



A Corpus-based Study of Human-translated vs. Machine-translated Texts: The Case of Ellipsis in English-Persian Translation

Milad Maleki *

English Department, Islamic Azad University, Isfahan (Khorasgan) Branch, Isfahan, Iran

Hossein Heidari Tabrizi

English Department, Islamic Azad University, Isfahan (Khorasgan) Branch, Isfahan, Iran

Abstract

The present study was intended to investigate the translation of different types of ellipsis in Persian and English using a descriptive method. Given the importance of the study of human translation (HT) and machine translation (MT), the present study focused on the ellipsis based on Halliday and Hasan's model (1976). In so doing, a bilingual parallel corpus called "Mizan corpus", including more than one million translated sentences, was compared with Google Translate, as an example of machine translation. The results of the study indicated that there were some mismatches in verbal and nominal ellipsis between English and Persian, but almost no mismatches were found for clausal ellipsis. There was a significant difference in the quality of human translation and machine translation (i.e. Google Translate) in favor of human translation. The findings also depicted that HT's quality was still higher than MT for all kinds of ellipsis. So, it was concluded that human translation is more dependable than machine translation.

Keywords: corpus, bilingual corpora, human translation, machine translation, ellipsis

INTRODUCTION

The corpus-based approach to study translation has become popular over the last decade, with a treasure of data now emerging from studies using parallel corpora, multilingual corpora and comparable corpora. Corpora, particularly bilingual corpora, prove to be valuable sources of information in translation research and education. As Varantola (1997) stated, more than 50% of the time of a translator is spend on consulting various sources to find the meaning of unknown words. In such a case, the utilization of computer-based bilingual corpora might increase the translation speed and its quality, since these corpora facilitate more near-native interpretations and strategies in both source and target texts (Aston, 1999). Moreover, they also assist novice translators to learn about overall patterns and find ways that help better

expressing of things, on condition that the corpora of texts are diverse and large enough (Zanettin, 1998). By itself, corpora serve as valuable resources in developing an understanding the principles and standards in different languages. To design and utilize translational corpora, it is crucial to consider the social and cultural contexts of the translations. In other words, considering social and cultural contexts, it can “provide a framework within which textual and linguistic features of translation can be evaluated” (Bernardini & Zanettin, 2004, p. 60).

According to Lin, Murakami, Ishida, Murakami and Tanaka (2010), MT has been significant research issue in the field of artificial intelligence for many decades. However, there are still huge discrepancies among human and machine translators. On the one hand, machine translations have always faced with limitations considering its quality as the consequence. These types of translation are not valid and dependable for translating documents necessitating high qualities. On the other hand, a bilingual person cannot be found everywhere for any purpose in the real world; moreover, the translations done by highly-trained translators will cost a lot in both labor and time. Moreover, MT is problematic compared with HT considering ambiguity and cultural transfer. Therefore, a human translator needs to revise the machine translated text. Google Translate (GT) is one of these systems which grossly differs from HT considering quality of the translation of different texts.

Human translators commonly have no severe problem handling ellipsis in source text, since they naturally recognize the meaning of elliptic sentences; consequently, they can easily detect the deleted material and find the missing words in their translation. MT, on the other hand, need some predetermined information to be accessible. This information need to be provided for MT systems; If not the existing gaps end in translation failure. English-Persian MT systems (e.g. Google Translate (GT)) are unable to resolve ellipsis.

Corpora and their processing software undoubtedly provided the translation scholars with effective tools for studying the exact nature of translation. Some translation theorists advised a note of caution in using corpora due to their limitations.

Alike the sentence-based systems, the existing MT systems do not consider the discourse and context while analyzing the source language. As a result, machine translation quality is utterly impaired (Zhang 2009). This study pinpoints, from quality and ellipsis type perspectives, the inadequacies of current MT systems in source language analysis compared with HT. The purpose is to analyze the features and functions of ellipsis and its influence on word and sentence meaning, which is more specifically done through the analysis of the differences between MT and HT. Therefore, the present study aims at comparing MT and HT in terms of verbal, nominal and clausal differences based on a corpus-base study.

