Machine Translation: Divergence in Hindi-English Language Pair

Venu Shree, Geetha Yadav, Iti Mathur, Nisheeth Joshi

Abstract: In translation divergence is a key element between two languages. This divergence problem arises when a sentence of a source text/document translated into target sentence and it is structurally different. Machine Translation needs to work upon the problems of divergence occurring between two languages. There are various Machine Translation developers who have already paid attention to this problem yet, it is not that easy to develop common strategies which can be applied for any language pair. Hence, there is a lot of study needed to be done in the field of translation divergence for different pairs of languages. In this present paper, a different pattern of translation divergence between two languages that is the Hindi-English language pair has been discussed. The various constructions for both the languages have been elaborated to identify the divergence in the Hindi English language pair. This will give us a way to come up with strategies to handle these situations and come up with correct translation. This paper is based on Sinha & Thakur (2005), who have based their study on Bonnie J. Dorr's (1994) classification. This paper thus focuses on the various categories of divergence in Hindi-English language pair.

Keywords: Language, Linguistic Divergence, Machine Translation.

I. INTRODUCTION

Language is a multifaceted system of human behavior. It plays an important role when it comes to communication between humans. Every human being can share his/her ideas, feelings, emotions or beliefs to each other as s/he can share common signs/codes that make up the language. There are many other means of communications used by human beings- winks, gestures, flags, nods, short-hand, horns, Braille, alphabet, smiles, mathematical symbols, sketches, morse code, dance, acting, miming, maps and so on. But the above systems are limited and somehow, they too depend upon language only. They are not as extensive, comprehensive, perfect and flexible as language is.It is difficult for us to think of a society without language. Language is so essential a form of communication between human beings that it gives shape to people's ideas, thoughts, and also guides and controls their actions. It is with the help of language that human thoughts and philosophy are conveyed from one generation to the other. It is worldwide in the sense that it is present everywhere in all activities. It is the most priceless thing one could have and it is as important as the air one breathes. In our day today life, all human beings are blessed with language and it is the specific property of humans only.

Venu Shree, Department of English and Modern European Languages, Banasthali Vidyapith Rajasthan, 304022, India.

Geetha Yadav, Department of English and Modern European Languages, Banasthali Vidyapith Rajasthan, 304022, India.

Iti Mathur, Department of Computer Science, Banasthali Vidyapith Rajasthan, 304022, India.

Nisheeth Joshi, Department of Computer Science, Banasthali Vidyapith Rajasthan, 304022, India.

Human beings are called "talking animals" -*homo loquens* because of the use of language.

Animals also have the system of language but it is limited to a very small number of messages like those of thirst and hunger, anger and fear etc. But when it comes to humans, the situation is totally different. We human beings can share infinite number of messages with each other. It is language that helps us to store knowledge, ideas, and thoughts and also it can be transferred to the next generation. Language is a common code that each and every human being uses to express herself/himself to others. There are a number of languages in the world which vary between 6,000 and 7,000. However, any specific estimate depends on language and dialects. Language finds expression both through Listening, Speaking, Reading and Writing. In order to get involved fully in school and society, one must learn the essentials- the structure, the rules, and the conventions of this system. It is through the process of cognition that language takes place in human mind. It is a process by which the sensory input is transformed, reduced, elaborated, stored, recovered and used.Language also extends to social life. The relationship between society and language is firmly secured. The connection between the two is deeply rooted. If any one of the two will not exist, then the other one will be affected. For peace and order in our society, language plays an important role. It is the primary tool for communication purposes, for showing authority and power, for attaining goals and objectives. But if it is not used in a proper way then it can also destruct the society. It must follow the consistency prevailing in the society to avoid divergence and to meet the boundary of individual differences, as each one of us has our own perception and point of view. Society however controls our languages by giving preferences as to what kind of language is acceptable and what not. We human beings must know how and where to use a particular kind of language. It might be possible that a certain kind of language is acceptable by a particular group and the same language is not acceptable by another group. This affects values in ways that have not been accurately understood. Language incorporates social values. However, social values are only the same as the linguistic values when the society is an unchanging one. Once society starts changing, then language changes produce special effects. Some of the most commonly approved definitions of language given by the experts in the field of linguistics are given below: Edward Sapir (1921) says: "Language is a purelyhuman and non-instinctive method of communicating ideas, emotions and desires by means of voluntarily produced symbols".

