

Divergence Patterns in Kashmiri –English Machine Translation: A View from Translation of Tenses

Sajad Hussain Wani *

Abstract

Kashmiri English language pair exhibits different kinds of translation divergences like addition, substitution, categorical divergence, conflation divergence. In addition there are many grammatical constructions and grammatical phenomena which exhibit considerable divergence. This paper presents an in-depth account of translation divergence across different tense types. Divergence across tenses happen due to many differences in grammatical phenomena and their expression like inflections, honorifics, proximate and remote kind of pasts. There is a one to many translation pattern between English and Kashmiri and this is illustrated for every tense type.

Keywords: Divergence, Machine Translation, Categorical Divergence, Honorifics, Proximate and Remote Pronouns.

Introduction

Translation is sometimes referred to as the most difficult task among the academicians. Translation is not a uni-dimensional process but is multidimensional interweaving linguistic, social, cultural and other factors. Translation is not only the transference of meaning between languages but is a negotiation between two cultures, between two mind sets and between two time periods. Translation can never reach to the level of absolute equivalence but is always accompanied by a change of some kind. The

* Department of Linguistics, University of Kashmir, Srinagar, India.

terms like absolute equivalence, near equivalence, linguistic loss, cultural loss are very common in the translation literature.

Machine translation, sometimes referred to by the abbreviation MT, is a sub-field of computational linguistics that investigates the use of computer software to translate text or speech from one natural language to another. At its basic level, MT performs simple substitution of words in one natural language for words in another. Using corpus techniques, more complex translations may be attempted, allowing for better handling of differences in linguistic typology, phrase recognition, and translation of idioms, as well as the isolation of anomalies. Automatic translation between human languages ('Machine Translation') is a Science Fiction staple, and a long-term scientific dream of enormous social, political, and scientific importance. It was one of the earliest applications suggested for digital computers, but turning this dream into reality has turned out to be a much harder, and in many ways a much more interesting task than at first appeared. Nevertheless, though there remain many outstanding problems, some degree of automatic translation is now a daily reality, and it is likely that during the next decade the bulk of routine technical and business translation will be done with some kind of automatic translation tool, from humble databases containing canned translations of technical terms to genuine Machine Translation Systems that can produce reasonable draft translations (provided the input observes certain restrictions on subject matter, style, and vocabulary).

Divergence is a common phenomenon in translation between two natural languages. Typically, translation divergence occurs when structurally similar sentences of the source language do not translate into sentences that are similar in structure in the target language (Dorr, 1993). As a consequence, dealing with divergence assumes special significance. Dorr (1993) categorizes translation divergences into two broad types. They are: (A) syntactic Divergences, (B) Lexical-semantic Divergences. They are further subcategorized as follows:

(A) Syntactic Divergence: i. Constituent order divergence, ii. Adjunction divergence, iii.

Preposition-stranding divergence, iv. Movement divergence, v. Null subject divergence, vi. Dative divergence and vii. Pleonastic divergence

(B) Lexical-Semantic Divergence: i. Thematic divergence, ii. Promotional divergence, iii. Demotional divergence, iv. Structural divergence v. Conflational divergence, vi. Categorical divergence and vii. Lexical divergence.

In Dorr (1994), she has examined the structure of the lexical-semantic divergences and proposed a LCS-based approach for their resolution. This classification takes into account various sources of differences between a set of translation languages and captures a large sets of translation divergences. The classification is based on the Government and Binding framework (Chomsky 1986, Jackendoff 1990) of linguistic theory which assumes a deep structure to capture the surface structure variations. The deep structure functions as the universal structure, i.e. applicable across languages. Thus both the classification and the resolution of the translation divergences are largely discussed from the perspective of the universal grammar. The classification captures the major grammatical issues in translation divergence across languages. However, it also misses a number of points that pertain to a particular set of translation languages.

The issue of divergence between a set of languages is associated with a number of factors ranging from linguistic to socio- and psycho-linguistic aspects of the languages involved. Although Dorr's classification takes into account many of the major linguistic factors associated with translation divergence, there still remains a number of points related to both linguistic and extra-linguistic factors that may exist in different sets of translation languages. Furthermore, the parameters of the classification does not take into account subtle semantic factors to the extent they are relevant for the classification of translation divergences in various languages.

