

# Linguistic Divergences in English to Bengali Translation

Niladri Sekhar Dash<sup>1</sup>

<sup>1</sup> Linguistic Research Unit, Indian Statistical Institute, Kolkata, India

Correspondence: Niladri Sekhar Dash, Linguistic Research Unit, Indian Statistical Institute, Kolkata, India.  
Tel: 91-033-2575-3281. E-mail: ns\_dash@yahoo.com

Received: November 8, 2012 Accepted: November 19, 2012 Online Published: January 25, 2013

doi:10.5539/ijel.v3n1p31

URL: <http://dx.doi.org/10.5539/ijel.v3n1p31>

## Abstract

Translation is an endeavour to automate all or part of the process of translation from one language to another. In general, translation is a complex task, which aims at preserving the semantic and stylistic equivalents of the source language texts into the target language texts. The most problematic area in manual and machine translation is the lexicon and the role it plays according to the context to create deviations. There are also cases where deviations occur owing to the divergence – one of the complex areas of investigation in translation. Divergence in translation normally arises when the sentences in the source language are realized differently in the target language. This paper seeks to discuss some of the major divergences that observed in English to Bengali translation. It also investigates how the different linguistic and extralinguistic constraints can play decisive roles in translation, resulting in divergences and other issues. The primary objective of this paper is to understand the types of divergence problems that operate behind English to Bengali translation the study of which is still in a state of its infancy. Proper identification and understanding of these problems are important in both manual and machine translation. Moreover, resolution of these problems is a pre-requisite for designing a robust machine translation system between the languages considered for the present study.

**Keywords:** translation, divergence, English, Bengali, concepts, lexical difference

## 1. Introduction

Divergence occurs when the sentences of the source language are not translated properly into similar forms in the target language. This is a common problem in translation between any two natural languages. In case of Machine Translation (MT), it is a complex phenomenon, and therefore, it is important to identify and categorize the divergences and solve these so that a system can generate accurate translation outputs in the target language.

Translational divergences occur at different levels and affect the overall quality of a translation. The present paper aims at providing a synoptic account on translational divergences as noted in Dorr (1994) with close reference to English-to-Bengali translation. It presents a brief account on the patterns of divergences that may arise in translation and provides some solutions to overcome the barriers.

The reason behind divergence mostly lies in incompatibility of one language with the other. This is a common phenomenon for all natural languages. There are situations when sentences in the source language need to be translated in entirely different forms in the target language. This makes the task of both manual and machine translation a difficult enterprise. A comprehensive study on divergence patterns, therefore, becomes necessary for designing an effective system of translation.

It has been argued that certain types of divergence are universal in the sense that they exist in all languages whereas some other types of divergence are specific to a particular pair of languages (Dorr 1994). It will be a useful contribution if one investigates the patterns and types of divergence across languages and language types as proper identification of the patterns and types will invariably enhance the performance level of a system as well as will increase the rate of accuracy in translation. Based on this hypothesis, in this paper we have tried to identify types of divergence noted in English-to-Bengali translation and locate the key elements that cause divergence in translation between the languages.

## 2. Early Works in This Area

English and Bengali, though considered to have evolved from the same source called 'Indo-European', have

developed in different directions and manners. Owing to large and long spatio-temporal differences across centuries, they now differ in many respects with regard to their surface forms, content, and composition. At present, they present a rich pool of data and information for the study of translational divergence registered in their forms and expressions. The present paper intends to harvest on this particular area of investigation keeping in mind their functional relevance in both manual and machine translation.

Till date considerable amount of research have been carried out on the nature and patterns of translational divergence between English and Hindi (Gupta and Chatterjee 2003, Goyal, Gupta and Chatterjee 2004, Sinha and Thakur 2004, Sinha and Thakur 2005a, Sinha and Thakur 2005b). Based on this studies efforts are made for developing automatic translation systems between this language pair: English and Hindi (Dave, Parikh and Bhattacharayya 2002). Perhaps no such studies have been made yet to understand the nature and patterns of divergence between English and Bengali although we observed that the amount of manual translation from English to Bengali is almost equal to that of English to Hindi. Therefore, keeping this requirement in sight we have tried to explore and understand the areas of divergence in translation between the languages.

In the following sections, we have tried to formalize the divergence types and patterns; observed how these divergences arise from different lexico-syntactic factors; and how we can address these issues keeping in mind various linguistic and extralinguistic factors embedded within respective languages. The functional relevance of the present study may be attested in manual and machine translation as well as in the acts of designing courses for teaching translation methods to the students and the professionals (Arnold *et al.* 1994). However, before that we need to understand what does translation mean and what are its basic goals of in the present world of information explosion and knowledge sharing.