According to Hall (1969), language is "the institution whereby humans communicate and interact with each other by means of habitually used oral-auditoryarbitrary symbols."

In the words of Noam Chomsky (1957), language is "a set of (finite or infinite) sentences, each finite in length and constructed out of a finite set of elements. (Syal&Jindal, 2011 p.4).As we all live in global world there is a huge demand for translation of texts/documents from one language to another language and especially in such a large multilingual society as India.Translation is the gist of a text and the following production of a subsequent text, similarly called a translation, which communicates the equivalent message in another language. The text that is translated is known as source text, and the manuscript into which it is translated is known as the target text. Translation is eventually a human action which enables individuals to swap information and feelings in spite of the various languages used. Linguistic diversity is reflected in more than five thousand languages existing in the world. It is trick for human beings to know and be familiar with all the languages of humankind. Thus, the method of translation has been adopted to converse the messages from one tongue to another. According to Munday (2016), "The English term translation, first attested in around 1340, derives either from Old French *translation* or more directly from the Latin *translatio* ('transporting'),itself coming from the participle of the verb *transferre* ('to carry over') (Munday, 2001p.7).The human translation process, for instance, may be described as:

    a. Decoding the meaning of the source text.
    b. Re-encoding the meaning in the target language"

There are many linguists who have defined translation in many ways and also described the process of translation. Nida and Taber (1969) in Munday (2001) have given the three stages of translation. The fig. given below shows the
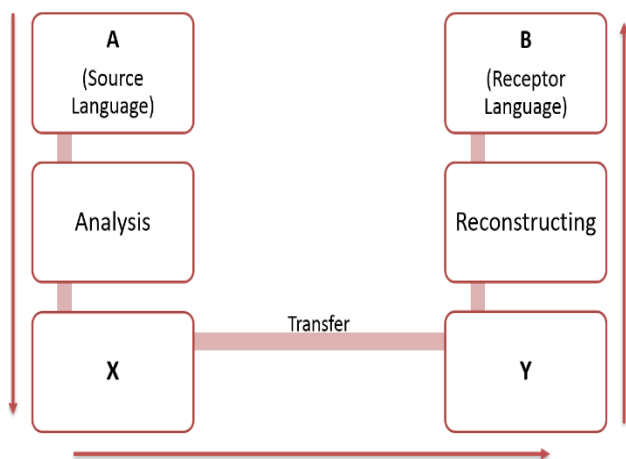
translation process.

Fig.1: Diagram of the Translation Process
(Munday, 2001p.63).

From the above figure, it is clear that source language is analysed the relationship based on grammar and the meaning of words and how they combine with each other. It can also be seen how it transfers the meaning from source language into target language. The structuring phase is the stage where the translator re-arranges or re-writes the text in such a way that the target product still remains meaningful and it can be read by everyone and the rules and styles of the target language are also maintained.

II. TYPES OF TRANSLATION

There are various types of translations. According to Brislin they are as follows:

a. *Pragmatic Translation-* This type of translation deals with message translation keeping in mind that the information is conveyed in its SL (source language) form and it does not get transferred with some other meanings or thoughts.b. *Aesthetic-Poetic Translation-* The second type of translation not only focuses on the conceptual meaning but also the deeper meaning as the beauty of the original writing, and its emotional feeling.

c. *Ethnographic Translation-* This type of translation focuses on the illustration of the cultural context of the languages i.e., the source as well as the target language. Translators must know and understand how the words are used in their original cultural context and must translate them in such a way that the words that have been used for cultural purposes must fit in the same way.

d. *Linguistic Translation-* The last one deals with the equivalence of meaning in the SL and TL (Target language). It deals with morphemes of the SL and grammatical form (yayang345.blogspot.com).

