Vibhakti Divergence between Sanskrit and Hindi

Preeti Shukla, Devanand Shukl and Amba Kulkarni

Department of Sanskrit Studies,
University of Hyderabad,
Hyderabad
shukla.preetidev@gmail.com,dev.shukl@gmail.com,apksh@uohyd.ernet.in

Abstract. Translation divergence at various levels between languages arises due to the different conventions followed by different languages for coding the information of grammatical relations. Though Sanskrit and Hindi belong to the same Indo-Aryan family and structurally as well as lexically Hindi inherits a lot from Sanskrit, yet divergences are observed at the level of function words such as vibhaktis. Pāṇini in his Aṣṭādhyāyī has assigned a default vibhakti to kārakas alongwith many scopes for exceptions. He handles these exceptions either by imposing a new kāraka role or by assigning a special vibhakti. However, these methods are not acceptable in Hindi *in toto*. Based on the nature of deviation, we propose seven cases of divergences in this paper.

Key words: Pāṇini, kāraka, vibhakti, semantics, translation divergence.

1 Introduction

Different languages may follow different conventions for coding the information of grammatical relations. This leads to divergence between languages at various levels in translation in general and in Machine Translation(MT) in particular. Translation divergence patterns have been discussed by Dorr [Dorr,1994], and based on those patterns considerable work is done for Indian Languages, viz., English-Sanskrit-Hindi MT [Goyal and Sinha, 2009], English-Sanskrit MT [Mishra and Mishra, 2008], English-Hindi MT [Dave *et al.,* 2002] to name a few.

Sanskrit and Hindi belong to the same Indo-Aryan family and Hindi inherits a lot from Sanskrit in terms of structure as well as lexicon. As such, when we look at the divergences between Sanskrit and Hindi, most of the cases discussed in Dorr's divergence turn out to be either of rare occurrence, or do not pose much problem as far as accessing the source text using machine tools is concerned. On the contrary divergences are observed at the level of function words such as vibhaktis. While both Sanskrit and Hindi are inflectional languages, Sanskrit is synthetic in nature while Hindi is analytic [Dwivedi, 2006]. In Sanskrit both *sup* and *tiṅ* suffixes are termed as vibhakti pratyayas.[1] *Sup* suffix is added to a nominal stem and *tiṅ* suffix is added to a verbal root. But in Hindi only the

[1] vibhaktiśca (suptiṅau vibhaktisanjñau staḥ -S.K.) (P-1.4.104)

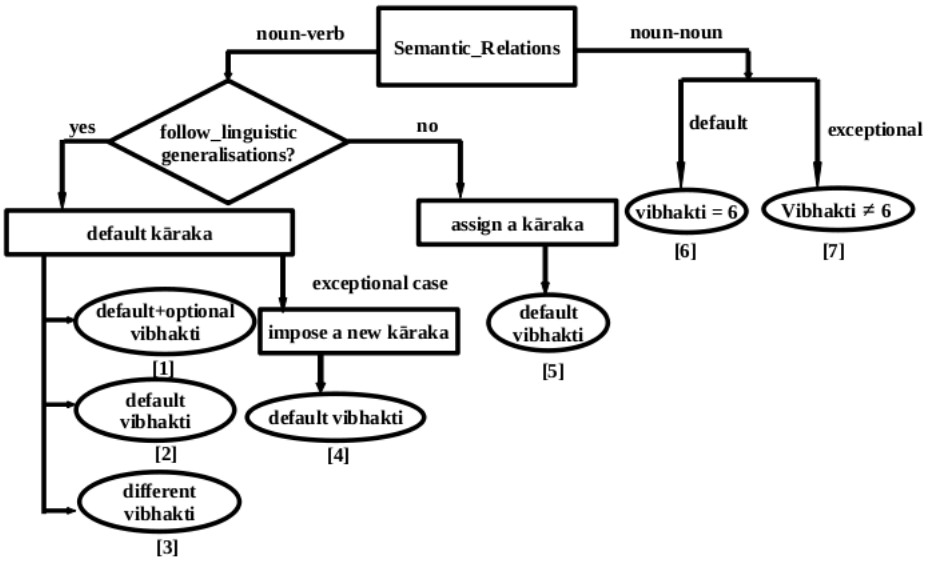
suffix added to the nominal stem is called as a vibhakti pratyaya. In this paper we discuss the divergences in the use of nominal suffixes (vibhakti pratyayas) between Sanskrit and Hindi.