### 3. What is Translation?

By simple definition, translation is an act of interpretation of meaning of a text in one language and subsequent production of an equivalent or nearly equivalent text in another language that is able to yield the same message found in the source language. Etymologically *translation* means ‘bring across’ or ‘carry across’. The term is derived from Latin *translatum* meaning ‘transfer’. The form *transfer* is again derived from *trans* ‘across’ and *ferre* ‘to carry/bring’. The Greek word *translation* is *metaphrasis* meaning ‘speaking across’. This has generated a new English word *metaphrase* meaning ‘literal translation’ or ‘word-for-word translation’. It has also produced another term in English *paraphrase* (< Gk. *paraphrasis*) ‘saying in other words’. While *metaphrase* indicates formal equivalence, *paraphrase* implies dynamic equivalence between the language pairs (Kasperek 1983: 83).

In practice, translation is a technique in which a translator takes input in the form of full sentences from the source language at a time and generates corresponding full sentences in the target language. The art of translation is as old as written literature. In almost all major languages, translation is treated as one of the most common methods of borrowing ideas and information from the neighbouring or distant languages. Therefore, it can be argued that translation from one language to another is one of oldest domains of applied linguistics in which people have been involved in the act of transferring information, ideas, and knowledge from one language to another (verbally or in written mode) to address their needs of various domains of knowledge. Translation has evolved over the centuries as an important profession as well as a discipline of theoretical and practical investigation.

### 4. Goal of Translation

In the present era of information exchange translation is treated as one of the key strategies that have can lasting impact on global information sharing, multilingual information interchange, cross-cultural knowledge sharing, and multilingual communication. A translator becomes an actor in the platform of cross-disciplinary knowledge exchange that registers huge functional relevance in global commerce and social progress.

In the context of knowledge-based society the question that stands tall is: why do we need to translate texts from one language to another when the possibility of sharing data and information through one or two major languages becoming the norm of the societies? Keeping aside the implication of linguistic imperialism hidden in this argument, we can identify the following reasons that motivate us to translate texts from one language to another:

- (a) The world of humanities, science and technology, and commerce and medicine requires accurate translation of information from one language to another.
- (b) Translation becomes necessary in those situations where there is real paucity in useful knowledge texts in languages.
- (c) Translation helps to promote international co-operation through sharing of information meant for global access.

- (d) Translation removes language barriers through immediate transfer of scientific, medical, agricultural, and technical information to the deserving people.
- (e) Translation of news and information is useful to address the rapid growth of internet, globalization, and localization.
- (f) In a multilingual country like India language barriers are dissolved through translation that enhances nationwide communication.
- (g) In World Wide Web there are more than million sites containing materials of different kinds in different languages. These materials are waiting to be translated for global access and utilization.
- (h) Demands for linguistic singularity, expansion of multilingualism, growth of literacy and readership, globalization of information, etc. also justify the needs for translation.
- (i) Translation is required for pure theoretical and applied research and for the purpose of commercial reasons.

Translation methods, during last thousand years, have undergone many modifications in goals and strategies. In each situation the prime objective has been to generate texts in the target language which may be accepted as maximally accurate and effective. The huge amount of text already translated as well as the mammoth amount of text still waiting to be translated urge for looking into the existing methods of translation as well as devising more useful and accurate methods of translation that can rightly account for the new types of text as well as new linguistic features reflected in the languages.

### 5. New Turn in the Journey

The availability of large digital as well as processed language corpora offers promising resources for developing new methods and processes for translation. Modern computers and text processing techniques may be used to access these corpora to elicit new properties from languages as well as for designing new methods of translation based on content or type of texts. This new approach becomes necessary, as it is understood that the methods and approaches used so far in the traditional approach of translation have lost much of their functional significance in modern text translation. Moreover, it is understood that methods and strategies of translation of scientific and technical text need to be different from methods and strategies used in translation of imaginative and literary texts. This argument also stands justified for translation of media texts, medical texts, commercial texts, and legal texts, etc.

It is already understood that translation is not a simple form-to-form (e.g., *word to word*) or frame-to-frame (i.e., *phrase to phrase* or *sentence to sentence*) replacement of texts (or part of it) from one language to another. By its default definition, it inherently demands for a kind of assurance that the texts of the source language are produced in the target language in such a manner that the translations are accepted as grammatically correct, linguistically valid, cognitively coherent, and conceptually acceptable. Moreover, the outputs in the target language should be nearest to the source language – in both sense and content – if not identical. Information embedded within the source language should not be lost in the target language and, reversely, no extra information, which was not originally present in the source language, should be tagged in the target language.

