

Discovering the Challenges of Machine Translation from Dari Language into English Language

DOI: <https://doi.org/10.47175/rielsj.v3i2.471>

| Mohammad Shouaib |

Parwan University,
Afghanistan

ABSTRACT

Machine translation involves the use of unnatural computer programs to translate from one language into another language without the involvement of human translators. It is not just about substituting words in one language for those in another. Machine translation involves the use of computational linguistic techniques to recognize and translate whole phrases while also considering the complex structure and figures of speech that are unique to each language. The current paper aims at casting a new light and exploring the problems of Machine Translator as an aid used by the students of translation to render from Dari into the English language. Also, the study aims at finding out such problems encountered by these students who study translation subject. To achieve the objectives of the study, students who were enrolled in the translation subject in the faculty of language and literature, English department for female students during the academic year 2020/2021 who were randomly selected by the researcher. A questionnaire with multiple choices was administered. The collected data were analyzed properly. The study has come out with spectra of results among them are the followings: The students who were the sample of the study faced various types of problems such as syntactic and semantic problems while using machine translation for rendering their given tasks. The rendition of target language used in the machine translator is inaccurate. And translation some specific cultural terms through a machine translation were out of context.

KEYWORDS

problems of machine translation; syntactical; semantic

INTRODUCTION

The globalization and growth of technological advancements touch every part of our lifestyles, fittingly; expressing information in several languages has grown to be one of the most important characteristics in communication. This needs substantial levels of rapidity and efficiency in translation facilities. Machine Translation is a complex and challenging research area because language translation in itself, is difficult. While human cognitive processes language interpretation or understanding, and translation on many levels, a machine processes data, linguistic form and structure, not meaning and sense. To fully understand the complexity of designing a machine translation system, it is necessary to analyse the functions of the human brain during a translation process. While human beings translate they usually start by attempting to decipher the source text on three levels:

- Semantic level: understanding words out of context, as in a dictionary.
- Syntactic level: understanding words in a sentence.
- Pragmatic level: understanding words in situations and context.



Structural differences between languages also present challenges for machine translation. Languages often differ in the basic word order of Subject-Verb-Object. English, SVO (Subject-Verb-Object) language, meaning that the verb usually comes in between the subject and object. In contrast, Dari - Persian SOV (Subject-Object-Verb). Therefore, developing a system for language pairs with different word orders such as English to Dari will be more difficult as compared to development for language pairs with the same word orders such as English to Dari language. Additionally, tenses that exist in one language may not exist in another language. English, for example, has explicit present progressive and present structure, whereas Dari has only one tense that encompasses both of these English structures.

Objectives and Questions of the Study

This paper identify and describe the problems faced when using a Machine Translator in translation from Dari language into the English language by the students of English classes majoring in translation in Parwan university Faculty of Language and Literature, English department. It also attempts to understand the causes of these problems and give some recommendations to help overcome them. In order to accomplish these objectives, the study clarified the following questions:

1. What are the problems that the students majoring in translation while using a Machine Translation?
2. Is the language used in the Machine Translator easy to understand?
3. Does Machine Translation imply the disappearance of a human translation?

Significance of the Study

A lot of research work carried out in Translation studies is on various literary problems that translators faced in translation in general, however practical studies that deal with Machine Translation in the Dari language are relatively small and only focus on one human translation. Therefore, this study is designed to fulfill a gap in literature since it purposes to identify and describe the problems faced by students when using a Machine Translator. The result of this study and recommendations for future research will help other scholars who wish to begin research on this issue. However, the results of this study cannot be generalized further than the selected sample.

Sample of the Study

In this paper, the sample selected for this research consists of 50 students from Parwan University Faculty of Language and Literature English Department who is study translation subject. There was clearly logic behind why this particular group of students was actually recommended as the participants. First, the actual requirements regarding selecting the learners were determined by their usage of the computer aid in their studies. Second, they were in the final year and taught in English. Third, they are native speakers of the Dari language. They were 50 female learners, aged between 21 and 25 years old.

Instruments of the Study

In this study, the researcher applied a questionnaire as a tool to gather data. The questionnaire was conducted, particularly to reach the objectives of this present research. The questionnaire included 10 specific questions. A letter which explained the scope, objectives of the study and the confirmed permission to carry out the questionnaire was presented to the participants. The questionnaire was created to associate the barriers that the learners encountered and the causes to their problems.