To investigate the purpose of the study and tackle the problem, three following research questions were raised:

- Is there any difference between human-translation (Mizan) and machine-translation (Google Translate) in translating nominal ellipsis from English to Persian?
- Is there any difference between human-translation (Mizan) and machine-translation (Google Translate) in translating verbal ellipsis from English to Persian?
- Is there any difference between human-translation (Mizan) and machine-translation (Google Translate) in translating clausal ellipsis from English to Persian?

REVIEW OF THE LITERATURE

Much early work in Corpus-based Translation Studies (CTS) set out to trail the research agenda put forward in Baker's seminal article in 1993 and investigated, on a scale that had not been possible before, those recurrent features that were thought to make translation different from other types of language production. These features, also called universals of translation, included the reported tendency of translated texts to be more explicit, use more conventional grammar and lexis, and be somehow simpler than either their source texts or other texts in the target language (Baker, 1996).

CTS is now recognized as a major paradigm that has transformed analysis within the discipline of translation studies. It can be defined as the use of the technologies provide by corpus linguistic to facilitate and clarify the translation process, which is ever more available through progresses in computer technology (Kruger, Wallmach, & Munday, 2013).

Human translation is influenced by the characteristics of source-target language transfer, cultural context and individual translators' translation ability (Bassnett & Lefevere 1992; Wong & Shen 1999). In the translation process, the ways of managing, decoding and recoding, equivalence (Gentzler 2001), loss and gain, and untranslatability (Bassnett 2002) will produce different versions of translation by different translators owing to various interpretation of source and target languages. Considering machine translation, nevertheless, the impact of the translator's skill and mastery will be ineffectual. However, a point worth mentioning us that the machine translation quality is still a concern. To assess the quality of machine translation, it is a requisite to make a comparison between the machine translation and the human translation in terms of the specifications of the source language i.e. its words, syntax, semantics, pragmatics and discourse. With such a comparison, we gain a general viewpoint on machine translation quality, compared with human translation and source language.

Owing to huge differences among languages (the differences between Persian and English languages in this study), some linguistic features such as syntax and semantics may undergo some changes in both human and machine translated texts. They may be ignored as well. One of the key items of this kind of changes is ellipsis. In the following section, ellipsis and its sub-components are described in details.

Concerning the merits and demerits of GT, Butler (2011) asserted that today GT is known as an application that is free and online and is the best of third party websites presenting a computerized translation of the content in any of the existing languages. Using GT is quick, easy and it provides satisfactory general content translation for more than 50 languages. It collects data and finds information on websites that were formerly inaccessible because of the language barrier. However, GT might misinterpret complex structures and provides imprecise translations while one uses it, may not be aware any errors and shortcomings.

Ellipsis

Ellipsis occurs when "something which is present in the selection of underlying (systematic) options is omitted in the structure-whether or not the resulting structure is in itself incomplete" (Halliday & Hasan, 1976, p.144). In crude terms, we can take as a general guide the notion that ellipsis occurs when something that is structurally necessary is left unsaid. Ellipsis, like substitution, is believed to embody the same fundamental relation between parts of a text (a relation between words or groups or clauses- as distinct from reference, which is a relation between meanings). A crucial issue of concern that is worth mentioning is the fact that where there is ellipsis there is a presupposition in the structure, that something is to be supplied, or 'understood', where understood is used in the special sense of "going without saying" (Halliday & Hasan, 1976). To put it differently, an elliptical item is one which, as it were, leaves specific structural gaps to be filled from somewhere else. This is precisely alike presupposition by substitution, however in substitution an explicit "counter" is utilized, e.g. one or do might function as a place-marker for the presupposed sections, while in ellipsis nothing is slotted in the gap (Halliday & Hasan, 1976). This might be the reason why ellipsis is characterized as substitution by zero. There are three types of ellipsis, considering the structural unit in which ellipsis occurs, namely nominal ellipsis, verbal ellipsis, and clausal ellipsis.

THEORETICAL FRAMEWORK

Halliday & Hasan (1976, p. 142) refer to ellipsis simply as 'substitution by zero'. Definition of ellipsis can be based on a well-known concept when it comes to this linguistic category; that something is left unsaid, or is completely omitted. However, it does not mean that what is not explicitly stated is not understood. On the contrary, ellipsis implies that if something is left unsaid can be understood, interpreted and reconstructed by simply examining the context.