2 Classification of Divergences

Pāṇini uses kāraka - a syntactico-semantic relation as an intermediary step to express the semantic relations through vibhaktis. The assignment of kāraka to various semantic categories is not one-to-one. Rama Nath Sharma (2002) observes-

> Pāṇini specifies his kāraka categories based upon the principle of sāmānya 'general', viśeṣa 'particular' and śeṣa 'residual'. The six categories are identified by general rules formulated based upon linguistic generalizations. Particular rules form exceptions to them. Usage which cannot be accounted for by the above two rule types is governed by rules relegated to the residual category. It is obvious that these exceptions are necessary to capture the peculiarities of usage falling outside the scope of the general rules.

Each kāraka in his system has a default vibhakti. But as is well-known, there are exceptions and hence there is no one-to-one mapping between the kāraka relations to vibhaktis. Pāṇini handles these deviations by employing two methods: (a) imposing a kāraka role and (b) assigning a special vibhakti. For example through the sūtra *ādhāro'dhikaraṇam* (P-1.4.45) a *locus* is mapped to *adhikaraṇa* and then it takes seventh case suffix by default (*saptamyadhikaraṇe ca* P-2.3.36). However the locus of the action related to the verbal roots *śīṅ, sthā and āsa* preceded by the upasarga *adhi* is termed as karma (*adhiśīṅsthāsāṃ karma* P-1.4.46) and then by the sūtra *karmaṇi dvitīyā* (P-2.3.2) this karma takes the second case suffix. Thus the deviation of the adhikaraṇa taking the second case suffix instead of seventh is handled by imposing a karma kāraka role in place of adhikaraṇa. Similarly, when the linguistic generalisations cannot be captured, he treats the cases as exceptional as in *rucyarthānām prīyamāṇaḥ* (P-1.4.33), where he assigns a sampradāna kāraka to the one who desires. In Fig 1, we summarise Pāṇini's way of mapping semantic relations to vibhaktis through kārakas.



1. Divaḥ karma ca (sādhakatamam, kārake) (P–1.4.43)
2. Karmaṇi dvitīyā (P–2.3.2)
3. Divastadarthasya (karmaṇi, ṣaṣṭhī śeṣe, anabhihite) (P–2.3.58)
4. Adhiśīṅsthāsām karma (ādhāraḥ, kārake) (P–1.4.46)
5. Krudhadruhērṣyānām yam prati kopaḥ (sampradānam, kārake) (P–1.4.37)
6. Ṣaṣṭhī śeṣe (P–2.3.50)
7. Pṛthagvinānānābhistṛtīyānyatarasyām (dvitīyā, pañcamī) (P–2.3.32)

Fig. 1. Semantic–Vibhakti_Mapping

This classification does not hold good as-it-is for Hindi. If a parallel grammar for Hindi in the Aṣṭādhyāyī style were available, it would have been a simple task to arrive at the divergence cases. In the absence of such a grammar, the grammar rules corresponding to exceptional cases and cases which can not be captured by linguistic generalisations, were all checked for corresponding vibhaktis in Hindi. Leaving aside those cases where vibhakti is the same in Sanskrit and Hindi and taking up those cases where it diverges, we found that the cases of divergences may be classified into seven types, such as -

1. **Optional Divergence :** Vibhaktis optionally found in Sanskrit but absent in Hindi. For instance, in Sanskrit the karma[2] of the verb is expressed by

[2] Roughly equivalent to the notion of *object* in western linguistics

the default second case suffix and optionally the fourth case suffix (bālakāḥ *vidyālayāya/vidyālayaṃ* gacchanti) but in Hindi it takes only the default second case suffix (bālaka *vidyālaya* jāte_haiṁ = 'Boys go to school').

