

DIVERGENCE IN KASHMIRI-ENGLISH MACHINE TRANSLATION: A STUDY IN PASSIVISATION PROCESS

Sajad Hussain Wani

INTRODUCTION

In English as in many other languages, the passive voice is a grammatical voice in which the subject receives the action of a transitive verb. Passive voice emphasizes the process rather than who is performing the action. Passive refers more generally to the verbs using this construction and the passages in which they are used. In English, a passive verb is periphrastic; that is, it does not have a one-word form, but consists of an auxiliary verb plus the past participle of the transitive verb. The auxiliary verb usually is a form of the verb to be, but other auxiliary verbs, such as *get*, are sometimes used. The passive voice can be used in any number of tenses. The process of changing an active verb into a passive one is called passivization. Passivization is a valence -decreasing process, and it is sometimes referred to as a de-transitivizing process, because it changes transitive verbs into intransitives. . One can still introduce the actor of a passive verb using a *by* phrase. When such a phrase is missing, the construction is called an agentless passive.

According to Merriam–Webster's Dictionary of English Usage, the passive voice should be used when the receiver of the action is more important than the doer, or when the doer is unknown, unimportant, or perhaps too obvious to be worth mentioning, as in these examples:

The dog was killed by the villagers.

The house was robbed last week.

The passive voice can also be used to make other changes to a sentence's emphasis, including emphasizing a modifying adverb or even the performer of the action. The passive voice is sometimes used to conceal the performer of an action or the identity of a person responsible for a mistake. The passive voice is often used in scientific writing because of the tone of detachment and impersonality that

it helps to establish. However, some scientific journals prefer writers to use the active voice. In general, the passive voice is used to place focus on the grammatical patient rather than the agent. This often occurs when the patient is the topic of the sentence. However, the passive voice can also be used when the focus is on the agent.

As for as Kashmiri is concerned; word order in Kashmiri resembles the one in German, Dutch, Icelandic, Yiddish and a few other languages. These languages form a distinct set and are currently known as Verb Second (V-2) languages. The word order generated by V-2 languages is quite different from Verb middle languages such as English. In a V-2 language for example, any constituent of a sentence can precede the verb. In a V middle language; only restricted constituents may precede the verb.

Traditionally passives in Kashmiri have been divided into two types:

1. REGULAR / PERSONAL PASSIVE

In the regular / personal passive the transitive verb is put in the infinitive accompanied with the auxiliary *yun* 'to come.' The auxiliary inflects for tense and agreement if any. The passive subject is marked nominative. Certain exceptional verbs, such as *la:yun*, 'to beat,' and *pra:run* 'to wait' that inherently mark their objects in the dative in the active version, retain the dative case on their passive subjects. The passivized nominative subject, but not the dative one, agrees with *yun*.

1.	raheel	chu	Sajadas	Parnavan	K
	rahel	be-mas-sing	sajad-dat	teach.pr	
Raheel teaches sajad / Raheel is teaching Sajad.					

Divergence In Kashmiri-English Machine Translation

2.	sajad	chu	pamavni	yivan	raheelni	zaeryi	<i>K</i>
	sajad	be-mas-sing	teach.inf. abl	come-pst	rahel-abl	by/through	
Sajad is being taught by/through Raheel							

3.	janan	lo:y	salimas	lo:ri	si:t	<i>K</i>
	ja:n-erg	beat-pst	salim-dat	stick-abl	with	
Jaan beat Salim with a stick						

4.	salim-as	a:v	layni	lo:ri	si:t	<i>K</i>
	salim-dat	come-pst	beat-inf-abl	stick-abl	With	
Salim was beaten with a stick.						

In the double transitive construction, the indirect object retains its dative case and the nominative noun (i.e., the former direct object) controls the agreement.

5.	sajadan	li:ch	raheelas	Chith	<i>K</i>
	sajad-erg	wrote	rahel-dat	Letter	
Sajad wrote a letter to Raheel.					