On the other hand, when we speak of ellipsis we do not take into consideration every case where there is some information that the speaker has to, so to say, 'supply' from its source. This would apply to almost every sentence ever spoken or written and would not be very useful in revealing the nature of the text. On the contrary, ellipsis involves debates about the phrases or clauses, whose structure is such that it can be presupposed by a certain language category that precedes or follows, and as such serves as a source of omitted information. Elliptical structure is one that leaves a particular

structural site empty and which can be filled in with the appropriate language category. This can be compared with the presupposition in substitution, except that in this case it is explicitly specified what has already been expressed in advance, such as, for example, *one* or *do*, while in the case of ellipsis nothing has been put in that position. Therefore, ellipsis can be described as, as it has previously stated, substitution by zero.

Table 1. English-Persian mismatches in different ellipsis (Mizan & GT)

	Mizan		Google Translate		Total
	Matches	Mismatches	Matches	Mismatches	
Nominal Ellipsis	13	87	10	90	100
Verbal Ellipsis	24	276	20	280	300
Clausal Ellipsis	244	6	43	257	250

In the beginning of survey, data had been collected from one million translated sentences in Mizan corpus. Concerning the research questions, the main focus was on verbal, nominal, and clausal ellipsis. In all cases, the data were extracted from our corpus and then analyzed in Google Translate as well.

A comparison of two analyses reveals that out of one million scanned sentences in Mizan corpus, 100 nominal ellipses were found. All the Persian translations were checked for any mismatches with their equivalents in English. According to Mizan, there were 87 mismatches between Persian and English sentences. This number of mismatches in GT was 90 out of 100.

Afterwards, the case of verbal ellipsis was investigated. Among the translated sentences, there found 300 cases in which verbal ellipsis was observed. After comparing all these sentences with Persian equivalents in Mizan and GT, there were 276 and 280 mismatches out of 300, respectively.

Finally, Mizan corpus analyzed for finding clausal ellipsis. There found 250 clausal ellipsis in which there were not any significant mismatches. There were only 6 sentences in which they've undergone mismatches in Mizan corpus. On the contrary, in GT, 257 sentences had mismatches between English and Persian translations in regard to clausal ellipsis.

METHODOLOGY

Instruments

There were two main instruments used in the study; Mizan English-Persian Parallel Corpus (Mizan, 2013) is a corpus that contains more than one million English sentences (often in classic literature) and their translations into Persian that have been provided by the Supreme Council of Information and Communication Technology in Iran. The corpus can be used in various applications, especially machine translation and natural language processing. Mizan parallel corpus has one million aligned Persian and English sentences. Using Mizan parallel corpus, Supreme Council of Information and Communication Technology developed a basic statistical translation system called

"Online Translator" in collaboration with Iran University of Science and Technology. The other instrument used in the study was Google Translate (GT)

Data Collection Procedure

Since the purpose of the present study was comparing human translation with machine translation, the researcher needed a parallel bilingual corpus which is valid and dependable. Google Translate is an example of machine translation which is widely used in research in machine translation; however, finding a human translated bilingual corpus was not an easy task. Nevertheless, in the recent literature of human translation, Mizan, a bilingual corpus of English-Persian developed by the Supreme Council of Information and Communication Technology of Iran, is commonly used as a reliable instance of human translation.

To find the instances of the three types of ellipsis, namely, nominal, verbal and clausal, the English sentences in Mizan corpus will be analyzed one by one to find the examples of the three abovementioned types of ellipsis, and then their Persian equivalent will be examined as well to see if the same ellipsis has been observed in the Persian version. As to the Google Translate, the sentences with the instances of ellipsis found in Mizan will be copied to Google Translate to find if there is any difference in the way the sentences are translated by Google Translate.

Data Analysis

Based on the taxonomy of cohesive theory provided by Halliday and Hassan (1976), the four elements of this theory, referencing, substitution and ellipses, conjunction, and lexical cohesion, built the categories for analysis. The findings of the comparison between English and Persian sentences in both Mizan and Google Translate will be presented in separate tables for each research question. In order to see whether there were any mismatches between English and Persian kind of ellipsis, a contrastive analysis was used.

