Google Translate Service: Transfer of Meaning, Distortion or Simply a New Creation?

An Investigation into the Translation Process & Problems at Google

Ahmed Abdel Azim ElShiekh

Associate Professor of Linguistics & Translation, Zarqa University, P.O. Box: 132014

Zarqa University, 13132, Zarqa, Jordan, On Secondment from Alexandria University, Egypt

Tel: 079-595-5838   E-mail: shishuj @gmail.com

Abstract

Internet users have the choice between several machine translation services that can automatically translate a given text or website in another language. Google Translate is one of the most popular services of this kind. It allows web users to translate text or websites into 51 languages. The present research aims at exploring the nature of the translation process provided by Google Translate Service with an eye on the demerits of the mechanisms used by Google as well as the possibility of refining the system used so as make a better use of it on the part of the average Internet user who is by no means a professional translator and, hence, may tend to take Google translation for granted even where it could be totally catastrophic. The scope of the research is confined to Translation from Arabic into English and vice versa. The data used in this study is based on short translation assignments that were presented by Zarqa University students in the Department of English & Translation in their Research Writing Course during the Summer Semester 2010 under the supervision of the research.

Keywords: Machine Translation, Semantic problems, Arabic vowelings, Google Translator Service

1. Introduction

1.1 Significance of the Research

Machine translation or MT, with its various forms and types, have increasingly become a subject of interest to both the layperson/average who seeks the translation of a given text/s as a means to another end and the translation specialist who attempts to conduct an academic research on MT as such. Apart from commercial MT programmes such as Wafi or Kafi, there have recently been an increasing number of Web Sites that offer the service of an automatic/machine translation of individual sentences or even whole texts. Internet users have the choice between several machine translation services that can automatically translate a given text or website in another language. Google Translate is one of the most popular services of this kind. It allows web users to translate text or websites into 51 languages. Google has recently announced that they had expanded their translation services by offering a translate gadget for webmasters, by which the latter can integrate the code of the Google Translate gadget into their website to offer visitors the option to automatically translate the text that is displayed on the website into a different language. In this research, however, the interest is confined to the general Google Translate Service which is available for all Internet users.

1.2 Objectives of the Research

The present research aims at (1) exploring the nature of the translation process provided by Google Translate Service with an eye on the demerits of the mechanisms used by Google and (2) investigating the possibility of refining the system used so as to make a better use of it on the part of the average Internet user who is by no means a professional translator and, hence, may tend to take Google translation for granted even where it could be totally catastrophic.

1.3 Scope of the Research

This research is confined to Translation from Arabic into English and vice versa in the field of advertisements, Koranic texts and literary works.

1.4 Tools & Data Collection

The data used in this study is based on short translation assignments that were presented by Zarqa University students in the Department of English & Translation in their Research Writing Course during the Summer Semester 2010 under the supervision of the researcher. The course in question is a final semester course according to Zarqa University regulations, and, hence, students are supposed to be on the verge of graduation and, therefore, reasonable competent in English and Arabic. This, however, is, indeed, not a decisive factor where the present research is concerned, since the students' role is basically confined to supplying the data rather than analyzing it. The analyses they presented in their research papers were only the product of undergraduate students on their first attempt to conduct an academic research in the field of translation per se. The source language texts in question are of three major types: Advertisements- Literary Styles- Religious texts (See Appendix 1). They consist of Arabic as well as English texts. The target language versions produced by Google Translate Service are also attached in Appendix 2.

2. A Snapshot at Theoretical Background

2.1 Machine Translation

Machine translation, in computational linguistics, publishing, and other fields, is a term that is used to refer to the use of computers to conduct large-scale translation operations. The electronic translation of one language into another or the electronic syntactic analysis of a text has been attempted since the mid 20th cent. However, the complexities of this type of operation, both practical and theoretical, have resulted in only a limited measure of success (Note 1).

A machine translation system usually consists of linguistic descriptions of the source and target languages (automatic vocabularies and formal grammars at all levels) and an algorithm (instructions for using the vocabularies and grammars, oriented only to their form), on the basis of which the translation itself is performed. The complete process of machine translation consists of the following principal stages: (1) analysis of the texts in the source language (search for words in the vocabulary, as well as morphological and syntactical analysis—that is, simulation of comprehension of the text), (2) conversion (transfer from the structure of the text in the source language to the structure of the text in the target language), and (3) synthesis of the text in the target language (syntactical and morphological presentation of the text—that is, simulation of construction of the text). In actual machine translation systems all these stages may be closely interrelated, and some may be absent.

The machine translation algorithm is usually performed by a general-purpose digital computer. The text produced as a result of the machine translation may be edited by a human, a "post-editor", who corrects mistakes and ambiguities in the translation. Here is an example, in general terms, of translation of the sentence "He was seen at 6 o'clock" from English into Russian according to the stages described above. In the analysis stage it is determined that "he" is the subject, "was seen" is the predicate (the verb "see" in the past indefinite, indicative mood, passive voice), and "at 6 o'clock" is an adverbial modifier of time. In the conversion stage, Russian translations are placed in correspondence to the English words and word combinations: "he" -» *on,* "see" -» *videt'* "6 o'clock" -> *shest' chasov*. Since the verb *videt'* is not used in the passive voice, the English passive construction is converted into a Russian impersonal construction: *on* becomes the direct object (*ego*) of the predicate (the verb *videt',* imperfective aspect, indicative mood, active voice, past tense, plural). In the synthesis state the case and prepositional markers of syntactical connections among words are worked out; in particular, the preposition "at" is translated into "*v* + accusative case" as an indicator of a time modifier, and *on* as a direct object is given the marker "accusative case." The word order is determined, and then the necessary forms of the words are shaped so that the result is *Ego videli v 6 chasov*. If the initial sentence had contained the English pronoun "it" instead of "he," ambiguity could have arisen during translation (without considering preceding sentences); *Ego videli* ... (if "it" is an airplane *[samolet,* masculine]), *Ee videli* . . . (if "it" is a rocket *[raketa,* feminine]), or *Eto videli*. (if "it" is an event or phenomenon, neuter). In this case the human post-editor can select the correct version (Note 2).

