translate* is a free online MT service that enables users to translate sentences, documents, and even websites instantly. It functions when a user types or pastes the material in the SL and *Google Translate* detects the ST and translates it into English as a default language. However, *Google Translate* has its drawbacks just as any other MT. Though it helps the reader to comprehend the content of ST, it does not usually produce accurate results. The accuracy depends on the language pair being used. For example, the accuracy of *Google Translate* increases substantially when translating a text from English into a European Language. Results of analyses were reported in 2010, showing that French into English translation is very accurate (Shen, 2010).

Another condition which could affect *Google Translate's* accuracy is the length of the sentence or paragraph of the ST. The future is just as bright as the present for online translation systems as Hutchins (2004:171) anticipates that “[U]sers of online translation systems (whether charged or free) will expect continued improvements, and this will be more likely with specialized services than with non-specialized ones.”

According to Wikipedia, *Google Translate* has gone through different stages of improvement (27 stages). In the first stage, the tool could translate the following pairs: English into German, English into Spanish, French into English, Spanish into English, and German into English. Arabic into English and English into Arabic was launched in the 5th stage in April 2006. It is worth noting that Google Inc. has recognized since an early stage that Arabic is one of the prominent languages which is targeted by millions of users around the globe. Arabic was introduced before many important languages in *Google Translate*.

*Google Translate* does not apply grammatical rules, since its algorithms are based on statistical analysis rather than traditional rule-based analysis. Indeed, the system's original creator, Franz Josef Och, the head of Google's Machine Translation group, has criticized the effectiveness of rule-based algorithms in favor of empirical approaches (Och, 2009).

### **2.9.1. Google's Approach**

Although Google won an international competition for English-Arabic and English-Chinese Machine Translation, translated texts can often include apparently absurd and obvious errors. This is due to the fact that *Google Translate* applies a Statistical Machine Translation approach (SMT) rather than a traditional approach. Its approach is summarized in their website as:

When *Google Translate* generates a translation, it looks for patterns in hundreds of millions of documents to help decide on the best translation for you. By detecting patterns in documents that have already been translated by human translators, *Google Translate* can make intelligent guesses as to what an appropriate translation should be. This process of seeking patterns in large amounts of text is called "Statistical Machine Translation". Since the translations are generated by machines, not all translations will be perfect. The more human-translated documents that *Google Translate* can analyze in a specific language, the better the translation quality will be. This is why translation accuracy will sometimes vary across languages.

*Google Translate* has been chosen in this research because I find myself in agreement with Aiken and others in many aspects. First, it is used more frequently, provides more language-pair combinations, and is probably more accurate overall (Aiken & Ghosh, 2009; Och, 2009). In addition, one study (Aiken, et al., 2009) compared four systems and found that *Google Translate* was best, followed by *Yahoo*, [X10](#), and *Applied Language*. Finally, an NIST comparison of 22 MT systems in 2005 (many not free or Web-based) found that *Google Translate* was

often first and never lower than third in the rankings using texts translated from Arabic into English and from Chinese into English.

## Chapter Three

### Legal Discourse Characteristics

#### 3.1. Preliminaries

In this chapter, the sub-language i.e. the legal discourse chosen in the assessment will be discussed. It is worth noting here that the current research does not opt for the legal language to be assessed haphazardly. In fact, the researcher favors to assess *Google Translate* on legal discourse for some valid facts. The foremost reason behind the selection is due to the formulaic nature, fixed terminology and the clear message such language usually possesses. Hatim (1997:38) states that “If Machine Translation is to be at all sustainable as a commercial enterprise, it will be in the area of restricted registers such as ‘initializing’ and ‘concluding’ legal documents.”

Therefore, it is supposed that MT researchers will easily avoid the different aforementioned ambiguities in automating the translation of legal texts. This fact is, however, considered to be over-tentative since it is based on theoretical reasoning rather than on solid facts and rigorous assessments.

The second reason for selecting legal language to be assessed is that up to the researcher’s knowledge, different assessments have been done on different MT systems where different sublanguages have been targeted (scientific, journalistic, weather reports, etc.) while no assessments have been executed on legal discourse in MT.

Finally, the demand for legal translation is on the increase around the world owing to globalization and the increased contact and exchange between peoples and states (Cao, 2007). The increasing international traffic of persons and goods as

well as the growing importance of international governmental and non-governmental institutions flourishes the business of translating legal texts.

To this end, it is worthwhile to present the characteristics of the legal discourse, its layout, and the different approaches to translating it as it could serve as an important theoretical background for MT researchers who are aiming at developing current systems as well as inventing new ones.

### **3.2. Characteristics of Legal Language**

Research into the English legal register has focused on syntactic and lexical characteristics. Findings show that structures are formulaic, and they abound with archaisms which often defy rules of modern writing. In addition, sentences are identified by being long (50 words in average), which significantly contributes to an impersonal style with typical legal vocabulary. Thus, Danet (1985:297-280) proffers the following features as characteristics of legal register: technical terms, common terms with uncommon meanings, archaic expressions, doublets, formal items, unusual prepositional phrases and a high frequency of *any*.

Syntactic complexity accounts for many hurdles lay persons are surmounted with in comprehending legal English. Gustafsson (1975) reports an average of 2.86 clauses per sentence, and nominalization is considered a prominent use in the legal register (Crystal & Davy, 1969). A high frequency of passive constructions and a high incidence of prepositional phrases are also a main feature in legal discourse which further complicates the legal language (Charrow & Charrow, 1979).

In addition to lexicon and syntax, Cao (2007) adds two more features which differentiate legal language from any other language, i.e. pragmatics and style of such language. As to pragmatics, law depends upon the performative function of language, i.e. legal utterances perform acts, creating facts, rights and institutions.

Hart (quoted in Cao, 2007) argues that words in legal language differ in meaning, import and effect depending on who utters them, where and when.

Style on the other hand refers to “the linguistic aspects of the written legal language and also the way in which legal problems are approached, managed and solved” (Smith 1995:190). Legal writing is characterized by an impersonal style, with the extensive use of declarative sentences pronouncing rights and obligations.

Mattila (2007:65-96) tackles the specificity of the legal language from a different perspective. He mentions eight characteristics of legal language in more details by devoting a full chapter in his book. It is worth noting here that Mattila discusses the universal characteristics which show themselves when under examination at the textual level. Semantic, lexical, syntactic and stylistic elements of language are under scrutiny. Since the evaluation of the current research is on the lexical and syntactic levels, the following characteristics can be of great theoretical value.

**3.2.1. Precision:** Accuracy and precision are considered essential characteristics of legal language. This essentially results from the requirement for legal protection and legal certainty where legal rules should avoid ambiguity in order to avoid the possibility of arbitrariness. As a result of such accuracy, legal language utilizes tautology as a rhetoric device to achieve its goal in transmitting legal messages with absolute clarity and without ambiguity. The inclusion of definitions of legal terms within a particular context is another feature of legal language to avoid the multiple meanings of different words (polysemy)

**3.2.2. Information (over)Load:** On the one hand, legal language should be as concise as possible to avoid laws and regulations that would be over-long

and unclear. On the other hand, legal language should avoid over-abstraction, in that way enabling decoding with minimum effort.

### **3.2.3. Universality and Aloofness**

**3.2.3.1. Abstraction and Hypothetical Character:** Modern law has an abstract character as it regulates entities that are mere mental creations: rights and duties. Law is based on experience drawn from the real world. However, it regulates hypothetical future cases which entails the use of conditionals and the word *if* is highly frequent, while the present tense dominates.

**3.2.3.2. Impersonality and Objectivity:** The use of the passive is a feature of legal language for the purpose of bringing the object of the action into the foreground while giving the actor a secondary role. Objectivisation is also apparent in drafting laws and particularly in the language of advocates. An advocate may write: “It appears that Article 27 of the law ... should be interpreted....” Instead of saying “it seems to me that Article 27 .....” (ibid:77)

**3.2.3.3. Neutrality:** The style of the legal language should be as neutral as possible due to the fact that the principal intention is to have an effect on the understanding rather than the feeling of the reader or listener.

**3.2.3.4. Metaphors:** In modern legal language, metaphors are rare; nevertheless, there are some exceptions to this fact. Their existence is attributed to the fact that some legal terms originate from metaphors. An example of metaphor in the legal language is “burden of proof, lion’s share, etc.)

**3.2.4. Systemic Character:** This entails that each element of the order forms part of a greater whole. An article forms part of a law, and a law forms part of legislation. The systemic character of the legal order appears clearly when

components are linked to each other by references which involve intertextuality.

**3.2.5. Structure and Formalism in Legal Texts:** The structure of legal texts is carefully elaborated. Amongst other things, logical disposition of these texts helps to place legal information in a hierarchy. A legal text moves from the abstract to the concrete, from the substantive to the procedural. The structure of the text should be consistent: the primary items are presented before secondary items, and general rules before special conditions and exceptions.

**3.2.6. Frequency of Initializations and Acronyms:** Legal language uses many initializations and abbreviations. Their use benefits the text in producing a shorter sentences which appear more easily to the reader, yet the text could appear more difficult to understand if the reader is unfamiliar with the initialization in use.

**3.2.7. Sentence Complexity and Diversity of Language Elements:** It is widely known that sentences in the legal language are traditionally very long and complicated. In recent years, though, language specialists have been seeking to improve the quality of legal language by shortening long sentences. Complicated expressions are added to legal text. For example, *at slow speed, in the event that, prior to, subsequent* are all used instead of *slowly, if, before, after* respectively. Legal language places less emphasis on verbs than ordinary language. This is because nouns are known to carry a more objective impression than a verb.

**3.2.8. Archaism and Solemnity:** As its name suggests, legal language is full of archaic terms and seriousness to elevate and dignify the language of such texts.

In reference to the complexity and rigidity of the legal language, Mattila (2007:35-36) states that

In legal language, it often occurs that communication fails because the message is hermetic, or closed. A legal message is sometimes formulated in such a complex way that a lay individual can hardly understand it. For example, this is often the case with laws and regulations, and with judicial decisions..... interpreting the legal text is a highly complex matter, notably because distilling meaning does not depend solely on linguistic arguments.

Observing the above-mentioned characteristics, one can anticipate the paramount importance of such theoretical data in devising machine systems specified for translating the legal language as well as improving current MT systems. Such theory should be indispensable to MT developers who aim at developing their outcomes.

### **3.3. Types of Legal Writing**

It is worth noting here that such characteristics of legal English derive from the language use in legal corpus in which the level of formality can be characterized as frozen or formal. Such corpus comprises of legislative language, administrative and testament language, jury instructions, and documents such as endowment-assurance policies, hire-purchase agreements, and insurance policies. Hiltunen (1990) distinguishes three types of legal writing:

**3.3.1.** Academic texts which consist of academic research journals and legal textbooks

**3.3.2.** Juridical texts covering court judgments or law reports

**3.3.3.** Legislative or statutory writings consisting of Acts of Parliament, contracts, treaties, etc.

### **3.4. Contracts as a Sub-Legal Text.**

These different sub-text types have their own peculiarities. It follows that translating each text type should be approached differently. The current research data refer to the last category where contracts will be the focal type of the legal literature. Since this study examines contracts in particular, they are worth discussing thoroughly. Sarcevic (2000:133-134) defines contracts as

agreements between two or more parties to exchange performances in a given situation for a specific purpose. The legal actions to be performed or not performed are set forth in the substantive provisions in the form of obligations, permissions, authorizations and prohibitions, all of which are enforceable by law.

