

Ergativity in Indo-Aryan

[Miriam Butt](#) and [Ashwini Deo](#)

[Universität Konstanz](#), [Stanford University](#)

Abstract

Many of the Indo-Aryan languages are characterized by morphological ergativity. This contribution briefly surveys the types of ergative patterns and the particular case markers that are employed in the Indo-Aryan languages of South Asia. The synchronic discussion is followed by a look at what has been proposed for the historical development of ergativity in these languages.

1. Ergative Patterns in Indo-Aryan

The ergative marker was first named as a special marker for subjects with reference to Caucasian languages such as Georgian (Dirr 1928). The same type of case marker had been noted for languages such as Basque and Greenlandic (Pott 1873), but was generally referred to as an "agentive nominative" in opposition to a "neutral nominative", i.e. what we simply call nominative or absolutive today. The semantic parameter of agentivity that had been noted consistently by the linguists of the last century in connection with the ergative has been replaced by a purely structural division in this century.

The standard formulation of the conception of ergativity goes back to Fillmore (1968). Plank (1979:4) concisely summarizes the basic idea as follows:

a. A grammatical pattern or process shows ergative alignment if it identifies intransitive subjects (Si) and transitive direct objects (dO) as opposed to transitive subjects (St).

b. It shows accusative alignment if it identifies Si and St as opposed to dO.

According to this idea, languages can be grouped into several types, based on the case marking displayed by subjects and objects. This is illustrated by the following table.

Clause Type	Language Type		
	Ergative	Accusative	Active
Transitive	Erg-Nom	Nom-Acc	Erg-Nom
Intransitive (Unaccusative)	Nom	Nom	Nom
Intransitive (Unergative)	Nom	Nom	Erg

As far as we know, all the Indo-Aryan languages which have an ergative case fall under the "active" type of language. We exemplify this on the basis of Urdu/Hindi. (Hindi is spoken in the northern part of India and is written with the Devanagari script. Urdu is spoken in Pakistan and the northern part of India and is written with the Arabic script. Despite these and some differences in vocabulary, the two languages are structurally identical.) (1) shows a transitive clause with an ergative, (2) an unaccusative intransitive where no ergative is possible, and (3) shows an unergative intransitive where the ergative is optional and generally signals a degree of agentivity.

(1) ram=ne gari cala-yi (hai)
 Ram.M.Sg=Erg car.F.Sg.Nom drive-Perf.F.Sg be.Pres.3.Sg
 'Ram has driven a/the car.' (Urdu/Hindi)

(2) ram/*ram=ne ga-ya
 Ram.M.Nom/Ram.M=Erg go-Perf.M.Sg
 'Ram went.' (Urdu/Hindi)

(3a) ram kHaNs-a
 Ram.M.Nom cough-Perf.M.Sg
 'Ram coughed.' (Urdu/Hindi)

(3b) ram=ne kHaNs-a
 Ram.M=Erg cough-Perf.M.Sg
 'Ram coughed (purposefully).' (Urdu/Hindi)

None of the Indo-Aryan languages exhibit "syntactic ergativity". Syntactic ergativity is found in some languages, such as the Australian language Dyirbal (Dixon 1994). Syntactically ergative languages encode the ergative pattern purely structurally. However, most languages are morphologically ergative in that pieces of the morphology serve to mark the ergative or active pattern. The Indo-Aryan languages fall under this class.

Agreement patterns vary from language to language. In Urdu/Hindi the verb can only agree with NPs which are direct arguments and which do not bear overt case marking. While most Indo-Aryan languages appear to follow this pattern by disallowing agreement with a non-nominative argument, some languages allow it. Nepali, for instance allows agreement with an ergative subject, while in Gujarati, the verb agrees with the direct object, whether accusative marked or unmarked, when the subject of the clause is ergative. Examples (4) and (5) illustrate the Nepali and

Gujarati patterns, respectively.

- (4) mai-le mero lugga dho-en
I=Erg I.M.Sg.Gen clothes.M.Pl.Nom wash-Perf.1.Sg
'I washed my clothes.' (Nepali)
- (5) ram-e gadi=ne jo-yi
Ram.M.Sg=Erg car.F.Sg=Acc see-Perf.F.Sg
'Ram has seen a/the car.' (Gujarati)

Ergative morphology in most Indo-Aryan languages furthermore shows a split along the lines of tense/aspect. We again illustrate this with an example from Urdu/Hindi. Here the ergative case marker *ne* is required by perfect verb morphology. The association of ergativity with perfect morphology is crosslinguistically well-established and is one of the factors that has contributed to the idea that ergative structures must arise out of passive constructions (see [section below](#)). However, not all Indo-Aryan languages are split-ergative. Assamese, for example, does not seem to exhibit such a split (Devi 1986).