In the existing literature, the issue of translation divergence for Hindi and English MT has not been exhaustively examined. Gupta et al (2003) and Dave et al (2001) discuss some of the translation divergences pertaining to English-Hindi MT and Hindi-English MT. Dave et al (2001) discusses the issue within the UNL-based Interlingua approach and only some of the

obvious types of divergences have been discussed. These works do not explore further areas of divergence. Similarly other scholars like R.K. Sinha and Anil Thakur (2004) have also given a very exhaustive account of translation divergence. All these researches have stressed on the need to study divergence patterns in other language pairs so that the accuracy in machine translation should reach to it's target goal.

Divergence patterns need to be studied in great detail as capturing these divergence patterns can be very helpful in developing algorithms which will be very useful in the development of machine translation systems in future. It is in this back drop that this paper deals with divergence which is observed when one translates between different tenses between Kashmiri and English language pair.

Methodology

This research paper forms part of a doctoral dissertation on morpho-syntactic divergence patterns in Kashmiri English machine translation. The methodology adopted is based on the principles of descriptive and structural Linguistics. The data has been collected from many informal and formal settings and data consists of about 30 hrs of recorded speech. Data has also been utilized from a variety of text books, magazines and newspapers. Some examples have been self created to illustrate a particular divergence pattern.

Analysis

Divergence in Kashmiri English language occurs due to many structural differences between Kashmiri and English languages although there are considerable similarities between these languages. The main difference between Kashmiri and English languages is that while English has lost most of its inflections, Kashmiri language is highly inflectional in nature. Kashmiri has comparably a greater degree of flexibility when compared to English language. While English is a gender neutral language; Kashmiri has gender for all the animate and inanimate nouns and nouns as well as verbs are marked with suffixes showing the gender of subject and object.

Verbs in Kashmiri show a complex agreement pattern. The primary agreement is obligatory. It is marked on the auxiliary if one is present, otherwise it is marked on the verb itself. The primary agreement is cued mostly to gender, number and person of the nominative case. Gender features are absent in the future tense. The nominative case is controlled by:

- (i) subjects of (a) present and future tenses, (b) past intransitives, and (c) passives;
- (ii) nominative/ absolutive direct objects of past and perfective tenses, and
- (iii) thematic objects of dative / psyche verbs.

Kashmiri is a honorific language and this is shown by second person plural marker and this leads to divergence of a sentence from Kashmiri to English. Similarly, Kashmiri makes a distinction between proximate and remote pronouns and this further leads to divergence across tenses in Kashmiri and English languages.

Translation of Tenses

Tenses show a one to many divergence pattern for every tense type. Kashmiri sentences are marked for the additional features of gender, honorifics and remote and proximate distinctions thus giving rise to one to many translation divergence. These tense paradigms can be used for any example based machine translation system in an effective manner. The one to many translation divergence is shown for all the tense types and subtypes in the tabulated form as below:

Paradigm for Present Indefinite Tense (6:20 paradigm)

1. I play cricket	<i>bI chus kirkaT ginda:n</i> (mas)
2. We play cricket	<i>As' chi kirkaT ginda:n</i> (mas)
3.I play cricket	<i>bI chas kirkaT ginda:n</i> (fem)
4. We play cricket	<i>As' cha kirkaT ginda:n</i> (fem)

