

# The redevelopment of Indo-Aryan case systems from a lexical semantic perspective

Miriam Butt · Tafseer Ahmed

**Abstract** The original case system found in Sanskrit (Old Indo-Aryan) was lost in Middle Indo-Aryan and then reinvented in most of the modern New Indo-Aryan (NIA) languages. This paper suggests that: (1) a large factor in the redevelopment of the NIA case systems is the expression of systematic semantic contrasts; (2) the precise distribution of the newly innovated case markers can only be understood by taking their original spatial semantics into account and how this originally spatial semantics came to be used primarily for marking the core participants of a sentence (e.g., agents, patients, experiencers, recipients). Furthermore, given that case markers were not innovated all at once, but successively, we suggest a model in which already existing case markers block or compete with newer ones, thus giving rise to differing particular instantiations of one and the same originally spatial postposition across closely related languages.

**Keywords** Case · Indo-Aryan · Diachrony · Semantic factors

## 1 Introduction

The original case system found in Sanskrit (Old Indo-Aryan; OIA) was lost in Middle Indo-Aryan (MIA). Many of the modern New Indo-Aryan (NIA) languages now have a system of case markers (mostly case clitics) that is similar to the OIA one in many respects. This paper tries to understand this development and suggests that: (1) a large factor in the redevelopment of the NIA case systems is the expression of systematic semantic contrasts; (2) the precise distribution of the newly innovated

---

M. Butt (✉) · T. Ahmed

Department of Linguistics, University of Konstanz, 78462 Konstanz, Germany  
e-mail: [miriam.butt@uni-konstanz.de](mailto:miriam.butt@uni-konstanz.de)

T. Ahmed

e-mail: [tafseer@gmail.com](mailto:tafseer@gmail.com)

case markers can only be understood by taking their original spatial semantics into account and how this originally spatial semantics came to be used primarily for marking the core participants of a sentence (e.g., agents, patients, experiencers, recipients). Furthermore, given that case markers were not innovated all at once, but successively, we suggest a model in which already existing case markers block or compete with newer ones, thus giving rise to differing particular instantiations of one and the same originally spatial postposition across closely related languages.

Taking Urdu/Hindi as a case study, we suggest that rather than taking *syntactic* factors to be primary for our understanding of case, lexical and clausal *semantic* considerations need to be factored in.<sup>1</sup> Section 2 first takes a global look at what is known in general about case change and posits that rather than thinking about change in primarily syntactic terms, for example in terms of a syntactic system of case alignment, what is needed instead is a close look at the semantics of case. Section 3 then goes on to provide evidence that semantic uses of case, in particular, semantically motivated case alternations are characteristic of South Asian languages in general and have been a systematic part of the language since OIA. Since the original inflectional case markers of OIA have been lost, but the NIA languages also employ semantically motivated case, the paper then goes on to try to understand how the new Urdu/Hindi case markers were drawn into the system (Sect. 4).

In particular, we look at the purported development of the modern ergative *ne* from the OIA instrumental and the introduction of the modern dative/accusative form *ko*. A consideration of the actual available diachronic data leads us to pursue a hypothesis in which the modern ergative is linked to a form that also gave rise to datives in related sister languages. In Sects. 4 and 5, we propose that this development follows quite naturally if: (1) one takes the lexical semantic content of the case markers involved seriously; (2) one understands that the new case markers slot into existing systems of semantic contrasts that are marked unambiguously or more explicitly through the case markers. In particular, we also suggest that the new case markers of modern Urdu/Hindi were drawn into the system in two different ways: (1) the dative/accusative *ko*, for example, was the result of a language internal reanalysis of an originally spatial postposition; (2) the introduction of the ergative *ne* was the result of language contact. Furthermore, borrowing of the dative version of *ne* was blocked by the prior existence of *ko*.

## 2 Language change

There are two very well known empirical observations with respect to case and language change. One of these is the observation that languages seem to change how they *align* or group their grammatical relations, i.e., *accusative* languages can turn into *ergative* languages and vice versa (Sect. 2.2). The other observation is that languages can start out with case systems and then go one of two ways. Either they

---

<sup>1</sup> Urdu is an Indo-Aryan South Asian language spoken in Pakistan as a national language and in India as one of the 18 official languages. Urdu and Hindi, which is primarily spoken in India, are structurally almost identical. The main differences lie in the lexicon (Urdu has borrowed much from Arabic and Persian, Hindi is more Sanskrit-based) and some parts of the phoneme system.

lose the original case marking system, presumably due to processes of erosion or language contact and subsequent reanalysis (also see Barðdal 2009 for a usage-based account) and fail to innovate a new one (e.g., Germanic and Romance languages), compensating instead with other parts of the grammar like a more rigid word order or a more articulated article system. The other route is to begin to press new forms into service, replacing the old ones, though not necessarily in a one-to-one relationship.

## 2.1 Redevelopment versus loss of case markers

The historical development by which new case markers are drawn into the system can be studied closely with respect to the Indo-Aryan languages in South Asia. All of these languages are ultimately descended from (versions of) Sanskrit, which had an inflectional declensional case paradigm. This case paradigm was much like the one that is found in Sanskrit's sister language Latin. An example is given in (1). The numbering in the left-most column provides the numbers by which the case markers are referred to in Pāṇini's grammar of Sanskrit (Katre 1987; Böhtlingk 1839–1840), the right-most column approximates the function of the case markers by using the names coming out of the well-established Graeco-Roman tradition (cf. Butt 2006).

(1) Number	Declension	Western Name
1	<i>devas</i>	nominative
2	<i>devam</i>	accusative
3	<i>devena</i>	instrumental
4	<i>devāya</i>	dative
5	<i>devāt</i>	ablative
6	<i>devasya</i>	genitive
7	<i>deve</i>	locative

Declension of Sanskrit *deva*- 'god' (adapted from Blake 2001, 64)

Of the original complex case marking system in Sanskrit only an inflectional ending *-e* survives today on nouns ending in *-a* in modern Urdu/Hindi.<sup>2</sup> The ending marks the oblique form of these nouns (e.g., *laṛka* versus *laṛke* 'boy'). However, modern Urdu/Hindi has innovated a series of new case markers, among them: *ne* 'ergative', *se* 'instrumental', *ko* 'dative/accusative', *k-* 'genitive'. These are not part of the inflectional morphology, but rather have the morphophonological status of clitics (Butt and King 2004).<sup>3</sup>

<sup>2</sup> The precise forms of the cases differed across noun classes in Sanskrit and were furthermore also subject to phonological processes. The Sanskrit case markers collapsed successively, so that more and more distinctions were lost until only a nominative/direct vs. oblique distinction was left (e.g., see Sen 1973, 68).

<sup>3</sup> One reviewer asks about adpositions. Since the majority of NIA case markers are drawn from former postpositions, a diachronic study of the development of new case markers must necessarily include a look at adpositions. However, we do not generally consider adpositions to be part of NIA case marking systems: adpositions have a very different syntactic and semantic distribution. See Butt and King (2004) for a detailed study of Urdu/Hindi.

Given the differing histories of Indo-Aryan versus Germanic and Romance languages with respect to case, a question that naturally arises is why a difference should exist: why does one language family innovate new case markers, while others do not?

One answer to this question could involve something along the lines of Kiparsky's (1987, 1988, 1997, 2001) idea of linkers, which posits that there are three *linkers* which serve to identify the arguments of a clause: position, case and agreement. Crosslinguistically, languages seem to make use of one, two or all of these linkers and for most languages one can observe an interacting system. English, for example, uses both position and agreement to identify structural arguments (the object is to the right of the verb, the subject is generally first in the sentence and agrees with the verb). Urdu uses case and agreement, but not position to identify its structural arguments.

Kiparsky assumes that over the course of time languages can change with respect to the main linkers they use. A sample scenario with respect to English is explicitly addressed in Kiparsky (1997). However, the possibility of changing the emphasis on one linker over another as part of language change does not by itself explain why case-rich free word order SOV languages like Old English and Latin have mutated into case-poor, fairly fixed word order SVO languages while other case-rich free word order SOV languages like Sanskrit have mutated into different versions of case-rich free word order SOV languages (e.g., Hindi/Urdu, Marathi, Gujarati, Punjabi, Nepali).

In some recent work, Hewson and Bubenik (2006) suggest that a relevant parameter is whether the languages develop *prepositions* versus *postpositions*. Languages which develop prepositions use these instead of case (their analysis is based mainly on spatial relations). Languages which instead make use of postpositions tend to assimilate those into a new case system. Furthermore, when articles are developed, a reliance on case becomes less likely.