RESULTS AND DISCUSSION

Investigating Research Question 1

This section addresses the research questions of the study and gives brief answers to research questions. Then the results will be compared with the previous findings in the field and similarities or differences will be discussed. This study is a modest contribution to the ongoing discussions about the quality of Google Translate as a machine translation. We concentrated on English to Persian translations done by Google Translate and Mizan Corpus. Google Translate has been evaluated by many researchers and compared to other Persian-English machine translation systems to indicate and show how well this system translates from Persian to English or vice versa (Mohaghegh, & Sarrafzadeh, 2009; Mohaghegh, Sarrafzadeh, & Moir, 2010, 2011); however, there might not be any study done like error analysis as human assessment to provide enough insight for errors and clearly show different types of errors made by Google Translate.

This might be the first study to assess the quality of Google Translate considering error analysis method presented by Keshavarz (1999).

It is apparent from Table 1 that in the case of nominal ellipsis, almost all sentences have mismatches between English and Persian translation.

Table 2. English-Persian mismatches in nominal ellipsis (Mizan)

No.	English Sentence	Persian Translation
1	He brings <u>his dog</u> , and Sam brings his too.	Ou Sagash ra avarad va Sam niz sage khod ra avarad
2	I caught the <u>first train</u> before you caught the second .	Man ghabl az in ke shoma be ghatar-e dovom beresid, be avari residam
3	Please help yourself with another <u>biscuit</u> . No, thank you. It was my fourth .	Lotfan yek biskooit digar bokhorid. Na mamnoon. In chaharomin biskooit bood.
4	You have a big <u>share</u> , but mine is a minor one .	To yek sahm ziadi dashti vali mal-e man nachiz bood.
5	All men have <u>dream</u> and I have my own too.	Hamey-e Ensanha Arezoo darand va man niz mostasna nistam.
6	Bill likes her <u>story</u> even though he hates other ones .	Bil dastan vey ra dost dasht dar halike az sayer dastanha motonafer bood
7	<u>The men</u> returned at midnight. Both were worn out.	Anha nesfe shab bargashtand. Har doye anha besyar khaste boodand.

With respect to nominal ellipsis, the findings presented in Table 2 reveals four main patterns of mismatch. The first mismatch is when the pronoun is replaced by the nominal it refers to. For example, in sentence one, his in English is replaced by "sage khod". The same could be seen in sentence 6. The second type of ellipsis was when the ordinal numbers of English were replaced by ordinal number + nominal. For example, in sentences 2 and 3, the ordinal numbers of "second" and "fourth" were replaced by "Ghatore Dovom" and "Chaharomin Biskooit" respectively. The third pattern was observed when the Persian translation differed from the English sentence. In sentence 5, "my own" is translated as "Man Niz Mostasna"; moreover, the positive English sentence is translated as negative in Persian. And finally, in sentence 7, the word "both" is translated as ".Har Doye Anha". In this sentence, a pronoun has been added to the sentence to make it comprehensible to Persian speakers.

Table 3. English-Persian mismatches in nominal ellipsis (GT)

No.	English Sentence	Persian Translation
1	He brings <u>her dog</u> , and Sam brings his too.	Ou sagash ra miavarad va sam khodash ra miavarad
2	I caught the first <u>train</u> before you caught the second .	Man gereftar-e ghatar-e aval ghabl az shoma gereftar-e dovom
3	Please help yourself with another <u>biscuit</u> . No, thank you. It was my fourth .	Lotfan khodetan r aba biskooit digar komak konad. Na motoshakeram. Chaharomin -e man bood.
4	You have a big <u>share</u> , but mine is a minor one .	Shoma yek eshtarak bozorg darid, ama man jozie ast.
5	All men have <u>dream</u> and I have my own too.	Hame mardan roya darand va man bish az had .
6	Bill likes her <u>story</u> even though he hates other ones .	Bil dastan-e ou r dost dasht hata agar az anhaye digar motonafer bood.
7	<u>The men</u> returned at midnight. Both were worn out.	Mardan dar nim-e shab bargashtand. Har do farsood-e boodand.

Investigating Research Question 2

The extracted cases of verbal ellipsis were then compared with their Persian counterparts. Analysis of the Persian translations provided by the Mizan and then GT was indicative of the fact that in those cases where Persian and English show some similar verbal ellipsis constructions, the human translator keeps the translation quite close to the original text, especially by retaining the ellipsis. However, in many cases, the elliptical forms are language-specific. In such cases, it is not possible to keep the ellipsis and the translator has to render the text in a non-elliptical form in order to provide the appropriate text, so to comply with Persian grammatical norms; that is, the gap resultant of verbal ellipsis in English is usually recovered by the antecedent verb or replaced by a verb.