2. **Exceptional Divergence :** Sanskrit has certain exceptional rules which block the default suffixes but Hindi uses only the default suffixes. For example, in Sanskrit the karma of the verb exceptionally takes the sixth case suffix (śakuniḥ *śatasya* dīvyati) but in Hindi it takes the default second case suffix (śakuni *sau_rupae* jītatā_hai = 'Shakuni wins hundred rupees').
3. **Differential Divergence :** Sanskrit and Hindi use different nominal suffixes, for example, in Sanskrit, the person against whom the feeling of hatred *druha* is directed takes the fourth case suffix (durjanāḥ *sajjanāya* druhyanti) while in Hindi it takes the fifth case suffix (durjana *sajjana_se* droha karte_haiṁ = 'The wicked hate the good').
4. **Alternative Divergence :** Sanskrit uses more than one vibhaktis but Hindi takes only few among them. For instance, in Sanskrit alternately the sixth as well as the seventh case suffix is used after a word in conjunction with *āyukta* (āyuktaḥ *haripūjane/haripūjanasya*) but in Hindi only the seventh case suffix is applicable (hari_kī_ *pūjā_meṁ* līna = 'Deeply absorbed in the worship of Hari').
5. **Non-kāraka Divergence :** Divergences at the level of non-kāraka nominal suffixes, such as upapada vibhaktis, sambandha vibhaktis, etc. For instance, in Sanskrit the word takes fifth case suffix when governed by the karmapravachanīya *prati* (pradyumnaḥ *kṛṣṇāt* prati asti) while in Hindi it takes the sixth case suffix (pradyumna *kṛṣṇa_ke* pratinidhi hai = 'Pradyumna is the representative of Krishna').
6. **Verbal Divergence :** Divergences due to the special demand of certain verbs. For example, the karma of the verb *ā+ruh* in Sanskrit (vānaraḥ *vṛkṣam* ārohati) takes the seventh case suffix in Hindi (bandara *peṛa_para* caṛhatā_hai = 'Monkey climbs on the tree').
7. **Complex-Predicate Divergence :** This divergence results when a Sanskrit verb is mapped to a complex predicate in Hindi. For instance, the karma of the verb *anu+sṛ* in Sanskrit (*rāmam* anusarati sītā) is expressed by the genitive case in Hindi (*rāma_kā* sītā anusaraṇa karatī_hai = 'Sita follows Rama').

These cases are elaborated in detail.

2.1 Optional Divergence

When a sūtra assigns optionally two different vibhaktis in Sanskrit but Hindi allows only one vibhakti, we term the resulting divergence as *Optional Divergence*.

Illustration:
divaḥ karma ca (sādhakatamam and kārake) (P-1.4.43) That kāraka which is supplemental in the accomplishment of the action of the verbal root

diva (to play) is termed as karma, in addition to karaṇa. e.g.,

(1) San: *akṣān/akṣaiḥ* dīvyati
Hin: *pāsoṁ_se* khelatā_hai
gloss: dice_with plays
Eng: '(He) plays with dice.'

In the above example (1), *akṣa* is the instrument of the verbal root *diva* expressed through the third case suffix derived from the general rule *sādhakatamaṃ karaṇam* (P-1.4.42) but due to the optional rule *divaḥ karma ca*, *akṣa* optionally takes the second case suffix. Whereas in Hindi this optional rule does not apply and thus Hindi allows only the third case suffix.