As our environment becomes more networked and connected internationally, the call for MT increases. Researchers predict that in the very near future English will no longer be the mother tongue of the majority of Internet users. According to *Wired*, Forrester Research estimates that by 2003 Americans will account for only one third of Internet users worldwide. ("Hello, World," *Wired*, May 2000).

Already the amount of material needed in different language versions is too vast for human translation alone, according to Systran, one of the oldest machine translation companies. MT is a long way from being able to replace

human translation and many experts feel it may never do so. But it can reduce the amount of work for human translators by taking over translations where accuracy is not essential, and by assisting humans with more important translation jobs.

MT offers some real advantages: according to Systran, MT is much faster than human translation (humans can translate 2000 - 3000 words a day, whereas Systran's MT software can translate 3700 words a minute). MT is much cheaper than human translation. MT software, indeed, has a much better memory than human translators: it can store translated documents and re-use phrases that have already been translated.

The accuracy of MT is much lower than competent human translation, but may be improved in various ways – for example, by making sure that spelling and punctuation are all correct in the original text. When used in conjunction with human translators, the main objective of MT is to provide a first draft which is then given to a human translator for editing and polishing. In that latter case, MT helps save much time, effort and money (Note 3).

2.2 Types of Machine Translation

There are basically three major types of machine translation programmes. These may be summed up as follows:

1- Fully automatic high quality translation (FAHQT)

2. Machine-aided human translation (MAHT)

3. Human-aided machine translation (HAMT)

2.3 Google Translate Service

Google Translate Service is one of the most popular computer-aided translation services, however, using an online-translator for individual lexical items, sentences and even full texts.

The following questions arise:

- How efficient and/or deficient are the target language texts produced by Google Translate Service?

- What are the most common problems and demerits that characterize that translation service?

- How does that translation application function?

The present research is an attempt to find answers to these questions.

2.4 Hypotheses of the Research

In accordance with the literature on machine translation problems whether on the lexical, syntactic, morphological or semantic levels, one would expect major problems first of all on the semantic level in particular wherever there is some type of ambiguity whether lexical or syntactic. Second, where some colloquial expressions are used, particularly in Arabic advertisements, it is expected that most MT programmes available would probably be confused by such terms. Finally, specialized technical terms may also present a problem to MT programmes that have been mainly designed to serve the average end-user in general. Otherwise, some varied problems of individual nature may crop up here or there. But what has the practical experience actually revealed? That is the basic concern of this paper.

3. Discussion & Findings

In this section, the researcher shall present a brief review of the major shortcomings observed in Google Translate Service translations of the various language registers under study. The review and discussion deals with each language register separately and on the grammatical/morphological level, the lexical/semantic level and, when required, on the stylistic/communicative level as well. The division and organization of this section is governed by the limitations of time and space as well as the scope of the present research as identified in the introduction.

3.1 Advertisements

The translation of advertisements usually represents a challenge for the translator, with the persuasive style of advertisements in general and the possibly different linguistic tools employed by different languages to achieve one and the same final effect; i.e., the

3.1.1 Problems on the Lexical/Semantic Level

This section discusses the semantic problems the machine translation programme has faced while translating the texts. The meaning of a word is for the most part based on its sense relationships towards other words surrounding it in a semantic field or by the 'role' it fulfills within the action described within a sentence. These sense relationships (synonymy, hierarchical Relationships: hyponymy and meronomy, antonymy, homonymy, polysemy, and collocation), which are either paradigmatic (Note 4) or syntagmatic (Note 5), are obviously a source of translation

problems as there will hardly be similar sense relationships between words of different languages even though they have the same meaning (Yule, 1996). Some semantic-translation problems are related to ambiguity or oligosemy (when a SL item has a particularly restricted range of meaning that it may not be possible to match this restriction in the TL) while others are related to collocations and/or the differences between the "referential", "cognitive" and/or "descriptive" meaning (i.e. the words as symbols which refer to objects, events, abstracts and relations)on the one hand, and the "contextual", "emotive", "evaluative", "expressive", "socio-expressive" and/or pragmatic meaning.

A) Ambiguity (Note 6)

'It is difficult to know, in translation, how accurately one ought to reproduce elements of vagueness or ambiguity apparently present in the original (Aquila, 1982: 11).

- In the first advertisement the machine translation programme considered the lexical item "أطْلِق" which is used as an imperative verb in the advertisement, as the simple past of the verb "يطلق".

- In the first part of the third advertisement, the machine translation programme considered the lexical item "عبّر" which is used as a command verb in the advertisement, as the noun "عَبْرَ". In the second part when the researcher used the full stop the machine translation programme considered it as a noun also but didn't use the same translation.

- The machine translation programme considered the lexical item "مين قدك" in the third advertisement as a colloquial item which is used for flirting in the target culture, while the Arabic expression is more general in nature.

- محتاج .= needy

- The Arabic lexeme محتاجis an ambiguous term, because in the colloquial variety it means "Do you need?" while in MSA it means "Needy". It was the second meaning that the Google MT programme has picked.

B) Cases of Unjustifiable Use of Transliteration

- The lexical item "المَي" in the second advertisement was transliterated into "Elmi" because it is a colloquial term that is not known to machine translation programme, so it considered that this lexical item a proper noun.

- The lexical item "بيعطيك" in the fifth advertisement was transliterated into "Biattiyk" .machine translation programme did not recognize the prefix "ب" and the suffix "ك", so it considered that this term a proper noun.

- The lexical item "مافي" in the sixth advertisement was transliterated into "Mafi" because it is a colloquial term that is not known to the machine translation programme. So it considered that this lexical item is a proper noun.