Obviously, there are many obstacles in the path of achieving high rate of accuracy in translation. These obstacles come not only from the areas of morphology, lexicology, syntax, and semantics, but also from the world of pragmatics, discourse, culture, history and geography. That means the task of translation asks for a kind of cognitive approximation between the languages where the goal is to find out ways and means to get as close as possible between the languages used in translation in as many ways as feasible.

The goal may be achieved through a long process of trial and error, since it requires regular evaluation of translation outputs, identification of translational errors, careful correction of errors, rigorous training of translators, modification of previous experiences, and augmentation of relevant linguistic knowledgebase. No doubt, the path is full of thorns. Yet it is tantalizing because it throws up challenges before us who want to achieve success in this area.

In recent years, it is observed that thousands of literary and non-literary texts are being translated in great enthusiasm across languages. To address diverse needs of translation amateur people in collaboration with professionals have taken part in this profitable and commercially viable enterprise. Also computer is introduced in the arena of online translation – one of the latest trends, which have been largely successful due to advancements made in computer technology, corpus linguistics, translation studies, language technology, and translation technology. This success, although partial, is achieved through active involvement of large number of

people from various fields of human knowledge for a long period of time with a determination for success through several stages of trial and errors.

## 6. Advantages of Translation

There are several advantages of translation. The direct advantage is that through this process knowledge, information and ideas of one language are transmitted to the people of another language. On the other hand, the indirect advantage is that through translation people of one language get scope to learn about the life, living, beliefs, history, geography, and culture of people of another language. Although such advantages are mostly noted in case of translation of literary and cultural texts related to history, culture, tradition, and other socio-cultural aspects, translation of non-literary and scientific texts generates large amount information and knowledge that contribute for the growth of collective wisdom of the target language community. In the present context of globalization of information and knowledge sharing, translation – both manual and automatic – occupies an important position in information exchange and knowledge building.

Since translation of literary and non-literary texts are involved in knowledge transfer, the issues related to translation and interpretation of texts as well as problems of translation of information and ideas occupy the centre of our attention. We, therefore, need to make serious attempts to understand types and techniques used so far in the work of translation and focus our attention on the goals of representing the content of a text of one language into another by different processes of translation that can serve purposes in the most suitable manners. The entire enterprise stands on successful realization of the basic truth that getting knowledge and information transferred accurately from one language to another is the main priority of translation and people across languages of the world are toiling hard to achieve this goal.

In essence, an authentic translation must take into account several constraints including the followings:

- (a) Situations in which a text is developed in the source language.
- (b) Context of use of words and phrases in the source text.
- (c) Rules of grammar and sentence in source and target language.
- (d) Writing conventions used in source and target language.
- (e) Information expressed in multiword units used in source text.
- (f) Intralinguistic knowledge inherent in the source text.
- (g) Extralinguistic information embedded in the source text.
- (h) Goal and the motive of the source text composer.
- (i) Discoursal and pragmatic information in source text.
- (j) Sociocultural knowledge and information of the source and target language.

While translation of scientific text often suffers from ‘information deficiency syndrome’ (in which the load of information of the source text is not adequately transferred to the target text), the translation of literary text often suffers from ‘spilling over syndrome’ (in which the idioms and fixed expressions used in the source text often proliferate into the target text). This generates a new set of lexical stock and other linguistic resources in the target language and these are made capable of expressing new ideas, concepts, and information obtained from the source language. For instance, a translation of English cultural text into Bengali generates a new text in Bengali as well as a set of new words, terms, idioms and phrases, which may or may not be conceptually identical to the objects and concepts of the source language.

Various theories and models are recommended over the centuries for providing solutions to the problems of translation. Similarly, many new techniques are proposed and implemented to deal with different linguistic problems and issues often faced in translation. As most of the problems are not yet solved for the benefit of the people engaged in translation English to Bengali, there is a rationale to address these issues with new set of examples and observations that may contribute in better comprehension of the problems and developing suitable strategies for overcoming these hurdles. In the following section, we refer to some of these problems, which are related to the patterns of divergence as defined in Dorr (1994).

## 7. Divergence in Translation

Translation is a highly complex task. It aims at preserving semantic and stylistic equivalents of the source text into the target text. The most problematic area in translation is the lexicon and the role it plays in the act of creating deviations in sense and reference based on the context of its occurrence in texts. Similar observation

stands valid for compound words, idioms and phrases – which also show semantic deviations in accordance with their usages in different kinds of text.