Alcaraz and Hughes (2002:126-127) go further by determining the five peculiarities of contracts which distinguish them from other legal documents.

- 3.4.1.** There must be an agreement between two parties, who may be individuals or groups, nonprofessionals or juristic experts.
- 3.4.2.** There must be valuable consideration given and received by each party. In other words, each party promises to give something in exchange for the other party's promise to give something else in return.
- 3.4.3.** The parties must intend their promises to be acted on and to be legally binding. Insignificant or vague actions are not constructible as contracts, nor are promises to undertake the impossible.
- 3.4.4.** The subject matter of the contract must not be illegal or "tainted with illegality"; so-called "contract killings" are not contracts in law.
- 3.4.5.** The contract must be freely entered into by both parties and both should be of equal bargaining power.

Contracts have been opted for in this research for the fact that "the form and style of contracts is fairly standardized, making them easy to translate once one is

familiar with that form and style” (Beyer & Conradsen, 1995:160). They postulate that what makes legal translation unproblematic is the inclusion of standard “boiler plate” clauses that can be easily translated by reference to previous or standard translations.

According to Wikipedia, the term “boiler plate language” describes the parts of a contract that are considered standard. Such language exists in pre-created documents such as pre-written complaint about a product opposed by groups that create such letters requiring the person requesting the action only to type or sign his or her name at the end.

Such formulaic nature of contracts in particular legitimizes their translations using machine (specifically designed for such purpose) as little ambiguity could be faced in rendering the different parts of the contract which are universally familiar in form and style.

### **3.5. The Difficulty of Translating Legal Texts**

It is commonly acknowledged that legal translation is complex, and it requires special skills, knowledge and experience on the part of the translator to produce such translation. This is due to the fact that legal translation can and often does produce not just linguistic but also legal impact and consequence, and because of the special nature of law and legal language (Cao, 2007). To Beyer and Conradsen (1995:175) “the greatest challenge facing the legal translator is to duplicate the meaning faithfully without changing it in any way.... which requires both linguistic skills and some basic understanding of legal concepts”.

Harvey (2002:177) considers legal translation as a category in its own right, and as “the ultimate linguistic challenge”. Translation in general is considered to be a creative task which only inspired translators would excel in their careers. In

contrast, legal translation rarely requires such kind of creativity as Durieux (1991) maintains that translators can and should be creative in technical texts, but is more cautious when it comes to legal texts, acknowledging that they are subject to special restrictions.

Danet (1985:11) postulates that creativity manifests itself differently in various areas of translation, making creativity a domain-specific attribute. However, he admonishes the legal translator that as “even slight changes in language may affect the substance, translators must always take account of legal factors when making linguistic decisions”. Nida observes that “sociolinguistic constraints that significantly interfere with translators’ creativity occur primarily in religious, legal, and classical texts” (Nida, 1998:127).

Farghal and Shunnaq (1992:203) state that “the translation of technical materials requires special expertise in the relevant discipline on the top of the general language competence”. In addition, legal translation is considered one of the most challenging endeavors not only for lay people who are willing to translate legal texts from one language into another, but also translators who find it one of the most challenging tasks in the career field. Thus, when encountered with a legal document, lay people and non-professional translators attempt to use different aids to help them. Legal dictionaries, if found, are not as efficient as ordinary dictionaries due to their terminological incongruency (Sarcevic 1989).

Language is not the only aspect which makes the legal language different, legal archaisms, new terminology for new areas of law, and technical terms are other elements that exist in the legal discourse which are not found in any other discourse. The above characteristics of the legal discourse turn it to be one of the

most arduous tasks for translators and interpreters respectively. Sarcevic (2000: 13) states that

unlike medicine, chemistry, computer science, and other disciplines of the exact sciences, law remains first and foremost a national phenomenon. Each national or municipal law, as it is called, constitutes an independent legal system with its own terminological apparatus and underlying conceptual structure, its own rules of classification, sources of law, methodological approaches, and socio-economic principles. . . . Due to differences in historical and cultural development, the elements of the source legal system cannot be simply transposed into the target legal system. As a result, the main challenge to the legal translator is the incongruency of legal systems.

Mattila (2007:37) contributes to the difficulty of legal translation by stating that

it is in the sphere of international cooperation where the risks attaching to transmission of legal messages are the greatest. Legal documents originating from a foreign State very commonly have to be translated, for example to be executed in the country where a debtor habitually resides. This task is highly difficult and errors often occur in legal translations.

Mattila stresses that such translation problems are aggravated when a need surfaces to operate through an intermediary language, before the final translation. An example could be when one is translating a text from French into English (as intermediary) and then into Arabic.

Despite the continued emphasis on preserving the letter of the law in legal translation, the basic unit of translation is not the word but the text. Since a text derives its meaning from one or more legal systems, legal translation is essentially a process of translating legal systems. Accordingly, it follows that, if legal translation is to be effective, the so-called search for equivalents cannot be reduced to a process of matching up 'equivalents'.

### **3.6. Translating Legal Texts using Machine Translation**

The question remains: will MT translation be able to achieve the task of translating such cumbersome language? As we have seen in the previous chapter, MT systems which draw on matching up equivalents of words or morphemes would most probably struggle and the result would most probably be a fiasco. On the other hand, SMT which depends on matching larger units of language would probably do a better job. This tentative assumption draws on the parallel corpora which already exist or to be created in the future that would not vary to a great extent since variations between translated texts would be to the minimum.

Different attempts to automate the process of translating legal texts took place. Mattila (2007:20) argues that “technical tools as aids to legal translation have been considerably developed over recent years. European Union translators regularly use automated translation tools and computer-aided methods of human translation.” He, however, stresses the need of human to control automated translation and use of terminological databanks in the final analysis of the translated text.

Another important fact is that SMT depends heavily on parallel corpora. The challenge of finding such corpora is the corner stone for any SMT. In this respect, the availability of different contracts professionally translated from English into Arabic and vice versa, statutes, United Nations documents, and agreements between individuals, companies and even countries among others enable MT researchers to have their hands on a valuable treasure that can be exploited to build efficient systems which draw on solid parallel corpora.

## Chapter 4

### Evaluation of Machine Translation Output

#### 4.1. Introduction

Evaluation has always been a major concept for researchers of MT. It is axiomatic that evaluation has proven complicated, traumatic, and misleading; yet both revelatory and useful. Its importance has been highlighted since the ALPAC report in 1966 aforementioned in Chapter Two of the current research. ALPAC had nine recommendations for future US research in translation, three of which directly recommended further work on evaluation. Such recommendations resonate nowadays and MT evaluation is important now as it used to be.

To this end, White (2003) believes that MT evaluation is significant for the fact that MT costs great deal of money to research, design, implement and complete a system with knowledge germane to the subject areas which the system is to translate. As a result, one needs to know whether the investment is worth making.

Although indispensable to the field, it should be noted that evaluating MT is a thorny task. This is due to the fact that evaluating translation is intrinsically hard. There can probably be more than one way to translate the same document, and translators may disagree on the best way to translate a document. As pointed out by Bowker (2000: 183), “the primary difficulty surrounding the issue of translation evaluation is its subjective nature – the notion of quality has very fuzzy and shifting boundaries”. The situation is aggravated in the case of MT and particularly SMT as Cancedda et al., (2009:2) postulate that “machine learning techniques typically rely on some kind of cost optimization in order to learn relationships

between the input and output data”. We can conclude that such techniques are relevant to machine and are very different to the ones adopted by humans.

White (2003) stresses the difficulty of MT evaluation which measures a certain attribute of something against a standard for that attribute. For this to happen, there need to be an identifiable “correct” or “best ideal” against which to compare the relevant attribute of the individual item being measured. The most obvious standard for MT, i.e. the “right” translation, is the very thing translation itself cannot provide. (ibid: 213)

## **4.2 Types of MT Evaluation**

Cancedda et al., (2009) suggest three different types of evaluation which require a quantitative measure of quality:

- 4.2.1** Assessing whether the output of an MT system can be useful for a specific application (absolute evaluation);
- 4.2.2** (a) Comparing systems with one another, or similarly (b) assessing the impact of changes inside a system(relative evaluation);
- 4.2.3** In the case of systems based on learning, providing a loss function to guide parameter tuning;

White (2003:223-232), on the other hand, proposes a descriptive model of evaluation types. He organizes the different types into the following categories:

*Feasibility tests:* an evaluation of the possibility for a particular feat to be accomplished at all, or for a particular approach, whether it has any actual potential for success after further research and implementation

*Internal evaluation:* unlike feasibility testing which is designed to show the potential of an approach, internal evaluations show that implementation of such an

approach can also extend beyond patterns into the language text that will actually occur in production.

*Declarative evaluation:* the purpose of declarative evaluation is to measure the ability of an MT system to handle text representative of actual end use. Analysis of errors should be sought for as a result.

*Usability evaluation:* The aim of usability evaluation is to measure the ability of a system to be useful to people whose expertise lies outside MT per se. In such research, these people may be translators, lawyers, law drafters, lay people, etc.

*Operational evaluation:* Operational evaluations answer the question “Is it worth it?” The primary factors to consider are all of the costs involved, against all of the benefits. Therefore, the purpose of such evaluation is to determine the cost-effectiveness of an MT system in the context of a particular operational environment.

*Comparison Evaluation:* Comparisons measure an attribute of a system against the same attribute of other systems. The purpose of comparison evaluations is to determine the best system, best implementation, or even the best theoretical approach for meeting current or future needs.

The current research is of a qualitative nature as it falls under the category of a feasibility test and declarative evaluation. It is hoped that the research will eventually demonstrate whether *Google Translate* can handle legal language and more specifically contracts. It is declarative based on the grounds that error analysis will be conducted to categorize such errors under lexical and syntactic sub-divisions. The quantitative evaluation is beyond the scope of this study as it could be carried out in further research.

### 4.3 Human vs. Automatic Evaluation of MT

MT evaluation could be performed by either human or machine depending on which type is being adopted as mentioned in 4.2. Humans usually rely on natural language and world knowledge, and their intervention can therefore be useful since their judgment of quality tends to be more accurate than any automatic measure (Cancedda et al., 2009). Moreover, Cancedda et al., favor fully automated evaluations on the types 4.2.2. (b) and 4.2.3. As the current assessment falls under type 4.2.1, human evaluation is the optimal alternative.

Nevertheless, human judgments tend to be subjective and have been shown to vary considerably from one assessor to another. As-Safi (2011:13) points out “An objective assessment of legal texts can be achieved contrary to literary texts, as literature is inherently subjective and more than one version or rendition are expected due to translation indeterminacy.”

White (2003:214) looks at the notion of ‘subjectivity’ from a different angle where he admits that while it is difficult to get around subjectivity, we could perhaps take advantage of it. To this end, he states:

despite the disagreement we are likely to have about translation correctness, we still strongly agree about linguistic intuitions in everyday life. We can talk to each other, read works that are hundreds of years old, order food, and so on, with very high confidence that, despite likely differences in our cultural or cognitive models of reality, we fully understand and agree about the meaning of the expressions and the event as a whole. Could we not capture these linguistic intuitions as means of measuring MT?