- The SL terms = وبس+مفيش are transliterated because these are colloquial terms that seem to be missing in the database of the Google MT programme. As for the second Arabic item in the list, i.e. = الالون ***, which is also an SMA lexical item, it was already misspelled in the original SL advertisement. Hence it was also transliterated instead of translated properly into "Colours".

- (أ) = الالوان

- وبس.= 1ups

- الالون .alolon =

- مفيش.mvic =

3.1.2 Problems on the Syntactic/Grammatical Level

This section discusses the syntactic problems the machine translation programme has faced while translating the texts.

- In the second advertisement, machine translation suggested that the agent is "المي" which transliterated into "Elmi" and the objective is "الغيرَة" which is translated into "jealousy".

- In the forth advertisement, machine translation did not recognize the bound morpheme "ك" which must be translated as "your".

- In the sixth advertisement, the whole sentence is not in the right order.

- In the translation of the Vodafone advertisement, the linkage verb is absent.

- Example: With service from Vodafone mobile Internet world of the Internet in your pocket

- The correct translation is = with service from Vodafone mobile Internet world of the Internet is in your pocket.

- To check this phenomenon and make sure it was not simply a coincidence, I presented the MT programme with another example of an Arabic structure that does not include a linkage verb on the superficial level, and again the result was that the linkage verb is absent in the TL text, The example was منى طالبة مجتهدة and it was translated into: Mona diligent student.

- (The correct translation is of course =Mona is a diligent student.)

- In the phrase "90 days is from the end of the grace period", we cannot use the verb "is" with plural noun "days" the correct is "days are".

3.2 Religious Texts

The religious texts included in this study are of two types; Hadith or Verbal Prophetic Tradition on the one hand and Koranic texts on the other hand. Both sub-registers present the best and most experienced translators with extremely serious problems. Hence, it is almost axiomatic that any machine translation programme is bound to be i9ncapable of producing an effective satisfactory translation of a Hadith or a Koranic text. Yet, the4 idea of including such registers or sub-registers was to compare and contrast the different types of difficulties involved in the translation of the three major registers under study. Indeed, a translation of a purely informative text may not present a real translation problem even for a mediocre automatic translation programme. But where expressive and or performative texts are concerned, translation difficulties and translation problems are not the exception as much as the norm.

3.2.1 Hadith (Sayings of the Prophet)

Hadith also spelled HADIT (ARABIC: NEWS, OR STORY) is the "record of the traditions or sayings of the Prophet Muhammad, revered and received as a major source of religious law and moral guidance, second only to the authority of the Qur'an, or scripture of Islam" (Encyclopaedia Britannica, 2000). Hadith is a typically problematic area for translation, with all its specialized terminology as well as all sorts of concepts that are particular to Islam and/or the Arabic culture in particular

3.2.1.1 Problems on the Lexical/Semantic Level

All in all, the automatic translation programme has failed to transfer the overall message displayed in the source language text. Following, however, is a brief survey of the major types of lexical and/or semantic problems involved.

A) Ambiguity

It was rather impossible for the machine translation programme to find equivalents to SL ambiguous items, whether the ambiguity is resulting from the actual original lexical item in the source text or from the inability of the MT programme to deal with Arabic vowelings. An example of the first case is the Arabic lexical item محارم which is a case of homonymy; it could mean female relatives but also means sacred boundaries drawn by God not to be violated by the believers. It is the second meaning that matches the context of the Hadith, The machine translation programme, though, opts for the first inappropriate meaning "female relatives". The other source of ambiguity that stems from the inability of the MT programme to deal with Arabic vowelings may be represented in the following. The Arabic حمى meaning protected area on the literal level is mistakenly read as حمّى meaning 'fever'.

B) Cases of Unjustifiable Use of Transliteration

As a well established translation tradition, regardless of whether the adopted approach is semantic or communicative in Newmark's terms, transliteration is resorted to when the source language item is either a proper noun or else representative of a source language concept that is clearly missing in the target language culture (Catford, 1965).

The Arabic lexical items الحرام and حلال were transliterated into Halal and Haram. This may represent a cultural problem, however, and, consequently shall be examined in rather more detail under the category of Cultural Problems per se.

C) Lexical Mismatches

The lexical item عَنْ in the phrase عَنْ أَبِيْ عَبْدِ اللهِ is translated as "*And*" *Abu Abdallah*, indicating the presence of a problem in the MT programme instructions with regard to the preposition عَنْ . The case remains an individual one in the TL text, and, as far as the researcher is concerned, it is inexplicable.

Similarly, the lexical item الشبهات in the context اتَّقَى الشُّبُهَاتِ was translated as *avoids doubtful matters* in a whole sentence, but when the researcher asked the Google Translate Service to translate it in isolation, it was rendered as

suspicions. In fact, the second option would have been quite a good contextual translation of the Arabic lexical item in question as used in the Hadith under study. There is no clear justification for the way the translation programme behaves with the very same lexical item in context versus in isolation. However, one can fairly conclude that the programme in question may be incapable of differentiating between different related meanings that fall within the range covered by the same lexical item. In other words, the programme is unable to deal with the semantic problem known as polysemy, which is hardly surprising, anyway. After all, all sorts of ambiguity, whether syntactic or lexical, represent serious translation problems for all MT software types**.**

3.2.1.2 Problems on Syntactic/Grammatical Level

The machine translation has not only failed to deal with Arabic vowelings, but also with the word order characteristic of English. Indeed, whole sentences seem to an Arabicized kind of English, echoing the word for word approach to translation according to Catford and/or Larson's very literal. In fact, the whole TL text consists of ungrammatical structures that seem to belong to English on a lexical level alone.