LITERATURE REVIEW

Researchers in the field of natural languages have made a serious attempt to back up manual translations by using machine translators. Consequently, Machine Translation therefore is considered as a valuable subject for researchers, profitable to developers and users. Researchers want to stratify their concepts to find out the dissimilarities that can be made by machine translators. By doing so, it will be easier for designers to identify the most challenging issues and make enhancements on the machine translators.

(Riazi 2003) said that translation of Dari sentences into English language was a problematic task. The difficulty comes from various sources, one being sentences in Dari language are too long. Another challenges the sentence structure. Dari phrase is actually syntactically unclear and complex, due to the usage of many grammatical relationships, word order and content along with conjunctions. Therefore, most of the studies in Dari / Persian Machine Translator mostly focus on the translation from English to Dari. The need to deal with the arrangement and the order of words in a machine translation from English language to Dari language. Also, offered hybrid-based strategy to handle those problems. Moreover, a couple of characteristics that had an impact on the ordering issue that were derived from the fact that various languages have different text orientation. remarkable differences between the Syntax of the Dari language and that of English language are another source of difficulty.

(Keshavrz M. H 1999) said that an important feature of Machine Translation is to maximize the meaning of text so that minimum attempts and fewer times are needed to comprehend the output. The operator should not put upwards too much effort to join the various elements of the translation. Moreover, an excellent Machine Translation should try to go for an additional step away from the essence level. Procedures required to be developed and improved so that the output can touch the excellent product possible with small editing needed.

In addition, (Keshavarz, M. H. 999) indicted that deletion and addition were problems that Machine Translation wants to look at so that its output is a reproduction of the source language text with no elements deleted or extra elements added. Spelling is another problem that requires attention.

Moreover, the problems of Machine translation from Dari language into English language as: first, non-vocalization is a problem of lexis that leads to a wrong choice of words in the target language and hence a major cause of interpretations. The second lexis problem is inadequate lexicon, rendering it producing completely wrong meaning of text for instance, the name of a place or a person. A third problem of lexis is words with multiple meanings, several Dari terms might have a couple of overlapping connotations in English language and the system want to determine which one to choose, for instance the term مرکز can mean center, position, rank, status. A fourth difficulty associated with lexis is having multiple senses; cultural features associated with the Dari language issuing constructs that literally mean 'friend of', 'mother of', and 'father of' to show possession.

Translation program, which usually translates a reasonably difficult English noun phrase into Dari language. In addition, Machine Translation approach is favorable and may be used to automate the translation of thesis headings within the computer science domain.

During translation of questions from Dari into English, several translation errors appeared which are of the type: wrong pronoun, wrong word order, wrong word sense and wrong transliteration. The decoded questions were fed into Answer Finder, which had a huge influence on its accuracy in returning correct answers. Answer Finder was greatly affected by the relatively reduced output of machine translation.



To overcome such problem first is to make some modifications to the question translation process to reduce the influences of translation by automatically editing remarkable regular errors using a regular written expression. Second, try constructing an interactive Machine Translation system by providing users more than one translation options to pick a more accurate option from.

(Feder, 2003) said that its recognizable from the common definitions of Translation Studies and machine Translation, that Translation Studies inspects translations, whereas, machine Translation are mechanical tools used to create translations. Translation Studies transacts with artistic and assessment, a component of the translation process, while Machine Translation concerns technical aspects and therefore, does not transacts with the translation process as a subjective and complex process involving for instance, cultural knowledge. Although the two fields have the same subject matter, which is text, they handle it differently. Translation Studies examines and evaluates text, creation and purpose of translation, whereas Machine Translation's emphasis is on how to help human translators in the creation of target text and on how to make this job easier and faster. These two components may be considered complementary, but their goals are obviously different.