These observed correlations are interesting as they point to headedness being a relevant factor for the redevelopment of case marking. They also, like Kiparsky's linkers, point towards a distribution of labor within a larger system, as the connection between the rise of articles and the loss of case marking seems to suggest. However, not all languages whose case marking system is lost/eroded immediately set about compensating the loss with the rise of positional licensing or the introduction of an article system. With respect to Indo-Aryan, most of the case marking had been lost by the time of Middle Indo-Aryan (200 BCE–1100 CE), but new case markers were not introduced until well into the New Indo-Aryan period (1100 CE–Present). And, indeed, some languages today still make do with just the oblique/non-oblique distinction that is left from the Old Indo-Aryan (Sanskrit) case system. How can this be explained?

The questions being asked in this section are fundamental ones. We will suggest that the answer to the questions lies in understanding the *semantics* motivating case systems. Indeed, Hewson and Bubenik's work already points out that a division of labor between case and other parts of the grammar, such as articles and adpositions need not be purely structural (as in Kiparsky's system), but could have a semantic

basis. Before providing evidence and more discussion with respect to this proposal, we briefly turn to the question of ergativity.

## 2.2 Ergativity and the alignment hypothesis

Many of the modern Indo-Aryan languages are ergative.<sup>4</sup> The ancestor language Sanskrit, however, was not. The development of an ergative case is generally assumed to be indicative of a *shift* or *alignment change* between grammatical relations and case marking in the entire case system of the language. This case shift has been proposed for a number of languages, namely: Inuktitut (an Eskimo language), Polynesian and Indo-Iranian (for a summary and references, see Butt 2006). The idea is that languages make a fundamental structural difference as to how they group or *align* the grammatical relations subject and object. *Accusative* languages like English, German, Latin or Sanskrit are taken to consistently mark their subjects in opposition to their objects. *Ergative* languages, on the other hand, mark their intransitive subjects the same as their objects, keeping a special marker, the *ergative* for transitive, agentive subjects (Dixon 1979, 1994; Plank 1979). The case shift between these two types of systems is generally taken to be mediated by either a passive or possessive structure that has been reanalyzed as an active verbal form.

There are several problems with this view, wide-spread though it is. For one, the only language family involving ergativity for which there is an actual deep historical record is Indo-Iranian and in this branch: (1) the case systems over time do not conform to the simple alignment picture posited in the literature; (2) the actual historical record for Indo-Aryan (a branch of Indo-Iranian) does not confirm the simple view of a change in alignment, but implicates a complex set of factors, many of them semantic (Butt 2001; Haig 2008). For another, the accusative versus ergative alignment hypothesis does not take into account phenomena such as psych verbs, where the subjects are often marked by datives or genitives, the effects of differential object and subject marking (e.g., Aissen 1999, 2003), the interaction of case with modality and the fact that modern structural case markers such as the ergative and accusative are generally drawn from original spatial expressions (Sect. 4).

In sum, the accusative vs. ergative alignment hypothesis simply does not match up with any known historical record (again, see Butt 2006 for some discussion). We suggest that the accusative versus ergative alignment hypothesis is deeply flawed precisely because it does not consider the vast and complex semantic functional load of case markers and that an understanding of this complex semantic picture is fundamental for understanding the documented historical changes.

---

<sup>4</sup> That is, they display so-called split-ergative patterns by which the ergative generally appears with a certain tense/aspect or is confined to certain type of NP. However, as is shown below (Sect. 3.2), the full range of ergative use and distribution in Indo-Aryan is not captured by standard notions of split-ergativity.

### 3 Case alternations in South Asian languages

This section first outlines the types of case marking there are in South Asian languages (Sect. 3.1), then presents a number of typical semantically motivated case alternations in modern South Asian languages (Sect. 3.2) and compares this to the situation in Old Indo-Aryan (Sect. 3.4). This brief survey and comparison shows that semantically motivated case alternations are not only a robust part of the modern systems, but were a systematic part of Old Indo-Aryan as well.

#### 3.1 Types of case marking

South Asian languages are made up of the genetically unrelated Indo-Aryan, Dravidian, Tibeto-Burman, Munda and Dardic languages. The region in which South Asian languages are spoken covers Afghanistan, Pakistan, India, Nepal, Bangladesh, Sri Lanka, Tibet and Bhutan. Despite their genetic unrelatedness, South Asian languages are known to form a *Sprachbund* (Masica 1976, 1991), presumably due to millenia of language contact. Within this Sprachbund, there seem to be roughly three different types of case-marking strategies:

1. Simple nominative vs. oblique contrast (e.g., Balochi)
2. Marking on nouns plus complicated system of clitics on the verb (e.g., Kashmiri).<sup>5</sup>
3. More elaborate system of a mixture of morphological inflectional case marking and clitics (most languages, e.g., Urdu/Hindi, Nepali, Bengali, Malayalam).

Additionally, the languages discussed here are all characterized by relatively free word order (major constituents of a clause can be scrambled to produce information-structural effects) and there is rampant pro-drop. Agreement patterns are wild and varied (see Subbarao 1999; Butt and Deo 2001) and do not have much to do with the case systems, i.e., there is no discernible systematic correlation. The case markings include at least: ergative (not always), nominative (generally unmarked), accusative, dative, instrumental, genitive, ablative and some locatives. In particular, many of the Indo-Aryan languages employ an ergative case, but none of the Dravidian languages do. All of the languages are syntactically accusative (i.e., all subjects always pattern together with respect to control, anaphora, coordination, etc.), even the ones with an ergative case. Interestingly, despite differences in the particular type of case marking employed, all types of languages show evidence of semantic contrasts expressed by case markers.

##### 3.1.1 Nominative versus oblique contrast

Balochi is a language spoken in Pakistan (Balochistan) and neighboring Iran. Unlike many South Asian languages, Balochi has no gender marking and also only has a nominative vs. oblique case system to work with. Despite this seeming limitation,

---

<sup>5</sup> We do not explicitly provide examples for Kashmiri, which shows the usual kind of ergative split according to agency and the NP hierarchy (Sharma 2001).



A further well-known instance of DOM is governed by animacy: animate objects tend to be marked, inanimate objects unmarked (cf. Blake 2001; Aissen 2003). An example of this typical DOM is given in (6) from Marathi.

- (6) a. **ti keel k<sup>h</sup>aa-t-e**  
 she.Nom banana.Nom eat-Pres-3.Sg.F  
 ‘She eats a banana.’ (based on Blake 2001, 128) Marathi
- b. **ti ravi=laa c<sup>h</sup>aḷ-ḷ-a**  
 she.Nom Ravi=Acc torture-Pres-3.Sg.F  
 ‘She tortures Ravi.’ (based on Blake 2001, 128) Marathi

Semantically motivated case alternations are, of course, not just confined to objects. A range of examples with Differential Subject Marking (DSM) are also found in South Asian languages. These are particularly interesting as they are not as well understood as instances of DOM. A typical example involving the ergative to mark control over an action is shown in (7).

- (7) a. **ram k<sup>h</sup>ās-a**  
 Ram.M.Nom cough-Perf.M.Sg  
 ‘Ram coughed.’ (Tuite et al. 1985, 264) Hindi/Urdu
- b. **ram=ne k<sup>h</sup>ās-a**  
 Ram.M=Erg cough-Perf.M.Sg  
 ‘Ram coughed (purposefully).’ (Tuite et al. 1985, 264) Hindi/Urdu

In Hindi/Urdu this ergative/nominative alternation to mark control is confined to unergative verbs. The language also includes the more well-known alternation by which ergative versus nominative case marking is governed by tense/aspect—in the case of Hindi/Urdu by past/perfect morphology. Rather than providing an example for Hindi/Urdu, we here illustrate the same contrast for the closely related language Punjabi (8), which also includes a split according to person, by which first and second persons are not overtly marked with the ergative case, but third persons are (9)

- (8) a. **ram muṇḍiā=nū mar-da ε**  
 Ram.M.Sg boy.M.Pl=Acc hit-Pres.M.Sg be.Pres.3.Sg  
 ‘Ram is hitting the boys.’ Punjabi
- b. **ram=ne muṇḍiā=nū mar-ia si**  
 Ram.M.Sg=Erg boy.M.Pl=Acc hit-Past.M.Sg be.Past.3.Sg  
 ‘Ram has hit the boys.’ Punjabi
- (9) a. **mē kampuṭar bec<sup>h</sup>-ia**  
 I.F/M computer.M.Sg sell-Past.M.Sg  
 ‘I (male or female) sold the computer.’ Punjabi
- b. **o=ne kampuṭar bec<sup>h</sup>-ia**  
 Pron.3.Sg.F/M=Erg computer.M.Sg sell-Past.M.Sg  
 ‘He/She sold the computer.’ Punjabi



Both types of ergative versus nominative splits are well-known and well documented across a range of languages (cf. Silverstein 1976; Plank 1979). However, the DSM illustrated in (10) is not. The example shows an ergative/dative alternation that corresponds to a difference between desire and obligation. Note that the dative is underspecified in the sense that it can express both desire and obligation (Bashir 1999; Butt and King 2003).