- (6) ram gari cala-ta (hai)
Ram.M.Sg.Nom car.F.Sg.Nom drive-Impf.M.Sg be.Pres.3.Sg
'Ram drives a car.' (Urdu/Hindi)
- (7) ram=ne gari cala-yi (hai)
Ram.M.Sg=Erg car.F.Sg.Nom drive-Perf.F.Sg be.Pres.3.Sg
'Ram has driven a/the car.' (Urdu/Hindi)

Another very common split crosslinguistically is the so-called NP-split, whereby only a subset of the nominals may display ergative morphology. Urdu/Hindi does not display this kind of a split, but it can be found in the closely related language of Punjabi, for example. In Punjabi, the first and second person pronouns are not marked for ergativity, whereas third persons are (pronouns as well as nominals).

- (8a) maiN kamm kita
I.Nom work.M.Sg.Nom do.Past.M.Sg
'I did some/the work.' (Punjabi, based on Bhatia 1993)
- (8b) o=ne kamm kita
Pron.3.Sg=Erg work.M.Sg.Nom do.Past.M.Sg
'He/She did some/the work.' (Punjabi, based on Bhatia 1993)

ami sita=ke dekh-lam
I-Sg.Nom Sita.F.Sg.Acc see-Perf.1.Sg
'I saw Sita.'
(Bengali)

Ergative case markers in Indo-Aryan take a variety of forms. Here we show a selection and compare the ergative markers to dative markers of the same languages in order to show that some of the forms appear to have a marked similarity (this preshadows some of the discussion in the [next section](#)).

	Dative	Ergative
	(subjects and objects)	(subjects only)
Hindi/Urdu	ko	ne
Punjabi	nuN	ne
Sindhi	kHe	OBLIQUE INFLECTION
Gujarati	ne	-e
Marathi	la	ne/ni
Bengali	ke	NONE
Oriya	ku	NONE
Assamese	ko/no	-e
Nepali	lai	le

Some of the ergative markers are inflectional. These are prefixed with an "-". All the others presumably function as clitics, as in Urdu/Hindi (Butt and King 2001, Sharma 1999).

3. The Origin of Ergative Case

There are several theories as to the origin of the ergative in Indo-Aryan. Most of the discussions take Urdu/Hindi as a representative language. In this section, we give a short overview of the proposals, along with some commentary.

Ergative as Passive The early (Western) linguistic literature on South-Asian languages (18th-19th century) refers to the ergative alternatively as an *agentive* or *instrumental*. Because the ergative in many languages has connotations of agency and shares features with an instrumental, the ergative construction was first analyzed as a passive in many languages (see Trask 1979:390 for some discussion). However, this view soon became a minority view due to detailed language-specific work, which showed that more often than not, ergatives were subjects of active sentences.

Passive/Participle to Ergative With respect to language change, the connection to a passive forms the basis for a hypothesis that ergative constructions arise from former passive constructions via a reanalysis of the following type:

NP_{instr} NP_{nom} V > NP_{erg} NP_{nom} V (adapted from Garrett 1990:265)

The precise morphology involved on the verb was a *-ta* participle in Sanskrit which has either been lost or retained as a glide or an *-e* in most of the modern Indo-Aryan languages. The Sanskrit *-ta* participle finds its origin in the Proto Indo-European deverbal adjective in **-to-*. In Sanskrit, the *-ta* formed a deverbal participle which agreed with a noun. This participle had passive interpretation with transitive verbs but active interpretation with intransitives and verbs of motion (Garrett 1990:263, Speijer 1886:280). It is indubitably the case that the modern ergative patterns occur primarily in conjunction with the verbal morphology descended from the original Sanskrit *-ta*. However, the precise nature of the original Sanskrit participle and its modern descendents remain the subject of debate.

Despite the possible active interpretations of the participle, the dominant idea for the development of ergativity in modern Indo-Aryan languages is one which sees a passive construction as being reanalyzed as an ergative. It is this proposal which has become accepted as common wisdom, despite many dissenting voices (e.g., Beames 1872, Kellogg 1893, Klaiman 1978, Zakharyin 1979, Andersen 1986, Hock 1986). Consider, for example, the quote from Dixon (1994) where this hypothesis is presented as textbook knowledge (also see Harris and Campbell (1995:263)).