Table 1 lists the sūtras which account for an additional optional vibhakti in Sanskrit either on account of imposed kāraka (Nos.1-3) or on account of special vibhakti assignment. In all these cases Hindi allows only the default kāraka and hence the default vibhakti.

Sr.No.	**Sūtra**		**Default**		**Optional**	
			Kāraka	**Vibhakti**	**Kāraka**	**Vibhakti**
1	1.4.43	divaḥ karma ca	karaṇa	3	karma	2
2	1.4.44	parikrayaṇe.	karaṇa	3	sampradāna	4
3	1.4.53	hṛkronyatarasyām	kartā	3	karma	2
4	2.3.12	gatyarthakarmaṇi.	karma	2	-	4
5	2.3.22	sañjño'nyatarasyāṃ.	karma	2	-	3
6	2.3.31	enapā dvitīyā	-	6	-	2
7	2.3.59	vibhāṣopasarge	karma	2	-	6
8	2.3.71	kṛtyānāṃ karttari.	kartā	3	-	6

Table 1. *Optional Divergence*

2.2 Exceptional Divergence

When sūtras give exceptional rules for cases by restricting the general rules but these rules are not applicable in Hindi, the divergence is termed as *Exceptional Divergence*.

Illustration:
adhiśīṅsthāsāṃ karma (ādhāraḥ and kārake) (P-1.4.46) That kāraka which is the locus of the verbal roots *śīṅ* (to lie down), *sthā* (to stand), *āsa* (to

sit), when preceded by the upasarga *adhi* is termed as karma. e.g.,

(2) San: bālakaḥ *paryaṅkam* adhiśete
Hin: laḍakā *palaṃga_para* sotā_hai
gloss: boy bed_on sleeps
Eng: 'Boy sleeps on the bed'.

In this example, the locus *paryaṅka* of the verbal root *adhi + śīṅ* is termed as karma and thus takes the second case suffix. But this rule does not apply in Hindi. So *paryaṅka* which is the locus and hence adhikaraṇa by the rule *ādhāro'dhikaraṇam* (P-1.4.45) takes the seventh case suffix. This applies to other verbs *sthā and āsa* with prefix adhi as well.

Table 2 lists the sūtras which account for non-default vibhaktis in Sanskrit by imposing a new kāraka (as in Nos.1-3) or by imposing an altogether different vibhakti for the kāraka. In all these cases, Hindi however takes the default kāraka and hence default vibhakti.

Sr.No.	**Sūtra**		**Default**		**Exceptional**	
			Kāraka	**Vibhakti**	**Kāraka**	**Vibhakti**
1	1.4.46	adhiśīṅsthāsāṃ karma	adhikaraṇa	7	karma	2
2	1.4.47	abhiniviśaśca	adhikaraṇa	7	karma	2
3	1.4.48	upānvadhyāṅvasaḥ	adhikaraṇa	7	karma	2
4	2.3.58	divaḥ tadarthasya	karma	2	-	6
5	2.3.61	preṣyabruvorhaviṣo	karma	2	-	6
6	2.3.64	kṛtvo'rthaprayōge	adhikaraṇa	7	-	6

Table 2. *Exceptional Divergence*

2.3 Differential Divergence

All those cases where Pāṇini's rule assigns a certain vibhakti either through kāraka or by direct case assignment and Hindi uses altogether different vibhakti are termed as the case of *Differential Divergence.*

Illustration:
apavargē tṛtīyā (kālādhvanoḥ atyantasaṁyoge) (P-2.3.6) The third case-suffix is employed after the words denoting the duration of time or place (*adhvan*), when accomplishment (*apavarga*) of the desired object is meant to be expressed. e.g.,

(3) San: *ahnā* anuvākaḥ adhītaḥ
Hin: *dinabhara_meṁ* anuvāka (veda kā eka bhāga) paṛha_liyā
gloss: whole day_in anuvaka (veda of one part) read
Eng: 'Anuvaka was perseverely and effectually studied (by him) in one day'.