- (10) a. **nadya=ne** zu ja-na hε  
Nadya.F=Erg zoo go-Inf be.Pres.3.Sg  
'Nadya wants to go to the zoo.' Urdu
- b. **nadya=ko** zu ja-na hε  
Nadya.F=Dat zoo go-Inf be.Pres.3.Sg  
'Nadya has/wants to go to the zoo.' Urdu

Indeed, the use of case to mark differences in modality is quite common in South Asian languages—some further examples from Bengali (an Indo-Aryan language, no gender distinctions and no ergative) and Malayalam (a Dravidian language, no ergative) are provided below.

- (11) a. **amma** kuṭṭiye aḍik'k'-aṇam  
mother.Nom child.Acc beat-want  
'Mother must beat the child.' Malayalam
- b. **ammak'k'ə** kuṭṭiye aḍik'k'-aṇam  
mother.Dat child.Acc beat-want  
'Mother wants to beat the child.' Malayalam
- (12) a. **avan** var-aam  
he.Nom come-may  
'He may come.' (possibility) (Butt et al. 2004) Malayalam
- b. **avanə** var-aam  
he.Dat come-may  
'He may come.' (permission) (Butt et al. 2004) Malayalam
- (13) a. **ami** toma=ke cai  
I.Nom you=Acc wants  
'I want you.' (Klaiman 1980, 279) Bengali
- b. **amar** toma=ke cai  
I.Gen you=Acc wants  
'I need you.' (Klaiman 1980, 279) Bengali

These and other types of DSM to express clausal semantic distinctions have yet to be investigated in some depth. In particular, they remain to be investigated within formal semantics.

### 3.3 Interim summary

To summarize then, several types of semantically motivated case alternations can be identified in modern South Asian languages, irrespective of the genetic type of the language, the particular morpho-syntactic realization of the case system and whether the language contains an ergative case or not. The semantics responsible for case alternations include well-known factors such as animacy, definiteness/specificity, the person hierarchy, tense/aspect, but also encompass phenomena that are less well understood with respect to case such as modality.

The small sample of case alternations presented in the previous section also showed that a particular case marker (i.e. the dative) is not always associated with the same range of semantics crosslinguistically. Rather, the particular meaning contributed to the clause appears to emerge out of the case system of the language—that is, datives, for example, seem to take on different ranges of functions crosslinguistically depending on what other case markers exist in the language and the system of semantic contrasts expressed by the language. In Sect. 4, the paper looks at a case study in Urdu/Hindi involving the introduction of the modern ergative, the instrumental and the dative/accusative in order to demonstrate how a given case marker can slot into the overall case system and take on differing functions, depending on its original semantic import and the currently existing case system.

Before turning to that, the remainder of this section takes a quick look at the situation in OIA in order to demonstrate that the modern semantically-motivated case alternations are actually an old part of the language, despite the fact that the individual case marking system itself was lost in MIA and new case markers were drawn into the system in NIA. The data from OIA in conjunction with NIA suggest that semantically-based case alternations are a very robust characteristic of the South Asian areal *Sprachbund* and this in turn partly explains why case-rich free word order SOV languages like Sanskrit have mutated into different versions of case-rich free word order SOV languages. In contrast, while semantically-motivated case alternations were also a part of Old English (cf. Allen 1995), this characteristic does not seem to have been of an areal nature, thus leading to the eventual loss of a rich case marking system.<sup>6</sup>

### 3.4 Looking back in time

A quick look at Sanskrit case marking patterns shows that the use of case alternations to express semantic distinctions was a systematic part of OIA. The data in this section provides just a piece of the larger picture, but should suffice for the purposes of this paper.

---

<sup>6</sup> One reviewer questions the use of “areal” here. The point is that Old English lived cheek to jowl with Celtic languages, as well as having contact to Romance. The situation continues with modern English. However, we have not seen claims about, or references to a *Sprachbund* consisting of Germanic, Celtic and Romance. Indeed, these language families continue to work very differently, in stark contrast to the genetically unrelated languages of South Asia, which function quite similarly in many respects.

### 3.4.1 Partitivity

Jamison (1976) provides a very nice discussion of the multiple uses of the accusative in Sanskrit and the older stage of the language, Vedic. From her discussion, it becomes clear that the accusative had semantic import (particularly with respect to causatives, see the next section), was sensitive to verb classes and participated in a case alternation that is known as the partitive alternation in languages like Finnish.<sup>7</sup>

- (14) a. pibā somam  
 drink.Imp soma.**Acc**  
 ‘Drink soma. (all of the quantity)’ Vedic  
 (Ṛgveda VIII.36.1, from Jamison 1976)
- b. pibā somasya  
 drink.Imp soma.**Gen**  
 ‘Drink (of) soma.’ (Ṛgveda VIII.37.1, from Jamison 1976) Vedic

As can be seen, when the accusative is used in (14a), the soma in question is drunk up entirely, while in (14b) the use of the genitive signals that only part of the soma was drunk.<sup>8</sup>

### 3.4.2 Case alternation in causatives

Causatives are another interesting area in which case alternations can be found. Sanskrit allowed for either an accusative causee as in (15a) or an instrumental causee as in (15b). The examples are taken from Speijer (Speijer, 1886, 49), who sees the difference as one of involvement of the causee. If the causee is merely the instrument by which the action is brought about, but is not involved/affected by the action, the instrumental is used. Otherwise the accusative is used.

- (15) a. mantrapūtam carum rājñīm prāśayat  
 consecrated.Acc porridge.Acc queen.Sg.**Acc** eat.Caus.Impf.3.Sg  
 munisattamaḥ  
 best-of-ascetic.Nom  
 ‘the best of ascetics made the queen eat a consecrated porridge.’  
 (Kathāsaritsāgar 9.10) Sanskrit
- b. tām śvabhiḥ khādayet rājā  
 Demon.F.Sg.Acc dog.Pl.**Inst** eat.Caus.Opt.3.Sg king.Nom  
 ‘Her the king should order to be devoured by dogs.’  
 (Mahābhārata 8.371) Sanskrit

Interestingly, the same type of alternation exists crosslinguistically in Romance languages and is found in NIA as well (Alsina and Joshi 1991). An example from Urdu/Hindi is given in (16).

<sup>7</sup> For a recent paper looking at case variation with respect to the OIA accusative, see Dahl (2009).

<sup>8</sup> *Soma* is a drink consumed as part of rituals.

- (16) a. saddaf=ne      masala      cak<sup>h</sup>-a  
 Saddaf.F=Erg    spice.M.Nom    taste-Perf.M.Sg  
 ‘Saddaf tasted the seasoning.’      Hindi/Urdu
- b. anjum=ne      saddaf=**ko**      masala      cak<sup>h</sup>-va-ya  
 Anjum.F=Erg    Saddaf.F=Acc    spice.M.Nom    taste-Caus-Perf.M.Sg  
 ‘Anjum had Saddaf taste the seasoning.’      Hindi/Urdu
- c. anjum=ne      saddaf=**se**      masala      cak<sup>h</sup>-va-ya  
 Anjum.F=Erg    Saddaf.F=Inst    spice.M.Nom    taste-Caus-Perf.M.Sg  
 ‘Anjum had the seasoning tasted by Saddaf.’      Hindi/Urdu

The difference between accusative and instrumental case has been analyzed as one of *affectedness* by Alsina and Joshi (1991) (also see Saksena 1980, for an in-depth discussion of the notion of affectedness and causees see Ackerman and Moore 1999): when the accusative is used, the causee is interpreted as being more involved/affected; when the instrumental is used, the causee is seen as being rather more incidental to the caused action.

In Urdu/Hindi this semantic alternation is found with just one verb class, which comprises just a handful of verbs. However, the fact that it existed in Sanskrit with the same type of ingestive verbs and now exists in Urdu/Hindi is interesting since it shows that the same type of alternation can be found in different stages of the language despite the fact that the original case marking of Sanskrit was lost in MIA and that NIA languages like Urdu/Hindi have drawn an entirely new set of case markers into the language. There thus seems to be something quite robust about this type of semantically motivated difference in causees.