(Hutchins, 2001) said that since the concepts of applying computers aids to render normal languages was initially suggested from the 1940s as well as the primary inquiries were started in the 1950s, translators have seen improvement possibly in contempt or in fear. Moreover, they have discarded the idea that everyone might even think that translation can be mechanized, or they are even scared that their own career could be taken over entirely by machines. On the other hand, there is no hesitation that computer-based translation devices are certainly not competitors to human being translators, however they generally assist hem in order to enhance efficiency, throughout a complex translation they have ever attempted. In this context (Hutchins, 2001) distinguished:

1. Machine Translation devices, which purposes to pledge the entire translation procedure, but whose production must positively be reviewed.
2. Machine Translation (translation Aids), which assist the particular expert translator.
3. Translation devices for the non-translator individual user, which create simple versions to help in understanding. These types of distinctions were not identified before the 1980s.

Moreover, (Hutchins, 2001) said that the major emphasis of Machine Translation study is to the development of systems that translate written scripts of scientific as well as technical nature, away from systems that translate literary and legal texts. In fact any kind of texts messages where style with presentation are essential elements of the message. On the other hand, there are apparent possible advantages even when the achievement is only partial.

(Salem et al., 2008) said that more to the problems involved in creating an efficient translation aids from Dari language into English language, the word order of Dari language creates hindrances to the language in translation process.

Moreover, (Salem et al., 2008) stated that Dari language has a great set of morphological features. These kinds of characteristics are generally available such as prefixes or suffixes that can entirely enhance the particular sense of the word. Furthermore, in Dari there are remarkable words that carry the definition of a complete sentence, for instance ما سفرمی کنیم which mean, in English we will travel. Dari free word order creates an enormous challenge to Machine Translators due to the vast possibilities to express the same sentence in English.

(Bowker and Ehgoetz, 2007) carried out a research to discover user approval regarding Machine Translation productivity, using time, cost and quality, as three variables for assessment. Moreover, (Bowker and Ehgoetz, 2007) asked experts of translation to judge three various target language texts of the identical source language texts. Three distinct target



language texts of every source language text were created: raw Machine Translation production, post edited Machine Translation output and human translation. To be able to improve the raw Machine Translation production, first, translate the source language texts with the Machine Translation system, recognize unfamiliar words and added entries of these terms to the Machine Translation dictionary.

(Keshavarz, M. H 1999) conducted a research to explore the relationship concerning quality and productivity of the post editing results from translation reminiscences as well as Machine Translation with regards to texts translated without any assistance. Quality has been assessed as the number of errors in the target language text. The mistakes were identified, measured along with processing speed was estimated as the number of source language words processed in each minute.

(Hutchins, W & Somers H. 1992) said that the investigation of the samples of raw Machine Translation production coming from various Machine translation systems which, depicted that the mistakes are a great deal in common. A number of mistakes happened in every language pairs, no matter the system applied. Moreover, the general classification of Machine Translation mistakes happening regardless of language pair and Machine Translation system. The primary mistake categories were grammatical, syntactic, lexical and errors due to imperfect input. Explored Machine Translation from the viewpoint of contrasting it with human being translation. (Fiederer and O'brien, 2009) carried out the investigation to see if Machine Translation production actually decreased the translation quality than human translation. They selected 30 phrases from an individual manual in English which were both translated by a human being and a machine. The majority of the evaluators preferred human translations. Class experiments making use of Google Translator. The main goal of the research was to motivate the learners to determine issues with their translation processes along with technology. The participants of the research were 19 second year Master's degree students. Quantitative data was obtained by means of figuring out the entire period required to produce the last translation. The findings of the experiment showed that there was no important variation in the period taken to create the production by the Machine Translation and without it, not worthy distinction relating to the language groups. More so, it had no systematic dissimilarity between the qualities of the translations as evaluated by the learners.

(Hutchins, W & Somers H. 1992) said that Machine translation is definitely an autonomous computer system with approaches and strategies which might be labeled as: First, the direct approach to be used within machine translation devices, requires at the least a linguistic concept. The direct method depends on parade fined source language and target language binomial in which every expression of the source language system is straightway connected to a similar component in the target language with a unidirectional association.

Second, the transfer technique focuses on the concept of level representation and consists of three levels; the study level, the transfer level and the generation level. The study level presents the original language text message linguistically along with an original language dictionary. The transfer level changes the outcome with the investigation level as well as determines the linguistic along with structural equivalents involving the pair of languages. This relies on a bilingual lexicon from the source language into target language. The generation level creates a new text within the second language on the basis of linguistic information from the original language through a second language dictionary.