3.2.1.3 Problems on the Stylistic/Communicative Level

The style in the machine translation text is clumsy and not organized. The suggested translation is more or less typical of a word-for-word approach. It operates on the lexical level rather than the sentence level, let alone the textual one. The inevitable result is that the translated version lacks the communicative effect of the original passage in the source language, even where the lexical equivalents are properly established and the semantic message is more or less transferred. The following are cases in point.

The phrase اتَّقَى الشُّبُهَات was translated as *avoids doubtful matters*, which renders the lexical meaning and the overall semantic message of the original satisfactorily yet falls quite short of the SL stylistic effect.

Another example is وَقَعَ فِي الشُّبُهَات which is translated into *falls into doubtful matters*, which also manages to transfer the semantic message but in a purely prosaic style.

3.2.1.4 Cultural Problems

Cultural problems arise when some concepts in the source text are totally missing in the TL culture or at least confused with similar yet far from identical ones. Among the most common techniques used to handle such problem is usually transliteration (ElShiekh, 2010). Thus, in case the concept concerned is not really missing the TL culture, it would be inappropriate to resort to transliteration. There are two cases in point in the Google machine translation programme text.

The first example is the lexical item حلال in the context إِنَّ الْحَلاَلَ بَيِّن which is not really a cultural specific term. It is rather a concept that is common to all religions and, hence, of a universal nature. Therefore, it should have been translated simply into "lawful" or "legitimate". In fact, when I presented the Arabic lexical item in question separately to Google MT programme, to my surprise, it offered the possible equivalent "Lawful", yet it transliterated it as "Halal" when dealing with it in context.

The second example is the lexical item الْحَرَام in the context وَإِنَّ الْحَرَامَ بَيِّنٌ which is also certainly not a cultural specific term. A communicative equivalent could be "the forbidden" or "the prohibited", but the machine translation programme dealt with it as if it were a missing concept and, consequently, transliterated it as *Haram*.

3.2.2 The Holy Koran (Koranic Texts)

The Qur'an or *Koran* is the holy book for Muslims, revealed in stages to the Prophet Muhammad over 23 years. Koranic revelations are regarded by Muslims as the sacred word of God, intended to correct any errors in previous holy books such as the Old and New Testaments (Hence, even the mere idea of translating the Koran was often regarded by several Muslim scholars as inappropriate and sometimes even blasphemous. Thus, it is expected that the translation of Koranic texts involves serious translation problems on all levels, whether linguistic or cultural.

3.2.2.1 Problems on the Lexical/Semantic level.

A) Cases of Unjustifiable Use of Transliteration

To opt for the transliteration rather than translation of religious terms often implies the intention of focusing on what divides different religions instead of what may bring them closer to each other (See ElShiekh and Saleh, 2011). Even when it is not intentional, as could be the case with the Google Translate Service, the effect is still more or less the same. However, the situation is even more problematic according to the data discussed in the present research. There are quite a few cases of probably 'innocent' or 'unbiased' yet almost totally unjustifiable use of transliteration. The machine translation programme transliterated the verbal phrase هَوَى (contextually meaning "it fell") into "Hui"; the programme resorted to transliteration instead of picking one of two possible meanings of the source language

item (it could literally mean "it fell" as well as "it loved"). Furthermore, the transliterated target language product also reveals that the programme could not identify the Arabic vowelings and, thus, read the last letter as representative of the /i/ sound rather than the /a:/. So it considered this word as a proper noun. According to Arberry, it should have been "Plunges" which refers to the "النجم".

The machine translation programme transliterated the verbal phrase   فَتَدَلَّى   as "Vtdly"; the programme resorts to transliteration again, as it does not understand the meaning of this Arabic expression which refers to the word "وحي" because of the attachment of the conjunction "ف" at the beginning of the word. According to Arberry, it should have been "Suspended".

The machine translation programme transliterated the verbal phrase   أَفَتُمَارُونَهُ as "Ovtmarunh", as it does not understand the meaning of this word which refers to the godless people who have a dispute with the prophet -peace upon him- because of the attachment of the conjunction "ف" at the beginning of the expression as well as the pronoun/morpheme at the end of the phrase. In accordance with Arberry's translation, it should have been "Dispute".

B) Ambiguity

The machine translation programme translated the Arabic lexeme   مِرَّةٍ   as "Once", it mixes between the meaning of "once" and the meaning of "very strong" & that is because this machine could not reading the vowel points & it could not understand that this word refers to the word "وحي". This is a case of absolute homonymy.

The machine translation programme translated the verbal phrase فَأَوْحَى   as "Inspired"; the programme cannot distinguish the difference between the meaning of "Inspired" and the meaning of "revealed". This is a case of polysemy.

The machine translation programme translated the Arabic negation marker   مَا   into the English "what" which is either an interrogative or a relative pronoun. Here again the programme is unable to disambiguate in accordance with the contextual meaning of a given item as the Arabic source language item is a case of absolute homonymy. The Arabic   ما   may serve as a negation marker, a question word or a relative pronoun. Only the linguistic and semantic context dictates which possible meaning to accept.

The machine translation programme translated the verbal phrase   ضَلَّ   into "lost" as in "lost something" whereas its contextual meaning in the verse should be "went astray".   This is another case of polysemy.

C) Lexical Mismatches

The machine translation programme transliterates the verbal phrase دَنَا, meaning "Drew near/close" into "DNA" which means that it deals with the Arabic item as if it were some kind of abbreviations.

The machine translation programme translates كَذَبَ in ما كذب الفؤاد ما رأى   into "false-hearted" instead of "heart lies". This may be due to the inability of the software to identify the Arabic as a verb; instead, it may have regarded it as an abstract noun or a verbal noun/gerund meaning (Lying/Lies).

3.2.2.2 Problems on the Syntactic/Grammatical Level

• 	Verse (1): The machine translation programme missed the singular pronoun "it" after the word "when" because it translated the word "هوى" as a noun not as a verb & that is because it doesn't understand that the word "هوى" refers to the agent which is "the star", so it doesn't need the singular pronoun "it".