6.	raheel- as	a:yi	chith	leikhni	sajadni	zaeryi	K
	raheel- dat	come -pst	Letter	write-inf- abl	sajad-inf- abl	through/ by	
A letter was written to Rahel by / through Sajad.							

2. CAPABILITY PASSIVE

Capability passive is impersonal in nature. It usually requires a negative or interrogative context. It uses the same verbal morphology as the one in personal passive. However, the capability passive usually retains the postpositional agent. Secondly, unlike the regular passive; the agent may be in the dative. Third, this passive ranges over both intransitive and transitive verbs. The sense encoded in capability passive is similar to that conveyed by the modal *hekun* (could):

7.	temsind	zaeryi	a:v	ni	leikhni	K
	he-gen-abl	by /through	come-pst	not	write-inf-abl	
He was unable to write.						

AIMS AND OBJECTIVES

Translation divergence has become one of the central topics in main stream linguistics in general and computational linguistics in particular. Divergence studies are being conducted throughout the world and in India as well. Divergence studies are promising in more than one way and it's multifarious nature can be very fruitful for the quality of machine translation output in the coming years. Most of the machine translation systems in past have suffered from the low quality of translation output and a detailed study of translation divergence in any language pair ensures a very high quality translation. The morpho-syntactic divergence patterns can to a very high degree increase the quality of machine translation in the modern world.

Divergence In Kashmiri-English Machine Translation

This paper talks about the phenomenon of Passivisation in Kashmiri with reference to English language and points out certain divergence patterns in different kinds of passive sentences in both languages. It then tries to map out the linguistic rules guiding the translation divergence between passivised sentences in the said language pair.

ANALYSIS

As discussed in the introduction section; traditional grammars have delineated two categories of passives in Kashmiri i.e; personal passives and capability passives but there is no general consensus about the nature of passivisation in Kashmiri. Some other scholars stress the fact that passivisation in Kashmiri is rare and follows another set of rules. It must be noted here that syntax is not the only marker of passive sense but extra-linguistic factors are also very important for passive voice and this also gets reflected in the supra-segmental features of the active and their passive counterparts.(a research area which needs to be studied). Although many grammars in Kashmiri have mentioned some facts about passivisation but a detailed account has never been presented. This paper provides a modest attempt in presenting Kashmiri passives in a more systematic and detailed way. The following type of passives were analysed in Kashmiri for the study of divergence patterns:

CANONICAL PASSIVE

The first type of passives which were analyzed can be called as the canonical passives, which are those passives which are formed on the analogy of English language. Passive constructions have a range of meanings and uses. The canonical use is to map a clause with a direct object to a corresponding clause where the direct object has become the subject. For example:

8.	mariya	chi	pasand	yivan	karni	suhailni	zaeryi	K
	mariya	be- sing- fem- pr	like	come- pr	do-inf- abl	suhail- abl	by/ through	
Mariya is liked by Suhail.								

The English equivalent of (8) is quite simple in English with direct object getting the place of subject. The semantic constraint can be the focus on the object "Mariya" and hence Passivisation. Another way of expressing the same thing is through (9) which though comparatively simple has the same meaning.

9.	mariya	chi	pasand	yivan	suhail-as	K
	mariya	be-pr-fem-sing	like	come-pr	suhail-dat	
Mariya is liked by Suhail.						

Another way of conveying the same concept is by placing focus on the Mariya taking it to the subject position without Passivisation.