Table 4. English-Persian mismatches in verbal ellipsis (Mizan)

No.	English Sentence	Persian Translation
1	I couldn't <u>write</u> with them in the same room with me, but I could with Harold.	Man nemitavanam ba anha dar yek otagh benevisam vali ba Harold mitavanam benevisam
2	You might <u>do it</u> , but I won't .	Shoma Momken ast in kar ra bokonid vali man anjam nemidaham .
3	She won't <u>laugh</u> , but Jake will .	Ou hargez nemikhandad vali jak khahad khandid .
4	Josh likes to <u>sleep late</u> , and Hillary likes to as well	Jash dost darad shabha dir bekhabad va hilari ham niz shabha dir mikhabad
5	The people never do , who say they will <u>help</u> .	In afraad hargez komak nakhahand kard , che kasi chenin harfi zade ast?
6	Tom <u>threw</u> a ball and Harry did too.	Tam va heri tup ra andakhtand .
7	The glass <u>shattered</u> , and the plate did too.	Livan va Boshghab har do shekastand .
8	Don't like to <u>bother</u> no one unless we have to, which we do , in your case	Nemikhahim kasi ra aziat konim magar inke majboor shavim ke dar mored-e shoma chenin kardim .
9	I did not see Edsels when you did.	Man edselz ra zamanik-e shoma ou ra didid nadidam
10	Tom likes Mary, and Harry Suzan.	Tam mari ra dost darad , vali hari soozan ra milkhahad.

Three general patterns were emerged from the analysis of Verbal ellipsis in Mizan Corpus. The first category of ellipsis was the replacement of auxiliary or modal with the main verb in its Persian Translation. For instance, in sentence 9, the auxiliary "did" is replaced by the verb "Didid". The same was true for sentences 1 where "could", modal verb, is completed by the main verb "Benevisam". In number 8 the auxiliary verb "do" is replaced by "Kardim".

In the third pattern, the ellipsis of the English sentences is totally ignored by Persian Translation. Therefore, it can be concluded that the third pattern is the opposite of the second pattern. As an example, in sentence 10, the two sentences are merged into one Persian sentence therefore the ellipsis of the English sentence is totally disappeared in the Persian translation. The same can be seen in sentences 6 and 7.

Table 5. English-Persian mismatches in verbal ellipsis (GT)

No.	English Sentence	Persian Translation
1	I couldn't <u>write</u> with them in the same room with me, but I could with Harold.	Man nemitavanesta ba anha dar haman otagh benevisam, ama man ba Harold mitavanam .
2	You might <u>do it</u> , but I won't .	Shoma momken ast an ra anjam dahid vali man nakhaham kard .
3	She won't <u>laugh</u> , but Jake will .	Ou nemikhahad bekhandad, ama jik khahad .
4	Josh likes to <u>sleep late</u> , and Hillary likes to as well.	Jash dost darad khabidan-e dir va hilari hamchenin .
5	The people never do , who say they will <u>help</u> .	Mardom hargez nemikonand . Ki migooyad anha komak khahand kard?
6	Tom <u>threw</u> a ball and Harry did too.	Tam top ra andakht va hari bish az had anjam dad .
7	The glass <u>shattered</u> , and the plate did too.	Shishe shekaste shod va zarf hamchenin .
8	Don't like to bother no one unless we have to, we do , in your case	Dost nadarim hich kas ra azar dahim magar dar morede shoma .
9	I did not see Edsels when you did .	Man Edsel ra zamanike shoma anjam dadid nadidam.
10	Tom likes Mary, and Harry Suzan.	Tam dost mery va hary soozan.

When the two languages present similar construction, Google translator (GT) also produces a quite reasonable translation. However, in cases where Persian does not allow ellipsis, GT fails to recover the gap left by zeroed material in the source text. Auxiliary verbs also pose some specific problems, as GT translate them into light or lexical verbs.

Investigating Research Question 3

For the last type of ellipsis, clausal ellipsis, the Persian translations were analyzed but almost no mismatches found with the English sentences.