"Jakobson has also mentioned three different kinds of translations and they are as follows:

a. *Intralingual translation, or 'rewording'*:'an interpretation of verbal signs by means of other signs of the same language'.

b. *Interlingual translation, or 'translation proper'*:'an interpretation of verbal signs by means of some other language'.

c. *Intersemiotic translation, or 'transmutation'*:'an interpretation of verbal signs by means of signs of non-verbal sign systems".

"Intralingual translation would occur, for example, when we rephrase an expression or text in the same language to explain or clarify something we might have said or written. Intersemiotic translation would occur if a written text were translated, for example, into music, film or painting. It is interlingual translation which is the traditional, although by no means exclusive focus of translation studies". (Munday, 2001p.9).Thus, it can be said that translation is a process of transferring words/messages from source language (SL) to target language (TL). Translation can also be done within the same language i.e. from one dialect to a different one. However, translation is a very common process when it comes to poetry field. According to poets, translation is a process of transformation, interpretation. It's true that translation can be defined in various ways but every process of translation is involved in some or the other expression of sense. A translation means a text which is different from its original one but it is important to have the sense or message same in both the texts.

Translation can be compared with clothes as the same kind of dress is sometimes worn in a different manner by two different people but for the same occasion.

Translation is not an easy process but it depends on what kind of documents one needs to translate. If a document is translated then it is easy as compared to a text or document it is necessary that the target text should match the word meanings, the grammar rules and also the structures of the sentences. When a text is translated then it gets success if it satisfies the needs of the target audience, in the form of its meaning, its sense and its structure.It is good to have knowledge of both the languages for the translator so that s/he can translate a text/document in an accurate manner. A translator has to focus on each and every part of the original text such as its meaning, sense and its culture as well. And the translator has to focus on her/his own creativity in arranging everything of the source text and re-writing it in the target text. S/he has to show her/his own skills of creativity which will reflect his/her own culture and personality. During the translation of a text, the translator keeps in mind the grammar of both the texts, that is, the lexicon, structure of the sentences. The most important point to keep in mind is the cultural context of the source text so that the translator understands the meaning of the text in an accurate manner and is able to translate the language and its proper contextual and cultural connotations and denotations too. And these rules are followed by a translator so that s/he can keep the meaning of the source text alive. It is important for a translator to translate a given text in the best possible manner. There are few qualities of a good translator which can be acquired by a little bit of hard work. An adequate translator would have the capability of understanding the state of mind of the writer of the original text. A good translator should always keep in mind that the essence or meaning of the original text should not get vanished from the text; the target text should look as similar as the original text or it should be very close to it. A good translator has experience of translating texts, and also has the knowledge about the necessity of knowing the contents/socio-cultural context of a text to make changes accordingly. A loyal translator should have the capability that whatever s/he has translated should look like the mouthpiece of its original form.A good translator should have an organized personality with little flexibility, and should have the knowledge of how to use technology so that s/he can pay attention to each and every point very closely and know what is there to adapt from the original text and what not. And the most important thing a translator should have is the detailed and clear knowledge of both the languages and the subject which s/he is working upon. If a translator is working on the life of any great personality and is translating a biography, s/he should have deeper knowledge about the life of that person. It is very important for a translator to have good language skills and command on vocabulary so that s/he can make good choice of words. The translators who have experience know how to use technical words in a proper way including the use of monolingual dictionaries, bilingual dictionaries and encyclopaedias etc. A translator should make use of these resources as soon as s/he starts writing or translating any text or document so that after a few attempts it will become easy for the writer/translator to use these glossaries /terms in a more effective way and it will also consume less time.