### 3.4.3 Case alternations in Pāṇini

Beyond the examples presented above, Pāṇini’s grammar of Sanskrit mentions 23 possibilities of case alternations (Katre 1987; Böhtlingk 1839–1840). Some of these alternations have to do with formal reasons conditioned by the morphophonology of the language. Others are governed by lexical semantics and still others are clearly expressing semantic distinctions along the lines already illustrated for Old and New Indo Aryan.

For example, take Pāṇini’s Rule 2.3.12, which states that with verbs of motion, if the motion is not an actual motion, then only the accusative can be used. That is, if a person named Ram goes to a village, the village can be marked either accusative or dative. But if only one’s thoughts “go” towards a village, the dative cannot be used.

**Rule 2.3.12:** The Dative and Accusative are used for verbs of movement, but the dative cannot be used if motion is abstract.

It is beyond the scope of this paper to list/analyze all of the alternations described by Pāṇini, but it should be clear that Sanskrit (and the older stage Vedic) was a language in which case was sensitive with respect to both lexical and clausal semantics.

### 3.5 Middle Indo-Aryan

Given the fact that semantically motivated case alternations existed in OIA and that they exist in the modern languages, it would be interesting to know what the situation in MIA was, where the original Sanskrit case system eroded massively, leading to just a direct versus oblique contrast by NIA. The relevant time line and stages of the language are given in (17).

- (17) A. Old Indo-Aryan  
1200 BCE—600 BCE (Vedic)  
600 BCE—200 BCE (Epic and Classical Sanskrit)
- B. Middle Indo-Aryan (Aśokan inscriptions, Pāli, Prakrits,  
Apabhraṃśa—Avahaṭṭha)  
200 BCE—1100 CE
- C. New Indo-Aryan (Bengali, Hindi/Urdu, Marathi and other modern  
North Indian languages)  
1100 CE—Present

Serious work remains to be done on MIA in order to investigate whether and what types of semantically motivated case alternations there may have been. Given modern languages like Balochi, which have just a nominative vs. oblique contrast and which nevertheless manage to express semantically motivated case alternations (cf. (2) and (3)), it would seem to be entirely possible that even the version of MIA which had been reduced to just a nominative versus oblique contrast contained semantically motivated case alternations.

Indeed, there are already some indications that it did. For one, different verb classes required different subject marking ('desire' with Nominative/Direct, 'please' with Oblique; Peterson 1998, 100). For another, there was a split-ergative system in place which was governed by aspect (Peterson 1998). This in and of itself may not be so interesting as tense/aspect splits tend to be seen as a morphosyntactic part of the language. A type of this basic structural tense/aspect and NP-split (first and second arguments are always unmarked) can also be found in the MIA Niya texts (where the language is probably Gandhari), but Jamison (2000) argues that semantic/pragmatic factors such as humanness and whether the agent is of a high rank (like a king) play a role in determining whether third person agents are marked with an instrumental or are unmarked (absolute/nominative).

### 3.6 Summary

In sum, both OIA and NIA use semantically motivated case alternations as a robust and essential part of the language. The same is true of other modern, non Indo-Aryan languages and there is some evidence that MIA may also have used case to make semantic distinctions. Given that the original inflectional case system of Sanskrit was lost almost completely during MIA, the question of how and why the new case markers were innovated presents itself. The next section therefore takes a closer look at the historical situation with respect to some case markers in modern Urdu.

#### 4 Modern innovation

Perhaps the most discussed case in Urdu/Hindi has been the ergative *ne*. A close second is the accusative/dative *ko*, which has a fairly complex synchronic distribution. It marks indirect objects, causees, subjects of psych verbs, and in general animate and specific objects (cf. Butt and King 2004). The *ko* is generally traced to the Sanskrit locative noun *kákshhe* ‘armpit, side’ (Kellogg 1893, 13). The origin of the ergative *ne*, however, to date remains obscure despite a general belief in the literature to the contrary: the text-book knowledge in Dixon (1994) and Harris and Campbell (1995), for example, presents a picture in which the modern ergative clitic *ne* is directly related to the inflectional Sanskrit instrumental. This analysis was apparently first proposed by Trumpp (1872, 113) and despite being debunked immediately and roundly by Beames (1872–1879), it has continued to persist. As such it is worth summarizing the pertinent arguments here, even though they have already been presented elsewhere (Butt 2001).

The instrumental is supposed to have emerged as an ergative through a reanalysis of an originally “passive” structure as in (18) into an active one.

- (18) *evam-uk-tā*                      *tu*              *haṃsena*              *damayantī*  
 so-say-Part.Nom.Sg              then              goose.Inst.Sg              Damayanti.Nom.Sg.F  
 1. ‘Then Damayanti was spoken to like that by the goose.’  
 2. ‘Then the goose spoke to Damayanti thus.’ Sanskrit Nalopākhyāna I,30

Constructions as in (18) with the adjectival participle in *-ta* are indeed the ancestors of the modern Urdu and Punjabi past/perfect clauses as in (7), (8b) and (20b). Furthermore, constructions with this adjectival participle could already be interpreted as an active past tense form in Sanskrit (Speijer 1886, 255, 294; Bynon 2005).

However, while the Sanskrit adjectival participle is clearly the ancestor of the modern Urdu (and Punjabi) past/perfect morphology, the original Sanskrit inflectional instrumental *-ena* is not likely to be the ancestor of modern Urdu *ne* because it had already changed into *ē* by MIA (1100 CE) (Beames 1872–1879, 266–272, Kellogg 1893, 130–132, see Butt 2001 for a summary and discussion). And, indeed, this *ē* can still be found in modern Urdu as *-e*, an oblique marker of masculine nouns in *-a*. Masica (1991) gives the table in (19) for MIA. The endings are simply listed in their surface forms as they occur distributed over several nominal paradigms. The modern form *ne* only began appearing as of 1400 CE.

(19)

	Singular	Plural
Nominative	<i>-u, a, am̐</i>	<i>-a, aī</i>
Accusative	[same as Nominative]	
Instrumental	<i>-em̐, im̐, he, hi</i>	<i>-e(h)ī, ehi, ahī</i>
Ablative	<i>-hu, ahu, aho</i>	<i>-hū, ahū</i>
Genitive/Dative	<i>-ho, aho, ha, su, ssu</i>	<i>-na, hā</i>
Locative	<i>-i, hi, him̐</i>	<i>-hī</i>

The following examples from Old Hindi show that while a typical “ergative” pattern existed in the middle ages whereby the verb agreed with the object just in the presence of perfect morphology, there was no overt ergative marker on the subject. The modern form in (20) would be *kabir=ne*, cf. the modern Urdu example in (21b), where the verb morphology is perfect, the subject is ergative and the verb agrees with the object.

(20)	masi	kāgad	chū-yo	nahī	kalam
	ink.Nom	paper.M.Nom	touch-Perf.M.Sg	not	pen.F.Sg
	gahī	nahi	hāth	jāro	juga
	take.Perf.F.Sg	not	hand	four.Pl	age.Pl
	māhātma	jehi	<b>kabir</b>	jan-ā-yo	nāth
	glory.Nom	who.Sg.Acc	<b>Kabir.M.Obl</b>	know-Caus- Perf.M.Sg	lord.Nom

‘Kabir touched not ink nor paper, he took not pen in hand; He made known the lord to whom is glory in the four ages.’ Old Hindi  
(Kabir, Sakhi 183; Beames 1872, 269)

(21)	a.	<b>ram</b>	gari	<b>ccla-ta</b>	(hε)
		Ram.M.Sg.Nom	car.M.Sg.Nom	drive-Impf.M.Sg	be.Pres.3.Sg
		‘Ram drives a car.’			Urdu/Hindi
	b.	ram=ne	<b>gari</b>	<b>ccla-yi</b>	(hε)
		Ram.M.Sg=Erg	car.M.Sg.Nom	drive-Perf.M.F.Sg	be.Pres.3.Sg
		‘Ram has driven a/the car.’			Urdu/Hindi

With respect to morphophonological factors quite a bit more can be (and has been) said about the unlikelihood of Sanskrit instrumental *-ena* being the direct ancestor of modern Urdu/Hindi *ne* (see Butt 2001 for an overview). However, no work to date has explored the underlying semantics. The next sections therefore take a look at the semantic functions of the instrumental in Sanskrit, compare these to modern Urdu and the sister language Punjabi, and conclude that a direct connection between the original Sanskrit instrumental and the modern Urdu ergative is not likely on semantic grounds either.