Table 5. English-Persian mismatches in clausal ellipsis (Mizan)

No.	English Sentence	Persian Translation
1	All the students <u>were watching TV</u> in the their dorms the same as teachers .	Hamey-e danesh azmoozan dar khabgaheshan mashghoole didane tevelizion boodand. Moaleman Anha ham haminkar ra mikardand .
2	Jennifer <u>has taken all the stuff</u> she needed and also her medicine.	Jenifer har chiz ra ke niaz dasht va hamchenin daroohayash ra gereft.
3	A: How did <u>they enter here</u> ? B: I'll show you how .	A: Anha chgoone varede inja shodand? B: Behet neshan khaham dad .
4	A: When <u>will you be leaving</u> ? B: Tomorrow .	A: Che zamani khahi raft? B: Farda .
5	A: I know <u>that his check is still valid</u> . B: I know too .	A: Midanam ke hanooz chekash motabar ast. B: Manam hamintor .
6	Our plane <u>has landed</u> , has it ?	Havapeimay-e ma forood amad, intor nist ?
7	A: I just heard John <u>has an operation</u> . B: He has ?	A: Shenideam ke jaan amal dashte ast. B: Jedi ?
8	A: Can you read this book <u>without your glasses</u> ? B: No, but I can just look at its photos .	A: Aya mitavani in ketab ra bedoon-e eynak bekhani? B: Na, vali be tasavirash negah mikonam.
9	A: When did <u>you mentor arrive</u> ? B: Yesterday .	A: Morabie shoma key resid? B: Dirooz .
10	A: <u>Are you coming</u> ? B: Yes .	A: To ham miayi? B: Bale .

Then all the sentences were translated via GT in order to investigate the way clausal ellipsis will be translated by a machine. But this time, GT has provided insufficient translation for this kind of ellipsis.

Table 6. English-Persian mismatches in clausal ellipsis (GT)

No.	English Sentence	Persian Translation
1	All the students <u>were watching TV</u> in their dorms the same as teachers .	Hameye danesh amoozan dar khabgah-e khod televizion didand hamanande moaleman .
2	Jennifer <u>has taken all the stuff</u> she needed and also her medicine.	Jenifer hame chiz ke niaz dasht ra gerefte ast va hamchenin pezehki-e khod ra.
3	A: How did <u>they enter here</u> ? B: I'll show you how .	A: Chegoone anha inja vared shodand? B: Man be shoma chegoone neshan midaham.
4	A: When <u>will you be leaving</u> ? B: Tomorrow .	A: Hengamike shoma tark mikonid? B: Farda.
5	A: I know <u>that his check is still valid</u> . B: I know too .	A: Man midanam ke chek-e khod hamchenan motabar ast B: Man bish az had midanam .
6	Our plane <u>has landed</u> , has it ?	Havapeimaye ma forood amad, an ra darad ?
7	A: I just heard John <u>has an operation</u> . B: He has ?	A: Man faghat shenidean jaan yek amaliat darad B: Ou darad ?
8	A: Can you read this book <u>without your glasses</u> ? B: No, but I can just look at its photos .	A: Mitavanid in ketab ra bekhanid bedoone eynak-e khod? B: Na ama man mitavanam be axs ha negah konam
9	A: When did <u>you mentor arrive</u> ? B: Yesterday .	A: Vaghti ke morabi shoma nayamad? B: Dirooz .
10	A: <u>Are you coming</u> ? B: Yes .	A: Dari miayi? B: Are

Summary of findings

This study was carried out on both human and machine translations to investigate their qualities. The data gathered in this study were analyzed comparatively based on Haliday and Hasssan's (1976) taxonomy.

- There were mismatches in accordance to verbal ellipsis in English-to-Persian translation.
- There were mismatches in accordance to nominal ellipsis in English-to-Persian translation.
- There were NO mismatches in accordance to clausal ellipsis in English-to-Persian translation.
- There is a significant difference in the quality of HT and MT (i.e. GT) in favor of HT.