Since the current research adopts human evaluation, it will neglect the different methods of automatic evaluation and focus on the different cases of human evaluation. Such selection does not by any means underestimate automatic evaluation as it has become an active research topic. It is important to notice here

that automatic evaluation cannot only benefit MT but also human translation evaluation.

With respect to human evaluation, White (2003) proposes three case scenarios which can serve as an evaluation method to be adopted by MT researchers:

**Case 1 (output only):** This is the simplest way in which an MT developer scrutinizes an output and indicates whether it is of good language. In this case, the following metric is used:

<p>Look at each sentence, one at a time; <b>EITHER:</b> the sentence is completely good English; <b>OR:</b> the sentence is degraded by up to <i>n</i> errors. <b>OTHERWISE</b> the sentence is wrong</p>
---

Figure 1. Case 1: counting errors.

This metric is used to score output sentences expressing either a quantitative measure (by sentence, document, or whole test set), a qualitative method (by characterizing the errors in some way), or both. The shortcoming of such case is that we know nothing about the input and therefore our characterization may not really help to improve the system as translation issues are completely neglected.

**Case 2 (input and output):** Here an MT developer looks at both the input and output, but the metric has to be changed as there are now two parameters: whether the output is fluent (intelligibility), and whether the information in the input is conveyed in the output (fidelity).

Look at both the input and the output of each sentence;  
**EITHER:** the sentence is a completely good translation  
it seems to be good English  
it seems to say just what the source language said;  
**OR:** the sentence is degraded by up to  $n$  errors (intelligibility);  
**AND/OR:** the sentence is degraded by up to  $m$  information errors (Fidelity).  
**otherwise** the sentence is wrong

### Case 2: intelligibility and fidelity.

The drawback of this method is the limitation of the portability and reusability of the measurement which requires special skills, i.e. translational skills that may be hard to find and commit for the task of applying these metrics. Such a shortcoming should not be of a concern for those who are competent in both languages and who have a fair background in translation. However, the main shortcoming of the above metrics is that “good” could vary from one evaluator to another who will simply use their intuition to make evaluative judgments which are prone to human frailties, and therefore such metrics could be highly subjective. Moreover, the distinction between the two parameters, i.e. intelligibility and fidelity is not clear cut. They are correlated: a completely unintelligible expression conveys no information.

**Case (3) one input, two outputs:** In this case an MT developer should be able to make changes in the system so that it can improve its coverage. Since this ability is somehow impossible in our research where no changes can be done on *Google Translate*, this case will be neglected.

Another distinction which should be brought to attention is the difference between the black box and the glass box evaluation. White (2003:241) states:

The black-box view is a look at the input and output without taking into account the mechanics of the translation engine. The glass-box view looks inside the translation engine to see if each of its components did what was expected of them in the course of the translation process.

The advantage of the black box view in the current research resides in the fact that the evaluation will be able to determine better the language coverage of the system which will serve to either confirm or refute our hypotheses about how the system tries to handle legal language. Moreover, such evaluations could serve as important data as it is entirely reusable for other systems.

The current research adopts the black box evaluation since the mechanism which *Google Translate* adopts is statistical rather than traditional where definite rules and algorithms are furnished. Therefore, the output will be assessed against two levels, i.e. lexical and syntactic levels. Sub-categories will be furnished based on the results of the assessment and the errors *Google Translate* makes. Moreover, the evaluation purpose is to evaluate the language from a legal perspective since the language under assessment is restricted to the legal discourse where the hybrid, intricate nature of this type of language can hardly be subjected to simplistic sets of criteria or categories of assessment. Among the criteria on which the assessment will be made are precision and ambiguity, because ambiguity may result in serious legal consequences engendered by different possible interpretations.

To find out a way out of this dilemma, the researcher decides to adopt certain criteria from a more elaborate model of assessment to overcome the shortcomings in the previous model. As-Safi (1996) presents a tri-criterion assessment which is proposed to diagnose errors/losses which usually stem from

the three constituent phases and pertains to the following parameters for assessment: first, accuracy checked on two planes: comprehension and equivalence; second, the language of the TT checked on three planes: punctuation and spelling(which could be ruled out as machine outcome will overcome spelling and punctuation mistakes), grammaticality and acceptability and choice of diction; third, literariness which lies outside the scope of present study hence it is replaced by legal technicalities.

#### **4.4 Methodology and Corpus**

The data of the current research comprise articles deduced from six different civil contracts which regulate different matters namely purchasing of real property contract, selling an anodizing plant contract, two leases of a dwelling unit, a lease for space in an office building, and a maintenance and service contract (Reuters). The selection is based on three factors: (1) whether the article is extracted from the preamble, body or conclusion as preambles and conclusions are formulaic in nature; (2) the Article length: it is widely known that the longer the Article, the more complex is its structure and therefore it is more complicated to be handled by the MT system; (3) each Article has a number of anticipated problems to the machine. The purpose of the current corpora is to be able to make a general claim about the ability of *Google Translate* to translate the infinite legal structures, based on a finite test set.

These contracts are English-authentic which can be found in Sabra (1995) *Translation of Contracts* where a number of contracts are translated from English into Arabic. The current research opts for authentic English contracts since machine software cannot pre-edit linguistic errors prevalent in STs, but choose

texts that are linguistically correct will eliminate erroneous translations resulting from the absence of ST pre-editing (Al-Dabbagh, 2013).

Moreover, the researcher intends to assess *Google Translate* against fourteen articles extracted from six different contracts rather than being limited to one contract. The main aim of this diversity is to try to be as objective as possible in assessing such a statistical machine which draws heavily upon its corpora, i.e. should that one contract be included or excluded from the system corpora, the findings will not reflect the actual status of the MT system. In this respect Jurafsky and Martin (2000:204) assert:

Suppose we are trying to compute the probability of a particular test sentence. If our test sentence is part of the testing corpus, it will have an artificially high probability. The testing corpus must not be biased by including this sentence.

In the following assessment, I will make use of corpora as “test examples” which are authentic sentences that contain anticipated linguistic or legal problems, i.e. articles selected from the six different contracts mentioned above. The test examples will be fed to the MT system (*Google Translate*) to see how it will deal with them.

All test examples in this thesis are collected and added in Appendix I to constitute a legal ‘test suite’ for English-Arabic translation along with the output of *Google Translate* and the proposed translation by the researcher. In addition, the researcher has opted for test suites after carrying out an attempt which shows that *Google Translate* alignment is done on the phrase level. In the experiment, the researcher has input a full contract at once while focusing on the rendition of one of its articles. The article under scrutiny was input out of its context and the result reveals that *Google Translate* identically translates both articles regardless of the

surrounding context. In addition, the assessment carried out in the current research is on the lexical and syntactic level.

Test suites are very important in MT development and enhancement. Rules in an MT system may be “too strict or too general to apply correctly in all circumstances” (Doug Arnold, et al., 1994: 176). Test suites help developers detect such errors easily. They can then modify the rules and run the test suite again to measure improvements.

I would like to mention some facts about the test suite I constructed in this thesis. First, it is useful only in testing legal errors since the sentences are extracted from legal contracts. It is not meant to measure translation performance in general. Second, it is designed for English-Arabic translation. Third, the purpose of test examples is not to compare *Google Translate* with other MT systems as this could be done in further research. The aim is to explore linguistic and legal problematic areas surmounted by *Google Translate*.

It must be pointed out that this evaluation, like any other one, is not, nor can it possibly be a comprehensive one. A complete assessment covering all aspects of lexical and syntactic errors made by the system requires many extended volumes, which is beyond the purpose of this research. The emphasis in the current research has been basically on salient drawbacks on both levels which hinder the overall process of translating legal texts by *Google Translate*.

#### **4.5 Qualitative Assessment**

In this section, the researcher opts for categorizing the problems which prevail during the assessment. Such categorization could serve the system developers as well as other researchers who aspire for developing the output of *Google Translate*.

The assessment is carried out on two main domains, i.e. the lexical and syntactic levels. The textual level is ruled out in the current assessment since articles rather than a full contract is fed into the system.

#### **4.5.1 Lexical Level Assessment**

##### **4.5.1.1 Polysemy and Homonymy**

One of the challenges that faces MT in general is the fact that many words in the English language are polysemous, i.e., they carry more than one meaning. It should be noted that polysemy is different from homonymy where homographs are etymologically unrelated words that happen to be represented by the same string of letters in a language. *bass the fish* is derived from Old English *barse* while *bass the voice* is derived from Italian *basso*. Conversely, Polysemes are etymologically and semantically related. *Line* in a *line of people* and *a line drawn on a piece of paper* are etymologically related and their semantic relation is obvious.

Computer applications that handle the content of natural language texts need to come to terms with polysemy as it has been considered a stumbling block in natural language processing (Ravin & Leacock, 2000). The study of polysemy in computational linguistics touches upon the problem of how to map expressions to their intended meanings automatically. The computer possesses the same resource for sense identification as human translators do, the context. Nevertheless, computers are “handicapped” because they can only interpret the context as strings of letters, words or sounds and not as meanings (ibid: 23). Bar-Hillel (1960) stresses that, although it is a trivial matter for an English speaker to assign the appropriate sense of *pen* (enclosure rather than writing device) in *the box is in the pen*, no computer can do so. We can disambiguate *pen* because our world

knowledge includes information regarding the relative sizes of toy boxes, writing implements, and play pens.

Despite the earlier warnings about the impossibility of automatic sense resolution without the characterization of world knowledge, there is a resurgence of interest in automatic sense identification, most notably with the aid of machine readable dictionaries and statistical analyses of large textual corpora (Ide & Veronis, 1998).

Different attempts to disambiguate senses in computational linguistics have taken place starting with a theory of semantic interpretation that posits structures that are contained in the mental lexicon, yet independent of world knowledge (Katz & Fodor, 1963). In the late 1980s, machine-readable dictionaries (MRDs) were proposed for sense disambiguation (Lesk, 1986). Lesk devised a simple method to link dictionary definitions if they share words in common. Finally, corpus-based disambiguation methods have been developed. A problematic aspect of this type of corpus-based approach is what Gale et al. (1992) call “the knowledge acquisition bottleneck”. In order to get materials for training a program on senses of a particular polysemous word, the corpus of contexts containing that word has to be manually partitioned into its different senses. Apart from being an assiduous task, the work put into manual tagging does not scale up, i.e. the partitioning required for one word will be of no use in disambiguating any other word and will not decrease the amount of manual effort required.

Homonyms, on the other hand, are different words with the same form in which dictionaries distinguish them by assigning to them distinctive numbers. For example, *bank* (1), one of whose meaning is “financial institution”, and *bank* (2), one of whose meaning is “sloping side of the river”, are generally regarded as

homonyms (Lyons, 1995:27). They have been classified by lexicographers as separate words and therefore were given separate entries in dictionaries.