We might thus expect a split ergative system conditioned by aspect or tense, where the ergative is found in perfect aspect or past tense, to be likely to have a passive origin.

This is precisely what happened in the Indic and Iranian branches of Indo-European (for which we do have written records and can be fairly certain about what happened, although different scholars have suggested diverse interpretations). [Dixon 1994:190]

The dissenting voices mentioned by Dixon range from an argument that Sanskrit as well as modern Hindi were basically "patient-oriented" and thus should both be considered ergative (Hock 1986), to the interesting observation that stativity may be the relevant factor that is coming into play (Trask 1979:397) in the sense that a deverbal stative predicate is made active via an integration into the inflectional paradigm into the language (see Deo 2001a for a case study of Marathi).

Direct Descendent of the Sanskrit Instrumental -*ina* Both the proponents of the passive-to-ergative view and the dissenters (with the exception of Deo) assume that the modern Urdu/Hindi ergative *ne* is a direct descendant of the original Sanskrit inflectional instrumental -*ina* (and allomorphs thereof). However, the historical facts clearly speak against such an analysis.

Researchers of the last century such as Beames (1872-79) and Kellogg (1893), are very clear on the idea that the modern Urdu/Hindi *ne* could not possibly be a descendant of the Sanskrit instrumental -*ina*. Kellogg essentially lists three main problems with the hypothesis that the ergative *ne* be descended from the Sanskrit -*ina*: erosion, timing, and usage.

The highly inflected case system of Sanskrit underwent a general collapse over the ages and the case endings eroded and fell together. According to Sen (1973:68), the instrumental -*ina*/-*ena* eroded to *eN* by Middle Indo-Aryan and fell together with what was left of the dative: *e*. It is generally agreed (e.g., Sen 1973, Beames 1872-79, Kellogg 1893) that this *eN/e* furnished the current oblique marker of Urdu/Hindi. We may further assume that this morpheme is the ancestor of the inflectional ergative morphology in Assamese and Gujarati (see the [table](#) above). However, this morpheme could not have been the ancestor of the other ergative forms in the table.

To take modern Urdu/Hindi *ne* as an example again: this form is often described as a postposition in the literature (e.g., Davison 2000, Mahajan 1990). We follow Mohanan (1994) in treating it as a clitic (see Butt and King 2001 for a detailed discussion). Furthermore, as Kellogg also points out, in synchronic terms the ergative *ne* behaves much like other postpositions (or clitics) which are known to have developed from nouns: *meN* 'in' and *par* 'on'. The synchronic and diachronic data therefore seem to point to a relatively normal path of development: the instrumental -*ina* eroded away and the ergative *ne* came into the language as a grammaticalized form of a noun. On the other hand, the commonly assumed

development from the inflectional morpheme *-ina* to a clitic (or postposition) *ne* would need to involve degrammaticalization, a highly unusual form of historical development and one which cannot be substantiated by the known synchronic and diachronic facts.

Another problem with the *-ina* to *ne* hypothesis is the relatively late appearance of the ergative in High Hindi. Beames (1872-79:267-271) surveys Old-Hindi writers like Chand, Kabir, Tulsi Das and Behari Lal and finds that he cannot trace the ergative *ne* back to more than 200-300 years ago (1600-1700). The writers he surveyed tend to use the oblique form *-e* (the old instrumental) of nouns/pronouns in constructions that today would be termed "ergative". The question then arises, if an "ergative" pattern based on the old instrumental was already in place, why then introduce a new marker into the language?

Beames (1872-79:270) traces the modern ergative *ne* to a dative form *neN* that was used in a dialect of Hindi spoken in provinces adjacent to the Moghul court during the reign of the Moghul Emperor Shah Jehan (1627-1658). Beames sees this time period as a very likely one because a change in administrative policies led to an influx of Hindu administrators, who might have put their stamp on the language of the court. Beames does not say which dialect the dative *ne* could have been borrowed from.

The precise origin of the non-inflectional forms of the ergative in the [table](#) above thus remains to be determined.

Language Contact Zakharyin (1979) ascribes the Urdu/Hindi ergative form *ne* to language contact with Tibeto-Burman, who use an ergative form *na*. However, this hypothesis does not explain why Nepali, a language which is geographically very close to the Tibeto-Burman languages, would employ *le* as an ergative marker (Devyani Sharma, p.c., August 2000).