CONCLUSION

Returning to the question posed at the beginning of this study, it is now possible to state that in the case of nominal ellipsis, Mizan could be considered as a reliable source for translation, whereas GT provides insufficient translation. The majority instances of verb ellipsis after auxiliary *be* occurred in a confirming answer to a previous statement; for which the adopted translation strategy was using the confirming statement. Regarding GT, it does not recover the gap resulting from verbal ellipsis. It also translates all the operators *do*, *be*, and *have*, as a lexical or light verb; it lacks person and number agreement between the subject and the verb; and the tense is not preserved. GT in dealing with verbs ellipsis after modal verb *can* operates fairly well; however it fails in dealing with other modal verbs: after *will* the translation is 'subject + modal verb the passive voice'; after *may* and *must/should* it gives the unnatural combination of subject and the modal verbs following it, as the gap needs to be recovered by the antecedent verb or replaced by a pro-verb. It seems that GT, in dealing with this particular verbal ellipsis, mostly produces inadequate translations. From the collected evidence, it was not possible to discover why GT only performs properly in some cases. The results indicate that the Persian human translator dealing with English verbal ellipsis predominantly adopts the strategy of recovering the zeroed verb from its previous occurrence in the discourse. More than 250 sentences with clausal ellipsis were found in Mizan; however, no cases of clausal ellipsis mismatch between English and Persian was observed in Mizan; the findings of this study about ellipsis were in line with the results of the previous studies in the literature, like Nourmohammadi (1988) and Hashemian (1989). Research into the translation process tries to analyse the psychological reactions of translators as they translate, using methods including Kussmaul's (1995) 'think-aloud protocols' and Jakobsen's (2003) Translog software for tracking translator's work patterns on the computer. The quantity of analysis of the finished result of translation is enormous, but not much is conducted in a systematic manner, despite efforts by such people as House (1977, 1997) to introduce functional analysis of translation, Baker (1998) and Laviosa (1998) to observe tendencies in translation using translation corpora, and attempts to establish 'universals' of translation (see Mauranen, 2004).

The main concern of the paper was to compare the quality of Google Translate and human translation. Summing up the results, it can be concluded that the difference between the frequencies of different types of ellipsis have some mismatches between English and Persian languages. Therefore, the types of ellipsis did not affect the quality of translation of machine translations. The single most important consideration in the quality of Google Translate was to help users decide if the Google Translate will best suit their needs and if they can trust on its translated outcomes. From the research that has been undertaken based on Keshavarz's model (1999) of error analysis, types of errors and their frequencies were identified to accomplish automatic metrics evaluations with the purpose of improving the systems.

Machine translation (MT) is developing increasingly such that possibility of replacing Human Translation (HT) does not seem unlikely. However, it is controversial to what extent machine can replace human beings in certain areas including translation. To shed light on this issue, this study was conducted to compare the efficiency of Google Translation (GT) and HT in relation to different kind of ellipsis posed by Halliday and Hassan (1976).

REFERENCES

- Baker, M. (1993). Corpus Linguistics and Translation Studies: Implications and Applications. In *Text and Technology: In Honour of John Sinclair*, ed. Mona Baker, Gill Francis & Elena Tognini-Bonelli. Amsterdam & Philadelphia: John Benjamins.
- Baker, M. (2007). Corpus-based Translation Studies in the Academy. *Journal of Foreign Languages*, 5.
- Halliday, M.A.K. & R. Hassan (1976). *Cohesion in English*. London and New York: Longman.
- Laviosa, S. (1998). The Corpus-based Approach to Translation Studies. *Special Issue of Meta*, 43, 4-7.
- Mauranen, A. (2004). Translation universals. In Brown, Keith (ed.), *Encyclopedia of Language and Linguistics*. Vol. 13. Oxford: Elsevier, 93-100.
- Supreme Council of Information and Communication Technology, Mizan English-Persian
- Parallel Corpus, (2013). Available: <http://dadegan.ir/catalog/mizan> [2014-01-01].
- Zhang, M. F. (2009). From Experience, Text to Deconstruction of the Concept of Translation – A Glimpse of Translation Textbooks in Western Countries. *The Journal of PLA Foreign Language Institute*. Luoyang, China.
- Zanettin, F. (2000). Parallel Corpora in Translation Studies. Issues in Corpus Design and Analysis. In M. Olohan. (ed.), *Intercultural Faultlines. Research Models in Translation Studies I. Textual and Cognitive Aspects*. Manchester: St Jerome, 105-118.