Lyons stresses that homonymy results in lexical and grammatical ambiguity. He illustrates such ambiguity in the following example:

*They found hospitals and charitable institutions.*

The above example can be construed as a present tense sentence containing a form of the verb *found* or, alternatively, as a past tense utterance containing a form of *find*. The ambiguity is partly lexical since the two lexemes hold different meanings, and is partly grammatical in so far as its ambiguity depends upon the grammatical non-equivalence of *found* construed as a form of *found* and of *found* construed as a form of *find* (ibid: 56).

By the same token, Hutchins and Somers (1992) believe that the first and most obvious is the problem of homonyms or, more specifically, homographs. Homographs are two or more words that are spelled alike but pronounced differently, such as *tear* (the byproduct of crying) and *tear* (rip). Choosing the right equivalent according to the context has proven to be problematic in the current research. This shortcoming could be attributed to the fact that *Google Translate* adopts a statistical approach in which the semantic analysis does not take place.

Al-Najjar (2004) stresses that homonymy as well as polysemy are the main obstacles that face MT in disambiguating word senses. He gives the example of *water* as a noun and *water* as a verb to indicate polysemy. *right* as an adjective to mean a direction opposite to left and *right* the opposite of *wrong* is furnished to refer to homonymy. He adds that acronyms constitute a problem to MT as there is no specialized lexicon that states what such abbreviations stand for.

It is significant to note that it is less difficult to automatically identify senses of homonyms than it is to identify senses of polysemes due to the fact that the former contexts will consist of quite different vocabulary, whereas contexts of polysemes may be quite similar. This fact coincides with the results of the current research analysis. Polysemy constitutes a major impediment to *Google Translate* in comparison with homonymy where scarce errors were noticed in the analysis.

### **Test Examples Analysis**

The Articles are scrutinized by checking the input and output along with a comparison of the suggested translation to try to locate where *Google Translate* fails to render the appropriate sense of the polysemous words. For the purpose of brevity, the research furnishes the sense *Google Translate* erroneously opts for and the one suggested by the translator by looking up the word from the dictionary (*Merriam-Webster's Advanced Learner's English Dictionary, 2008*). Whenever necessary, legal dictionaries are consulted to look up words that pertain to the legal terminology (*Faruqi's Law Dictionary and the Law Dictionary online*). Eleven lexical items (polysems and homonyms) are examined in which *Google Translate* fails to render the correct sense according to their contexts:

#### **deposit** (Article 2):

1. An amount of money that is put in a bank account (إيداع)
2. Money that you give someone when you agree to buy something. (a deposit shows that there is an agreement between a buyer and seller)  
(عربون)
3. Money that you pay when you rent something and that you can get back if you return the thing or leave it in good condition (تأمين)

The word *deposit* is mentioned two times in the Article. *Google Translate* has construed the Article by opting for the third sense rather than the second at first, and repeats the error at the end of the Article by opting for the first entry instead of the second which has resulted in changing the overall meaning of the sentence.

**property** (Article3):

1. Something that is owned by a person, business, etc. (ملكية)
2. A piece of land often with buildings on it that is owned by a person, business, etc. (عقار)
3. A special quality or characteristic of something (خاصية)

The word *property* is mentioned three times. In the first two instances, *Google Translate* fails in assigning the right sense of the word *property* as it opts for the first sense instead of the second. In third instance, the system chooses the third sense instead of the second which does not relate by any means to the subject matter of the Article. Such erroneous sense disambiguation affects the comprehensibility of the Article.

**maintain** (Article 3):

1. To cause something to exist or continue without changing (يبقي أو يحافظ)
2. To keep in good condition by making repairs, correcting problems, etc. (يصون)

The Article pertains to a lease of an apartment. *Google Translate* fails to render the right sense which is the second above.

**said** (Article 4):

This is a homonym as “*said*” is listed in two different entries in the dictionary:

1. Said (1) is past tense and past participle of SAY(1) (قال)
2. Said (2) is an adjective which is always used before a noun. It is used formally in law which means: mentioned or referred to before (المذكور سابقا)

“*Said*” as a lexical item is used very often in the legal discourse; nevertheless, *Google Translate* fails to render it as an adjective and translates it as a verb which changes the whole structure of the output let alone its overall meaning.

**counter** (Article 5):

1. counter(1): a person or a device that counts something (عداد)
2. counter(2): as a reaction against (رد فعل)

*Google Translate* opts for the first definition of “counter” instead of the second which is mentioned in an expression: *payable over its counters* (قابل للدفع). Such rendition does not convey the appropriate sense of the word in its context.

**code** (Article 8):

1. A set of laws or regulations (قانون)
2. A set of letters, numbers, symbols, etc. that is used to secretly send messages to someone (رمز/شيفرة)

The context in which the word “codes” is mentioned in Article 8 carries the meaning of laws and regulation (building and housing codes). However, the system erroneously chooses the meaning of the second entry as an equivalent to the word “codes”.

**utility** (Article 9):

1. The quality or state of being useful (فائدة)

2. A service (such as a supply of electricity or water) that is provided to the public-usually plural. (خدمات)
3. A company that provides electricity, water, etc. (شركة خدمات)

The word “*utility*” is mentioned three times in Article 9. *Google Translate* is successful in handling the sense of the word “*utility*” at the head of the Article; however, it fails in rendering it in the middle and end of the Article. The system erroneously opts for the first sense listed in the dictionary i.e. the quality or state of being useful. By the end of the Article, *Google Translate* renders the homograph “*utility*” as an adjective “a utility tool” (اداة) which does not only reflect the incorrect meaning of the sentence but also creates an incomprehensible sentence by all means.

**rate** (Article 9):

1. The number of times something happens or is done during a particular period of time (معدل)
2. An amount that is paid: a price or amount to be paid that is set according to a scale or standard (رسم)

The word “*rate*” carries the sense of the second entry rather than the first according to the context it appears in.

**furnish** (Article 9):

1. To provide (a room or a building) with furniture.(يؤثث)
2. To supply or give to (someone) something that is needed or wanted (يزود).

*Google Translate* is not able to translate the word into its correct sense (the second).

**qualification** (Article 10)

1. A special skill or type or experience or knowledge that makes someone suitable to do a particular job or activity (مؤهل)
2. (Formal) something that is added to a statement to limit or change its effect or meaning, e.g. they will agree to the deal only with the qualification (stipulation) that they get 40 percent of the profits. (شرط)

The word “*qualification*” should be construed in this legal context as stipulations. Failing to render the correct sense of the word results in changing its legal binding effect.

**remedy** (Article 14) the word *remedy* can appear as a noun and a verb. *Google Translate* is able to render its sense as a verb since it has only one entry in the dictionary, i.e. it is not polysemous. However, as a noun it has two different senses:

1. A medicine or treatment that relieves pain or cures a usually minor illness (علاج)
2. A way of solving or correcting problems (طريقه لحل)

By consulting a legal dictionary (the Law Dictionary), we find that there is a legal sense which is restricted to the legal genre:

3. Remedy is the means by which the violation of a right is prevented, redressed, or compensated. (تعويض)

*Google Translate* fails to render the legal sense of the word “*remedy*”, and therefore fails to furnish the legal terminology in the output.

#### 4.5.1.2 Couplings

Couplings are parallel structures in which two synonymous or quasi-synonymous words usually belonging to the same form or class are put together. These words could be nouns, verbs, adjectives or even prepositions.

In translating couplets, there is no clear-cut strategy in approaching such expressions, i.e. whether to keep the couplets in the TL or to render them as single words. Although couplets are a feature of the legal language in Arabic, some of them cannot be rendered as such, and a one word is utilized instead. Sabra (1995) lists couplets which could be translated to their counterparts in Arabic with no deletion

acknowledge and confess	يقر و يعترف
act and deed	عمل و فعل
including but not limited to	يتضمن و لا يقتصر على
fraud and deceit	غش و خداع
null and void	لاغ و باطل
safe and sound	مأمون و سليم
save and except	فيما عدا و باستثناء
shall be and remain	يكون و يظل
shun and void	يتحاشى و يتجنب
unless and until	مالم و الى ان

The list below shows couplets which cannot be rendered as such and are rendered as single words:

by and between	بين
deem and consider	يعتبر
each and every	كل
free and clear of	خالي

from(date)... <u>to and including</u> (date)	و لغاية
keep and maintain	يحافظ على
made and entered	تحرر
made and signed	تحرر
terms and conditions	شروط
true and correct	صحيح
upon and subject to	تخضع لـ

The above parallel expressions are of paramount importance for *Google Translate* to exist in its lexicon as the researcher fed the above expression and found out that the system was unable to handle most of the above couplets.

### **Test example analysis**

The fourteen Articles were examined to locate instances of couplets for the purpose of verifying how *Google Translate* handles such expressions.

#### **terms and conditions** (Article 1)

*Google Translate* fails in rendering the couplet “*terms and conditions*” into one word. It rather repeats the word “شروط” twice which affects the comprehension of the whole Article.

#### **upon and subject to** (Article 1)

“Upon and subject to” is a couplet which should be rendered into a verb plus a preposition “تخضع لـ”. *Google Translate* renders the preposition “*upon*” as a single lexical item into “عليها”. As a result, the comprehensibility of the whole Article is affected.

### **Including but not limited to (Article 6)**

*Google Translate* erroneously translates the above couplet into “بما في ذلك و لا ”تقتصر على”. Couplets are parallel structures which entail that both lexical items should be of the same part of speech. The system fails in maintaining parallelism where the first word in the couplet is translated as a preposition while the second is translated as a verb. A suggested translation is added by the researcher to the first chart.

### **To and including (Article 7)**

The couplet above is used to specify the end of the term, and it should be translated into one word “لغاية”. *Google Translate* is unable to identify the beginning and end of the duration in Article 7 which could be due to the absence of such couplet in its corpora.

#### **4.5.1.3 Legal Adverb (here/there+ preposition)**

As mentioned in Chapter Three, legal register is replete with technical terms and archaic expressions. In the current subcategory, the assessment sheds light on the expressions with “here” and “there” connected with a preposition to form an adverb. The words ‘hereinafter’ and ‘hereby’ are adverbs. Other similar adverbs are commonly used in legal documents and official statements. In the following example (Mckay et al., 2005:73) the adverb ‘*hereunto*’ means “to this document”.