Third, the axis language approach is around the notion of producing a text message free without any specific language. This specific rendering purposes as being fairly neutral, common axis which is distinctive completely from both source and target language. Theoretically, this technique reduces the machine translation procedure to two steps:



evaluation and production. The study of the source text guides to a conceptual rendering, the different elements which might be united through the production component within the equivalents in the second language. The study on this approach is related to artificial intelligence and representation of knowledge. The systems in line with the concept of a pivot language tend not to intention at straight translation, but alternatively reformulate the original text message from the crucial information.

(Keshavarz 1999) pointed out that Dari language has always been a challenge for machine translation because of its rich and morphological complex features. Moreover, Dari / Persian has a variety of word forms and word orders which make it possible to express any sentence in various forms. Furthermore, the existence of several dialects and the fact that the word order is not usually identical to the source language and target languages, this leads to the opportunity of having more than one meaning for the same sentence. The exactness of any machine translator is generally evaluated by matching results to human judgments translation. One of the techniques used to assess machine translation systems is called Bilingual Evaluation Understudy which was introduced in the study which, claimed to be language independent and highly correlated with human evaluation.

Currently and in future, uses of Machine Translation are limited to significance translation, or a quick translation for smart users, when individual translation is actually out of question as a result of time and other issues. The Machine Translation is intended at serving to learn user transacting with transitory texts, generally speaking, they assist communication in many circumstances.

It is really understandable that human being translators must react undesirably in order to accept the idea of Machine Translation. This is partly simply because their own particular traditional education has made all of them to assume a top standard regarding functionally modified or innovatively translated literary texts, and they find the Machine Translation results improper.

The encouraging aspect associated with enhanced communication through Machine Translation, for the human being translator, is that it stimulates curiosity about texts in unidentified languages with individuals who would previously have merely ignored their reality. In the long run, this inquisitiveness can only lead to a request for better human being translation. In fact, it is possibly true to say that English is a bigger threat to multilingualism and the translator than Machine Translation.

RESULTS AND DISCUSSION

Machine translator represents an actual barrier to the students in translating from the language to the English language. Answers from the questionnaire were presented in the following categories;

How often do you use the Machine Translator?

The descriptive analysis for the question; “how often do you use the Machine Translator?” is shown in Table 1 and Figure 1. According to the frequency test for this question, a majority of the respondents, that is 50 percent, said they used Machine translation everyday while only 15 percent of them said they used Machine Translation a few times in a week. However, only 20 percent among the 50 participants used Machine Translation a few times in a fortnight. From the total participants only 7 percent used the Machine Translation once in a while. Figure 1 shows all the results from the questionnaire of the 50 participants.

Table 1. How often do you use the Machine Translator?

		Frequency	Percent	Valid Percent	Cumulative Percent
	Everyday	27	50.9	54.0	54.0
	A few times in a week	8	15.1	16.0	70.0
	A few times in a fortnight	11	20.8	22.0	92.0
	Once in a while	4	7.5	8.0	100.0
	Total	50	94.3	100.0	
Total		53	100.0		

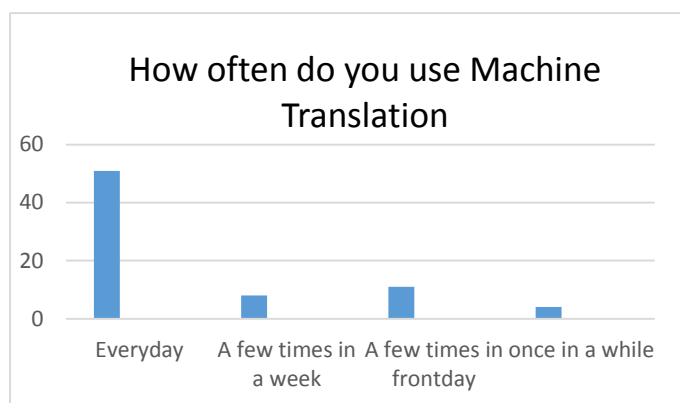


Figure 1. Does the Machine device Translator serve the purpose of its creation?