In example (3), the person not only studied *anuvāka* but completely understood and memorised them. Attainment of knowledge is the fruitful result and hence the word *ahnā* which denotes the duration of time takes the third case-suffix. But in Hindi it takes the seventh case-suffix.

Table 3 lists the sūtras which describe cases where Hindi vibhakti differs from the Sanskrit vibhakti.

Sr.No.	**Sūtra**		**Sanskrit**		**Hindi**
			Kāraka	**Vibhakti**	**Vibhakti**
1	1.4.34	ślāghahnuṇsthā.	sampradāna	4	2,4,6
2	1.4.35	dhāreruttamarṇaḥ	sampradāna	4	6
3	1.4.37	krudhadruherṣyāsūyā.	sampradāna	4	5,7
4	1.4.38	krudhadruhorupasṛṣṭa.	karma	2	5,7
5	1.4.39	rādhīkṣyoryasya.	sampradāna	4	6
6	2.3.6	apavargē tṛtīyā	-	3	7
7	2.3.17	manyakarmaṇyanādare.	karma	2,4	6
8	2.3.23	hetau	-	3	4
9	2.3.37	yasya ca bhāvena.	-	7	6
10	2.3.43	sādhunipuṇābhyām.	-	7	4
11	2.3.67	ktasya ca varttamāne	-	6	3

Table 3. *Differential Divergence*

2.4 Alternative Divergence

In case of *Alternative Divergence* Sanskrit allows more than one case suffixes and Hindi takes only a few of them, and rarely an altogether different case suffix.

Illustration:
dūrāntikārthaiḥ ṣaṣṭhyanyatarasyām (pañcamī) (P-2.3.34) When in conjunction with *dūra* (distant,far), and *antika* (near) and their synonyms, the sixth case suffix is optionally employed and alternately the fifth. e.g.,

(4) San: *grāmasya/grāmāt* vanaṃ dūram asti

Hin: *gānva_se* jaṅgala dūra_hai
gloss: village_from forest far is
Eng: 'Forest is far from the village'.

In example (4), in Sanskrit *grāma* takes optionally the sixth case suffix as well as the fifth when combined with *dūra* as well as *antika* but Hindi takes only the fifth case suffix.

Table 4 lists the sūtras where Sanskrit takes more than one vibhaktis but Hindi takes only a few of them.

Sr.No.	**Sūtra**		**Sanskrit**	**Hindi**
			Vibhakti	**Vibhakti**
1	2.3.32	pṛthagvinānānā.	2,3,5	5,6
2	2.3.34	dūrāntikārthaiḥ.	5,6	5
3	2.3.36	saptamyadhikaraṇe ca	2,3,5,7	0
4	2.3.39	svāmīśvarādhipati.	6,7	6
5	2.3.40	āyuktakuśalābhyāṃ.	6,7	7
6	2.3.41	yataśca nirdhāraṇam	6,7	7
7	2.3.44	prasitotsukābhyāṁ.	3,7	2,7
8	2.3.72	tulyārthairatulopamābhyāṁ.	3,6	6
9	2.3.73	caturthī cāśiṣyāyuṣya.	4,6	2,6

Table 4. *Alternative Divergence*

2.5 Non-kāraka Divergence

Certain words known as *upapada* demand specific vibhaktis called *Upapada Vibhaktis* for nouns with which they combine. These are all typically exceptions to the *ṣaṣṭhī śeṣe* (P-2.3.50). In Hindi, however, in most of the cases sixth case suffix is used.

Illustration:
sahayukte apradhāne (tṛtīyā) (P-2.3.19) The word denoting *apradhāna* in conjunction with *saha* and its synonyms takes the third case suffix. e.g.,

(5) San: *rāmeṇa* saha sītā vanaṁ gacchati
Hin: *rāma_ke* sātha sītā vana jātī hai
gloss: rama_with sita forest goes
Eng: 'Sita goes to forest with Rama'.