Creativity is the most important thing a translator should have. The translator who is experienced will carry source text into target text with much more creativity and will make it possible that the target audience should enjoy the text the same way as the audience of the source text would have enjoyed. And though translators use different structures from the original text yet they are used in such a way that they look appealing to the target audience.There is always an important point for a translator related to how serious s/he is towards the work s/he is doing. A good translator should have in depth knowledge of the source text so that s/he can do justice with the source text and translate it most effectively in the target text. A translator should keep the habit of reading newspapers, journals, classics, digital material in routine as these may help to have a greater command on the language.Another important part of translation is linguistic competence. A translator should have the knowledge of grammar and linguistic levels such as syntactic, semantic and pragmatic features. A translator has the right to convey an idea in her/his own way that is different from the source text but it's not that s/he can use her/his own idea or personal opinion in the target text. Moreover, the translator should keep in mind the style being used for the target text should sound natural and appealing to the audience of the translated text.For the translators it is also important that they should have knowledge about the latest developments of words and phrases of the source text because every now and then new terminology keeps emerging.To sum up, it can be said that the act of translation is not as easy as it sounds and one should have patience and full dedication towards the work one is doing. A translator should go through the work many times before submitting the final. There should be an assurance in the mind of the translator that the quality which s/he is finalizing is the best one for the target audience (osou.ac.in/eresources).So, the above discussion makes it clear as to what qualities a good translator should have and also the importance of translators in contemporary times.

Recent years have witnessed the development of new type of translators called the machine translators, where the work of the human translators is done by a machine.Machine Translation can be defined as the learning of scheming of the systems which translate one individual tongue into a different one. These systems take results in one natural tongue and translate it into a new human tongue. In this compound method, it is mandatory for the translators to collect a cluster of things into cognizance. As for illustration, an individual is able to evaluate a translation value that how universal and easy it sounds in the target language. Translation should be done in such a way that the reader may not even think that the source work was written in a different tongue and was translated. The original sense of the source must be preserved. The tone of the manuscript, contextual references, slang,

cultural and specific expressions are some of the few elements that are taken into account by the translators for good results. Machine translators first analyse the construction of each and every phrase or term of the source manuscript.

They this arrangement is split down into elements that can be easily translated and is recomposed in terms of a similar structure in the target language. The machine translated texts are poor in quality than an individual's translation and this is for the reason that the formation of every tongue is unlike and it's not taken into account by most of the translators. Word to word translations by some translators, in some cases, provide terrible results. The translation may not sound natural and it may comprise of many grammatical mistakes, lose its original meaning and may not be well structured. Moreover, depending on the context, some words can have different possible meanings in the target language. Due to this fact, human analysis is incomparable. Since the machine translators don't have a global vision, the perfect understanding of machine text becomes inevitable. Most often, the human translators are the native speakers of the target language. This is the reason why, unlike machine translators, human translators are well versed with all the subtleties, expressions, specificities, grammar, slang, cultural connotations and other such aspects of a language which allow them to keep the sense and the quality of the source text and give the finest, most precise and suitable translation.Most human translators may concentrate on a particular field of knowledge in which they have a translation degree and may have a qualified background. The understanding of human translators in the area of assignment expedites the translation process and provides them an edge over machine translators which are unable to provide accurate translation of specific technical terms. The source text is sometimes left as it is if the machines are unable to provide equivalent terms. Thus, for non-crucial projects which do not require perfect translation but only a general understanding, machine translation can be used. On the contrary, for significant assignments, which will be presented before universal audience, human translation is much more reliable and indispensable to provide and to ensure the ideas an individual wants to convey and can be absolutely understood by the audience. Translation of the texts can be processed by several tools that are free as well as paid and people belonging to all kind of fields of study and professions can now access them very easily. As we all know, machine translation is also important in our day-today life, but there is an important issue or we can say a problem that restricts MT systems to obtain correct translation and that is Translation Divergence.Translation Divergence is a complex topic in MT. When syntactically and lexically similar sentences of L1 do not completely translate into the sentences that are similar in syntactical and lexical structures in the target language L2, then language divergence comes into being. The qualities of translations are affected at different levels according to the degree of complexity involved in a particular language pair.