10.	mariya	chi	pasand	suhail-as	K
	mariya	be-pr-fem-sing	like	suhail-dat	
Suhail likes Mariya./ Mariya is what Suhail likes					

The examples (8) is a typical case of canonical passive as in English where a "by" phrase becomes necessary in certain contexts. Examples (2), (4), (6) and (7) are typical examples of canonical passives in Kashmiri which have been used by the traditional grammars. What is argued here is that canonical passives of both personal and capability type which have been used by traditional grammars are somehow unnatural and have been constructed on the analogy of English and Hindi and are somehow restricted to written form of Kashmiri. Besides, the word *zaeryi* (through) / *dast* (involving hand) is ambiguous as it connotes agency rather than the agent. By agency is meant the means and not necessarily the agent. English "by" has no translational equivalent in Kashmiri and this provides the key difference between passivisation in Kashmiri and English. This can be illustrated by the following example:

11.	sofi	a:yi	janni	zaeryi	marni	K
	sofi	come-pst	john-abl	through/*by	kill-inf-abl	
Sofi was killed through/by john						

Divergence In Kashmiri-English Machine Translation

Example 11 is ambiguous having two interpretations which are that either John is the agent who killed Sofi directly or John is the cause or involved in her killing, so that we can say that using canonical passive in Kashmiri is ambiguous and hence is used in the context where John is agency and not the agent. It is because Kashmiri has another mechanism at the morphological level which shows the agent in a more clear and un-ambiguous way. The mechanism involved is that of flexible word order in Kashmiri as the object can be brought into focus without actual Passivisation. The corresponding structure of (11) in Kashmiri is (12):

12.	sofi	maer	jann-an	K
	sofi	kill-pst	john-dat	
John killed Sofi				

The Kashmiri sentence (12) takes Sofi (direct object) to the subject position without any change in the meaning. Semantically Sofi becomes focused but in English such a construction is not possible; in English if direct object has to be the focus in context; the only way is through Passivisation but in Kashmiri rich morphology of the language compensates for Passivisation. The dative -an in Kashmiri makes the movement of direct object very easy so that it can be moved to any position; be it sentence initial, middle or final.

Thus one reason for the less usage of canonical passives in Kashmiri is the rich morphological case system which can bring in focus any of the constituents of a sentence by shifting it to initial position. Such a movement is not possible in English and hence there is a greater occurrence of the canonical passives as no other syntactic operation can bring the focus on the constituent except the canonical Passivisation. Thus regarding the canonical passives in Kashmiri, we can say:

1. Due to the absence of a translation equivalent of "by" in Kashmiri; the Kashmiri canonical passives are of less occurrence and when these occur; there is an ambiguity connoting agency rather than agent involved in the process.
2. Another reason for the less occurrence of the canonical passives in Kashmiri can be attributed to the rich morphological system of Kashmiri language which offers movement of constituents and hence semantic focus without Passivisation and hence less occurrence of canonical passives which are frequent in English.

Since by marker has no close equivalent in Kashmiri, zaeryi (through) has been used which shows agency rather than agent. "daes" on the other hand connotes agent but is rarely used. This zaeryi is also used in literary works on the analogy of English and is a marked sentence in Kashmiri. Sometimes the use of zaeryi (through) / daes (by hand) becomes impossible and in such cases we have no choice but to drop agent whereas it is not so in English. In such examples again; the canonical passives are replaced by the flexible word order providing the focus on any constituent other than the agent.

NON-CANONICAL PASSIVES:

One non-canonical use of passive is to promote an object other than a direct object. It is usually possible in Kashmiri to promote an indirect objects as well. For example:

13.	sajad-as	a:v	gula:b	dyeni		<i>K</i>
	sajad-dat	come-pst	rose	give-inf-abl		
Sajad was given a rose.						

Thus in the above non-canonical use of passive, indirect object Sajad was raised to the subject position instead of direct object gula:b (rose). In this respect Kashmiri language resembles English and other dechticaetiative languages.

Another non-canonical use of passivisation in Kashmiri is where the object of postposition is promoted to the subject position. In such sentences where the object of postposition is promoted to the subject position; the passive construction follows more naturally than the corresponding active construction. e.g;

14.	kamr-as	manz	a:v	haersi	khyeni	<i>K</i>
	room-dat	in	come-pst	harisa	eat-inf-abl	
The harissa was eaten in the room.						