There are also cases where deviations occur owing to divergence – one of the complex areas of investigation in translation. Divergence in translation arises when the sentences in source language are realized in a different manner in the target language. It is necessary to understand how different linguistic and extralinguistic constraints play decisive roles in translation resulting in divergences and other issues. Proper identification and understanding these problems are important in both manual and machine translation. Moreover, resolution of such problems is a pre-requisite for generating good translation in target language.

Divergences occur at different levels and severely affect the quality of a translation. Dorr (1994) proposes ways to look into this aspect of translation in minute details between any two languages involved in translation. Based on this we focus on various divergences observed in English-to-Bengali translation as well as explore other issues related to problems that arise in translation between languages.

The underlying principles for divergences are attributed to those cases where the basic concept of a sentence in the source language is expressed and distributed over different words or a series of different words in the target language (Dash 2004). As a language-dependent phenomenon, divergence may affect quality of a translation and reduce authenticity of a translated text. Although all types of translational divergence are not observed in all language pairs used in translation, certainly there are some translation problems, which may be considered universal and true to all language pairs.

Based on large number of empirical examples found in English, Spanish, and German corpora, Dorr (1993) has classified divergences into two broad types and defined them accordingly:

- (a) Syntactic Divergence, and
- (b) Lexical-semantic Divergence.

#### 7.1 Syntactic Divergence

Syntactic divergence relates to the sentential properties associated with each language and these are usually independent of the actual lexical items used in the sentence. Dorr (1993) records the following types of syntactic divergence that may arise from structural and syntactic aspects of source and target language texts.

- (a) **Constituent order divergence:** This relates to the feature of word order distinctions between the languages used in translation.
- (b) **Adjunction divergence:** This relates to different patterns of use of adjunct structures in the language pairs used in translation.
- (c) **Preposition-stranding divergence:** It relates to proper selection of governors in formation of sentences in the languages, and
- (d) **Pleonastic divergence:** It refers to ‘null-subject’ use in the source or target language.

Besides these major syntactic divergences, there are some minor divergence types, such as, *movement divergence*, *dative divergence*, etc., which are not so frequent in usage in language (Dorr, Ayan and Habash 2004).

#### 7.2 Lexical-Semantic Divergence

Lexical-semantic divergence refers to those cases that are determined lexically. These are characterized by certain linguistic properties, which may be attributed to the lexicon itself (Gupta and Chatterjee 2001). There are seven sub-types of lexical-semantic divergence that includes *conflational divergence*, *structural divergence*, *categorical divergence*, *promotional*, *demotional divergence*, *thematic divergence*, and *lexical divergence* (Dorr 1993). Later, they are reorganized as six major types, such as, *light-verb construction*, *manner conflation*, *head-swapping*, *thematic divergence*, *categorical divergence*, and *structural divergence* (Dorr 2002). These are discussed in brief in the following sub-sections with reference to English to Bengali translation.

##### 7.2.1 Conflational Divergence

Conflational divergence is a process of incorporation of components that are necessary to express the appropriate sense of the verb of the source language into the target language. In reality, it is the translation of a single word unit of the source language into two or more words in the target language, which may include combination of a ‘light-verb’ and another meaning-bearing element in the target language. The source of this kind of divergence lies in the variation in selection of verbs in the source language and in the target language. Consider, for instance, the English verb *stab*, which is often cited as the most elegant example of this kind.

(1a) English: He stabbed John.

(1b) Bengali: se Janke churi diye mār̄la.

(1c) Bengali: Se Janke churi mār̄la.

In the above example (1a), English verb *stab* is a single word unit, which carries the entire meaning of an action clubbed with information of tense and aspect. This single word unit is translated into Bengali as a combination of a string of words as *churi diye mār̄la* (1b), since Bengali does not have a single-word unit equivalent to the English word *stab*. Therefore, when the English verb incorporates an instrumental adjunct within it, Bengali uses an explicit instrumental case marker (i.e., *diye*) to produce an equivalent translation for the English sentence.

However, the English word *stab* can also be translated into Bengali as *churi mār̄la* (1c) where the instrumental case marker *diye* is removed from the word pair *churi mār̄la*. This is possible because deletion of the instrumental case marker *diye* does not hamper the role of *churi*, which acts as an instrument in the act of killing. The manner of divergence in the Bengali translation can be expressed through the following diagram (Fig. 1).