If we talk about the language pair Hindi-English, it is important to know that Hindi and English differ in many aspects and hence translation in this language pair presents a rich source for the study of translation in Machine Translation. These languages also show differences from the point of view of syntactic and lexical perspective that need to be properly examined. The language divergence between Hindi and English can be considered as SOV (subject-object-verb) and SVO (subject-verb-object) classes of languages.

III. METHODOLOGY

In this paper, we examine the different areas of translation divergences from Hindi to English machine translation perspectives. We take Dorr's (1994) classification of translation divergence as the point of departure to examine the topic of divergence in Hindi and English language pair.

Dorr (1994) has identified seven classes of translation divergences (as quoted in Sinha and Thakur [2005:245-264]). These classes are:

a. *Thematic Divergence-* This type of divergence occurs when there are two different sentences of two different languages and the structure of verb of both the languages is different from each other. It occurs when a verbal object of a sentence in one language becomes the subject of the main verb in another language after translation.

Eg. मनीषा कावेरी को पसंद करतीहै|

Kaveri is liked by Manisha.

In the above example it is seen that Kaveri in English sentence is a verbal object but in Hindi Kaveri is the subject of the main verb.

b**.** *Promotional Divergence-* This is the second of the seven divergences and it is also known as Head-Swapping divergence. This type of divergence occurs when there is a difference of syntactic structure. And the syntactic structure in one language is affected in another language. In other words, it can be said that when an adverbial element or adverbial phrase in one language becomes the main verb in the other language then it is called promotional divergence and the opposite case results in Demotional divergence.

Eg. of Demotional Divergence

गाना बजाया गया|

Music was turned on.

In the above sentence it is seen that verbal element 'बजाया गया' becomes adverbial element 'turned on' in English.

d. *Promotional Divergence-* Many studies are available on the issue of divergence between the English-Hindi language pair and they include Jindal [2014]; Sinha&Anil [2005]; Dorr [1994]; Gupta& Chatterjee [n.d]. All these studies have discussed Promotional/Demotional divergence but given examples only of Promotional divergence and not on Demotional divergence. Promotional divergence in English-Hindi language pair becomes Demotional in the case of Hindi-English language pair. The examples of Promotional divergence in the Hindi-English language pair thus have become difficult to find.

d. *Structural Divergence-* When it comes to structural divergence PP (Prepositional Phrase) in one language is realized by a NP (Noun Phrase) in another language.

Eg. वह एक सुहानी सी सुबह थी।

It was a pleasant morning.

In the above sentence it is seen that in the Hindi sentence 'सुहानी सी सुबह' is a PP and when it is translated to English it becomes NP in English 'a pleasant morning'.

e. *Conflational Divergence and Inflational Divergence-* These occur when there are two or more than two words in one language which are realized by one word in the other language and the opposite case is referred to as inflational divergence.

Eg. हड़बड़ा कर वो उठ बैठे।

Perplexed, he sat up.

In the above sentence it is seen that हड़बड़ा कर are the two words in Hindi but when it gets translated into English it becomes one word that is 'perplexed'.

f. *Categorial Divergence-* This divergence is about the category. It occurs when the category of a word gets changed. In other words, it can be explained as when a predicate is verbal element in one language but adjectival in the other language.

Eg. राजा को ईर्ष्या होने लगी।

The king started feeling jealous.

In the above sentence 'ईर्ष्या होने लगी' is verb but when it gets translated into English it becomes adjective that is 'feeling jealous'.

g. *Lexical Divergence-* This is the last category of divergence and it occurs when there is a lack of exact words to translate an idea/concept or a phrase from one language to the other.