Historical Stability Butt 2001, on the other hand, proposes that *ne* might be a reduced form of the Sanskrit locative *janiye* 'for the sake of, because of, caused by' (based on a suggestion by Aditi Lahiri, p.c., December 1999). The semantics of this form are compatible with both agentive and dative (goal/benefactive) uses and as such this noun may have given rise to both dative and ergative forms in the [table](#) shown above. Butt further proposes that the ergative pattern in modern Urdu/Hindi is an instance of historical stability rather than an example of a radical accusative-to-ergative shift. This is based on the idea that the original Sanskrit -

ta participle already formed an "ergative" pattern in the sense that the logical subject was marked nominative with intransitives, inceptives and verbs of motion, but instrumental with all others. This pattern has simply been retained in the modern language, though instantiated through new case morphology. This idea is consonant with Hock's 1986 claim that both Sanskrit and Hindi were essentially patient oriented and that in terms of this basic property, no historical change has taken place.

Ergativity as Licensed by the Development of IP In contrast, Deo 2001a argues that the historical development of ergativity in modern Indo-Aryan languages can be explained by a cross-linguistically attested historical shift: the development of a more articulated phrase structure in the form of an IP. This study looks at syntactic and morphological data from Old Indo-Aryan, Middle Indo-Aryan and Marathi, one of the modern Indo-Aryan languages.

The analysis rests on the crucial assumption that loss of morphology in a language (or language family) triggers phrase structural changes. The introduction of functional categories in compensation for the reduced functionality of inflectional morphology has been argued to be a general tendency in Indo-European syntactic change and this has been discussed in detail for Middle English and Greek (Kiparsky 2000, Kiparsky 1996). The hypothesis that Indo-Aryan also conforms to a similar trajectory is supported by the following syntactic and morphological changes in the language family, exemplified by a representative language, Marathi.

- The verb 'be' which functioned as both the copula and the auxiliary verb, was optionally used in Sanskrit, the older language. This verb becomes obligatory in both these contexts. In other words, there is a general constraint in the language on the overt presence of a finite, inflected verb, corresponding with I or Infl.
- This development also affects negation patterns. The negative particle in Sanskrit, a free particle that may occur with or without an auxiliary, is affixed to the auxiliary (via an intermediate process of cliticization) in Marathi. The auxiliary is no longer optional in negated clauses, but obligatorily present. This change is shown in examples (10) and (11). Example (10) exemplifies a zero-copula negative clause in Sanskrit, while in example (11), we see the changed pattern in Old Marathi. The negation particle is an affix on the auxiliary, which is obligatorily present in a negated nominal clause.

(10) na shoko na jara tasyam
 NEG affliction.M.Nom.Sg NEG old age.F.Nom.Sg her.F.Loc.Sg

`There (is) no affliction or old age in her.'

(Sanskrit)

(11) shabdu bole mukhe pari dnyaan naahi
word.M.Nom.Sg. speak.Pres.M. 3.Sg. mouth.N.Inst.Sg but knowledge.N.Nom.Sg be.NEG.Pres. 3.Sg
`(He) speaks with his mouth, but there is no knowledge.' (Old Marathi)

•The participle, on which the modern ergative clause is based, was a nominal, participial category in Sanskrit, inflecting for number and gender. However, in the modern languages, this form is integrated into the finite verbal paradigm. In Marathi, this participle was incorporated into the general verbal paradigm of the language via two processes.

- The development of person marking on the verb in addition to gender and number marking.

- The stem morphology of the participial form changes so that it uses the allomorph used by the other members of the verbal paradigm.

This process of incorporation into the verbal paradigm of the language suggests a shift of the categorial status of the participial from nominal to verbal or its finitization. This process is extremely important for the ergative reanalysis of the Old Indo-Aryan participial clause and the corresponding emergence of morphological ergativity in the modern Indo-Aryan languages.

The facts discussed here constitute independent evidence for the assumed phrase structural change, development of the IP, in the Indo-Aryan languages. Most of these facts, viz. obligatory copula, changed negation patterns, and changes in stem morphology are attested in other Indo-Aryan languages such as Urdu/Hindi, Punjabi and Gujarati also.

- The introduction of a new functional category, the IP, also explains the emergence of the ergative as a structural case in structural terms. The projection of an IP creates spec, IP, a position that licenses subjects. The movement of oblique (instrumental), passive agents in this position allows them to be re-interpreted as subjects. In other words, these oblique noun phrases are licensed positionally, rather than by case, as arguments of the verb. This account, thus, structurally motivates the passive to ergative shift in Indo-Aryan. This change can also be fitted under the more general change in Indo-Aryan of licensing non-nominative subjects.