5. You play cricket	<i>tsI chukh kirkaT ginda:n</i> (mas,sng)
6. you play cricket	<i>toh' chiv kirkaT ginda:n</i> (mas, pl/ hon.sng)
7. You play cricket	<i>tsI chakh kirkaT ginda:n</i> (fem,sng)
8.You play cricket	<i>toh' chavI kirkaT ginda:n</i> (fem,pl)
9. He plays cricket.	<i>yi chu kirkaT ginda:n</i> (prox)
10. He plays cricket.	<i>hu chu kirkaT ginda:n</i> (remI)
11. He plays cricket.	<i>su chu kirkaT ginda:n</i> (remII)
12. She plays cricket	<i>yi chi kirkaT ginda:n</i> (prox)
13. She plays cricket	<i>hə chi kirkaT ginda:n</i> (remI)
14. She plays cricket	<i>sə chi kirkaT ginda:n</i> (remII)
15. They play cricket	<i>yim chi kirkaT gind:an</i> (mas, prox, pl/hon. sng)
16. They play cricket	<i>hum chi kirkaT ginda:n</i> (mas, remI,pl/hon. sng)
17. They play cricket	<i>tim chi kirkaT ginda:n</i> (mas, remII, pl/hon.sng)
18. They play cricket	<i>yimI cha kirkaT ginda:n</i> (fem, prox,pl)
19. They play cricket	<i>humI cha kirkaT ginda:n</i> (fem, rem1,pl)
20. They play cricket	<i>timI cha kirkaT ginda:n</i> (fem, remII,pl)

From the above given paradigm for present indefinite tense; it becomes clear that 6 forms of a sentence in present indefinite in English has about 20 forms in Kashmiri and thus it can be said that for Kashmiri English language pair a (6:20) paradigm is observed when translating between these two languages in the present indefinite tense. The Kashmiri sentences are divergent because different inflectional forms are utilized for marking of gender , number, honorific marking, and different pronouns are utilized for marking third person proximate, remote I (rem I) and remote II (rem II) pronouns. The one to many paradigms for other tense types are shown as follows:

Paradigm for Past Indefinite Tense (6:13 Paradigm)

1. I played cricket	<i>m'e g'und kirkaT</i> (mas,fem)
2. we played cricket	<i>asI g'und kirkaT</i> (mas,fem)
3. You played cricket	<i>tse g'unduth kirkaT</i> (mas,fem,sng)
4. you played cricket	<i>tshi g'undvI kirkaT</i> (mas,fem,pl/hon. sng)
5. He played cricket	<i>yem' g'und kirkaT</i> (prox)
6. He played cricket	<i>hom' g'und kirkaT</i> (remI)
7. He played cricket	<i>tAm' g'und kirkaT</i> (remII)
8. She played cricket	<i>yemi g'und kirkaT</i> (prox)
9. She played cricket	<i>homi g'und kirkaT</i> (remI)
10. She played cricket	<i>tAmi g'und kirkaT</i> (remII)

11. They played cricket	yimav g'und kirkaT (mas,fem, prox,pl/hon.sng)
12. They played cricket	humav g'und kirkaT (mas,fem, remI)
13. They played cricket	timav g'und kirkaT (mas, fem,remII,pl/hon.sng)

Paradigm for Future Indefinite Tense (6:16 Paradigm)

1. I shall play cricket	bi gindI kirkaT (mas,fem)
2. We shall play cricket	As' gindav kirkaT (mas,fem)
3. You will play cricket	tsI gindakh kirkaT (mas,fem,sng)
4. You will play cricket	tchi gindIv kirkaT (mas,fem,pl/hon. sng)
5. He will play cricket.	yI gindI kirkaT (prox)
6. He will play cricket.	hu gindI kirkaT (remI)
7. He will play cricket	su gindI kirkaT (remII)
8. She will play cricket	yI gindI kirkaT (prox)
9. She will play cricket	hɔ gindI kirkaT (remI)
10. She will play	sɔ gindI kirkaT (remII)
11. They will play cricket	yem' gindan kirkaT (mas, prox, pl/hon. sng)

12. They will play cricket	<i>hum gindan kirkaT</i> (mas, remI, pl/hon. sng)
13.They will play cricket	<i>tim gindan kirkaT</i> (mas, remII, pl/hon. sng)
14.They will play ricket	<i>yimI gindan kirkaT</i> (fem, prox,pl)
15. They will play cricket	<i>humI gindan kirkaT</i> (fem, remI,pl) mn
16. They will play cricket	<i>timI gindan kirkaT</i> (fem, remII,pl)

Paradigm for Present Continuous Tense (6:20Paradigm)