Eg. रोहित की माँ रो पड़ी।

Rohit's mother broke into weeping.

So, here in the above sentence it is seen that the Hindi 'रो पड़ी' gets translated to "broke into" which means 'ज़बरदस्ती घुस आना' but for translation the word "broke into" has been used.

IV. CONCLUSION

Thus, we have discussed the seven major divergences and based on the above discussions it can be said that translation or machine translation from Hindi to English is not an easy task and it involves much more complexity and more so when it comes to translations which are literary in nature, i.e. involve the use of literary language. It is very difficult to obtain correct translation when it comes to the language pair Hindi-English translation. A translator needs to examine each and every detail of the entirely different grammatical structure, along with its socio-cultural context in both the languages, that is, Hindi and English.

REFERENCES

1. Dorr, Bonnie J. "Machine translation divergences: A formal description and proposed solution." *Computational Linguistics* 20, no. 4 (1994): 598-610.
2. Gupta, Deepa, and Niladri Chatterjee. "Study of divergence for example-based English-Hindi machine translation." *STRANS-2001, IIT Kanpur* (2001): 43-51.
3. Jindal, Shishpal "Divergence in English –Punjabi machine translation". International Journal of Soft Computing and Artificial Intelligence 2, no. 2 (2014): 109-110.
4. Munday, Jeremy. *Introducing translation studies: Theories and applications*. Routledge, 2013.
5. Sinha, R. Mahesh K., and Anil Thakur. "Translation Divergence in English-Hindi MT." In *the Proceeding of EAMT Xth Annual Conference, Budapest, Hungary*, pp. 30-31. 2005.
6. Syal, Pushpinder, and Dharam Vir Jindal. *An introduction to linguistics: Language, grammar and semantics*. PHI Learning Pvt. Ltd., 2007.

Web sources

1. http://shodhganga.inflibnet.ac.in>bitstream (Retrieved 15July 2018)
2. http://osou.ac.in/eresources/clt-01-unit-01-pdf (Retrieved 31 July 2018)
3. http://yayang345.blogspot.com/2014/12/types-of-translation-according-to.html (Retrieved 11 Aug 2018)

Author's Profile

Venu Shree is a Research Scholar in the department of English and Modern European Languages, Banasthali Vidyapith. She did her masters in English literature and M.Phil (ELT) from the department of English and Modern European Languages, Banasthali Vidyapith. Her research interests are Computational Linguistics, Machine Translation, Translation, and Linguistics.

Geetha Yadav is currently working as an Associate Professor at Banasthali Vidyapith in the Department of English and Modern European Languages. She has an experience of 20 years of teaching at UG and PG level, including 5 years of M.Phil. (ELT) and guiding research scholar for Ph.D. She has presented several papers at various national and international seminars and conferences. She has also delivered invited lectures at various institutions and has been a member of the Board of Studies of some institutions. She has published three papers in various anthologies. Her field of ongoing research is narratology.

ItiMathur is an Assistant Professor at Banasthali Vidyapith. Her areas of interest are computational semantics and ontological engineering. Besides this, she is also involved in the development of MT engines for English to Indian languages. She is one of the experts empanelled with TDIL program, Department of Electronics and Information Technology (Deity), Govt. of India, a premier organization that oversees Language Technology Funding and Research in India. She has several publications in various journals and conferences and also serves on the program committees and editorial boards of several conferences and journals.

Nisheeth Joshi is an Associate Professor in Department of Computer Science, Faculty Mathematics and Computing. He has done his PhD in the area of Natural Language Processing. Being involved in teaching and research for over 10 years, he has developed the art of explaining even the most complicated topics in a straight forward and easily understandable fashion. He also has vast experience of handling large scale research projects sponsored by various funding agencies. This has helped him in developing practical insights of complex AI systems. He has delivered invited talks in several conferences, workshops and training programs. He has authored several papers in international journals and conferences on various topics in AI and NLP.