#### 4.1 The instrumental

A rough overview of the functions of the Sanskrit instrumental is given in (22) (based on Pāṇini’s Grammar of Sanskrit).

#### (22) Functions of the Sanskrit Instrumental

Instruments

Agents with participials and nominalizations (e.g., as with the adjectival participial in *-ta* above)

Causes and indirect causees (‘sadness through/because of a girl’)

Lexically required by some verbs like ‘sacrifice’, ‘be occupied with’

Temporal (completion within a time span) and Spatial Expressions

Comparison

If the old Sanskrit instrumental was really reanalyzed as modern ergative *ne*, then what happened to all the other uses of the instrumental—why did only the agents of participials use become reanalyzed as the ergative *ne*? Recall from (16) that Urdu also has an instrumental case, namely *se*. This case marker is generally traced back to either Sanskrit *sam* ‘with’ or to the locative singular noun *sañge* ‘in attachment to’ (Kellogg 1893, 132). In the modern language, it encompasses the functions sketched in (23) (see Khan 2009 for more details).

(23) **Functions of the Urdu Instrumental *se***

Instruments

Agents of passives

Expressions of (dis)ability (‘Nadya cannot walk’)

Non-affected and indirect causees (cf. (16))

Comitative/Sociative (e.g., ‘speak with’)

Lexically required with certain verbs (‘love’, ‘see’)

Temporal and spatial expressions with the meaning of source (ablative)

Made of Material (‘made of steel’)

Comparison

Manner

A comparison of the Sanskrit and modern Urdu uses would suggest that in terms of semantics, Urdu *se* is the more likely successor of the original Sanskrit instrumental. However, even with respect to this, the picture is not as clear cut. Consider the closely related sister language Punjabi, which divides up the semantic space occupied by Urdu *se* among two different case markers (other languages like Sindhi make even further distinctions, using three different markers).

**Punjabi**

– *nal* ‘with’

Instruments

Comitative/Sociative (e.g., ‘speak with’)

Manner

Made of Material (‘made of steel’)

– *tō* ‘from’

Agents of passives

Expressions of (dis)ability

Non-affected and indirect causees

Temporal and spatial expressions with source meaning (ablative)

Comparison

Indeed, Old Urdu/Hindi also did not use just one marker for the range of meanings of modern *se*, but several different ones: *sō/sē/siti* vs. *te/t<sup>h</sup>e/t<sup>h</sup>ē*. According to Shirani (1987a), the distinctive use of these markers was probably due to language contact with Persian (the language of the court at the time): *te* (from MIA *tahī* ‘from there’, Oberlies 1998) was presumably a direct translation of Persian *az* ‘from’; the use of *se* a direct translation of Persian *ba* ‘with’. Modern Urdu has dispensed with this distinction, allowing *se* to take up a larger semantic space.



This historical development alone indicates that *se* cannot be related directly back to the original Sanskrit instrumental. The current range of uses of *se* are similar to that of the Sanskrit instrumental, but this is more likely due to the fact that certain types of spatial semantics (e.g., ‘with’, ‘at’) crosslinguistically give rise to certain types of case marking functions (instrument, comitative, etc.).

#### 4.2 The ergative

To summarize the results so far—a direct connection to the original Sanskrit instrumental cannot be established for either the modern ergative *ne* or the modern instrumental *se*. Rather, *se* originates from an originally spatial expression meaning ‘with’ (Beames 1872–1879; Kellogg 1893). So what about the origin of *ne*? Beames 1872–1879; Kellogg 1893 propose the participial form *lage* of the verb *lag* ‘stick to’ as a possible ancestor, but are not able to back up their claim convincingly.

A comparison with a sister language proved to be helpful in the previous section, so perhaps a look at closely related languages might also prove to be helpful with respect to understanding the origin of *ne*. A quick survey shows that the form *ne* is used as an ergative marker in Urdu/Hindi, Punjabi and Marathi. However, perhaps rather surprisingly, it is also used **both** as an ergative and a dative/accusative marker in Rajasthani (Allen 1960) and Gujarati (Tessitori 1914) (the actual forms are *ne* or *nai* and are also used as an ablative).

Similarly, in Haryani, a very similar form *nae* is used for ergative (25), dative (26) and locative (24) uses. Haryani is of particular interest because it was spoken in the environs of Delhi, which was one of the major residences of the Moghul court and which in turn was where the majority of Old Urdu was spoken and written in the middle ages.

- (24) sāāj<sup>h</sup>      **nae**  
 morning    **at**  
 ‘in the morning’ (Singh 1970, 180) Haryani
- (25) saḍ<sup>h</sup>      **nae**    buḍ<sup>h</sup>:iaa    ki    j<sup>h</sup>ūprīi    kii    kun    mae    laat    maaryi  
 Sadhu    **Erg**    old-lady    Gen    cottage    Gen    corner    in    leg    hit  
 ‘The Sadhu kicked the corner of the old lady’s cottage.’  
 (Singh 1970, 180) Haryani
- (26) yaah    bi    raam    pyaarii    **nae**    e            de    diye  
 this.Pl    too    Ram    Piyari    **Dat**    Particle    give    give.Imp  
 ‘Give these to Ram Piyari too.’ (Singh 1970, 180) Haryani

It is very likely that the Haryani form *nae* is closely related to the Urdu *ne*. Indeed, other work on this language cites the form as *ne* (Shirani, 1987b) and also notes the dual use as ergative vs. object marker, as in (27a). The modern Urdu equivalent of (27a) is (27b).

- (27) a. man=ne                      sahab=ne                      mar-a  
 Pron.1.Sg=Acc/Dat    Sahib.M.Sg=Erg    hit-Perf.M.Sg  
 ‘The Sahib hit me.’ (Shirani 1987b)                      Haryani
- b. muj<sup>h</sup>=ko                      sahab=ne                      mar-a  
 Pron.1.Sg=Acc/Dat    Sahib.M.Sg=Erg    hit-Perf.M.Sg  
 ‘The Sahib hit me.’                      Urdu

### 4.3 An dative/ergative connection

One interpretation of the data from Haryani, Rajasthani and Gujarati is that the modern ergative *ne* also must be analyzed as having a spatial origin. In particular, there would seem to be a significant connection to a dative or locative meaning. This idea can be found as far back as Beames (1872–1879, 270), who suggests that the modern ergative *ne* is related to a dative form *nē* and that this dative was used on subjects in a dialect of Hindi spoken in provinces adjacent to the Moghul court. He speculates further that during the reign of the Moghul Emperor Shah Jehan (1627–1658) a change in administrative policies led to an influx of Hindu administrators, who influenced the language of the court.<sup>9</sup>

In more recent times, Montaut (2003, 2006, 2009) independently also draws parallels to dative marking. For one, she claims that ergative and dative alignment types are similar and sees ergative alignment as a stative predication of localization. For another, she points out that the *n-* forms in several languages, such as Konkan and Bhili, have the spatial/dative meaning ‘to’. She finds the most convincing etymology for modern Urdu/Hindi *ne* to be the one advanced by Tessitori (1913, 1914), who traces it back to the Apabhramśa form *kaṇṇahī*, which in turn is related to the Sanskrit locative of ‘ear’, *karne*, and which was realized as *kanhaiN* (or *kanhai*, *kanhi*, *kanhali*, *kaṇi*) in Old Rajasthani and mostly meant ‘aside, near’.

Tessitori (1913) argues for a general process by which initial syllables with *k-* could be deleted, thus leading to the modern *n-* forms. Additionally, he goes through the Old Western Rajasthani texts available to him and shows that both the *k-* forms (e.g., *kanhi*) and the, presumably derived, newer *n-* forms (*naī*, *naī*) were used side by side with exactly the same meaning. He also shows that the contexts the *n-* forms are used in the newer texts are exactly the same as the contexts of the *k-* forms in the older texts.

In terms of sound change and the material involved, Tessitori’s reconstruction thus seems much more convincing than Beames’ idea that the modern *ne* should be connected up with the participial form *lage* of the verb *lag* ‘stick to’. The reason that Beames fixed on *lage* as the origin for *ne* is a semantic one as it is ‘used with a very wide range of meanings, and with great laxity of application,—as is natural from its meaning, which may, without violence, be diverted to many uses. Thus, in Old-Bengali it is used in the sense of “on account of”, “for”’ (Beames, 1872–1879, 265). That is, Beames sees the semantic etymology as being promising because one

<sup>9</sup> Note that Shah Jehan is otherwise better known for building the Taj Mahal.