*In testimony whereof I have hereunto set my hand and affixed my seal of office in the City of London, this sixth say of June two thousand and eleven.*

Mckay et al., (2005:74-80) list the most common adverbs used in the legal register. The researcher furnishes suggested translations to these adverbs into

Arabic by consulting *Faruqi's Law Dictionary* since such parallel data are of great value to MT systems adopting statistical approaches:

hereby	as a result of this statement, by this means	بموجب هذا النص/نتيجة لهذا النص
herein	in this document, in this matter	الوارده (المتضمنة) في هذه الوثيقة
hereinafter/hereafter	in a later part of this document, starting from this time	فيما بعد في هذه الوثيقة
hereinbefore/heretofore	in a preceding part of this document, earlier in this document	سابقاً(في ما تقدم) في هذه الوثيقة
hereof	of this document	من هذه الوثيقة
hereto/hereunto	to this matter or document	لهذه (الوثيقة)
hereunder	below, later in this document	تالياً/ذيلًا هذه (الوثيقة)
hereupon	in consequence of this, at this point	بموجب هذا/بناء عليه
herewith	together with this document	مرققاً بطيه

Chart 1: adverbs using “here”+ preposition

thereafter	from the point in time	بعد ذلك (التاريخ), فيما بعد
thereby	as a result of that, by that mean	نتيجة لذلك, و عليه
therefor	of that	لذلك(الشيء)
therein	in that document, in that respect	الوارده في (الوثيقة)
thereinafter	later in this document	لاحقا في (الوثيقة)
thereinbefore	earlier in this document	سابقا في (الوثيقة)
thereof	of that	تابع له/ لها
thereto/thereunto	in accordance with that stated below in the document	بموجب ما ذكر اسفل (الوثيقة)
theretofore	before that time	قبل ذلك
thereunder	in accordance with that stated below in the document	بموجب ما ذكر اسفل (الوثيقة)
thereupon	in consequence of that; immediately after that	على ذلك, بناء عليه, بعد ذلك مباشره
therewith	with that, soon after that	بالتزامن مع

Chart 2: adverbs of “there”+preposition

### Test Example Analysis

As mentioned above, legal contracts excessively use the adverbs listed above for the purpose of precision which is a requirement of such text type.

**hereunto** (Article 4): *Google Translate* opts for translating the adverb into “هنا”. However, the system should avoid the use of references and instead should

utilize repetition in accordance with the characteristics of the legal discourse. A proposed translation to the adverb is “هذا العقد”.

**herein** (Article 5): *Google Translate* repeats the error listed above in not mentioning the document in which such conditions apply. It instead translates it as “الواردة” rather than “الواردة في هذا العقد”.

**thereof** (Article 6): the adverb “*thereof*” is mentioned twice in the Article. In the first instance, *Google Translate* translates it as “ذلك” which is correct, but it fails to make the legal addition which is necessary to avoid ambiguity in the Article. The adverb can be translated to “ذلك الحدث أو الظرف”. In the second instance, the adverb is translated by *Google Translate* as “منه” instead of “من إستمرار الحدث”.

## 4.5.2 The Syntactic Level Assessment

“If words are the foundation of speech and language processing, syntax is the skeleton” (Jurafsky and Martins, 2000:285). From their quotation, one can anticipate the importance of syntax in the field of computational linguistics and more specifically in MT. In the current research, the researcher assesses the output on sub-syntactic categories where *Google Translate* fails to handle certain syntactic structures.

### 4.5.2.1 Morphological Parsing

Morphology is the study of the way words are built up from smaller meaning-bearing units, A morpheme is often defined as “the minimal meaning-bearing unit in a language” (ibid, 2000:59). For example, the word *fox* consists of a single morpheme while the word *cats* consists of two: the morpheme *cat* and the morpheme *-s*. The problem of recognizing that *cats* breaks down into the two

morphemes is called **morphological parsing** which applies to many affixes other than plurals such as English verb forms ending in *-ing* (*walking, sending, congratulating, etc.*).

Morphology is an indispensable knowledge for language processing. This is due to the fact that English morphemes are productive, i.e. the *-ing* can be applied to every verb and similarly the *-s* can be added to most nouns. The idea of listing all nouns and verbs with their derivations can be quite inefficient should the MT system developer decide to abandon morphological parsing.

The following table (ibid: 65) shows how parsing is carried out on English morphology. The goal is to take input forms (left column) and produce output forms containing the stem of each word as well as assorted morphological features (right column).

input	Morphological Parsed Output
cats	cat+ N + PL
cat	cat+ N + SG
cities	city + N + PL
geese	goose +N + PL
goose	(goose + N + SG) or (goose+V)
gooses	goose + V + 3GS
merging	merge + V + PRES-PART
caught	(catch + V + PAST-PART) or (catch + V + PAST)

It should be noted that some input forms (such as *caught* or *goose*) above will be ambiguous between different morphological parses. The parser is not able

to decide on the proper parse. In fact, disambiguation requires some external evidence such as the context in which the word falls.

In the current research, there are instances in which *Google Translate* fails to disambiguate certain lexical items which result in erroneous syntactic renditions.

### **Test Example Analysis**

**desires** (Article 1):

The word “*desires*” is repeated twice in Article one as a verb “يرغب”. It is ambiguous since it could be a verb preceded by a singular subject (desire + V + 3GS) or it could be a plural noun (desire + N+ PL). *Google Translate* is able to disambiguate the word at first, but it fails to do so in the second where it translates it into a noun “رغبات”.

**credit** (Article 12):

The morpheme *credit* could be a verb (credit +V) or a noun (credit + N + SG); hence it poses ambiguity to the morphological parser of the machine system. *Google Translate* fails to translate “credit” as a verb (ينسب), and translates it erroneously into a noun (الإئتمان).

#### **4.5.2.2 Concord**

Concord refers to “the requirement where the forms which agree with one another must be of the same subdivision of the form-classes to which they belong” (Robert, 1964:200). In other words, certain words are required to take forms which correspond in specific ways with certain other words (Robins, 1964:249). Concord features are essential and should be carefully applied to ensure the generation of

sound sentences in the TL. Since concord applies to the TL, it should fulfill the specific requirements of this language.

In a contrastive analysis study, As-Safi (1972) discusses concord and its peculiarities in English and Arabic. He divides concord into three subcategories: concord of person, gender, and number. In addition, he touches upon concord-deviances or peculiarities in both languages.

Concord rules are complicated and such complication becomes more obvious when attempting to design an MT system between languages that have great morphological variations and big differences in concord requirements. The English verb, for example, indicates in the present tense whether the subject is singular or plural in the 3<sup>rd</sup> person, but it does not give any information about gender. In other tenses, the English verb is completely neutral. English adjectives and determiners are also, to a great extent, neutral to number and/or gender. This poses a problem when translating English into Arabic, a language that is highly sensitive to agreement features. Arabic verbs, adjectives, and most determiners are highly reactive to the noun they modify, whether singular, dual, plural, feminine, masculine, human or non-human.

To achieve accurate concord in the output, MT systems can draw the information they need in some cases from the SL. In the SL, lexical items, especially proper nouns and titles, should be defined whether they are feminine or masculine, singular or plural. The SL should tell us that *Jack* is masculine while *Jill* is feminine. In other cases the information needed for agreement is derived from the TL. In the TL, lexical items, especially common nouns, should be defined, whether they are feminine or masculine, singular, dual or plural. The TL will tell us that "كتاب", which is the equivalent to *book*, is masculine while "كراسة",

the equivalent to *notebook*, is feminine. A more complicated case is when the information needed for concord cannot be drawn either from the SL or the TL. In this case the context is the determining factor. The system looks for clues in previous or subsequent sentences; otherwise it will take the marked form of masculine.

In the current research, the outputs of the fourteen articles are analyzed to pinpoint cases of erroneous concord cases.

### **Test Example Analysis**

#### **Affected party (Article 6):**

*Google Translate* makes an error in concord in number as it translates “*affected party*” as "الطرف المتضررين" instead of “الطرف المتضرر” which does not show agreement in number between the noun and its modifier.

#### **Paid by tenant (Article 9):**

In the above phrase, *Google Translate* fails to apply the concord rule of gender. *Tenant* “المستأجر” is a masculine noun in Arabic and the verb has to agree with its subject. The system opts for the verb “تدفع” which is in the feminine form and therefore does not agree with its subject. It should opt for the masculine form of the verb “يدفع”

#### **Use Premises for permitted use (Article 11):**

The problem is prevalent in the translation of the word *premises*. It is a plural noun; however, it should be translated into a singular noun since it is defined in Merriam-Webster as:

Premises (plural): a building and the area of land that it is on.

From the definition above, one can conclude that the word *premises* refers to *one* building and therefore should be rendered as the singular noun “المبنى” instead of “المباني”. However, *Google Translate* renders it as a plural noun. To add insult to injury, the system is not able to show agreement in number between the noun and its anaphoric reference in which a singular pronoun is rendered as “به” instead of the plural form “بها”.

#### **A breach which cannot be remedied (Article 14)**

*Google Translate* fails to show concord in gender as regards the anaphoric reference of the noun breach (خرق) which is masculine in the Arabic language. Instead, the system utilizes the feminine reference which is attached to its verb “علاجها”. The masculine form of the verb should be used instead “علاجه”.

#### **4.5.2.3 Modality: shall/may**

The modality effect is particularly crucial in legal contexts, as the accuracy of information relating to location, interaction and manner of interaction is perhaps even more important than in any other context (Brennan and Brown 1997; Brennan 1999).

*Shall* is used in official documents to show an obligation, command, promise, etc. For example, "*All payments shall be made by the end of the month*", *shall* here is different from the auxiliary verb which indicates the future tense." *May* is used to refer to the possibility that someone *may* do something in a certain way, or that something *may* be done in a certain manner. For instance, "The Second Party *may* assign this Agreement to the third party without a prior written consent of the First Party". *May not* is used to indicate the opposite as in "The Second Party *may not* assign this Agreement to the third party without a prior written consent of the First Party" (Fakhouri: 2008).

Sabra (1995:31) states that the verb preceded by *shall* is usually translated into the present tense, and he stresses that it should not be used as a verb which indicates future. He lists some examples:

1. Amoco shall be exempted from all customs and duties.  
تعفى أموكو من كل الرسوم الجمركية.
2. The arbitration shall be held at Paris.  
يعقد التحكيم في باريس.
3. The life insured shall pay to the Insurance Society every subsequent premium in due time.  
يدفع الشخص المؤمن عليه الى شركه التامين كل قسط تال في موعده المستحق.

By the same token, Hatim (1997:30) stresses that the modal “shall” is not the future shall but rather an empty legal “shall”. Thus, it should be rendered as the Arabic present simple verb which is intrinsically timeless and legally binding. However, he states that “shall” could be rendered as the verbs “يتم” or “يجب” depending on the legally “obligatoriness” the context requires. It should be noted that the verb “يتم” is used mainly in the case of rendering an agentless passive. For example:

1. An amount of \_\_\_\_\_ JDs shall be paid upon signing the Contract (Saqf Al-Hait, 2009).  
يتم دفع مبلغ \_\_\_\_\_ دينار عند توقيع العقد.
2. In normal circumstances, the period of notice shall not be less than one month (Hatim et al., 1995:168).  
و في الحالات العادية يجب أن لا تقل مدة الشعار عن شهر واحد

On the other hand, *may* as a modal in legal discourse should be rendered as "يجوز" and the negative form of it is "لا يجوز". "يحق" and "لا يحق" are equivalent to

“يجوز” and “لايجوز” respectively and therefore the two verbs can be used interchangeably (Hatim, 1997:45).

Sabra (1995:32) lists some examples on the matter above:

1. The minister of petroleum may enter into concession agreements with foreign companies for petroleum exploration.  
يجوز لوزير البترول ان يبرم اتفاقيات امتياز مع شركات اجنبيه للبحث عن البترول.
2. The United Nations may hold fund, gold, or currency of any kind.  
يجوز للامم المتحده ان تحتفظ بأموال أو ذهب أو عمله من أي نوع.
3. The Contractor may not assign the contract to a third party without the prior written consent of the Employer.  
لا يجوز ان يتنازل المقاول عن العقد للغير بدون موافقه خطيه مسبقه من صاحب العمل.