In example (5), in Sanskrit it is seen that *sītā and vana* are directly related to the verb in the form of kartā and karma but *rāma* is not directly related to the verb and hence is called *apradhāna*. The association of *rāma* with *sītā* is denoted by the word *saha* due to which *rāma* takes the third case suffix from the above rule. But in the case of Hindi, *rāma* in conjunction with *saha* takes only the sixth case suffix.

Table 5 lists the sūtras where Sanskrit generally takes the second case suffix barring the sixth case suffix but Hindi does not take this exemption into account and takes the sixth case suffix.

Sr.No.	**Sūtra**		**Sanskrit**	**Hindi**
			Vibhakti	**Vibhakti**
1	1.4.84	tṛtīyārthe	2	6
2	1.4.85	hīne	2	6
3	1.4.87	upo'dhike ca	2	6
4	1.4.89	lakṣaṇetthambhūtā.	2	6
5	1.4.90	abhirabhāge	2	6
6	1.4.95	atiratikramaṇe ca	2	5,6
7	2.3.4	antarā'ntareṇayukte	2	6
8	2.3.5	kālādhvanoratyantasanyoge	2	7
9	2.3.10	pañcamyapāṅparibhiḥ	5	2
10	2.3.11	pratinidhipratidāne ca yasmāt	5	6
11	2.3.16	namaḥ svastisvāhā.	4	2,4,6
12	2.3.19	sahayukte apradhāne	3	6
13	2.3.26	anyārāditarartedik.	5	5,6

Table 5. *Non-kāraka Divergence*

2.6 Hindi Specific Divergences

All the divergences we covered so far were specific to the Aṣṭādhyāyī, and some of them might be attributed to the characteristics of Sanskrit. The default vibhakti of karma in Sanskrit is dvitīyā. But in Hindi, the karma takes vibhakti other than dvitīyā. These divergences may be attributed to the nature of Hindi. On the face of it divergences may be classified into two classes depending upon whether the Sanskrit verb maps to a simple Hindi verb or a complex predicate.

(a) Verbal Divergence
Consider the example,

(6) San: vānaraḥ *vṛkṣam* ārohati
Hin: bandara *peṛa_para* caṛhatā_hai
gloss: monkey tree_on climbs
Eng: 'Monkey climbs on the tree'.

Here, the karma of the verbal root *ā + ruh* is *vṛkṣa* whereas in Hindi the karma *vṛkṣa* takes seventh case suffix. The important question from Hindi grammar point of view then is - whether vṛkṣa should be termed as a karma as in Sanskrit or should it be termed as an adhikaraṇa? If we term it as a karma, then the treatment will be close to Sanskrit but one needs to impose a rule accounting for the divergence in the suffix. If we term it as an adhikaraṇa, then we deviate from the Sanskrit grammar but can take an advantage of the default vibhakti mapping.
There are many such instances, and these should be treated carefully by Sanskrit scholars who know Hindi well, taking insights from how Pāṇini handled such cases.

(b) Complex Predicate Divergence
If a verb in Sanskrit maps to a complex predicate[3] in Hindi, then the karma of the verb takes the sixth case suffix. e.g.,

(7) San: *rāmam* anusarati sītā
Hin: *rāma_kā* sītā anusaraṇa karatī_hai
gloss: rama_of sita follows
Eng: 'Sita follows Rama'.

Here in example (7), *rāma* which is the karma is expressed with the genitive case in Hindi. This change is a systematic one which may be explained through the rule *kartṛkarmaṇoḥ kṛti* (P-2.3.65). In Hindi *anu + sṛ* is translated as a complex predicate *anusaraṇa_karanā* where *anusaraṇa* is the kṛdanta and *karanā* is the main verb.