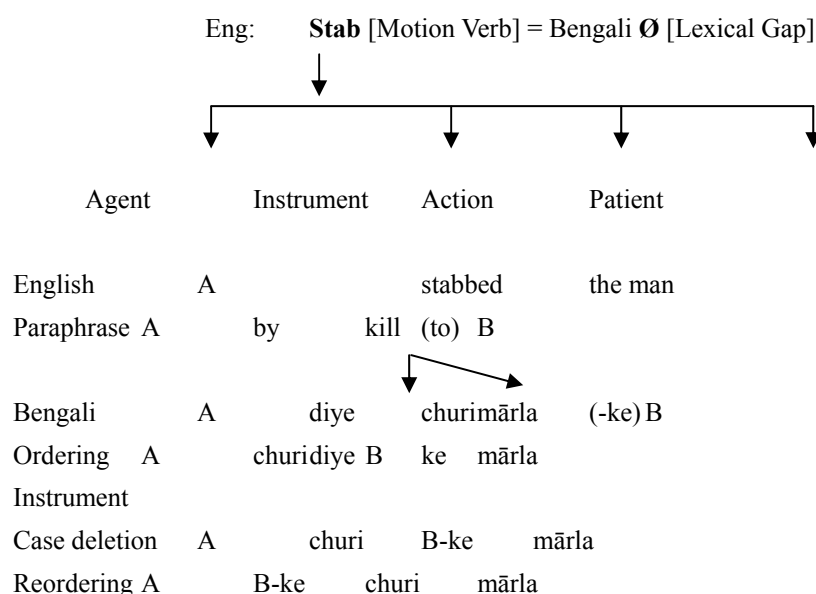


Figure 1. Conflational divergence from English to Bengali along with case marker deletion in translation

### 7.2.2 Structural Divergence

Structural divergence is the realization of the verbal object as a Noun Phrase (NP) in the source language and as Prepositional Phrase (PP) in the target language and vice-versa. Here the incorporated arguments like *subject* and *object* of the source language are often realized as the PP in the target language, and such mismatches in translation arise due to introduction or elimination of function words (i.e., prepositions). For instance, consider the following example:

(2a) English: John entered the house.

(2b) Bengali: Jan gharer madhye prabeś karla.

(2c) Bengali: Jan ghare dhukla.

The English verbal object *house* (2a) is realized as a noun phrase, which is, in actuality, a part of the PP (*into the house*). However, due to absence of the preposition (i.e., *into*), the phrase *the house* is realized as a NP, which is, however, rendered into Bengali as a postpositional phrase (i.e., *gharer madhye*).

In Bengali, most of the arguments and adjunct NPs of the verb are realized with an overt postposition, whereas in English these are often represented as simple NPs. Thus, a Bengali output sentence structurally deviates from the English sentence in translation in a rule-based manner: a verbal object is a NP in English, but a postpositional phrase in Bengali.

Structural divergence is also observed when an NP argument in the source language is realized as a PP adjunct/oblique NP in the target language. For example, the English sentence *Pearl attended the class* may be translated in Bengali as *pārl klāše upasthit chila*. In the English sentence the noun phrase *the class* is a verbal object, which has become a PP adjunct in Bengali translation (e.g., *klāše*).

### 7.2.3 Categorical Divergence

Categorical divergence is characterized by the use of different parts-of-speech in the source language as well as in the target language. It arises when part-of-speech of a word in the source language changes during the process of translation to generate an entirely different lexical category of words in the target language. In such a case, the mismatch in translation normally occurs due to the change in the parts-of-speech of words of the concerned languages.

(3a) English: I am hungry.

(3b) Bengali: āmār khide peyeche.

(3c) Bengali: āmi k ṣudhārta.

In the English sentence (3a) the predicate *hungry* is an adjective, which is translated into Bengali as a noun (i.e., *khide*) as shown above (3b). Although this type of divergence is quite common in translation from English to Bengali, the interesting thing is that such divergence is possible to avoid if the English sentence is translated into Bengali as *āmi k ṣudhārta*, where a new feature does arise. The English adjective remains an adjective in Bengali, but the copula (i.e., *am*) of the English sentence becomes null in Bengali.

### 7.2.4 Promotional Divergence

Promotional divergence refers to the promotion of a logical modifier of the source language to the head position in translation in the target language. Due to this reason, it is called **Head Swapping Divergence** in which a modifier is realized as an adverbial phrase in the source language but as a main verb in the target language. It implies that an adverbial element in the source language is exposed by a verbal element in the target language. Consider the following example:

(4a) English: Mary usually comes here.