• 	Verse (2), (3): The machine translation programme uses the conjunction "and" in verse number (2) and repeats it in verse number (3) instead of the conjunction "neither" in verse number (2) and the other conjunction "nor" in verse number (3).

• 	Verse (5): The machine translation programme does not read the first word "عَلَّمَه" which refers to God.

• 	Verse (7): The machine translation programme uses the singular pronoun "he" instead of "it" because it reads the last agent referred to in "صاحبكم" and relates the pronoun "he" to this agent "صاحبكم", it does not understand that the Koran in this verse means the agent to be"النجم" not "صاحبكم".

• 	Verse (10): The machine translation missed the word "then" at the beginning of the translation, and it uses the article "the" before the word "servant" instead of the possessive adjective "his".

3.2.2.3 Problems on the Stylistic/Communicative Level

• 	Verse (9): The whole translation is wrong, The machine translation programme is never translated this verse as what God means of it, machine translation is completely different from the right one & the right meaning.

• 	Verse (12): The machine translation programme dose not recognize the interrogative prefix at the beginning of the lexical item "أفتمارونه" and, thus, translates it as a noun "ovtmarunh", so it missed the right meaning of the verse.

3.3 Literary Texts

Thee third language register dealt with in the present paper is that of literary writing, which also presents a real challenge to the professional translator, let alone machine translation software. The text in question is an article written by late Dr Zaky Naguib Mahmoud, an Egyptian thinker rather than a creative man of letters. Yet the article under study here is typically a piece of literature. The writer assumes the role of a typically old fashioned hypocritical Arab that boasts all sorts of fake glories of the east and rejects all kinds of advancement and intellectual progress offered by the west simply because they come from abroad. The text, thus, is more of a dramatized satire rather than an argumentative article. The writer resorts to the use of metaphorical expressions, dramatic irony and implicit statements, all of which are typically representative of literary styles.

3.3.1 Problems on the Lexical/Semantic Level

This part discusses the semantic problems that the Google machine translation programme has encountered while translating the literary text that has been chosen by a Zarqa student from a graduation research made by Mona Ahmed Saleh at Alexandria University, Egypt June 2008. The original research by MS Saleh was a daring and an ambitious yet quite successful attempt at both translation a long article (12 Pages) by the late Egyptian scholar and thinker Zaky Naguib Mahmoud from Arabic into English and commenting on the translation problems involved, based on her study of the Theories of Translation course (4th Year). As the research got a degree of A (72 out of 75), the researcher gave Zarqa University students a copy of that research to see how an undergraduate student could produce such a good academic work as well as learn how to tackle translation problems in a systematic way, even in an undergraduate research. All the students doing the course at Zarqa made some use of the research one way or the other, but one student in particular (Zaid Kashur), thought of using a small part of the original source language text as a sample of literary styles to see how the Google Translate Service shall deal with it.

A) Transliteration

To begin with, most of the lexical items were transliterated; because the MT programme is unable to recognize a lexical item when an additional morpheme is attached to it, whether as a prefix or a suffix. The first striking semantic-translation problem is the lexical item "عبيط" which is used in its colloquial sense. The choice of such an Egyptian-spoken lexical item with the standard lexical item "جنة" is deliberate, as it combines what is farcical and ridiculous with the sublime, which serves perfectly the satirical tone of the article. The first lexical item of the title which is strictly speaking a proper MSA item (جنة) meaning (Paradise or Heaven), while the second lexeme (العبيط) is used in its colloquial Egyptian sense meaning (Idiot or Fool). With his satirical style, the author, Dr Zaky Naguib Mahmoud, attempts to combine the sublime, both culturally and linguistically (جنة) with the mundane and even ridiculous colloquial item (عبيط). This combination is meant to reflect the dualism that Arabs and/or Muslims suffer from. In the world of fancy, they believe that they live in heaven and that they are bound to go to Paradise after death because they "are the best". Zaky Naguib Mahmoud, who is both an Arab and a Muslim, however, maintains that, in reality, Arabs and Muslims in the second half of the twentieth century, the time he wrote that article, are nothing but fools (As found in Saleh, 2008). Thus, this may look hard to be recognized for many of translation students or either translators in accordance to choose between the two lexical items *Fool and Idiot, the machine translator did not recognize one of them at all but it transliterated it into* Elabeet which considered it as a proper noun. The next transliterated lexical item is" بغاث", the MT programme was unable to recognize the lexical item whether in context or even when the researcher presented it in isolation. Also the same problem is related to the lexical item "تفري"; here the MT programme did not recognize that the lexical item is the main verb of the sentence; the word was translated into Mincing in the authorized translated text. Other words were transliterated too, "الكريم" and the phrase " ورشاد من قادتي". The first word is an adjective and the other رشاد is a noun in the source text. The MT programme considered these words as proper nouns. Another confusing problem attached to the word "الله" this word is problematic not only for the MT programme but also for most of translators; the MT programme as well as several translators translate it into "Allah" considering that the concept does not exist in other religions and/or cultures, whereas it indeed does exist in the Christian and Jewish faiths and must be translated into God (the supreme being). The last problem related to words proceeded with a conjunction changing them into transliterations, "بعزلتي", "بوقدها", "فيتوهم"and "فيغفو", these words will be discussed in the Syntactic/Grammatical level.