The ablative and dative meanings are fairly close with verbs of asking, speaking, etc. in South Asian languages in general (Khan 2009, 81). Once an ablative meaning is established, a further path of development to an agentive meaning is quite reasonable via an agent as source metaphor.<sup>12</sup> Consider the German preposition *von* ‘from’ in its agentive use in (30).

- (30) Das Gesetz wurde von ihm umgesetzt.  
 the.N.Sg.Nom law.N.Sg.Nom was from he.M.Sg.Dat realized  
 ‘The law was brought into force by him.’ German

Let us therefore assume that Tessitori’s path of reanalysis (or something very akin to it) is correct. Namely, a MIA form is pressed into service as a locative postposition. This postposition is in turn pressed into service as a new case marker and its semantic domain encompasses dative and accusative as well as ablative and agentive (ergative) meanings.

We therefore have a new case marker *ne/nae* that can in principle function both as an ergative (agentive) and dative. This is the situation in (at least) Haryani, Gujarati and Rajasthani. However, this is not the situation in modern Urdu/Hindi, where the *ne* exists as an ergative, but not as a dative/accusative. The role of dative/accusative marking is instead filled by *ko* (cf. (4), (27), (29a)) and there are even parts of the language where *ne* and *ko* express a semantic alternation involving modality (DSM; (10), repeated here in (31)).

- (31) a. **nadya=ne** zu ja-na hε  
 Nadya.F=Erg zoo go-Inf be.Pres.3.Sg  
 ‘Nadya wants to go to the zoo.’ Urdu  
 b. **nadya=ko** zu ja-na hε  
 Nadya.F=Dat zoo go-Inf be.Pres.3.Sg  
 ‘Nadya has/wants to go to the zoo.’ Urdu

In order to understand how these differences in historical development could have come about, a look at the attested development of *ko* in Urdu is clearly necessary. This is done in the next subsection.

#### 4.4 Dative/accusative *ko*

In Urdu, the accusative/dative *ko* now marks recipients/goals in ditransitives, experiencers (subjects of psych verbs) and specific objects (see Ahmed 2006 for a comprehensive survey). Beames (1872–1879, §56) reconstructs the Urdu *ko* to the locative of Sanskrit *kaksha* ‘armpit, side’ → Old Hindi *kākha*, accusative *kākham* →

<sup>12</sup> The role of metaphor in language change is discussed in, for example, Sweetser (1991) for perception verbs, modals, conjunction and conditionals. For discussions involving case marking and metaphors, see Lakoff and Johnson (1980) and Stolz et al (2006). An in-depth discussion of the role of metaphors in the development of new case markers would lead us too far afield, but see Khan (2009) for some discussion with respect to South Asian patterns in general.

*kahũ* → *kõ* → *ko*. In this case, Beames' reconstruction is generally accepted in the literature.

Dative and accusative uses of *ko* can be identified as far back as 1200 CE. The writer Chand, for example, uses versions of *ko* as a dative goal ('give a gift to the Brahmins') as well as with the object of 'seek' (Beames 1872–1879, 254; Kellogg 1893, 130–131). The same is true of Baba Farid (1173–1266), who wrote in Multan (now in Pakistan). A dative use is shown in (32), a dative experiencer use in (33), and an accusative with 'seek' in (34).

- (32) *jindu kũ samj<sup>h</sup>ai*  
 life Acc/Dat teaches  
 '(it) teaches to life' (Verse 1, from Khan 2001, 142) Old Urdu/Punjabi
- (33) *farid mē janya duk<sup>h</sup> muj<sup>h</sup> ko*  
 Farid I know grief/pain I.Obl Acc/Dat  
 'Farid, I know I have grief . . . (lit. grief is to/at me)' Old Urdu/Punjabi  
 (Verse 81, from Khan 2001, 226)
- (34) *ḍ<sup>h</sup>onden diye suhag kũ*  
 seek give husband Acc/Dat  
 '(you) are seeking a husband . . .' Old Urdu/Punjabi  
 (Verse 114, from Khan 2001, 263)

The use of *kõ/kũ* (and various other spellings) to mark indirect and direct objects continued and spread over the centuries so that it was to be found very commonly, especially in the dative use, by about 1600 (Beames 1872–1879, 253–254).

Interestingly, around 1800 *ko* is also found systematically with directed motion verbs such as *cal* 'walk/go', *poāc* 'reach' and *ja* 'go' and alternates systematically with another new case marker *mē* 'in' so that *ko* is generally used with endpoints where there is no guarantee that they have been attained or if they are abstract. This is illustrated with respect to past vs. future tense (35), habituals (36), and concrete vs. abstract endpoints (37).

- (35) a. *ek vilayat mē poāce*  
 one city in reached  
 'reached a city' (Dehalvi 1804) Old Urdu
- b. *is manzil ko kab poāco-ge*  
 this destination Dat/Acc when reach.2-Fut.Pl  
 'When will (you) reach this destination?' (Dehalvi 1804) Old Urdu
- (36) a. *dili mē ga-ye*  
 Delhi in go-Perf.M.Pl  
 '(they) went to Delhi' (Dehalvi 1804) Old Urdu

- b. ek qafila sodagarō ka damiḡq  
 one caravan.M.Sg merchants Gen.M.Sg Damascus  
**ko** ja-ta t<sup>h</sup>a  
 Dat/Acc go-Impf.M.Sg be.Past.M.Sg  
 ‘A caravan of merchants used to go to Damascus.’ (Dehalvi 1804)  
 Old Urdu
- (37) a. ek goristan **mē** poāce  
 one graveyard in reached  
 ‘reached a graveyard’ (Dehalvi 1804) Old Urdu
- b. apne haq **ko** poāc kar  
 self right Dat/Acc reach having  
 ‘after having attained one’s right’ (Dehalvi 1804) Old Urdu

Providing an explanation for the development of all of the current uses of Urdu/Hindi *ko* is part of on-going work and far beyond the scope of this paper. With respect to this paper, the following aspects are of interest. For one, *ko* entered the language about 400 years before *ne*. The earliest occurrences of *ko* have been traced back to about 1200. In contrast, the first attested instances of *ne* date back to only about the 1600s (during the reign of Shah Jehan). For another, both *ne* and *ko* originated as spatial postpositions, with *ne* meaning ‘near’ as well as ‘from’ and *ko* being used for goals/recipients (dative) as well as abstract locations (dative experiencer) and unattained or abstract goals or endpoints, cf. the verb ‘seek’ in (34) and the semantic alternation with ‘in’. This semantic alternation is reminiscent of the one found in Sanskrit (Sect. 3.4.3), whereby the dative could only be used with concrete goals. Finally, *ne* and *ko* differ with respect to their history in Urdu/Hindi in that older forms of *ko*, for example Old Hindi *kākha*, can be found in older texts, but no older form of *ne* exists in older Hindi/Urdu texts. This points to a situation in which *ko* was brought in as a postposition and then slowly developed into a case marker within Urdu/Hindi, but *ne* was introduced into the language via contact with another language.

## 5 Case blocking and case spread

At this point, we summarize the diachronic and synchronic picture as it appears to us. OIA had an inflectional case marking system that was eroded over the centuries during MIA. A new set of postpositional markers was drawn into the system during MIA, some of which have developed into case markers (usually clitics) in NIA. The meaning ranges of the MIA postpositions was not generally restricted to just a narrow semantic domain. Rather, as is typical for adpositions in general, they showed a flexibility in meaning. This flexibility in meaning allowed one and the same form to fulfill more than one case marking function. For example, the Urdu instrumental *se* ‘with’ functions as an instrumental, a comitative, or ablative (cf. (23)). Another example is the form *ne*, for which dative/accusative, ablative, locative and agentive/ergative uses are attested in Old Western Rajasthani.

Which function each of these NIA case markers took on in each of the individual NIA languages seems to have been governed by language particular factors which still need to be explored in more detail. Balochi, for example, has not innovated any new case markers. Some languages like Rajasthani or Haryani use *ne* as both a dative/accusative and an ergative, other languages like Urdu use different forms to fulfill these functions (*ne* versus *ko*).

What is common to all South Asian languages, however, is the use of case marking to express semantically motivated alternations. This was the point made in Sect. 3 and it was made for OIA as well as NIA. The point of this observation is a significant one. We see the overall propensity of South Asian languages to express semantically-based case marking alternations as the primary reason to draw new case markers into the system. That is, a new case marker will develop in order to express a semantic contrast and will often display an intricate interaction with the lexical semantics of the verb.