The descriptive analysis for the question; “does the Machine device Translator serve the purpose of its creation?” is in Table 2 and Figure 2. According to the frequency test for this question, among the 50 participants, 32 percent who used the Machine device translator said it needs improvement, while only 28 percent of them said that it serves the purpose of its creation to some extent. However, only 15 percent among the 50 participants who used the Machine Translation said yes the Machine translation serve the purpose of its creation. From the total participants only 18.9 percent who used the Machine Translation said no it does not serve the purpose of its creation.

Table 2. Does the Machine device Translator serve the purpose of its creation?

		Frequency	Percent	Valid Percent	Cumulative Percent
	Yes	8	15.1	16.0	16.0
	No	10	18.9	20.0	36.0
	To some extent	15	28.3	30.0	66.0
	Needs improvement	17	32.1	34.0	100.0
	Total	50	94.3	100.0	
Total		53	100.0		

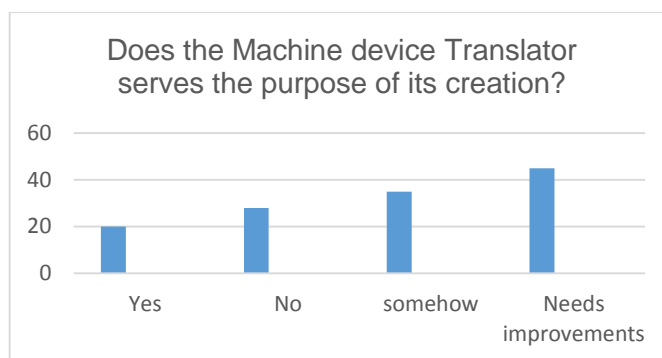


Figure 2. Does the Machine device Translator serves the purpose of its creation?

Is the language utilized in the Machine Translator simple to be comprehended?

The descriptive analysis for the question; “is the language utilized in the Machine Translator simple to be comprehended?” is shown in Table 3 and Figure 3 According to the frequency test for this question, among 50 participants, the majority of the respondents, 64 percent said no, because the language used in the Machine Translator is not easily comprehensible. While only 13 percent of them said the language used in the Machine Translator easily understood to some extent. However, only 2 percent among the 50 participants who used the Machine Translation said yes the language used in the Machine Translator is easily understood. From the total participants only 15 percent who used the Machine Translation said that language used in the Machine Translator needs improvement. Figure 3 shows all the results from the questionnaire of the 50 participants.

Table 3. Is the language utilized in the Machine Translator simple to be comprehended?

		Frequency	Percent	Valid Percent	Cumulative Percent
	Yes	1	1.9	2.0	2.0
	No	34	64.2	68.0	70.0
	Somehow	7	13.2	14.0	84.0
	Needs improvement	8	15.1	16.0	100.0
	Total	50	94.3	100.0	
Total		53	100.0		

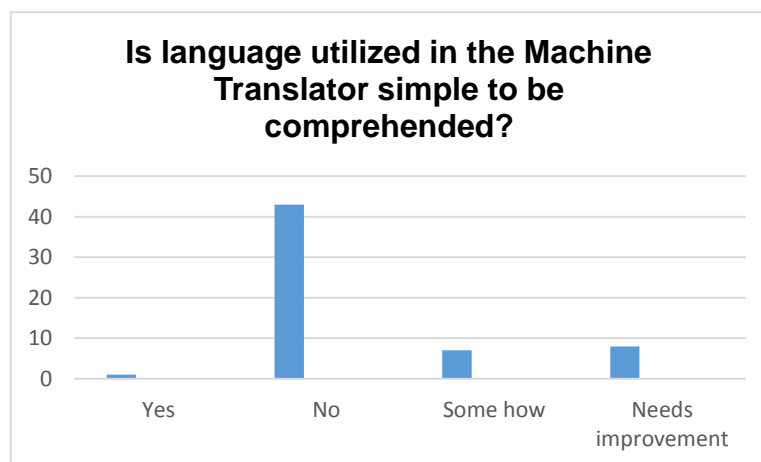


Figure 3. Is the language utilized in the Machine Translator simple to be comprehended?

What type of information do you generally search for in a Machine Translator?