B) Ambiguity

The second part of the discussion is about ambiguous words. The first word is "نسجتها" which has the given meaning "spun", that is related to clothes and ropes but in the source text it is related to dreams that the writer made up in his mind. Another sentence full of ambiguous words," تهب فيها النسائم عليلة بليلة " , the verb "تهب" was translated into "windy" which is adverb where it should be translated into a verb which is blew, the word "عليلة" was translated into "sick" considering it as a sick person where it is a pre-modifier to breezes, the last word in the sentence "بليلة" the

problem here is the entrance of the conjunction which changed its spelling in the MT programme the MT programme has to read the word without a conjunction and read it as "ليلة" which means "night", the word "بليلة" also a pre-modifier to breezes. To continue with an exceptional word where there is no explanation for mistranslating it, "جنتي" it is easy to recognize that this word means Heaven or Paradise in any case, but it was translated into Committees, first of all, the word is plural where it is singular in the source text, it is capitalized which has no explanation either; also it is easy for the MT programme to translate the word "Committees" into "لجان" which is totally unrelated to the context. Another word was misread which "الهرم" the MT programme is read it as "Haram" which translated it into "Pyramid" but it is read as "Harim" which means an old falcon as it refers to the source text. Another related problem in relation to the following sentence "أنا في جنتي السمح الكريم" is that it was translated into "I'm in the Holy Committees tolerant" the word holy usually refers to religious thing such as Quran or a place but not to a person, the MT programme did not recognize that it refers to the subject who is generous, the other word is tolerant which refers to "السمح" but it is attached with the word "الكريم" which they are post modifiers to the subject. Another sentence" his father slaughtered camel and camel" the word "يذبح" translated into Slaughtered which is correct as a word itself but in literary Arabic called "sacrifice" as someone sacrifice a camel or a goat out of charity or as an offering or service to God, the same sentence includes a repeated word "camel" without recognizing that the second occurrence in the source text refers to a female-camel which is "الناقة".

4. Conclusions & Recommendations

4.1 Conclusions

The conclusions of the present research may be fairly divided into two main types; first, those that are quite expected and are almost typically characteristic of all types of machine translation software, and, second, those conclusions that were not as much expected to reach as far as the researcher is concerned and that are also more or less applicable to the software in hand rather than all or even the majority of automatic translation software in general.

First: The General Type

- Google Machine translation programme is unable to identify or recognize   Arabic vowelings

- Google Machine translation programme is unable to recognize morphemes/endings attached to verbs in different conjugations

- Google Machine translation programme deals better with simple sentences in informative types of style.

- Google Machine translation programme does not recognize rhyme

- The programme often functions as a bilingual dictionary rather than a proper MT programme, resulting in a more or less word for word translation

The MT programme employed by Google is unable to deal with the semantic phenomenon of ambiguity, which is a common characteristic feature of all machine translation programmes.

Second: The Particular Type

- Google MT programme does not recognize Arabic vowelings, which is an additional demerit which is avoided by a few other machine translation programmes, especially those that are not simple free but commercial software such as Al Wafi and/or Al-Kafi. This, indeed, needs to be rectified.

- The unexpected and inexplicable inability of the programme in question to give the appropriate equivalent of a lexical item in context even though it does provide a proper translation of it in isolation, such as in the case of حلال and الحَرَام . It is quite interesting in this respect to note that despite the needless transliteration technique adopted by Google MT programme in dealing with حلال and الحَرَام it has properly and communicatively translated the Arabic الله into God. On the other hand, a few professional translators resort to transliteration either because they are Muslims who deny the followers of other religions the right to believe in الله or because they are non-Muslims who would like to deal with Islam as a pagan creed that has a pagan god called Allah (See ElShiekh, 1991).

- An unexpected and an inexplicable case of over-translation ahs has also occurred as the Google MT programme introduced the English lexical item Bukhari while there was no mention at all of البخاري in the SL text. It is worth mentioning here, however, that the Hadith in question does appear in Bukhari. This may imply that the Google MT programme does not depend on a set of fixed dictionaries as much as a flexible database that may end up in offering some extra information that do not, strictly speaking, come in the SL text though it has some relevance to it.

- It is also interesting to note that the Google MT programme suggested "tissues" as a translation of the Arabic محارم when given the Arabic lexical item in isolation. This indicates that the database available for the programme sometimes includes equivalents for colloquial Jordanian dialectical expressions while it lacks more basic MSA terms. This is certainly another problem that has to be addressed in the future for the improvement of the Google MT translation software.

4.2 Recommendations

- To sum up, the researcher suggests that the experts in charge of the Google MT programme may provide it with the necessary instructions to let it cope with Arabic vowelings, which, in turn, shall help the programme avoid some otherwise unnecessary ambiguities.

- The researcher also recommends supplying the programme with the necessary data to help it deal with simple Arabic prepositions such as عن that was wrongly translated into "and".

- Last but not least, there should be more focus on the use of grammatical structures in English; otherwise, the Google programme seems to function basically as a bilingual dictionary rather than a proper MT programme.

- Finally, the researcher hopes this research could stimulate more experienced researchers in the field of machine translation and computational linguistics to attempt a further exploration of the topic under study, and, thus, come up with more important findings and recommendations.

References

Arnold, D. J., Lorna Balkan, Siety Meijer, R. Lee Humphreys, & Louisa Sadler. (1994). *Machine Translation: an Introductory Guide*. Blackwells- NCC, London.

Bell, Roger. (1992). *Translation and Translating*. Longman Group Ltd, London, UK.

Catford, J. C. (1965). *A Linguistic Theory of Translation.* OUP, Oxford, UK.

Conway, Kyle, & Bassnett, Susan (ed.). (2006). *Translation in Global News*. Proceedings of the conference held at the University of Warwick 23 June 2006.

ElShiekh, A., & Saleh, Mona (2011). Translation versus Transliteration of Religious Terms in Contemporary Islamic Discourse in Western Communities. *International Journal of English Linguistics*, Vol. 1, No. 2, September 2011. http://dx.doi.org/10.5539/ijel.v1n2p141

Encyclopaedia Britannica. (2000). CD Edition. Copyright©1994-2000, Encyclopaedia Britannica, Inc.

Halliday, M., & Hassan, Ruqaiya. (1976). *Cohesion in English*. Longman, London.