Divergence In Kashmiri-English Machine Translation

In case of English; it is possible only in some cases to promote the object of a preposition.

e.g;

15.	masl-as	mutlak	aayi	kath	karni	K
	problem-dat	About	come-pst	talk	do-inf-abl	
The problem was talked about						

In the above example; the Kashmiri construction is an unmarked one and is of regular occurrence. In English equivalent here, the preposition is "stranded"; that is, it is not followed by an object. In Kashmiri, the objects of postpositions are frequently brought to the subject position whereas this is not so in English language.

It is possible to promote a content clause that serves as a direct object. In this case, however, it typically does not change its position in the sentence, and an expletive "ye" (it) takes the normal subject position. Kashmiri resembles English in this type of non-canonical Passivisation:

16.	ye	chu	vanni	yiva:n	Ki	su	a:v	K
	it	be-pst-mas	say-inf-abl	come-pst	That	he	came	
It is said that he came								

A stative passive or a resultative passives also exist in Kashmiri just like English. A stative passive rather than describing an action, describes the result of an action. e.g;

17.	byuol	ou:s	hokhmut	K
	seeds	be-pst-mas-sng	dried	
The seeds were dried				

This example (17) has two different meanings, roughly the following:

1. [Someone] dried the seeds./ The seeds were dried by someone.
2. The seeds were not fresh.

The former meaning represents the canonical, eventive passive; the latter, the stative passive.

Adjectival passives in Kashmiri are not true passives; they occur when a participial adjective is used predicatively . e.g;

18.	yi	ou:s	taya:r	karni	a:mut	k
	this	be-pst-mas-sing	prepare	do-inf-abl	come-pst-pc	
This was prepared.						

In example (18) "tayar karni a:mut" acts as an adjectival passive and is used predicatively. There are many adjectival passives in Kashmiri. The nature of adjectival passives and their formation is different in English and Kashmiri. A wide variety of examples of adjectival passives can be cited which are formed as a result of complex morphological processes in Kashmiri; a discussion of which is beyond the scope of this paper. In some cases, the line between an adjectival passive and a stative passive may be unclear.

In a few cases, passive constructions retain all the sense of the passive voice, but do not have immediate active counterparts. For example:

Divergence In Kashmiri-English Machine Translation

19.	khandar	a:sun	chu	Ni	sahal	k
	marriage	be-inf-mas-sing	be-pr-mas-sing	Not	easy	
To get married is not easy./ Getting married is not easy.						

20.	zuv	dyun	chu	ni	yara:n	k
	life	give-inf-mas-sing	be-pr-mas-sing	do-inf-abl	friendship /easy.	
To sacrifice life is not easy./ Sacrificing life is not easy.						

21.	theek	karun	chu	ni	sahal	k
	correct	do-inf-mas-sing	be-pr-mas-sing	do-inf-abl	easy	
To correct is not easy.						

22.	teiz teiz	karun	chu	ni	mumkin	k
	quick quick	do-inf-mas-sing	be-pr-mas-sing	do-inf-abl	possible	
To do so quickly is not possible./ Being so quick is not easy.						

In Kashmiri many such passive constructions are utilized where a non-canonical passive construction is formed with the help of compound and conjunct verbs which do not have a close translational equivalent in English. Compound and conjunct verbs and their frequent usage is an unmarked feature in Kashmiri and a large number of compound and conjunct verbs are used in place of many non-canonical passive constructions in English.

Examples from 19-21 represent different types of conjunct verbs which are used in non-canonical passive sense which have no absolute translation equivalents in English except infinitival and participial constructions. In 19 and 20; conjunct verb consists of a noun and a verb, in 21 it consists of an adjective and a verb and in 22 it consists of an adverb and a verb. The only way to translate these frequently used conjunct verbs is through infinitival and participial constructions of English.