It is necessary to study these divergences in karma vibhakti in detail. Sanskrit has approximately 2000 verbal roots and many more with the *upasargas*. Mapping the corresponding verb-frames in Sanskrit to the Hindi verb-frames may give some hints towards these divergences. Though it is a voluminous task, any MT system requires this study.

3 Conclusion

The cases of divergences may be summarised then as:
(A) Divergences originating from Sanskrit Grammar (as shown below)

[3] A complex predicate consists of a noun followed by a light verb such as karanā, honā, etc. e.g. vivāha karanā, snāna karanā, etc.

Sr.No.	Cases	Sanskrit	Hindi
1	Optional	Sanskrit uses optional vibhakti in addition to default vibhakti.	Hindi allows only the default vibhakti.
2	Exceptional	Sanskrit uses different vibhakti than the default (Pāṇini blocks the default vibhakti by treating it as an exceptional case).	Hindi uses the default vibhakti.
3	Differential	Pāṇini imposes certain kārakas or vibhaktis which cannot be captured through the semantic generalisations.	Hindi uses different vibhaktis.
4	Alternative	Sanskrit uses more than one alternative vibhakti.	Hindi uses only one of them.
5	Non-kāraka	Sanskrit uses more than one alternative vibhakti.	Hindi uses the sixth vibhakti.

(B) Divergences due to the idiosyncrasy of Hindi:

1. Special vibhakti expectancy of verbs
2. Complex Predicate

These seven cases of divergence may prove useful in resolving the ambiguities at the level of nominal suffixes in Machine Translation and may serve as a platform for looking at vibhakti level divergences between Sanskrit and other Modern Indian Languages.

Acknowledgements
This work is a part of the Sanskrit Consortium project entitled 'Development of Sanskrit computational tools and Sanskrit-Hindi Machine Translation system' sponsored by the Government of India. We gratefully acknowledge Prof.Lakshmi Bai for the discussions at various stages.

References

1. Acharya Ramshastri: *Saṁskṛta śikṣaṇa saraṇī.* Acharya Ramshastri Jñānapīṭha (Hindi) (1998)
2. Cardona, George: *Pāṇini and Pāṇinīyas on śeṣa Relations.* Kunjunni Raja Academy of Indological Research, Kochi (2007)
3. Dorr, Bonnie: *Classification of Machine Translation Divergence and a Proposed Solution.* Computational Linguistics 20(4) (1994)
4. Dave, S., Parikh, J., Bhattacharya, P.: *Interlingua Based English-Hindi Machine Translation and Language Divergence.* Journal of Machine Translation (JMT) 17 (2002)
5. Dwivedi, Kapil Deva: *Bhāṣā vijñāna evam Bhāṣā śāstra.* Viśvavidyālaya Prakāśana, Varanasi (2006)

6. Goyal, Pawan and Sinha, R. Mahesh K.: *Translation Divergence in English-Sanskrit-Hindi Language Pairs*. In Amba Kulkarni, and Gérard Huet, editors, Sanskrit Computational Linguistics, Proceedings of Third International Symposium, Hyderabad. Springer-Verlag LNAI 5406 (2009)
7. Mishra, Vimal and Mishra, R. B.: *Study of Example Based English to Sanskrit Machine Translation*. Journal of Research and Development in Comp Sc. And Engg Number 37, Jan-June (2008)
8. Sharma, Ram Rang and Sharma, Malti: *Vaiyākaraṇa Siddhānta Kaumudī of Śrī Bhaṭṭoji Dīkṣita (Kāraka Prakaraṇam)*. Bharatiya Vidya Prakashan, Varanasi-Delhi (Hindi) (1997)
9. Sharma, Rama Nath: *The Aṣṭādhyāyī of Pāṇini*. Vols 1-3. Munshiram Manoharlal, Delhi 2nd Ed (2002)
10. Vasu, Śrīśa Chandra: *The Siddhānta Kaumudī of Bhaṭṭoji Dīkṣita*. Vol I. Motilal Banarsidass, Delhi (2003)