The researcher has attempted to feed the above examples into *Google Translate* and found out that the system was able to translate the sense of *may* as “يجوز” in the first sentence while erroneously translated *may* in the second and third example as “قد”.

### **Test Example Analysis**

The articles are analyzed to locate instances where the modal “*shall*” or “*may*” are mentioned. *Google Translate*’s output is examined to determine whether the system is able to handle the modality in the legal context.

“*Shall*”, as a binding modal, is mentioned ten times in the fourteen articles. None of which is *Google Translate* able to render into a simple Arabic present tense of the verb to which the modal is attached. *Google Translate* opts for utilizing a different verb in different contexts such as:

**buyer shall pay (Article 2):**

*Google Translate* adds the verb “يتعين” instead of using the simple present tense of pay “يدفع”.

**the party shall give notice (Article 6)**

*Google Translate* opts for the verb “تتاح” where the simple present of “give” should be utilized instead of “يقدم”. The verb “تتاح” does not carry the sense of *shall* where the action is compulsory rather than voluntary.

**the duration shall be from (Article 7)**

*Google Translate* erroneously mentions the verb “يكون” and “تكون” in the same sentence “تكون مدة هذه الاتفاقية يكون من”. Such repetition hinders the comprehension of the Article.

**landlord shall comply with (Article 8)**

*Google Translate* translates “shall” as “يجب” which is appropriate. However, it does not use the preposition “على” which affects the overall comprehensibility of the sentence.

**charges shall be paid by tenant (Article 9)**

*Google Translate* uses the verb “يتم دفع” instead of “يدفع”. The input is agentive passive and should be rendered into active in the Arabic sentence (Hatim, 1997:227). *Google Translate* fails in following this rule and maintains an agentive passive in the structure of the output which defies the norm of the Arabic language and which is characterized by being “normally shunned as awkward” (ibid:31).

**tenant shall use the premises (Article 11)**

As in Article 8, the system makes use of the verb “يجب” as an equivalent for “shall”. Nevertheless, the verb employed along with a preposition “على” followed by the subject should be used in order to form a well-formed syntactic phrase in Arabic “يجب على المستأجر”. *Google Translate* fails to construct such a phrase and instead utilizes an incomprehensible and ill-formed syntactic phrase “يجب استخدام المباني المستأجر”.

**subscriber shall credit any report (Article 12)**

*Google Translate* uses the verb “يقوم” as an equivalent to *shall* whereas the simple present of “credit” “ينسب” should be used.

**subscriber shall indemnify (Article 13)**

*Google Translate* chooses the verb “يجوز” which is considered a grave flaw in legal translation since the equivalent chosen does not hold the binding feature of the modal *shall*; yet it conveys the sense of possibility rather than obligation.

**reuters may terminate this agreement (Article 14)**

*Google Translate* opts for “قد” as an equivalent to the modal *may* in this context. The particle “قد” signifies a probability of something to happen. The system should opt for “يجوز” or “يحق” to indicate that it is legally permitted to do something.

After having examined *Google Translate* on the lexical and syntactic levels, one should be cognizant of the different errors that could occur in using the system in translating contracts. Such errors are systematic which entails that they are most probably imminent when end users choose *Google Translate* as an option in

translating legal documents. Having said that, the Research and Development Department at the giant corporation should take the above findings and analysis into consideration in order to improve the performance of their tool in general and in the field of legal translation in particular.

## Conclusions and Recommendations

### Conclusions

The evaluation conducted on the legal discourse and more specifically contracts shows that *Google Translate* is not successful in producing outputs which are fully comprehensible to the target reader nor does it render the accurate legal sense exhibited in the inputs. The purpose of the evaluation is twofold. First of all, the main aim is to figure out whether the system is able to handle legal texts for different segments (feasibility test). The second aim is to furnish an error-analysis by determining the prominent drawbacks the system suffers from on two main levels, i.e. the lexical and syntactic levels (declarative evaluation).

From a feasibility test perspective, *Google Translate* does not perform well in translating legal discourse specifically contracts, regardless of the type of contracts under assessment. It is axiomatic to say, however, that legal translators, law librarians, and lay people have to reconsider their strategies should they decide to depend on *Google Translate* in translating legal documents. Because of the peculiarity and importance of precision the legal discourse has, people who are aiming at translating legal texts are cornered in an ultimate alternative which is “professional human translation”.

In spite of the erroneous results *Google Translate* produces, it can still deliver the gist of the article. This could be helpful for lawyers and lay people who are willing to know the content of the text before paying a substantial amount to get it translated professionally.

Beyond the basic task of identifying the subject matter of a text, *Google Translate* is of little use in a law library. Law librarians and lawyers who use

*Google Translate* should be fully aware of the program's shortcomings, which means they should always consider the translations questionable at best. When precision is vital, *Google Translate* should not be used at all.

From a declarative perspective, the researcher aims at providing error analysis on two levels, i.e. the lexical and syntactic levels. Under the lexical level, *Google Translate* is assessed on three different subcategories: Polysemy and homonymy, legal couplets and legal adverbs. On the syntactic level, *Google Translate* is assessed on morphological parsing, concord or agreement and finally on legal modality.

The analysis shows that *Google Translate* makes several errors on the lexical level. The system fails to disambiguate several senses in the legal context. Technically speaking, the system is not successful in handling legal terminology and expressions especially when it comes to legal couplets and adverbs of *here* and *there* attached to a preposition. On the syntactic level, the system shows deficiencies in its morphological parser where some words are translated erroneously into different parts of speech. In addition, the system fails to adhere to concord rules of the TL. Finally, the legal modality is not handled properly by *Google Translate* which fails to form the functional equivalence of such modals in the assessment.

## Recommendations

Although complicated and traumatic, MT remains one of the prominent research areas in the field of computational linguistics. Therefore, researchers should not refrain from delving into such a thorny area which shows a lot of potential in the future. It is widely known that research on such a topic is meager in the Arabic language. For MT to improve systems translating from and into Arabic, the necessity of rigorous research should be stressed in the Arab countries through universities teaching translation by designing special courses to students to raise their awareness in such a vital field. The collaboration of different fields such as mathematics, IT, engineering along with linguistics and translation studies is highly advisable since such work necessitates the team work of the mentioned specialties to come up with promising results.

Since MT systems consume a lot of effort and money, it is of great importance to conduct several assessments to make sure that the system is running effectively, and that it meets the need of its consumers. The findings in the current research could serve *Google Translate* well in different aspects. Should *Google Translate* take the analyzed errors into consideration in the current research, outcomes will show substantial improvements.

The idea of limiting the scope of translation in *Google Translate* could help the system handle different text types separately. The corporation should set out in their agenda the necessity to enable its users to choose among different text types the one the user is intending to translate. Since it is a SMT, it could group its corpora depending on the text type, i.e. scientific, legal, journalistic, medical, etc. By doing this, the probability of ambiguity will decrease sharply.

The need to conduct further research on corpora studies is significant in the field. The presence of comparable and parallel corpora in Arabic designed for SMT will help such systems to improve the quality of their translation as it is widely known that the availability of corpora is the backbone of any SMT.

The test set the researcher devised is of paramount importance for future research. The test set could be utilized to assess different MT systems in a comparative study. Moreover, the researcher opts for the most salient errors the system makes. Further research could be done on different categories such as word order, noun phrases, modifier displacement, definiteness, passive voice, tenses, among others, which lie beyond the scope of the present research.

Finally, legal translation is a practice which should be approached meticulously due to the precision required in such practice. End users, whether lawyers, law drafters or lay people should reconsider their options when confronted with legal texts that need to be translated. At the moment and even in the near future, MT is suitable to figure out the gist of the text rather than using it to produce legal official documents.

## **Bibliography**

- Aiken, M., & Ghosh, K. (2009). Automatic translation in multilingual business meetings. *Industrial Management & Data Systems*, 109(7), 916-925.
- Aiken, M., Ghosh, K., & Wee, J. (2009). An evaluation of the accuracy of online translation systems. *Communications of the IIMA*, 9(4), 67-84.
- Alcaraz, E. & Hughes, B. (2002). *Legal Translation Explained*. Manchester: St. Jerome Publishing.
- Al-Salman, S. (2004). The Effectiveness of machine translation. *International Journal of Arabic-English Studies*, 5(1), 145-160.
- Abdul-Hameed, Y. (2008). Evaluation of machine translation systems: The translation quality of three Arabic systems. *In proceedings of the XVIII FIT World Congress*. Shanghai: China, 1339-1352.
- Abu-Al-Sha'r, M. A., & Zughoul, R. M. (2009). Arabic-English-Arabic online machine translation: An analytical and evaluative study. *Turjuman*, 18(2), 33-83.
- Al-Dabbagh, U. (2010). *Free Online Machine Translation Services as Providers of Gist Translation: A Fiction or a Reality*. Paper presented at the 2nd Jordan International Conference on Translation, 30th November. Amman, Jordan: Petra University.
- Al-Dabbagh, U. (2013). Assessing the Quality of Free Online Machine Translation Services. *International Journal of Translation-IJT*, 25(2).
- Arnold, D., Balkan L., Humphreys L., Meijer S. & Sadler L. (1994). *Machine Translation: An Introductory Guide*. Manchester: Blackwell.

- Arnold, H. (2003). Why translation is difficult for computers? In H. Somers (Ed.), *Computers and Translation: A Translator's Guide* (pp. 119-142). Philadelphia: John Benjamins Publishing Company.
- As-Safi ,A.B. (1996). Towards an Objective Assessment of Literary/Belletristic Translation. *Translatio*, 15(1), 5-20.
- As-Safi, A.B. (1972). Concord and Peculiarities in English and Arabic, a Contrastive Analysis. *Bulletin of College of Arts*. Basra: Basra University 1(7) 1-37.
- As-Safi, A.B. (2011). *Translation Theories, Strategies and Basic Theoretical Issues*. Amman: Dar Amwaj.
- Bar-Hillel, Y. (1960). The Present Status of Automatic Translation of Languages. *Advances in Computers* 1, 91-163.
- Bennet, S., & Gerber, L. (2003). Inside Commercial Machine Translation in H. Somers (Ed.), *Computers and Translation: A Translator's Guide* (pp. 175-191). Philadelphia: John Benjamins Publishing Company.
- Bennet, S. (2003). The relevance of linguistics for Machine Translation in S. Harold (Ed.), *Computers and Translation: A Translator's Guide* (pp. 143-161). Philadelphia: John Benjamins Publishing Company.
- Beyer V., & Conradsen, K. (1995). Translating Japanese Legal Documents into English. In M. Moris (Ed.), *Translation and the Law* (pp145-179). Philadelphia: John Benjamins Publishing Company.
- Bowker, L. (2000). A Corpus-Based Approach to Evaluating Student Translations. *The Translator*, 6(2), 183–210.
- Brennan, M., & Richard B. (1997). *Equality before the Law: Deaf People's Access to Justice*, Durham: Deaf Studies Research Unit, University of Durham.