(4b) Bengali: prāyi Meri ekhāne āse.

(4c) Bengali: Meri prāyi ekhāne āse.

(4d) Bengali: ekhāne Meri prāyi āse.

In English sentence (4a), the modifier *usually* has been used as an adverbial phrase, which has been shifted to the head position in Bengali translation (4b). Alternatively, we can have also a Bengali translation like *Meri prāyi ekhāne āse* (4c) and *ekhāne Meri prāyi āse* (4d) where the position of the modifier can be kept unchanged without affecting the meaning of the sentence. Therefore, in these cases, at least, there is no promotional divergence between the two languages.

On the other hand, for an English sentence like *The discussion is on*, the possible Bengali translation will be *ālocanā calche*. In this example, the English modifier *on* is an adverb, but has been realized as a main verb (i.e., *calche*) in the Bengali translation.

### 7.2.5 Demotional Divergence

Here the reverse process takes place when the main verb (i.e., *likes*) of an English sentence (5a) is realized as a noun in Bengali translation (5b) as the following example reveals. This kind of divergence is normally characterized by demotion of a head verb to a modifier position as the following examples show.

(5a) English: He likes to eat sweets.

(5b) Bengali: se miṣṭi khete pachanda kare.

(5c) Bengali: se miṣṭi khete bhālobāse.

While the form *like* is used as a main verb in English, it is used as a noun (i.e., *pachanda*) in Bengali translation (5b). However, the absence of a primary verb in the Bengali sentence is made up with addition of a helping verb *karā* 'to do' with the noun and this verb carries appropriate tense, person, and number information to become maximally equivalent in translation to the sentence of the source language.

Another interesting example demotional divergence in translation from English to Bengali can be revealed in an English sentence like *It suffices*, which is translated in Bengali as *eṭāi yatheṣṭa*. This shows that while in English

sentence the word *suffices* is used as a main verb, it is used as an adjectival modifier (e.g., *yātheṣṭa*) in Bengali translation.

#### 7.2.6 Lexical Divergence

This type of divergence arises due to absence of an exact translation map between the source and the target language. For instance, a word in an English sentence given below (6a) deviates from its literal translation of a word in the Bengali sentence (6b). Such lexical mismatches are attributed to lexeme-phrase substitution or lexical fusion.

(6a) English: Dacoits broke into the house.

(6b) Bengali: ḍākāterā jor kare ghare ḍhukla.

(6c) Bengali: ḍākāterā darjā bheṅge ghare ḍhukla.

Since there is no appropriate parallel phrasal verb in Bengali for its English counterpart *break into*, Bengali translation uses a verb form (i.e., *ḍhokā* ‘enter’) along with an overt phrasal element *jor kare* ‘by force’ to generate the required sense. This has been a regular method of translation because Bengali does not have suitable phrasal verbs that can stand as good translational equivalents for some English phrasal verbs. This type of divergence may overlap with *conflational* and *inflational* divergences mentioned above.

In case of lexical divergence what is lexically realized as main verb in the source language may be realized as a different verb in the target language (Saboor 2010). For example, the English sentence *The lady ran into the room* may be translated in Bengali as *Mahilāṭī dauṛe ese ghare ḍhuke paṭlen*. While *runs* occurs as a main verb in the English sentence, in Bengali translation a compound verb (i.e., *ḍhuke paṭlen*) is used for expressing the sense. One could have used the term *ḍhukēn* (a single verb form) in the Bengali translation, but in that case, the actual connotative sense of *run* in the source language could not have been properly expressed in the Bengali translation.

On the other hand, if slight change is made in the structure of the English sentence, as in, *The lady ran into her old friend*, the appropriate Bengali translation may be *purāṇo bondhur sāthe mahilāṭī haṭhāt dekhā haye gela*. In the English sentence *run* also occurs here as a main verb but with as a member of a group verb with an idiomatic sense. Therefore, in Bengali translation one has to use a phrase like *dekhā hala* or *dekhā haye gela* for capturing the suddenness of an unexpected encounter of the agent of the sentence.

In both the cases the most striking thing is that in Bengali translation the English phrasal verb *runs into* has been realized by two different verb forms in Bengali (i.e., *ḍhuke paṭla* or *dekhā hala*) for expressing the meaning of the English sentences and in both cases different adverbial forms (i.e., *dauṛe ese* and *haṭhāt*) have been used as supporting sense-expressing elements. This implies that divergence not only relates to differences in lexical mapping but also differences in structural mapping between the two languages.