Badkismati aasin (having badluck), afva asin (being a rumour), shararat aasun (to be angry), shararat thaavin (being angry), boguz tha:vun (to be jealous) are some of the conjunct verbs which are used in Kashmiri very frequently and which can be said to carry a non-canonical passive form without any immediate active counterparts.

The modality in Kashmiri is expressed by the explicators a:sun 'to be,' pyoun 'to fall,' pazun 'to be proper,' lagun 'to get hurt/ to need,' and gatshun 'to go.' The explicators take all the gender, number, tense, and case markers. It must be clear here that these models are also the explicators which form a component of the conjunct verbs and hence again increase the frequency of non-canonical passives in Kashmiri. Thus modal auxiliaries also contribute to the Passivisation mechanism in Kashmiri language.

Pronominalization in Kashmiri is again a phenomenon which can account for some of the non-canonical passivisation processes in Kashmiri. Only a detailed study can reveal further intricacies of the relation between passivised elements and pronominalization.

In Kashmiri language passives can be studied at many other levels like double passives, alone past participles, ergative verbs, reflexive verbs, gerunds and nominalizations. Thus Kashmiri has retained different mechanisms of passivisation process which needs to be studied. Hence, we can conclude that passivisation has unique mechanisms in Kashmiri which are as different to English as they are similar to it. A further study of Passivisation process is expected to reveal many other interesting results.

DIVERGENCE PATTERNS IN PASSIVISATION AND THEIR RELEVANCE TO KASHMIRI ENGLISH MACHINE TRANSLATION

1. From our study; it becomes clear that canonical passives in Kashmiri are in most cases the forced interpretation of the term as Kashmiri due to flexible word order and rich morphology can bring both agent and patient (for that matter any constituent) to the subject position. The case markers in Kashmiri can indicate the agent and patient and can bring the focus according to the need.
2. Another divergence pattern which can be pointed out here is that between conjunct verbs and infinitival constructions. Most of the infinitival constructions in English are mapped on to conjunct verbs in Kashmiri (Examples 19,20,21).
3. Most other types of passive constructions in English like infinitival, adjectival, content clause in English are mapped on to conjunct verbs in Kashmiri. As it has been pointed out earlier that compound and conjunct verbs are frequently used in Kashmiri; hence these have taken the function of Passivisation.
4. Modal verbs also account for many passives in English language. Kashmiri modal usage is mapped on by many passives in English language. The reason for this correlation is the overlap between explicators of conjunct verbs and the modals in Kashmiri.
5. pronominalization in Kashmiri also accounts for many passive constructions in Kashmiri as these carry a sense of passiveness putting more emphasis not on agent but some other and keeping subject position occupied with some other element rather than the agent.
6. Many simple verbs in English are mapped on to conjunct verbs in Kashmiri which carry passive sense.e.g; jump(vuth tulen), expect (umeed Karin), release (aza:d karun) are some of the verbs which are replaced by conjunct verbs in Kashmiri.

CONCLUSION

From the above studies of Passivisation in Kashmiri and English; it becomes clear that English and Kashmiri have their unique properties and unique mechanisms for different processes which are related to the Passivisation as a whole. Whereas English uses various syntactic processes for Passivisation; Kashmiri does so in terms of certain morphological processes due to it's rich morphology

.Passivisation process is carried on in their own unique ways. Thus syntactic mechanism of Passivisation in English is mapped on to morphological mechanisms in Kashmiri and divergence observed is that between syntax and morphology. The above stated rules for the divergence patterns can be incorporated in any machine translation systems translating between Kashmiri and English in the coming future. Such a minute linguistic analysis and the study of divergence between two languages is bound to increase the quality of machine translation output of any machine translation system; given that the major challenge at present in many translation systems is linguistic rather than any other.