1. I am playing cricket	<i>bl chus kirkaT ginda:n</i> (mas)
2. We are playing cricket	<i>As' chi kirkaT ginda:n</i> (mas)
3.I am playing cricket	<i>bl chas kirkaT ginda:n</i> (fem)
4. We are playing cricket	<i>As' cha kirkaT ginda:n</i> (fem)
5. You are playing cricket	<i>tsI chukh kirkaT ginda:n</i> (mas,sng)
6. you are playing cricket	<i>toh' chiv kirkaT ginda:n</i> (mas, pl/ hon.sng)
7. You are playing cricket	<i>tsI chakh kirkaT ginda:n</i> (fem,sng)
8.You are playing cricket	<i>toh' chavI kirkaT ginda:n</i> (fem,pl)
9. He is playing cricket.	<i>yi chu kirkaT ginda:n</i> (prox)

10. He is playing cricket.	<i>hu chu kirkaT ginda:n</i> (remI)
11. He is playing cricket.	<i>su chu kirkaT ginda:n</i> (remII)
12. She is playing cricket	<i>yi chi kirkaT ginda:n</i> (prox)
13. She is playing cricket	<i>hə chi kirkaT ginda:n</i> (remI)
14. She is playing cricket	<i>sə chi kirkaT ginda:n</i> (remII)
15. They are playing cricket	<i>yim chi kirkaT gind:an</i> (mas, prox, pl/hon. sng)
16. They are playing cricket	<i>hum chi kirkaT ginda:n</i> (mas, remI,pl/hon. sng)
17. They are playing cricket	<i>tim chi kirkaT ginda:n</i> (mas, remII, pl/hon.sng)
18. They are playing cricket	<i>yimI cha kirkaT ginda:n</i> (fem, prox,pl)
19. They are playing cricket	<i>humI cha kirkaT ginda:n</i> (fem, remI,pl)
20. They are playing cricket	<i>timI cha kirkaT ginda:n</i> (fem, remII,pl)

Paradigm for Past Continuous Tense (6:20 Paradigm)

1. I was playing cricket	<i>bI o:sus kirkaT ginda:n</i> (mas)
2. We were playing cricket	<i>As' A:s' kirkaT ginda:n</i> (mas)
3.I was playing cricket	<i>bI A:sIs kirkaT ginda:n</i> (fem)
4. We were playing cricket	<i>As' a:sI kirkaT ginda:n</i> (fem)

5. You were playing cricket	<i>tsI o:sukh kirkaT ginda:n</i> (mas,sng)
6. you were playing cricket	<i>toh' A:sivI kirkaT ginda:n</i> (mas, pl/hon.sng)
7. You were playing cricket	<i>tsI A:sIkh kirkaT ginda:n</i> (fem,sng)
8.You were playing cricket	<i>toh' a:sIvI kirkaT ginda:n</i> (fem,pl)
9. He was playing cricket.	<i>yi o:s kirkaT ginda:n</i> (prox)
10. He was playing cricket.	<i>hu o:s kirkaT ginda:n</i> (remI)
11. He was playing cricket.	<i>su o:s kirkaT ginda:n</i> (remII)
12. She was playing cricket	<i>yi A:s' kirkaT ginda:n</i> (prox)
13. She was playing cricket	<i>hə A:s' kirkaT ginda:n</i> (remI)
14. She was playing cricket	<i>sə A:s' kirkaT ginda:n</i> (remII)
15. They were playing cricket	<i>yim A:s' kirkaT gind:an</i> (mas, prox, pl/hon. sng)
16. They were playing cricket	<i>hum A:s' kirkaT ginda:n</i> (mas, remI,pl/hon. sng)
17. They were playing cricket	<i>tim A:s' kirkaT ginda:n</i> (mas, remII, pl/hon.sng)
18. They were playing cricket	<i>yimI a:sI kirkaT ginda:n</i> (fem, prox,pl)
19. They were playing cricket	<i>humI a:sI kirkaT ginda:n</i> (fem, remI,pl)
20. They were playing cricket	<i>timI a:sI kirkaT ginda:n</i> (fem, remII,pl)

Paradigm for Future Continuous Tense (6:16 Paradigm)