Hatim, B., & Mason, J. (2003). (4th edition). Discourse and the Translator. Longman, Singapore. ("Hello, World," Wired, May 2000). [Online] Available: http://www.bbc.co.uk/religion/religions/islam/texts/quran_1.shtml

Hutchins, W. John, & Somers, Harold L. (1992). *An Introduction to Machine Translation.* ACADEMIC PRESS Harcourt Brace Jovanovich, Publishers, London.

Larson, M. (1984). Meaning-Based Translation: A Guide to Cross-Language Equivalence. U.S.A, University Press of America, Inc.

Leech, J., & Svartik, J. (1994). *A Communicative Grammar of English*. (2nd edition). Longman, Singapore.

Newmark, P. (1981). *Approaches to Translation.* Prentice Hall, U.K.

Newmark, P. (1988). *A Textbook of Translation*. Prentice Hall, U.K.

Translation and Interpretation. [Online] Available: www.diplomacy.edu/language/Translation/machine.htm

Universal Translators. Wired, May 2000. A listing of machine translation research and development hubs worldwide.

Venuti, Laurence. (1998). The Scandals of Translation. London, Routledge. http://dx.doi.org/10.4324/9780203269701

Venuti, Laurence. (2008). *The Translator's Invisibility*. (2nd edition). London, Routledge. http://dx.doi.org/10.4324/9780203269701

Widdowson, H. (1983). *Learning Purpose and Language Use*. OUP, UK.

Wolf, Michaela. (2002). Culture as Translation and Beyond: Ethnographic Models of Representation in Translation Studies. In Theo Hermans (Ed), *Cross-cultural Transgressions: Research Models in Translation Studies II: Historical and Ideological Issues*. Manchester, UK.

د. عطية، فوزي (1986)، *علم الترجمة- مدخل لغوي*. القاهرة، دار الثقافة الجديدة.

د.محمد عناني. (2003). *الترجمة الأدبية بين النظرية والتطبيق*. الطبعة الثانية، الشركة المصرية العالمية للنشر ـ لونجمان، القاهرة.

Appendix 1

<u>Source Language Texts</u>

<u>أ- الإعلانات</u>

1. بيبسي أطلق حماسك

2. المي قتلتها الغيرة

3. عبر مين قدك

4. عبر مين قدك

5. ماونتن ديو قوي قلبك

6. الدنيا صيف الاسعار عالكيف ... البسة لكافة الاعمار اكسبنسف للالبسة

7.BlackBerry بكل الألون وكمان ممكن تدفع بالتقسيط واحصل على فودافون الخط برقم ذهبى هدية

8. محتاج انترنت؟

عايز تدخل على الانترنت و مفيش كمبيوتر؟

دلوقتي مع فودافون ممكن تدخل على الانترنت من أى مكان و في أي وقت...

مع خدمة موبايل انترنت من فودافون عالم الانترنت في جيبك

9. لو فاتتك فترة السماح

الحق رجع خطك ببلاش

اطلب 858 لإعادة الشحن

و سيتم تشغيل الخط بعد الشحن مجانا   مع فودافون انت الكسبان ...!

هذا العرض سارى لعملاء فودافون الكارت بعد اقصى 90 يوم من إنتهاء فترة السماح

10   دوبلناها .

وتكتكناها

وزودناها

خطك الجديد الي رح تحكي فيه بالثانية من الثانية الاولى

وكل حكيك اليوم دوبلناه

وتكتكناها مع كل شحن بطلعلك رسائل محلية ببلاش

وزودناها مع مفاجآتنا طووووول الوقت

انضم النا وحول علينا بالضغط على   # 148 *

لمعرفة قيمة البونص اضغط   # 117 *

ولازم يكون عمرك تحت ال 25   سنة

<u>ب- نصوص دينية</u>

<u>1- حديث شريف</u>

عَنْ أبيْ عَبْدِ اللهِ النُّعْمَانِ بْن بشِيْرٍ رضي الله عنهما قَالَ: سَمِعْتُ رَسُوْلَ اللهِ صلى الله عليه وسلم يَقُوْلُ: (إنَّ الحَلالَ بَيِّنٌ وإنَّ الحَرَامَ بَيِّنٌ وَبَيْنَهُمَا أُمُوْرٌ مُشْتَبِهَات لاَ يَعْلَمُهُنَّ كَثْيْرٌ مِنَ النَّاسِ، فَمَن اتَّقَى الشُّبُهَاتِ فَقَدِ استَبْرَأ لِدِيْنِهِ وعِرْضِه، وَمَنْ وَقَعَ فِي الشُّبُهَاتِ وَقَعَ فِي الحَرَامِ كَـالرَّاعِي يَرْعَى حَوْلَ الحِمَى يُوشِكُ أنْ يَقَعَ فيْهِ. ألا وَإنَّ لِكُلِّ مَلِكٍ حِمَىً. ألا وَإنَّ حِمَى اللهِ مَحَارِمُهُ، ألا وَإنَّ في الجَسَدِ مُضْغَةً إذا صَلَحَتْ صَلَحَ الجَسَدُ كُلُّهُ وإذا فَسَدَت فَسَدَ الجَسَدُ كُلُّهُ ألا وَهِيَ القَلْبُ) متفق عليه

2- وَالنَّجْمِ إِذا هَوَى (1) مَا ضَلَّ صَاحِبُكُمْ وَمَا غَوَى (2) وَمَا يَنطِقُ عَن الهَوَى (3) إنْ هُوَ إلَّا وَحْيٌ يُوحَى (4) عَلَّمَهُ شَدِيدُ القُوَى (5) ذُو مِرَّةٍ فَاستَوَى (6) وَهُوَ بِالأُفُقِ الأَعْلَى (7) ثُمَّ دَنَا فَتَدَلَّى (8) فَكَانَ قَابَ قَوْسَيْنِ أوْ أدْنَى (9) فَأوْحَى إلَى عَبْدِهِ مَا أوْحَى (10) مَا كَذَبَ الفُؤَادُ مَـا رَأى (11) أفَتُمَارُونَهُ عَلَى مَـا يَرَى (12)