- Brennan, M. (1999). Signs of Injustice. *The Translator*, 5(2), 221–246.
- Brown, P. F., Della Pietra, S. A., Della Pietra, V. J., & Mercer, R. L. (1993). The Mathematics of Statistical Machine Translation: Parameter Estimation. *Computational Linguistics*, 19(1), 263-311.
- Butler, C. (1985). *Computers in Linguistics*. Oxford: Basil Blackwell Ltd.
- Cancedda, N. Dymetman, M. Forster, G. & Goutte, C. (2009). A Statistical Machine Translation Primer. In C. Goutte, N. Cancedda, M. Dymetman, & G. Foster (Eds.), *Learning Machine Translation* (pp1-45). Massachusetts: MIT Press.
- Cao, D. (2007). *Translating Law*. Clevedon: Multilingual Matters.
- Charrow, R. P. & Charrow, V. R. (1979). *Making Legal Language Understandable: A Computational Approach*. Oxford: Oxford University Press.
- Crystal, D., & Davy, D. (1969). *Investigating English Style*. London: Longman.
- Danet, B. (1985). *Legal Discourse: Handbook of Discourse Analysis*. London: Academic Press.
- Doug, A. Balkan, L. Lee Humphreys, R. Meijer, S. Sadler, L. (1994). *Machine Translation: An Introductory Guide*. Manchester: Blackwell.
- Durieux, C. (1991). “Liberté et créativité en traduction technique”. In M. Lederer & F. Israel (Eds.), *La liberté en traduction* (pp. 169–179). Paris: Didier Erudition.
- El-Desouki, A., Abd Elgawwad, A., & Saleh M. (1996). A Proposed Algorithm for English-Arabic Machine Translation System. *In Proceedings of the 1<sup>st</sup> KFUPM Workshop on Information and Computer Sciences (WICS)*. pp 66-72
- El-Saka, T., Rafea, A., Rafea M., & Madkour, M. (1999). English to Arabic Knowledge Base Translation Tool. *In Proceedings of the 7th International Conference on Artificial Intelligence Applications*, Cairo. Pp 66-72

- Fakhouri, M. (2008). *Legal Translation as an Act of Communication: the Translation of Contracts between English and Arabic* (Unpublished master's Thesis). An-Najah National University, Nablus, Palestine.
- Fargal, M. & Shunnaq, A. (1992). Major Problems in Legal Translation. *Babel*, 38(4), 203-210
- Faruqi's Law Dictionary English-Arabic* by Harith Faruqi Beirut: Librairie du Liban
- Gale, W., Church, K.W., & Yarowsky, D. (1992). A Method for Disambiguating Word Senses in a Large Corpus. *Computers and the Humanities*, 2(6), 415-439.
- Gustafsson, M. (1975). *Some Syntactic Properties of English Law Language Translation*. London: Academic Press.
- Harvey, M. (2002) Incongruity of Company Law Terms. *Translation Journal*. Retrieved from <http://www.translationdirectory.com/article1185.htm>.
- Hatim, B. (1997). *English-Arabic/Arabic-English Translation: A Practical Guide*. London: Saqi Books.
- Hatim, B., Shunnaq, A., & Buckley R. (1995). *The Legal Translator at Work: A Practical Guide to Arabic English Legal Translation*. Irbid, Jordan: Dar Al-Hilal for Translation and Publishing.
- Hiltunen, R. (1990). *Chapters on Legal English. Aspects Past and Present of the Language of the Law*. Helsinki: Suomalainen Tiedeakatemia.
- Hutchins, W.J., & Somers, H.L. (1992). *An Introduction to Machine Translation*. London: Academic Press.
- Hutchins, W.J. (1994). Machine Translation History. In R. Asher (Ed.), *the Encyclopedia of Languages and Linguistics* (pp. 2322-2332). Oxford: Pergamon Press.

- Hutchins, J., Hartmann, W. & Ito, E. (2005). *Compendium of Translation Software*. Retrieved from <http://ourworld.compuserve.com/homepages/WJHutchins/Compendium-11.pdf>.
- Ibrahim, M. (1991). *A Fast and Expert Machine Translation System Involving Arabic Language*, Ph. D. Thesis, United Kingdom: Cranfield Institute of Technology.
- Ide, N., & Veronis, J. (1998). Introduction to the Special Issue on Word Sense disambiguation: the state of the art. *Computational Linguistics*, 24(1), 1-40.
- James, C. (1980). *Contrastive Analysis: Applied Linguistics and Language Study* London: Longman
- Jurafsky, D., & Martin D. (2000). *Speech and Language Processing: An Introduction to Natural Language Processing, Computational Linguistics, and Speech Recognition*. New Jersey: Prentice Hall.
- Katz, J.J., & Foder, J.A. (1963). The Structure of a Semantic Theory. *Language*, (3)9, 170-210.
- Lederer, M. (2003). *Translation: The Interpretive Model*. Manchester: St. Jerome Publication.
- Lee, S. Y. (2004). Morphological Analysis for Statistical Machine Translation. *In the Proceedings of the Human Language Technology, Conference of the North American Chapter of the Association For Computational Linguistics*, pp. 57-60
- Lesk, M. (1986). Automatic Sense Disambiguation: How to Tell a Pine Cone from an Ice Cream Cone. *In Proceedings of the 1986 SIGDOC Conference*. New York: Association for Computing Machinery, pp. 55-75.
- Lyons, J. (1995). *Linguistic Semantics: an Introduction*. Cambridge: Cambridge University Press.

- Maalej, Z. (1994). English-Arabic Machine Translation of Nominal Compounds. *In Proceedings of the Workshop on Compound Nouns: Multilingual Aspects of Nominal Composition*, Geneva.
- Manning, D.C., & Shutze, H. (2002). *Foundation of Statistical Natural Language Processing*. Massachusetts: The MIT Press.
- Mattila, H. (2007). *Comparative Legal Linguistics*. Burlington: Ashgate Mellinkoff.
- Mckay W.R., Helen E., & Grant B. (2005). *Legal English: How to Understand and Master the Language of Law*. London: Pearson Education Limited.
- Merriam-Webster's Advanced Learner's English Dictionary* (2008) Springfield, Massachusetts: Merriam-Webster Incorporated.
- Mokhtar, H.M.O., Darwish, N.M., & Rafea, A. (2000). *An Automated System for English-Arabic Translation of Scientific Text (SEATS)*. In proceedings of MT2000: Machine Translation and multilingual applications in the new millennium, the British computer society (BCS), London, pp1-5.
- Nida, E. A. (1998). Translators' Creativity Versus Sociolinguistic Constraints. *In Proceedings of the 9th International Conference on Translation and Interpreting*, Prague, September 1995, 127–36. Amsterdam/Philadelphia: John Benjamins.
- Nirenburg, S. (1987). Knowledge and Choices in Machine Translation. In S. Nirenburg (Ed.), *Machine Translation: Theoretical and methodological issues* (pp. 1-21) Cambridge: Cambridge University Press.
- Och, F. (2009). 51 Languages in *Google Translate*. Google Research Blog. <http://googleresearch.blogspot.com/2009/08/51-languages-in-google-translate.html>.
- Rafea, A., Sabry, M., El-Ansary R., & Samir, S. (1992). A Machine Translator for Middle East News. *In Proceedings of the 3rd International Conference*

- and Exhibition on Multi-lingual Computing*, the Centre for Middle Eastern and Islamic Studies, University of Durham, Durham, Uk, pp. 55-60.
- Ravin, Y., & Leacock, C. (2000). Polysemy: Theoretical and Psycholinguistic Study of Jury Instructions. *Columbia Law Review*, (7)9, 1306-1374.
- Robert, A. (1964). *Introductory Linguistics*. Philadelphia: Chilton Company.
- Robins, R.H. (1964). *General Linguistics: An Introductory Survey*. London: Longman.
- Saqf Al-Hait, A. (2009). *The Reliable Guide to Legal Translation*. Amman: Dar Al- Thaqa for Publishing and Distributing.
- Sarcevic, S. (1989). Conceptual Dictionaries for Translation in the Field of Law. *International Journal of Lexicography*, 2(4), 279-290.
- Sarcevic, S. (2000). *New Approach to Legal Translation*. The Hague: Kluwer Law International.
- Shalan, K. (2000). Machine Translation of Arabic Interrogative Sentence into English. In *Proceedings of the 8th International Conference on Artificial Intelligence Applications*, Cairo: Egyptian Computer Society (EGS), pp. 473-483.
- Shen, E. (2010). [Comparison of Online Machine Translation Tools](http://web.archive.org/web/20110210034229/http://tcworld.info/index.php?id=175). Retrieved from <http://web.archive.org/web/20110210034229/http://tcworld.info/index.php?id=175>.
- Smith, S. A. (1995). Cultural Clash: Anglo-American Case Law and German Civil Law in Translation. In Marshall Morris (Ed.), *Translation and the Law* (pp. 179–200). Amsterdam: John Benjamins.
- Somers, H. (2003). Introduction. In H. Somers (Ed.), *Computers and Translation: A Translator's Guide* (pp. 1-13). Philadelphia: John Benjamins Publishing Company.
- Suliati Arabic Corpora (2013). Retrieved from [http://www.comp.leeds.ac.uk/latifa/arabic\\_corpora.htm](http://www.comp.leeds.ac.uk/latifa/arabic_corpora.htm).

- The Law Dictionary* (2013). Featuring Black's Law Dictionary Free Online Legal Dictionary 2nd Ed. Retrieved from <http://thelawdictionary.org/>
- Tucker, B. A. (1978). Current Strategies in Machine Translation Research and Development. In S. Nirenburg (Ed.), *Machine Translation: Theoretical and methodological issues* (pp. 22-41). Cambridge: Cambridge University Press.
- White, J.S. (2003). How to Evaluate Machine Translation. In H. Somers (Ed.), *Computers and Translation: A Translator's Guide* (pp. 211-245) Philadelphia: John Benjamins Publishing Company.
- Wikipedia (2013). Retrieved from <http://en.wikipedia.org/wiki/Boilerplate>

## المراجع العربية

- الوسيطي, سليمان (2005). "التفاعل ما بين الإنسان و الآله" متاحة على شبكة الأنترنت على [www.acatap.htmlplanet.com/arabization-j/accessories/Jour-4.html](http://www.acatap.htmlplanet.com/arabization-j/accessories/Jour-4.html)
- عبدو, د (2007) "المقاربه بين الأسم الموصول و الضمير في اللغة الأنجليزية و العربية و كيف يستفاد من النتائج في الترجمة الآلية", *بحوث أطلس للدراسات و البحوث*. المجلد (2) العدد (1) 91-128.
- النجار, ماجد (1999) "علم اللغة و الترجمة الآلية", *دراسات الترجمة*, 1.
- النجار, ماجد (2004) "بناء معجم لغوي للترجمة الآلية من اللغة الأنجليزية ألى اللغة العربية", *بحوث أطلس للدراسات و البحوث*. المجلد (2) العدد (2) 80-115.
- الصافي, عبد الباقي (2004) "الترجمة الآلية: الواقع و التطورات" *بحوث أطلس للدراسات و البحوث*. المجلد (2) العدد (2) 207-227.
- صبره, محمود (1995) *ترجمة العقود*. مصر: دار الكتب القانونية.