1. I shall be playing cricket	<i>bl a:si kirkaT ginda:n</i> (mas,fem)
2. We shall playing cricket	<i>As' a:sav kirkaT ginda:n</i> (mas,fem)
3. You will be playing cricket	<i>tsI a:sakh kirkaT ginda:n</i> (mas,fem,sng)
4. you will playing cricket	<i>toh' A:siv kirkaT ginda:n</i> (mas,fem, pl/hon.sng)
5. He will be playing cricket.	<i>yi a:si kirkaT ginda:n</i> (prox)
6. He will be playing ricket.	<i>hu a:si kirkaT ginda:n</i> (remI)
7. He will be playing cricket.	<i>su a:si kirkaT ginda:n</i> (remII)
8. She will be playing cricket	<i>yi a:si kirkaT ginda:n</i> (prox)
9. She will be playing cricket	<i>hə a:si kirkaT ginda:n</i> (remI)
10. She will be playing cricket	<i>sə a:si kirkaT ginda:n</i> (remII)
11. They will be playing cricket	<i>yim a:san kirkaT gind:an</i> (mas, prox, pl/hon. sng)
12. They will be playing cricket	<i>hum a:san kirkaT ginda:n</i> (mas, remI,pl/hon. sng)
13. They will be playing cricket	<i>tim a:san kirkaT ginda:n</i> (mas, remII, pl/hon.sng)

14. They will be playing cricket	<i>yimI a:san kirkaT ginda:n</i> (fem, prox,pl)
15. They will be playing cricket	<i>humI a:san kirkaT ginda:n</i> (fem, remI,pl)
16. They will be playing cricket	<i>timI a:san kirkaT ginda:n</i> (fem, remII,pl)

Paradigm for Present Perfect Tense (6:15 Paradigm)

1. I have played cricket	<i>me chu kirkaT g'undmut</i> (mas,fem)
2. we have played cricket	<i>asi chu kirkaT g'undmut</i> (mas,fem)
3. You have played cricket	<i>tse chuth kirkaT g'undmut</i> (mas,fem,sng)
4. you have played cricket	<i>t̄hi chuvI kirkaT g'undmut</i> (mas,fem, pl/hon.sng)
5. He has played cricket.	<i>yem' chu kirkaT g'undmut</i> (prox)
6. He has played cricket.	<i>hom' chu kirkaT g'undmut</i> (remI)
7. He has played cricket.	<i>tAm' chu kirkaT g'undmut</i> (remII)
8. She has played cricket	<i>yemi chu kirkaT g'undmut</i> (prox)
9. She has played cricket	<i>homi chu kirkaT g'undmut</i> (remI)
10. She has played cricket	<i>tami chu kirkaT g'undmut</i> (remII)
11.They have played cricket	<i>yimav chu kirkaT g'undmut</i> (mas,fem, prox, pl/hon. sng)

12. They have played cricket	<i>humavchu kirkaT g'undmut</i> (mas,fem, remI,pl/hon. sng)
13.They have played cricket	<i>timav chu kirkaT g'undmut</i> (mas,fem, remII,pl/hon. sng)

Paradigm for Past Perfect Tense (6:15 Paradigm)

1. I had played cricket	<i>me o:s kirkaT g'undmut</i> (mas,fem)
2. we had played cricket	<i>asi o:s kirkaT g'undmut</i> (mas,fem)
3.You had played cricket	<i>tse o:suth kirkaT g'undmut</i> (mas,fem,sng)
4. you had played cricket	<i>t̄hi o:svI kirkaT g'undmut</i> (mas,fem, pl/hon.sng)
5. He had played cricket.	<i>yem' o:s kirkaT g'undmut</i> (prox)
6. He had played cricket.	<i>hom' o:s kirkaT g'undmut</i> (remI)
7. He had played cricket.	<i>tAm' o:s kirkaT g'undmut</i> (remII)
8. She had played cricket	<i>yemi o:s kirkaT g'undmut</i> (prox)
9. She had played cricket	<i>homi o:s kirkaT g'undmut</i> (remI)
10. She had played cricket	<i>tami o:s kirkaT g'undmut</i> (remII)
11.They had played cricket	<i>yimav o:s kirkaT g'undmut</i> (mas,fem, prox, pl/hon. sng)
12. They had played cricket	<i>humav o:s kirkaT g'undmut</i> (mas,fem, remI,pl/hon. sng)