سورة النجم- القرآن الكريم

<u>ج- نصوص أدبية</u>

من "جنة العبيط" بقلم د. زكي نجيب محمود

أما العبيط فهو أنا، وأما جنتي فهي أحلام نسجتها على مر الأعوام عريشة ظليلة، تهب فيها النسائم عليلة بليلة، فإذا ما خطوت عنها خطوى إلى يمين أو شمال أو أمام أو وراء، ولفحتني الشمس بوقدتها الكاوية، عدت إلى جنتي، أنعم فيها بعزلتي، كأنما أنا الصقر الهرم، تغفو عيناه، فيتوهم أن بغاث الطير تخشاه، ويفتح عينيه، فإذا بغاث الطير تقري جناحيه، ويعود ليغفو، لينعم في غفوته بحلاوة غفلته.

Appendix 2

Target Language Texts

A- Advertisements

1. Pepsi launched Enthusiasm

2. Elmi killed her jealousy

3. By Sweet

4. Across. Sweet

5. Mountain Dew strong Heart

6. Sifk the Hi ….and back colors Atif; ….live your life with our own … what Btantsy trips allover the world identified Travel & Tourism tel. 5811855

7. Now you with Vodafone ups enjoy the best prices of all Alolon ® Blackberry and also pay in installments as possible and get the Vodafone line number golden gift

8. Needy Internet?

Want access to the Internet and Mvic computer?

This moment with Vodafone as possible access to the Internet from anywhere and at any time…

With service from Vodafone mobile Internet world of the Internet in your pocket

9. If she missed you the grace period

The right returned your line Bblash

Request 858 for the re- shipment

And operating the line will complete after the shipment

For free

With Vodafone you are Al Ksban!...

This presentation is applicable to Vodafone agents the card after

the end of 90 days is from the end of the grace period

10. Dopellnaha

And Tktknaha

And we provided

Welcomed the new line is to tell it from the first second per second

Every day Hkik Dopellnah

And Tktknaha with each shipment of messages Btalaalk local freebee

And given them to surprise us with Toowol time

Elena joined us and on pressing * 148 #

To know the value of the bonus, press * 117 #

And we have to be you are under 25 years

B- Religious Texts

1. And Abu Abdullah al-Nu'man ibn Bashir - may Allah be pleased with him - said:" I heard the Messenger of Allah

- peace be upon him - said The Halal and Haram between and which, if not take more people, it avoids doubtful matters of religion and his honor who falls into doubtful occurred in the no man's land Bukhari sponsors about the fever is about to enjoy himself when not if every king has not even God protect female relatives not even in the flesh, chew if shepherd who pastures around and if it is corrupt, namely heart "

2. By the Star when Hui (1) What lost your companion and seduce (2) and speaks of passion (3) It is only a Revelation revealed" (4) mighty in power (5) is once sat up (6) While he was up to (7) and then DNA Vtdly (8) was just around the corner (9) revealed to the servant what inspired (10) is false-hearted to the view (11) Ovtmarunh what he sees (12)

The Star Sura- The Holy Quran

C- Literary Texts

The Elabeet it's me, either the committees are the dreams spun over the years Arbor shady, windy breezes sick the night, if the step by step to the right or northern or in front or behind, and Vanni Sun Boukdha caustic, I returned to the committees, grant the Bazlte. As if I Falcon of the pyramid, fall asleep in his eyes, that Vitohm Bgat bird fear, and opens his eyes, if Bgat Tafri bird wings, back Vighafoo, the casket of the sweetness of slumber slumber.

I'm in the Holy Committees tolerant, which inherited the generosity parents and grandparents, it is the others was his father slaughtered camel and camel to feed each of hunger and destitution? The others belonged to Hatem Karim shelter this item? Were the characteristics of my forefathers to go with the air in vain, or is taking place in the veins with blood? Behold, I will bow to the miserable compassionate and you do not give him, and dissolve the patient sadness, and you do not Auasih, and decide by the envious say that the owners need me begging and tender and the needy palms to contract on the air, the heart of loving the good of the poor penny spent fast, P) Ed melted kept him not to cause a type that is lost, I seek refuge in Allah from person to understand the language of charity, shark, M, I swear that dominated the material cut into the whole world, and without the mercy of the Lord, and Rashad of Kadti. For you today in the flooded hath mercy!

Notes

Note 1. *The Columbia Electronic Encyclopedia*® Copyright © 2007, Columbia University Press. Licensed from Columbia University Press. All rights reserved. www.cc.columbia.edu/cu/cup

Note 2. http://encyclopedia2.thefreedictionary.com/Machine+Translation

Note 3. http://www.diplomacy.edu/language/Translation/machine.htm

Note 4.   i.e., established by the position of a word within a field.

Note 5. i.e., established by the position of a word within a field.

Note 6. There are two types of ambiguity:

a. Homonymy- Absolute homonymy is the case of two or more lexical items that (1) Are identical in form but have different unrelated meanings. (2) All their forms are identical. (3) All their identical forms are grammatically equivalent.

If the second and/or third condition maybe missing, it is a case of partial homonymy. Examples in English "Bank/s = أنشأ) Found (present simple) & وجد) Found (past tense) = Absolute homonymy. Found (present simple) أنشأ = مصرف، مصارف Bank/s & ضفة، ضفاف Partial homonymy". Examples in Arabic بيوت بيت وجمعه = منزل) house) أبيات وجمعها بيت = شعر) a line of poetry).

b. Polysemy- It is the case of one lexical item that covers a wide range of related meanings. Example in English (uncle, aunt). Example in Arabic (اختطاف = kidnap/ highjack).