#### 7.2.7 Thematic Divergence

Thematic divergence occurs when the argument of a verb exchanges its thematic role from one language to another (Habash 2002). Therefore, thematic divergence is characterized by a kind of divergence that arises from differences in realization of argument structure of a verb in the source and the target language (Jackendoff 1983, Lytinen and Schank 1982, Schank and Abelson 1977). Such mismatch in thematic roles arises in translation due to the event of **syntactic actant permutation**, as the following examples show.

(7a) English: I like Liza.

(7b) Bengali: āmi lijāke pachanda kari.

(7c) Bengali: āmār lijāke bhālo lāge.

In the English sentence (7a), the theme of the sentence is the object (i.e., *Liza*) of the verb (i.e., *like*). In Bengali one can have two alternative translations. In the first case (7a), one can have almost equal translation, where there is no change in the role of the object of the verb (i.e., *pachanda kari*), although word-order of the sentence is changed from SOV to SVO. In the second case (7b), however, the subject *āmi* changes into its possessive form (i.e., *āmār*) and a new compound verb (i.e., *bhālo lāge*) is introduced to capture the basic sense of the English sentence, while the object remains unchanged in its role in the sentence. Thematically, however, there is notable change in the translation. While in the first Bengali translation (7a) *āmi* is used in direct case as an agent of the action; in the second sentence (7b), *āmār* is used in oblique case as an experiencer. Thus, the object of the verb in English becomes the subject of the main verb in Bengali, while the subject of the English sentence becomes just an experiencer in the Bengali translation.

## 8. Beyond Divergence

The discussions presented above show that there are at least seven major sub-types of lexical-semantic divergence, which are quite normal and natural with regard to translation from English to Bengali. These divergences, therefore, need to be attested and addressed carefully at the time of translating English texts in Bengali. In manual translation, a translator can adopt a suitable approach or strategy based on the content, type, and context of texts of the source language to accomplish the goal. On the other hand, in case of machine translation, a rule-based approach (vis-à-vis) algorithm may be required as an alternative to deal with the sentences containing such divergences (Dash 2009).

After looking at the types of divergence cited above, we argue that while existing types explain certain divergences relating to English-to-Bengali translation, there are other translational divergences, which also seek close investigation and analysis for developing well-investigated rules and strategies for fruitful translation between the languages. We have classified these divergences into three broad types:

- (a) Syntactic problems: These relate to structure of sentences,
- (b) Lexical problems: These relate to usage and choice of words, and
- (c) Semantic-cum-sociocultural problems: These relate to meaning of words and unique cultural aspects of respective languages.

Minute analysis of examples show that most of the problems are related to word order, phrase order, use of particle, determiner, temporal expression, replicative words, lexical choice, echo words, reduplication, honorification, pleonastic usage, and pleonastic marker, etc. Due to limitation of space, these are not discussed here but are elaborated elsewhere in details with reference to examples obtained from English and Bengali language corpora (Dash 2012).

Although there are many explicit cases of divergence from English to Bengali translation, the issues that are addressed in this paper, cannot be ignored in the works of manual and machine translation. Arguably, the different types of problems presented here cannot be explained and accounted for just by referring to the existing types of divergence and convergence discussed at length in literatures of translation. We need to find out different methods and solutions to solve these problems.

## 9. Conclusion

Some important types of divergence related to translation from English to Bengali are discussed in this paper keeping in mind both manual and machine translation from English to Bengali. Following the works of Dorr (1993, 1994) and other works broadly related to this area (Dave *et al.* 2001, Gupta *et al.* 2003, Sinha and Thakur 2005), we have tried to explain how various types of divergence will crop up when one tries to translate English sentences into Bengali.

The paper mainly talks about the problems and limitations of translation from the structural point of view. It explains different patterns of occurrences across the languages considered for this phenomenon. There are also problems relating to cultural transmission, which are manifested in case of selection of the optimal construction among the possible and equally competing choices. Several early studies have referred to these problems as well as suggested various explanations relating to the structural limitations or constraints of the source and target languages – as the case may be in different contextual frames (Lyttinen and Schank 1982, Schank and Abelson 1977). This paper, due to its specific goal, has not tried to address these issues and problems.

The examples presented here are simple sentences, which are, in essence, comparatively less complex than the complex sentences; but analysis of which are essential for gathering insights and guidelines for dealing with complex sentences. This paper is just a tip of the iceberg, which we want to measure with a distant goal for developing automatic translation strategies from English to Bengali. This study may also help us in our final mission for developing systems for automatic or human aided machine translation system between the languages.