13.They had played cricket	<i>timav o:s kirkaT g'undmut</i> (mas,fem, remII,pl/hon. sng)
----------------------------	---

Paradigm for Future Perfect Tense (6:15 Paradigm)

1. I shall have played cricket	<i>me a:si kirkaT g'undmut</i> (mas,fem)
2. we shall have played cricket	<i>asi a:si kirkaT g'undmut</i> (mas,fem)
3.You will have played cricket	<i>tse a:seth kirkaT g'undmut</i> (mas,fem,sng)
4. you will have played cricket	<i>tchi a:svI kirkaT g'undmut</i> (mas,fem, pl/ hon.sng)
5. He will have played cricket.	<i>yem' a:si kirkaT g'undmut</i> (prox)
6. He will have played cricket.	<i>hom' a:si kirkaT g'undmut</i> (remI)
7. He will have played cricket.	<i>tAm' a:si kirkaT g'undmut</i> (remII)
8. She will have played cricket	<i>yemi a:si kirkaT g'undmut</i> (prox)
9. She will have played cricket	<i>homi a:si kirkaT g'undmut</i> (remI)
10. She will have played cricket	<i>tami a:si kirkaT g'undmut</i> (remII)
11.They will have played cricket	<i>yimav a:si kirkaT g'undmut</i> (mas,fem, prox, pl/hon. sng)

12. They will have played cricket	<i>humav a:si kirkaT g'undmut</i> (mas,fem, remI,pl/hon. sng)
13.They will have played cricket	<i>timav a:si kirkaT g'undmut</i> (mas,fem, remII,pl/hon. sng)

Paradigm for Present Perfect Continuous Tense (6:13 Paradigm)

1. I have been playing cricket	<i>m'e vo:t kirkaT ginda:n</i> (mas,fem)
2. we have been playing cricket	<i>asI vo:t kirkaT ginda:n</i> (mas,fem)
3. You have been playing cricket	<i>tse vo:tuy kirkaT ginda:n</i> (mas,fem,sng)
4.you have been playing cricket	<i>tchi vo:tvI kirkaT ginda:n</i> (mas,fem,pl/hon. sng)
5.He has been playing cricket	<i>yemis vo:t kirkaT ginda:n</i> (prox)
6. He has been playing cricket	<i>homis vo:t kirkaT ginda:n</i> (remI)
7. He has been playing cricket	<i>tamis vo:t kirkaT ginda:n</i> (remII)
8. She has been playing cricket	<i>yemis vo:t kirkaT ginda:n</i> (prox)
9. She has been playing cricket	<i>homis vo:t kirkaT ginda:n</i> (remI)
10. She has been playing cricket	<i>tamis vo:t kirkaT ginda:n</i> (remII)
11. They have been playing cricket	<i>yiman vo:t kirkaT</i>

	<i>ginda:n</i> (mas,fem, prox, pl/hon. sng)
12. They have been playing cricket	<i>human vo:t kirkaT</i> <i>ginda:n</i> (mas,fem, remI,pl/hon. sng)
13. They have been playing cricket	<i>timan vo:t kirkaT</i> <i>ginda:n</i> (mas,fem, remII,pl/hon. sng)

Paradigm for Past Perfect Continuous Tense (6:13 Paradigm)

1. I had been playing cricket	<i>m'e o:s vo:mut kirkaT</i> <i>ginda:n</i> (mas,fem)
2. we had been playing cricket	<i>asI o:s vo:mut kirkaT ginda:n</i> (mas,fem)
3. You had been playing cricket	<i>tse o:suy vo:tmut kirkaT ginda:n</i> (mas,fem,sng)
4. you had been playing cricket	<i>tchi o:svI vo:tmut kirkaT</i> <i>ginda:n</i> (mas,fem,pl/hon. sng)
5. He had been playing cricket	<i>yemis o:s vo:mtut kirkaT ginda:n</i> (prox)
6. He had been playing cricket	<i>homis o:s vo:tmutt kirkaT ginda:n</i> (remI)
7. He had been playing cricket	<i>tamis o:s vo:tmut kirkaT ginda:n</i> (remII)
8. She had been playing cricket	<i>yemis o:s vo:tmtut kirkaT ginda:n</i> (prox)