LIST OF ABBREVIATIONS

nom	= Nominative case.	dat	= Dative case
gen	= Genitive case.	erg	= Ergative case.
poss	= Possessive case.	mas	= Masculine
fem	= Feminine.	sing	= Singular.
pl	= Plural.	pr	= Present.
prg	= Progressive.	pst	= Past .
fut	= Future.	pp	= Preposition.
K	= Kashmiri.	abl	= ablative
pc	= participle	inf	= infinitive

REFERENCES

- Bhat, Roopkrishen. 1987. *A Descriptive Study of Kashmiri*. Delhi: Amar Prakashan.
- Brown, P. F., Pietra, S. A., Pietra, V. J. D., Pietra, D. and Mercer, R. L. 1993. "The Mathematics of Statistical Machine Translation: Parameter Estimation". *Computational Linguistics* 19 (2). 263-311.
- Dorr, B. J. 1993. *Machine Translation: A View from the Lexicon*. Cambridge, MA: MIT Press.

Divergence In Kashmiri-English Machine Translation

- Dorr, B. J. 1994. "Classification of Machine Translation Divergences and a Proposed Solution". *Computational Linguistics*. 20(4).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, College Park: USA.
- Gupta, D. and Chatterjee, N. 2002. "Study of Similarity and its Measurement for English to Hindi EBMT". *Proceedings of STRANS-2002*, IIT Kanpur.
- , 2003a. "Divergence in English to Hindi Translation: Some Studies". *International Journal of Translation* 15: 5-24.
- , 2003b. "Identification of Divergence for English to Hindi EBMT". *Proceedings of the MT SUMMIT IX, Orleans, LA*: 141-148.
- Jakobson, Roman. 1959. "On Linguistic Aspects of Translation". *Translation Studies Reader*. (2nd Edition). L. Venuti. New York: Routledge, 2000: 113-119.
- Kachru, Braj B. 1969 a. *A Reference Grammar of Kashmiri*. Urbana: University of Illinois.
- Koul, Omkar N. 1977. *Linguistic Studies in Kashmiri*. New Delhi: Bahri publications.
- Koul, Omkar N. and Wali, kashi. 2009. *Modern Kashmiri Grammar*. Delhi: Indian Institute of Language Studies.
- Shauq, Shafi 1983. "A Contrastive Study of Some Syntactic Patterns of English and Kashmiri with Special Reference to Complementation and Relativization". Unpublished Ph.D. dissertation: University of Kashmir.
- Sinha, R.M.K. 2002. "Translating News Headings from English to Hindi". Banff, Canada: *Proceedings of 6th IASTED International Conference on Artificial Intelligence and Soft Computing (ASC2002)*.
- Sinha, R.M.K. 2004. "An Engineering Perspective of Machine Translation: Angla Bharti - II and Anu Bharti - II Architecture". *Proceedings of International Symposium on Machine Translation NLP and TSS*. New Delhi: The McGraw-Hill Companies: 10-17.
- Sinha, R.M.K. and Thakur, A. 2005a. "Translation Divergence in English-Hindi MT". *The Proceeding of EAMT Xth Annual Conference*. Budapest, Hungary, 30-31 May.

S. Dave, Parikh, J. and Bhattacharyaa, P. 2001. "Interlingua-Based English-Hindi .Machine Translation and Language Divergence". *Machine Translation* 16(4):251-304.

Vikas, O. 2001. "Technology Development for Indian Languages". *Proceedings of Symposium on Translation Support Systems STRANS-2001*, IIT Kanpur.

Wali,Kashi and Omkar N. Koul 1997. *Kashmiri: A Cognitive-Descriptive Grammar*. London and New York: Routledge.

Watanabe, H.1992. "A Similarity-Driven Transfer system". *Proceeding of the 14th COLING* : 770-776.

Watanabe, H., Kurohashi, S. and Aramaki, E.2000. "Finding Structural Correspondences from Bilingual Parsed Corpus for Corpus-Based Translation". *Proceedings of COLING-2000*. Saarbrucken. Germany: 320.