This paper hopes to open up new avenues for future research in the area of translation divergence between English and Bengali not yet been studied in close details. We can also think of initiating similar studies between Hindi and Bengali to recognize as well as understand what kinds of problem, divergence, and convergence exist between the languages. Once we succeed to address these problems, then developing a translation system from Hindi to Bengali (or vice versa) cannot be a distant dream. What is clearly realized is that translation is a complex process where identification of patterns, classification of divergence and convergence, and finding out appropriate solutions are much-needed requirements for facilitating the process of translation – both manual and automatic.

## Acknowledgements

This is a revised version of the paper presented in the *Translators' Orientation Programme for Knowledge Text Translation* at the *Dept. of Linguistics, University of Calcutta, Kolkata, India*, 8-10<sup>th</sup> February 2012 conducted by the *National Translation Mission, Central Institute of Indian Languages, Ministry of Human Resource Development, Govt. of India*.

## References

- Arnold, D. J., Balkan, L., Meijer, S., Humphreys, R. L., & Sadler, L. (1994). *Machine Translation: An Introductory Guide*. London: NCC Blackwell.
- Dash, N. S. (2004). Issues involved in the development of a corpus-based machine translation system. *International Journal of Translation*, 16(2), 57-79.
- Dash, N. S. (2005). *Corpus Linguistics and Language Technology: With Reference to Indian Languages*. New Delhi: Mittal Publications.
- Dash, N. S. (2007). *Language Corpora and Applied Linguistics*. Kolkata: Sahitya Samsad.
- Dash, N. S. (2008). Context and contextual word meaning. *SKASE Journal of Theoretical Linguistics*, 5(2), 21-31.
- Dash, N. S. (2009). Linguistic tasks on translation corpora for developing resources for machine translation. Paper presented at the 31<sup>st</sup> All India Conference of Linguists (AICL-2009). Central University of Hyderabad. India.
- Dash, N. S. (2012). *Applied Linguistics* (In Press).
- Dave, S., Parikh, J., & Bhattacharayya, P. (2001). Interlingua based English-Hindi machine translation. *Journal of Machine Translation*, 16(4), 251-304. <http://dx.doi.org/10.1023/A:1021902704523>
- Dorr, B. J. (1992). The use of lexical semantics in interlingual machine translation. *Machine Translation*, 7(1), 135-193.
- Dorr, B. J. (1993). *Machine Translation: A View from the Lexicon*. Cambridge: MIT Press.
- Dorr, B. J. (1994). Machine translation divergences: A formal description and proposed solution. *Computational Linguistics*, 20(4), 597-633.
- Dorr, B. J., Ayan, N., & Habash, N. (2004). Divergence unravelling for word alignment of parallel corpora. *Natural Language Engineering*, 1(1), 1-17.
- Goyal, S., Gupta, D., & Chatterjee, N. (2004). A study of Hindi translation patterns for English sentences with 'have' as the main verb. *Proceeding of iSTRANS*, 1, 46-51.
- Gupta, D., & Chatterjee, N. (2001). Study of divergence for Example-Based English-Hindi machine translation. *Proceedings of the STRANS 2001*, IIT-Kanpur, 132-139.
- Gupta, D., & Chatterjee, N. (2003). Identification of divergence for English to Hindi EBMT. *Proceeding of MT Summit-IX*, 41-48.
- Jackendoff, R. S. (1983). *Semantics and Cognition*. MA. MIT: Cambridge University Press.
- Kasperek, C. (1983). The Translator's Endless Toil. *The Polish Review*, 28(2), 83-87.
- Lytinen, S., & Schank, R. (1982). *Representation and Translation*. (Technical Report 234), Department of Computer Science, Yale University, New Haven, CT, 1982.
- Schank, R., & Abelson, R. (1977). *Scripts, Plans, Goals, and Understanding*. Hillsdale, NJ: Lawrence Erlbaum Associates, Inc.
- Sinha, R. M. K., & Thakur, A. (2004). Synthesizing verb form in English to Hindi translation: Case of mapping infinitive and gerund in English to Hindi. *Proceeding of iSTRANS*, 1, 52-55.
- Sinha, R. M. K., & Thakur, A. (2005a). Translation divergence in English-Hindi machine translation. *Proceeding of EAMT Xth Annual Conference*, Budapest, Hungary, 30-31 May, 245-254.
- Sinha, R. M. K., & Thakur, A. (2005b). Divergence patterns in machine translation between Hindi and English. *Proceeding of MT Summit X*. Phuket, Thailand, 12-16<sup>th</sup> September, 346-353.