9. She had been playing cricket	<i>homis o:s vo:tmut kirkaT ginda:n</i> (remI)
10. She had been playing cricket	<i>tamis o:s vo:tmut kirkaT ginda:n</i> (remII)
11.They had been playing cricket	<i>yiman o:s vo:tmut kirkaT ginda:n</i> (mas,fem, prox, pl/hon. sng)
12. They had been playing cricket	<i>human o:s vo:mut kirkaT ginda:n</i> (mas,fem, remI,pl/hon. sng)
13. They had been playing cricket	<i>timan o:s vo:mut kirkaT ginda:n</i> (mas,fem, remII,pl/hon. sng)

Paradigm for Future Perfect Continuous Tense (6:13 Paradigm)

1. I shall have been playing cricket	<i>me a:sI vo:tmut kirkaT ginda:n</i> (mas,fem)
2. We shall have been playing cricket	<i>asI a:sI vo:tmut kirkaT ginda:n</i> (mas,fem)
3. You will have been playing cricket	<i>tse a:sI vo:tmut kirkaT ginda:n</i> (mas,fem,sng)
4.You will have been playing cricket	<i>t̄hi a:sI vo:tmut kirkaT ginda:n</i> (mas,fem,pl/hon. sng)
5.He will have been playing cricket	<i>yemis a:sI vo:tmut kirkaT ginda:n</i> (prox)
6. He will have been playing cricket	<i>homis a:sI vo:tmut kirkaT ginda:n</i> (remI)

7. He will have been playing cricket	<i>tamis a:sI vo:tmuT kirkaT ginda:n</i> (remII)
8. She will have been playing cricket	<i>yemis a:sI vo:tmuT kirkaT ginda:n</i> (prox)
9. She will have been playing cricket	<i>homis a:sI vo:tmuT kirkaT ginda:n</i> (remI)
10. She will have been playing cricket	<i>tamis a:sI vo:tmuT kirkaT ginda:n</i> (remII)
11. They will have been playing cricket	<i>yiman a:sI vo:tmuT kirkaT ginda:n</i> (mas,fem, prox, pl/hon. sng)
12. They will have been playing cricket	<i>human a:sI vo:tmuT kirkaT ginda:n</i> (mas,fem, remI,pl/hon. sng)
13. They will have been playing cricket	<i>timan a:sI vo:tmuT kirkaT ginda:n</i> (mas,fem, remII,pl/hon. sng)

Conclusion

The divergence is noticeable when one starts translating simple sentences containing a subject and object. Different types of divergences are observed between Kashmiri and English languages as were observed by Dorr including addition, replacement, substitution, promotion, thematic and other types of divergence. However, the fact remains that divergence has to be studied in much detail for the said pair of languages as there are many areas of study where a considerable amount of time and energy needs to be spent for the study of the divergence patterns.

Verbs in Kashmiri show complex agreement and conjugation patterns and this has many implications for translation between different tense types. Kashmiri sentences are marked for additional features of gender, honorifics and remote and proximate distinctions giving rise to a *one to many* divergence when translating across tenses. For example while translating a present definite tense; one can find that a sentence having a

total of six forms in English is translated in Kashmiri to about twenty forms giving rise to a (6:20) pattern of MT divergence. Past indefinite shows a (6:20) pattern whereas future indefinite shows a (6:16) pattern. The translation across other tense types is also highly divergent with one to many divergence with present continuous and past continuous showing a (6:20) pattern with a (6:16) pattern for future continuous tense. Present perfect tense, past perfect tense and future perfect tense show a (6:15) pattern whereas present perfect continuous tense, past perfect continuous tense and future perfect continuous tense show a (6:13) pattern. From a preliminary study of the two languages, it becomes obvious that for designing a machine translation system for Kashmiri English language pair; a number of approaches need to be combined (a hybrid approach) as an approach based on examples alone or a rule based approach alone can't yield the required output.