The descriptive analysis for the question; “what type of information do you generally search for in a Machine Translator?” is shown in Table 4 and Figure 4. According to the frequency test for this question, among 50 participants, a majority of the respondents, 39 percent said they used Machine Translation to increasing their vocabulary. While only 35 percent of them said they used Machine Translation to understand the meanings of the terms when translating a text. However, only 13 percent among the 50 participants who used the Machine Translation said they understood the meanings of the technical terms. From the total participants only 5 percent used the Machine Translation to learn new words. Figure 3.4 shows all the results from the questionnaire of the 50 participants.

Table 4. What type of information do you generally search for in Machine Translator?

		Frequency	Percent	Valid Percent	Cumulative Percent
	For understanding the meanings of the technical terms	7	13.2	14.0	14.0
	To learn a new word	3	5.7	6.0	20.0
	To understand the meanings of the terms when translating a text	19	35.8	38.0	58.0
	For increasing your vocabulary	21	39.6	42.0	100.0
	Total	50	94.3	100.0	
Total		53	100.0		

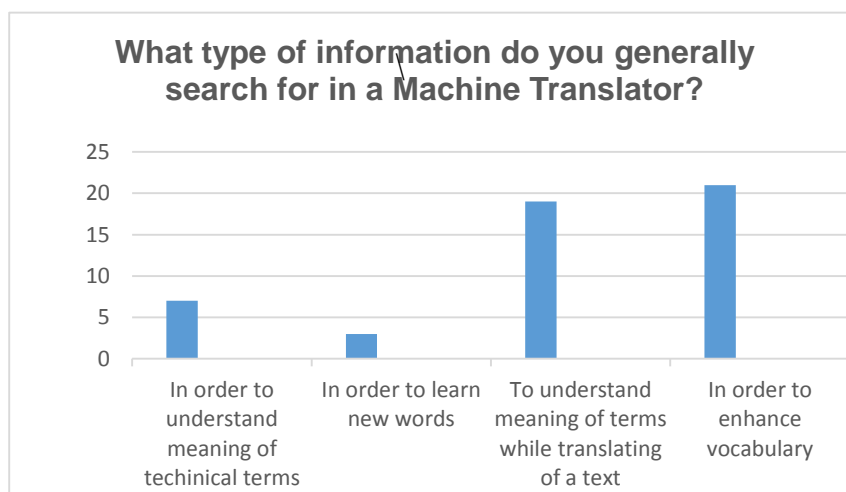


Figure 4. What type of information do you generally search for in Machine Translator?

Findings disclosed that Machine Translation represent actual obstacles for the students in Parwan University, in translating from Dari/ Persian language into English language. The frequency tests show that a majority of the students used the machine translator every day. Due to various development and present dynamic world there is need for a more effective means of translation, since there are not many human translators, or because individuals and organizations do not see translation as a complex activity demanding a high level of experience.

In answering the question whether or does the Machine Translator serve the purpose of its creation, a majority of the students, 32 percent of them said that the machine translation need some improvement. However, Machine Translation helps the students to make their job



more efficient and faster. Therefore, the answer for the question “Does the Machine Translator serve the purpose of its creation?” showed that the Machine Translation needs to be improved. The idea grew to become extensively assumed that, the aim of Machine Translation has to be improved to completely automate to create quality translations. The usage of human being help was regarded as a temporary arrangement design as reported by (Bar-Hillel, 1960).

The question whether or not the language used in the Machine Translator was easily understood, a majority of the students, 64 percent said no, because the language used in the Machine Translator was not easily understood in the translated script. Machine Translation cannot give a proper translation, because some terms are peculiar to a specific field and have no equivalent in target language. Most at times, the text translates well, and it can be easily comprehended, but other times, there are mistranslated expressions or sentences that do not follow proper syntax and can prevent understanding. The raw machine translation product is not considered a high-quality translation equivalent to what a human translator can produce, because the translation will need to go through different degrees of editing by a human translator before it can be used or publicly distributed.

Moreover, in answering the question, “what kind of information did you frequently look up at Machine Translator?” a majority of the respondents, 39 percent said they used Machine Translation to increasing their vocabulary, while translating the word or the term from Dari Language into English language. Machine Translation grants instant translations between dozens of various languages. It can translate terms or sentences between any alliances of various languages.