# Appendices

## Appendix 1

The test set (input) along with Google Translate output and the proposed output by the researcher

Input (authentic SL)	Output (GT translation)	Proposed Output
1. Seller desires to sell the Property to Buyer and Buyer desires to purchase the Property from Seller upon and subject to the terms and conditions set forth in this Agreement.	البائع يرغب في بيع العقار للمشتري والمشتري لرغبات المشتري ولشراء العقار من البائع عليها وتخضع لشروط والشروط المنصوص عليها في هذا الاتفاق.	يرغب البائع ببيع العقار للمشتري و يرغب المشتري بشراء العقار من البائع وفقا للشروط المبينة في هذه الاتفاقية.
2. Upon execution of this Agreement, Buyer shall pay to Escrow Holder (hereinafter defined) as a deposit against the Purchase Price, the sum of \$..... (the Deposit).	عند تنفيذ هذا الاتفاق، يتعين على المشتري دفع لحامل الضمان (الذي يعرف فيما بعد) كتأمين ضد ثمن الشراء، مبلغا وقدره ..... (إيداع).	فور تنفيذ هذا الاتفاق، يدفع المشتري لحامل السند العالق للملكية (المحدد بياناته في ما بعد) مبلغ ..... \$ كعربون لسعر الشراء (العربون).

<p>3. Seller has made no commitment to any tenant, governmental or quasi-governmental entity or other person or entity which commitment relates to the Property or imposes upon Seller or the successors or assigns of Seller any obligation to pay or contribute property or money or to construct, install or maintain any improvements on or off the property.</p>	<p>البائع لم يبد التزاما على أي المستأجر، جهة حكومية أو شبه حكومية أو أي شخص آخر أو كيان التزام يتعلق الملكية أو تفرض على البائع أو المتنازل لهم أو خلفاء من أي التزام البائع عن دفع أو المساهمة أو الممتلكات أو الأموال لبناء وتركيب أو الإبقاء على أي تحسينات أو إيقاف الخاصة.</p>	<p>لم يقدم البائع اي تعهد لأي مستأجر أو جهة حكومية او شبه حكومية او شخص آخر أو جهة أخرى لها علاقة بالعقار او يفرض على البائع او من يخلفه اي التزام بدفع او المساهمة بممتلكات أو بناء أو تركيب أو صيانة اي تحسينات على العقار او خارجه.</p>
<p>4. In witness whereof the said parties have hereunto set their hands and seals the day and year first above written.</p>	<p>وإثباتا لذلك قال الأطراف توقيعي هنا في أيديهم والأختام واليوم والسنة أعلاه أول من كتب.</p>	<p>وأثباتا لما تقدم، صادق الطرفان هذا العقد بأمضائيهما في اليوم و السنة المذكورين في بداية العقد أعلاه.</p>
<p>5. The Letter of credit is to be confirmed by an acceptable prime U.S. bank and payable over its counters at sight against presentation of documents as described below and with the conditions allowed herein.</p>	<p>خطاب الاعتماد هو تأكيد من أحد البنوك المقبولة الولايات المتحدة رئيس وتدفق على مدى عدادات لها في الافق ضد تقديم المستندات كما هو موضح أدناه وفقا للشروط الواردة يسمح.</p>	<p>يتم المصادقة على خطاب الاعتماد من بنك امريكي رئيسي مقبول ويكون الخطاب قابل للدفع مباشرة فور تقديم المستندات الموضحة أدناه و حسب الشروط المسموحة في هذا العقد .</p>

<p>6. As soon as any event or circumstance of Force Majeure takes place, the party so affected shall give notice thereof to the other party, giving full details thereof, including but not limited to the nature of the event or circumstance, the date of its having occurred, the anticipated duration thereof, the likely effect of the Force Majeure event or circumstance, and so on.</p>	<p>بمجرد حصول حدث أو ظرف القاهرة يحدث أو ظرف القوة القاهرة يحدث، تتاح للطرف المتضررين من ذلك تقديم إخطار عن ذلك إلى الطرف الآخر، وإعطاء التفاصيل الكاملة منها، بما في ذلك ولكن لا تقتصر على طبيعة الحدث أو ظرف، واليوم الذي حدثت بعد والمدة المتوقعة منه، والأثر المحتمل لقوة القاهرة أو ظرف، وهلم جرا.</p>	<p>مجرد حصول حدث أو ظرف ذي قوة القاهرة، يقدم الطرف المتضرر إخطار للطرف الآخر بذلك الحدث أو الظرف بحيث يتضمن ولا يقتصر على طبيعة الحدث أو الظرف و تاريخ حدوثه و المدة المتوقعة من استمرار الحدث والآثار المترتبة على القوة القاهرة و غير ذلك.</p>
<p>7. The duration of this rental agreement shall be from August 27 1983 to and including 12 p.m. on May 31 1984.</p>	<p>تكون مدة هذه الاتفاقية يكون من تأجير 27 أغسطس 1983 لوبينهم 12 من مساء يوم 31 مايو 1984.</p>	<p>يبدأ هذا الاتفاق التأجيري من تاريخ 27 أغسطس 1983 و لغاية الساعة الثانية عشرة ظهرا يوم 31 مايو 1984.</p>
<p>8. Landlord shall comply with the requirements of applicable building and housing codes materially affecting health and safety.</p>	<p>يجب المالك الامتثال للمتطلبات المعمول بها وبناء المساكن رموز يؤثر تأثيرا جوهريا على الصحة والسلامة.</p>	<p>يلتزم المالك بمتطلبات قوانين البناء و الاسكان و التي تؤثر تأثيرا جوهريا على الصحة و السلامة.</p>

<p>9. Utilities and services: charges of electricity, gas, water, sewer and garbage shall be paid for by Tenant, Tenant hereby acknowledges that Landlord, or the person authorized to enter into this Rental Agreement on Landlord's behalf, has fully explained to Tenant the utility rates, charges and services for which Tenant will be required to pay, other than those to be paid by tenant directly to the utility company furnishing the service.</p>	<p>المرافق والخدمات: يتم دفع رسوم الكهرباء والغاز والمياه والصرف الصحي وجمع ونقل النفايات من قبل المستأجر، المستأجر يقر بموجب هذا المالك، أو الشخص المفوض للدخول في هذا الاتفاق المستأجر نيابة عن المالك، فقد أوضح تماما المستأجر معدلات فائدة والرسوم والخدمات التي سوف تكون هناك حاجة المستأجر على الدفع، غير تلك التي يجب أن تدفعها المستأجر مباشرة إلى الأداة المساعدة تأثيث الشركة الخدمة.</p>	<p>المرافق و الخدمات: يدفع المستأجر مصاريف الكهرباء و الغاز و الماء و المجاري و النفايات و يقر المستأجر بأن المالك أو الشخص المفوض من قبل المالك قد وضّح للمستأجر مصاريف و رسوم خدمات المرافق الواجب دفعها من قبل المستأجر بالإضافة الى الرسوم المدفوعة من قبل المستأجر مباشرة لشركة الخدمات المزوده للخدمة.</p>
<p>10. Summary of Basic Terms: As used in this Lease, the following terms shall have the meaning set forth below subject to the qualifications, adjustments and exceptions set forth elsewhere in this Lease.</p>	<p>ملخص لشروط الأساسية: كما هو مستخدم في هذا العقد، يكون للكلمات التالية المعاني المبينة أدناه تخضع لتعديلات والمؤهلات والاستثناءات المنصوص عليها في أي مكان آخر في هذا العقد.</p>	<p>ملخص للمصطلحات الأساسية: كما هو موضح في هذا العقد يعني بالمصطلحات التالية المعاني المبينة أدناه و الخاضعة للشروط والتعديلات و الاستثناءات المبينة في كافة اجزاء العقد.</p>

<p>11. Tenant shall use the premises only for the permitted use and shall not use or permit the Premises to be used for any other purpose.</p>	<p>يجب استخدام المباني المستأجر فقط لاستخدام مسموح به، ويجب ألا تستخدم أو تسمح المباني لاستخدامها في أي غرض آخر.</p>	<p>يستخدم المستأجر المبنى للغرض المسموح به فقط و يمنع استخدامه أو السماح بأستخدامه لأي غرض اخر.</p>
<p>12. Where subscriber is licensed to disseminate the service, Subscriber shall whenever possible credit any report supplied by Reuters and republished by Subscriber as "Reuters".</p>	<p>حيث تم ترخيص مشترك لإعادة نشر الخدمة، يقوم المشترك كلما أمكن أي تقرير الائتمان المقدمة من رويترز ونشرها من قبل المشترك بأنه "رويترز".</p>	<p>في الأحوال التي يرخص فيها للمشارك بأعادة نشر الخدمة, ينسب المشارك الى رويترز كلما كان ذلك ممكنا أي خبر كان مصدره رويترز و أعاد نشره المشارك</p>
<p>13. Subscriber shall indemnify and hold harmless Reuters for any loss or damage or bodily injury caused to or by the equipment except where such loss, damage or bodily injury is caused solely by the negligence of Reuters or its contractors.</p>	<p>يجوز للمشارك الإضرار، وتعويض رويترز عن أي خسارة أو ضرر أو إصابة بدنية تلحق أو بواسطة المعدات إلا إذا كانت تسبب مثل هذه الخسارة أو الضرر أو الأذى البدني فقط عن إهمال من رويترز أو المتعاقدين معها</p>	<p>يعوض المشارك رويترز عن أي خسارة مادية أو إصابة جسدية ما عدا الأصابات الجسدية أو الخسائر الناتجة عن الأهمال من قبل رويترز أو مقاوليها.</p>

<p>14. In the event of any breach of this Agreement by Subscriber which cannot be remedied or is not remedied within thirty days of subscriber being requested to do so, Reuters, in addition to any other remedy it may have, may terminate this agreement immediately and recover the Equipment(if supplied by Reuters).</p>	<p>في حال وجود أي خرق لهذه الاتفاقية من قبل المشترك والتي لا يمكن علاجها أو لم يتم علاجه خلال ثلاثين يوما من تاريخ مشترك المطلوبة للقيام بذلك، رويترز، بالإضافة إلى أي علاج آخر قد تكون لديها، قد إنهاء هذه الاتفاقية فورا واسترداد المعدات (إذا قدمتها رويترز).</p>	<p>في حال أي خرق لهذا الأتفاق من قبل المشترك ولايمكن علاج هذا الخرق أو يتم علاجه خلال ثلاثين يوما من تاريخ تبليغ المشترك بذلك فإن لرويتز الحق بفسخ هذا الأتفاق بالإضافة الى اي تعويض تراه مناسبا و أسترجاع المعدات (أذا تم تزويدها من قبل رويترز).</p>
--	--	--

## Appendix 2

### Buckwalter transliteration chart

هـ	h	ذ	*	ل	l
أ	a	ر	r	م	m
آ	ya	ز	z	ن	n
واو	w	س	s	هـ	h
ا	a	ش	sh	و	w
ق	q	ص	s	ي	y
ا	a	ض	d	ي	y
ب	b	ط	t	ف	f
پ	p	ظ	z	ن	n
ت	t	ع	e	ك	k
ث	th	غ	g	ا	a
ج	j	ا	a	u	u
ح	h	ف	f	i	i
خ	x	ق	q	ع	e
د	d	ك	k	و	o