Old Urdu/Hindi *ko* marked spatial goals, but also unattained or abstract endpoints and alternated systematically with *mē* ‘in’ in a manner reminiscent of a pattern that existed in Sanskrit. Similarly, modern Urdu/Hindi *ko* alternates with the instrumental *se* on a subtype of causatives in exactly the same manner as was found in Sanskrit (Sect. 3.4.2).

From our perspective, the available evidence points to a situation in which new case markers are drawn into the system so that semantic contrasts can be made more explicit by providing a marker which signals the alternation in an unambiguous manner. Urdu/Hindi *ko* was drawn into the system early on to mark goals, but also to mark unattained or abstract endpoints. Agents were not explicitly marked as such in Old Hindi/Urdu, but agency was marked indirectly via the ergative agreement pattern in some situations (transitives, verb perfect), as illustrated in (20). Our hypothesis is that when Old Urdu/Hindi came into contact with a language (or languages) such as Haryani, where the agent was marked explicitly, it was natural to adopt this explicit marking into the language. Given that the expression of systematic semantic contrasts via case marking is a natural part of South Asian languages, it was also natural to expand the domain of the new agentive/ergative marker to mark volitionality/control in examples such as (38), repeated from (7).

- (38) a. **ram**                      k<sup>h</sup>ās-a  
       Ram.M.Nom      cough-Perf.M.Sg  
       ‘Ram coughed.’ (Tuite et al. 1985, 264)                      Hindi/Urdu
- b. **ram=ne**                k<sup>h</sup>ās-a  
       Ram.M=Erg      cough-Perf.M.Sg  
       ‘Ram coughed (purposefully).’ (Tuite et al. 1985, 264)                      Hindi/Urdu

Montaut (2003, 2006, 2009) provides several sets of further examples in which the *ne* is used to contrast with other case markers in the system in order to express greater volitionality/responsibility (using the instrumental *se* rather than *ne* indicates





## 6 Conclusion

This paper has argued that semantic factors need to be given central stage in order to achieve a perspicuous understanding of the development of new case markers in NIA after the old inflectional case system of OIA was eroded during MIA. It was demonstrated that case markers have been used systematically to express a range of semantic contrasts (DCM; Differential Case Marking) throughout the attested history (Sect. 3), with some of the semantic contrasts still being expressed in much the same way today as they were over 2000 years ago, despite the fact that the original case system was lost.

Section 4 looked at the sources of some of the modern Urdu/Hindi case markers, concentrating on the ergative *ne* and the dative/accusative *ko* in particular. The sources for the modern case markers were found to be originally spatial terms with meanings such as ‘near’, ‘at’, ‘from’ and ‘with’. Overall, this paper has taken a *lexical semantic* approach to understanding case in South Asian languages. The modern case markers are seen as contributing independent lexical semantic information to the clause, thus triggering the semantically-based case alternations. The range of their functions and semantic import are seen to follow directly from the original spatial postpositional meaning. However, as discussed in Sect. 5, not every language necessarily uses the same inventory of case markers—case markers can be drawn in both by language-internal reanalysis of an existing spatial postposition or via language contact. While a range of semantic meanings and functions may in principle be available, a new case marker may be constrained by the range of functions already expressed by existing case markers. This is taken to be the case with Urdu/Hindi *ne*, which in contrast to some sister languages, only expresses agentivity. That is, it is only realized as an ergative, but not as dative/accusative.

There is a range of other case markers in Urdu/Hindi as well as in related Indo-Aryan languages that have not been discussed in this paper. Much detailed work still remains to be done in order to arrive at a complete understanding not only of how the modern NIA case marking systems were developed, but also of how they work synchronically (analyses and descriptions for many NIA remain woefully inadequate). This paper has suggested that in order to arrive at a more complete understanding of case in its synchronic use and its diachronic development, we need to focus on the semantics associated with case markers, particularly the structural case markers such as ergative, accusative and dative. The syntactic and semantic role of case needs to be understood in the context of the entire case system of the language, and semantically motivated case alternations provide an interesting window on the semantics of case. Such case alternations tend to be overlooked in grammars, descriptions and analyses of languages. This paper has tried to suggest that more attention should be directed towards identifying and analyzing such alternations crosslinguistically.

For example, the first evidence for the modern ergative *le* in Nepali comes from a text from 1389. Here, and in other subsequent texts, the ergative consistently appears with a verb meaning ‘receive’ and marks the receiver. It only spread later to mark control/agency on transitive subjects (Poudel 2008). Thus, again, the existing

diachronic evidence in Nepali points to a close semantic connection between ergatives and datives (also note that the Nepali *le* is probably related to the Marathi dative *lai*), which remains to be explored and understood in detail.

We conclude that we need to understand the semantics of case better, not just in Indo-Aryan, but crosslinguistically. Indeed, crosslinguistic support for our argument comes from recent work on Germanic, where the question of ergative versus accusative alignment does not arise, but where semantic factors have been neglected to the detriment of explanatory power, e.g., see Barðdal (2001).

**Acknowledgements** The work presented in this paper was conducted as part of project A24 on South Asian case of the Sonderforschungsbereich (SFB) 471 at the University of Konstanz, funded by the Deutsche Forschungsgemeinschaft (DFG). SFB 471 at the University of Konstanz. We are heavily indebted to work done within the project by Scott Grimm, Tikaram Poudel and Karin Schunk. We would also like to thank three anonymous reviewers for extremely valuable comments, as well the audiences of workshops/conferences/colloquia in Australia, Germany, England, India and the U.S.A., where parts of this work have been presented in various forms. In particular, we would like to thank Ashwini Deo, Helen de Hoop, Klaus von Heusinger, Aditi Lahiri and Frans Plank for stimulating discussions and Cathryn Donohue and Jóhanna Barðdal for making this paper happen.

## References

- Ackerman, F., & Moore, J. (1999). Syntagmatic and paradigmatic dimensions of causee encodings. *Linguistics and Philosophy*, 22(1), 1–44.
- Ahmed, T. (2006). Spatial, temporal and structural usages of Urdu *ko*. In M. Butt & T. H. King (Eds.), *On-line proceedings of the LFG06 conference*. Stanford: CSLI Publications. <http://csli-publications.stanford.edu/site/ONLN.shtml>.
- Aissen, J. (1999). Markedness and subject choice in Optimality Theory. *Natural Language and Linguistic Theory*, 17, 673–711.
- Aissen, J. (2003). Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory*, 21, 435–483.
- Allen, W. (1960). Notes on the Rājasthānī verb. *Indian Linguistics*, 21, 4–13.
- Allen, C. L. (1995). *Case marking and reanalysis: Grammatical relations from Old to Early Modern English*. Oxford: Oxford University Press.
- Alsina, A., & Joshi, S. (1991). Parameters in causative constructions. In L. Dobrin, L. Nichols, & R. Rodriguez (Eds.), *Papers from the 27th regional meeting of the Chicago Linguistic Society (CLS)* (pp. 1–16).
- Barðdal, J. (2001). *Case in Icelandic—a synchronic, diachronic and comparative approach*. Ph.D. thesis, Lund University.
- Barðdal, J. (2009). The development of case in Germanic. In J. Barðdal & S. Chelliah (Eds.), *The role of semantic, pragmatic and discourse factors in the development of case* (pp. 123–159). Amsterdam: John Benjamins.
- Bashir, E. (1999). The Urdu and Hindi ergative postposition *ne*: Its changing role in the grammar. In R. Singh (Ed.), *The yearbook of South Asian languages and linguistics* (pp. 11–36). New Delhi: Sage Publications.
- Beames, J. (1872–1879). *A comparative grammar of the modern Aryan languages of India*. Delhi: Munshiram Manoharlal (republished 1966).
- Blake, B. (2001). *Case*. Cambridge: Cambridge University Press.
- Böhtlingk, O. (1839–1840). *Pāṇini's Grammar*. Delhi: Motilal Banarsidass (republished in 1998).
- Butt, M. (2001). A reexamination of the accusative to ergative shift in Indo-Aryan. In M. Butt & T. H. King (Eds.), *Time over matter: Diachronic perspectives on morphosyntax* (pp. 105–141). Stanford: CSLI Publications.
- Butt, M. (2006). *Theories of case*. Cambridge: Cambridge University Press.
- Butt, M., & Deo, A. (2001). Ergativity in Indo-Aryan. *KURDICA*, 5(3). <http://www.cogsci.ed.ac.uk/>.