In addition, the question could you find all the words that you were looking at Machine Translator? A majority of the respondents, 73 percent said most of the times they found all the words they were looking for using the Machine Translator. The problem of machine translators can translate words, but they cannot translate meaning. By definition, a machine translator will never comprehend the definition of anything. Sometimes they are some words or terms that are difficult to find them while translating because they are peculiar to particular languages.

Language is something that just individuals will be capable to completely comprehend and translate. Machine Translator arranged to produce automated translation to reasonably impressive levels, but machine translation will never be able to compete with human translators. The various forms, circumstances, cultures and differences included in the language are merely a few basic items that machines can't understand.

There are quite a few characteristics across languages that help to highlight why individuals will continuously have the upper-hand over machines when it approaches to translation. Several words really don't translate well between languages.

Finally, Idioms slangs and some words which has various meaning were problematic for the Machine Translation because they did require knowledge of English culture and they are peculiar to cultural specific concepts. The overall percentage score calculated for these words was nearly seventy one percent and this provided supporting evidence that these words were not easy to translate.

In the process of translation, automated or human the sense of specific cultural concepts text in the source language must be fully restored in the target language while translating. On the surface this appears straightforward, it is far more complicated. Translation is not a mere word-for-word replacement. A translator must explain and investigate all of the components in the text and know how every word may affect another. This needs great expertise in target language cultural. The greater difficulty rests in how machine translation can provide publishable quality translations.



The current Dari language usually is well-known as having arrangement asymmetries that are sensitive to word order effects. Most of these asymmetries were caused by a range several of effects problem first, through the investigation of the problems at the source language and second, the particular generation of difficulties in the target languages. Languages usually are different in the agreement demands. A number of such languages as Dari language need person, gender, number along with case agreements. Machine translation process grows by utilizing a number of strategies determined by their particular issues and difficulty.

CONCLUSION

The outcome of the investigation shows that the Machine Translation activity from Dari language into English language faced many obstacles in the translation process. In contrast, a language is simply a subject important for mankind; the idea of language poses enormous difficulties for machine Translation. The real reason for this is the practically infinite variety in a natural language. The words as well as rules along with how they can be linked together vary considerably from language to language. Although each and every language have common structures, commonly named deep structures. Simple translation applications depend on surface structure and they render one word after another. Several aspects promote the incorrect creation of machine translation. Human natural language is complicated, vague, ambiguous and imprecise. Words having more than one meaning, sentences with grammatical structures having several meanings, the identification of pronouns and other grammatical difficulties lead to translation software to fail.

Several deficiencies in the production of Machine Translation have been presented in this paper, due to either inadequate interpretation of the users or faulty generation of the target language words. Totally automated, great quality machine translation has not yet been attained. Still there is a lot that we can do to enhance the quality of Machine Translation production and expand its utility. In this paper, we have displayed the need to handle machine translation problems when translating from Dari language to English language.

REFERENCES

- Abdul Mehdi Riazi (2003). Translation Journal and Author Shiraz University, Iran
- Ashtani Aghlili Ali Akbaar. (2004). Translation of Literary Texts, Tehran.
- Beau Grrand (1998). How to Teach Translation. England
- Bowker, L., and Ehgoetz, M. (2007). Exploring User Acceptance of Machine Translation Output: A Recipient Evaluation.
- Bassnet A. (1990). Translation of Literary Subjects. Tehran.
- Clippel G. (2004). Theory of Translation. Cambridge University, England.
- Fiederer, R., and O'brien, S. (2009). Quality and machine translation: a realistic objective. *The Journal of Specialised Translation*. 1152-74.
- Hurchins, W. J. & Somers, H. L. (1992), An Introduction to Machine Translation. London: Academic Press.
- Herbert W. (1999) Methods of Translation. Cambridge University, England
- Keshavarz, M. H. (1999), Brief View on Google Translate machine. Tehran: Rahnama Publication.
- Pasargade Alddin (2002). Art of Translation. Shiraz University, Iran
- Rahimi, A. & Yazdanbakhsh, A. (2009). Dictionary of Spoken Phrases, Tehran, Jungle.
- Tajalli Ghaffar. (2005). Idioms and Metaphorical Expression in Translation. Tehran
- Wala Abdul Rahman. (2000). Basics of Translation. Kabul Press.
- Wren & Martin. (2004). Grammar and Composition. Delhi