List of Abbreviations:

mas = masculine	fem = feminine	sng = singular
pl= plural,	hon. = honorific	prox = proximate
remI = remote I	remII = remote II	

References

- Bhat, R. 1987. *A Descriptive Study of Kashmiri*. Delhi. Amar Prakashan.
- Bhatia, T. K. 1995. *Negation in South Asian Languages*. Patiala. Indian Institute of Language Studies.
- Bhatt, R.M. 1999. *Verb Movement and the Syntax of Kashmiri: Dordrecht*. Kluwar Academic Press.
- Chomsky, N.1965. *Aspects of Theory of Syntax*. Cambridge: MIT Press.

-----1981. *Lectures on Government and Binding*. New York: Mouton de Gruyter.

----- 1995a. *The Minimalist Program*. MIT Press: Cambridge.

Dorr, B. J. 1993. *Machine Translation: A View from the Lexicon*. Cambridge. MIT Press:

Dorr, B.J. 1994. Classification of Machine Translation Divergences and a Proposed Solution. *Computational Linguistics* 20(4), pp.597-633.

Dorr, B. J., Jordan, P. W. and Benoit, J. W. 1998. *A Survey of Current Paradigms in Machine Translation*, Technical Report LAMP-TR-027, UMIACS-TR-98-72, CS-TR-3961. University of Maryland, USA, College Park.

Dorr, B. J., Pearl, L., Hwa, R. and Habash, N. Y. A. 2002. DUSTer: A Method for Unraveling Cross-Language Divergences for Statistical Word Level Alignment. Proceedings of the Fifth Conference of the Association for Machine Translation in the Americas. AMTA-2002.

Grierson, G.A.1932. *A Dictionary of the Kashmiri language*. B.R. New Delhi: Publishing Corporation.

Gupta, D. and Chatterjee, N.2002a. Study of Similarity and its Measurement for English to Hindi EBMT. *Proceedings of STRANS-2002*. IIT Kanpur, pp.132-139.

..... 2002b. A Systematic Adaptation Scheme for English-Hindi Example-Based Machine Translation. *Proceedings of STRANS-2002*, IIT Kanpur, pp. 141-148.

Kak, A.A . and Wani, S.H. 2006. Strategies of Neutrality and Code Mixing Grammar. In *Recent Studies in Nepalese Linguistics*. Proceedings of the 12th Himalyan Languages Symposium & 27th annual Conference of Linguistic Society of Nepal, November 26-28, 2006. K. Rai, etal. (ed.) LSN: Nepal,pp.467-487.

.....2009. Code Switching, Code Mixing and Teaching Process. *In Teaching English as a Second Language: A New Pedagogy for a New Century.* (ed) Vyas, M and Patel, Y. L. New Delhi: PHI Learning Private Limited, pp.309-318.

..... 2009. An Evaluation of Kashmiri English Code Mixing in Myers Scotton Models with Special Reference to MLF and 4M model. *In Language Vitality in South Asia.* Fatihi, A.R(ed). Aligarh: Department of Linguistics, AMU, pp. 227-237.

Koul, O. N. 197. *Linguistic Studies in Kashmiri.* New Delhi: Bahri publications.

Koul, O. N. and Wali, K. 2009 . *Modern Kashmiri Grammar.* Delhi : Indian Institute of Language Studies

Sinha, R.M.K. 2002. Translating News Headings from English to Hindi. *Proceedings of 6th IASTED International Conference on Artificial Intelligence and Soft Computing* , ASC2002. Banff: Canada.

Wani. Sajad H.2010. Divergence in Kashmiri English Machine Translation: A Study in Passivisation Process. *In Interdisciplinary Journal of Linguistics IJL Vol.3*, pp. 235-250.

.....2011. *Code Mixing Constraints; From Grammatical to the Minimalist framework: A Study in Kashmiri-English Mixed Code.* M.phil dissertation published online in Asian EFL Journal available online at <http://www.asian-efl-journal.com/Thesis/Thesis-Wani.pdf>. Retrieved on 3-03-2012.