- Butt, M., & King, T. H. (2003). Case systems: Beyond structural distinctions. In E. Brandner & H. Zinsmeister (Eds.), *New perspectives on case theory* (pp. 53–87). Stanford: CSLI Publications.
- Butt, M., & King, T. H. (2004). The status of case. In V. Dayal & A. Mahajan (Eds.), *Clause structure in South Asian languages* (pp. 153–198). Berlin: Kluwer Academic Publishers.
- Butt, M., King, T. H., & Varghese, A. (2004). A computational treatment of differential case marking in Malayalam. In *Proceedings of the international conference on natural language processing (ICON) 2004*. Hyderabad.
- Bynon, T. (2005). Evidential, raised possessor and the historical source of the ergative construction in Indo-Iranian. *Transactions of the Philological Society*, 103(1), 1–72.
- Chatterji, S. K. (1926). *The origin and development of the Bengali literature* (Vol. II). Calcutta: D. Mehra, Rupa & Co (1975 edition).
- Dahl, E. (2009). Some semantic and pragmatic aspects of object alternation in early Vedic. In J. Barðdal & S. Chelliah (Eds.), *The role of semantic, pragmatic and discourse factors in the development of case* (pp. 23–55). Amsterdam: John Benjamins.
- Dehalvi, M. A. (1804). *Bagh-o-Bahaar*. Calcutta: Fort William College.
- Dixon, R. M. W. (1979). Ergativity. *Language*, 55, 59–138.
- Dixon, R. M. W. (1994). *Ergativity*. Cambridge: Cambridge University Press.
- Farrell, T. (1995). Fading ergativity? A study of ergativity in Balochi. In: D. C. Bennett, T. Bynon, & B. G. Hewitt (Eds.), *Subject, voice and ergativity* (pp. 218–243). London: School of Oriental and African Studies, University of London.
- Haig, G. L. (2008). *Alignment change in Iranian languages: A construction grammar approach*. Berlin: Mouton de Gruyter.
- Harris, A. C., & Campbell, L. (1995). *Historical syntax in cross-linguistic perspective*. Cambridge: Cambridge University Press.
- Hewson, J., & Bubenik, V. (2006). *From case to adposition: The development of configurational syntax in Indo-European languages*. Amsterdam: John Benjamins.
- Jamison, S. (1976). Functional ambiguity and syntactic change: The Sanskrit accusative. In *Papers from the parasession on diachronic syntax, 12th regional meeting of the Chicago Linguistic Society* (pp. 126–135).
- Jamison, S. (2000). Lurching towards ergativity: Expressions of agency in the Niya documents. *Bulletin of the School of Oriental and African Studies*, 63, 64–80.
- Katre, S. M. (1987). *Astādhyāyī of Pāṇini*. Delhi: Motilal Banarsidass (republished in 1989).
- Kellogg, S. H. (1893). *Grammar of the Hindi language* (2nd edn.). Delhi: Munshiram Manoharlal Publishers Pvt. Ltd. (reprinted 1990).
- Khan, M. A. (Ed.). (2001). *Aakhyaa baba farid ne*. Lahore: Punjabi Adabi Board.
- Khan, T. A. (2009). *Spatial expressions and case in South Asian languages*. Ph.D. thesis, University of Konstanz.
- Kiparsky, P. (1987). *Morphology and grammatical relations*. Stanford: Stanford University.
- Kiparsky, P. (1988). *Agreement and linking theory*. Stanford: Stanford University.
- Kiparsky, P. (1997). The rise of positional licensing. In A. van Kemenade & N. Vincent (Eds.), *Parameters of morphosyntactic change* (pp. 460–494). Cambridge: Cambridge University Press.
- Kiparsky, P. (2001). Structural case in Finnish. *Lingua*, 111, 315–376.
- Klaiman, M. H. (1980). Bengali dative subjects. *Lingua*, 51, 275–295.
- Lakoff, G., & Johnson, M. (1980). *Metaphors we live by*. Chicago: Chicago University Press.
- Masica, C. (1976). *Defining a linguistic area: South Asia*. Chicago: The University of Chicago Press.
- Masica, C. (1991). *The Indo-Aryan languages*. Cambridge: Cambridge University Press.
- Mohanan, T. (1994). *Argument structure in Hindi*. Stanford: CSLI Publications.
- Montaut, A. (2003). Oblique main arguments in Hindi/Urdu as localizing predications: Questioning the category of subject. In P. Bhaskarao & K. Subbarao (Eds.), *Non nominative subjects* (pp. 33–56). Amsterdam: John Benjamins.
- Montaut, A. (2006). The evolution of the tense-aspect system in Hindi/Urdu: The status of the ergative alignment. In M. Butt & T. H. King (Eds.), *Proceedings of the LFG06 conference* (pp. 365–385). Stanford: CSLI Publications.
- Montaut, A. (2009). Ergative and pre-ergative patterns in Indo-Aryan as predications of localization: A diachronic view of past and future systems. In A. Faithi (Ed.), *Language vitality in South Asia* (pp. 295–325). Aligarh: Aligarh Muslim University.
- Oberlies, T. (1998). *Historische Grammatik des Hindi: Die Genese seines morphologischen Systems aus dem Mittel- und Altindischen*. Verlag für Orientalische Fachpublikationen.

- Peterson, J. M. (1998). *Grammatical relations in Pāli and the emergence of ergativity in Indo-Aryan*. München: LINCOM Europa.
- Plank, F. (Ed.). (1979). *Ergativity: Towards a theory of grammatical relations*. New York, NY: Academic Press.
- Poudel, T. (2008). *Diachronic and semantic aspects of ergativity, talk held at a workshop 'A non-canonical perspective on case'*. Konstanz: University of Konstanz.
- Saksena, A. (1980). The affected agent. *Language*, 56(4), 812–826.
- Sen, S. K. (1973). *Proto-New Indo-Aryan*. Calcutta: Eastern Publishers.
- Sharma, D. (2001). Case clitics and person hierarchy effects. In P. Sells (Ed.), *Formal and empirical issues in Optimality Theoretic syntax* (pp. 225–256). Stanford: CSLI Publications.
- Shirani, M. (1987a). Sabras az mulla vajhi. In M. M. Shirani (Ed.), *Maqaalaat-E-Hafiz Mahmood Shirani (Papers of Hafiz Mahmood Shirani)*. Lahore: Majilis-e-Taraqqi-e-Adab.
- Shirani, M. (1987b). Urdu ki shakh haryani zaban mein talifat. In M. M. Shirani (Ed.), *Maqaalaate-Hafiz Mahmood Shirani (Papers of Hafiz Mahmood Shirani)*. Lahore: Majilis-e-Taraqqi-e-Adab.
- Silverstein, M. (1976). Hierarchy of features and ergativity. In R. M. W. Dixon (Ed.), *Grammatical categories in Australian languages* (pp. 112–171). Canberra: Australian Institute of Aboriginal Studies.
- Singh, J. D. (1970). *A descriptive grammar of Bangru*. Kurukshetra: Kurukshetra University Press.
- Speijer, J. S. (1886). *Sanskrit syntax*. Delhi: Motilal Banarsidas (republished 1973).
- Stolz, T., Stroh, C., & Urdze, A. (2006). *On comitatives and related categories: a typological study with special focus on the languages of Europe*. Berlin: Mouton de Gruyter.
- Subbarao, K. (1999). Agreement in South Asian languages and minimalist inquiries: The framework. In *Proceedings of the international symposium on language contact, convergence and typology in South Asia*. Tokyo: The Institute for the Study of Language and Cultures of Asia and Africa.
- Sweetser, E. (1991). *From etymology to pragmatics: Metaphorical and cultural aspects of semantic structure*. Cambridge: Cambridge University Press.
- Tessitori, L. (1913). On the origin of the dative and genitive postpositions in Gujarati and Marwari. *Journal of the Royal Asiatic Society of Great Britain and Ireland*, 553–567.
- Tessitori, L. (1914). Notes on the Grammar of the Old Western Rājasthāni with special reference to Apabhramśa and to Gujarā and Mārwaṛī. *Indian Antiquary*, XLIII, 181–216, 225–236, Chap. 3.
- Trumpp, E. (1872). *Grammar of the Sindhi language*. Osnabrück: Biblio Verlag (reprinted 1970).
- Tuite, K. J., Agha, A., & Graczyk, R. (1985). Agentivity, transitivity, and the question of active typology. In W. Eilfort, P. Kroeber, & K. Peterson (Eds.), *Papers from the parasession on causatives and agentivity at the 21st regional meeting of the Chicago Linguistic Society (CLS)* (pp. 252–270).