Data from Middle Indo-Aryan shows that there was an intermediate period of layering in Middle Indo-Aryan during which transitive clauses could show both passive and ergative properties (Bubenik 1996, Andersen 1986). For instance, it

has been observed that the ergative or the passive interpretation of a clause depends primarily on the word-order of the clause. These facts may be explained if it is assumed that the IP was not an obligatory projection in every clause in MIA, but could be projected in some clauses. In NIA, however, this clause is fully reinterpreted as an active ergative clause,

This analysis attempts to provide a unified account of the diachronic emergence of morphological ergativity in the Indo-Aryan languages, relating it to a number of other phrase structural changes in these languages across time.

4. Summary

The differing perspectives on the development of the ergative in Indo-Aryan sketched above show that there is no simple standard "textbook" story, as often suggested in the literature. The modern ergative did not come from a straightforward passive construction, but from a participle with nominal status. This participle was later integrated into the verbal paradigm.

The Sanskrit instrumental case marker is not the ancestor of the ergative in most of the modern Indo-Aryan languages and the true ancestor of modern forms such as *ne/ni* or *le* remains to be identified conclusively. In the modern South Asian languages, as in Sanskrit, the lexical semantics of a verb play a large role in the determination of case marking. As such, much more research needs to be done on the change or possible stability of the lexical semantics of verbs and verb classes (for some research on dative subjects in Marathi, see Deo 2001b). In addition, South Asian case marking systems need to be investigated more closely and the distribution of case markers such as dative, accusative, instrumental, or genitive in relation to the ergative (when it is present) should be taken into account. The Eastern Indo-Aryan languages raise a different kind of a problem: why was the original "ergative" pattern of Sanskrit abandoned? Does it have to do with the nature of the integration of the nominal participle into the verbal paradigm? Does the particular agreement morphology that was retained or generated play a role (the ergative language Marathi added person agreement, the accusative language Bengali lost gender agreement)? A detailed comparative investigation of several different modern Indo-Aryan languages would undoubtedly shed more light on these questions. Furthermore, comparative data from the Indo-Iranian branch might prove to be interesting, perhaps even crucial in understanding the structure and development of ergativity in Indo-Aryan better.

Bibliography

- Andersen, Paul Kent. 1986. Die *ta*-Partizipialkonstruktion bei Asoka: Passiv oder Ergativ? Zeitschrift für vergleichende Sprachforschung 99:75-95.
- Beames, John. 1872-79. A Comparative Grammar of the Modern Aryan Languages of India. Republished 1966. Delhi: Munshiram Manoharlal.
- Bhatia, Tej. 1993. Punjabi: A cognitive-descriptive grammar. London: Routledge.
- Bubenik, Vit. 1996. The Structure and Development of Middle Indo-Aryan Dialects. Delhi: Motilal Banarasidass.
- Butt, Miriam. 2001. A Reexamination of the Accusative to Ergative shift in Indo-Aryan. In M.Butt and T.H.King *Time over Matter: Diachronic Perspectives on Morphosyntax*, 105-141. Stanford, California: CSLI Publications.
- Butt, Miriam and Tracy Holloway King. 2001. The Status of Case. In Veneeta Dayal and Anoop Mahajan (eds.) *Clause Structure in South Asian Languages*. To appear.
- Davison, Alice. 2000. 'Dependent structural case' as a consequence of VP structure. *Texas Linguistics Forum* 42(2).
- Deo, Ashwini. 2001a. The *-ta* form in Indo-Aryan. Stanford University, Ms.
- Deo, Ashwini. 2001b. The Diachronic Origin of Dative Subjects. Stanford University, Ms.
- Devi, Jayantimala. 1986. Ergativity: A Historical Analysis in Assamese. PhD Thesis, University of Delhi.
- Dirr, Adolf. 1928. Einführung in das Studium der kaukasischen Sprachen. Leipzig: Verlag der Asia Major.
- Dixon, R. M. W. 1994. Ergativity. Cambridge: Cambridge University Press.
- Fillmore, Charles J. 1968. The case for case. In Emmon Bach and R.T. Harms (eds.) *Universals of Linguistic Theory*, 1-88. New York: Holt, Rinehart and Winston.
- Garrett, Andrew. 1990. The Origin of NP Split Ergativity. *Language* 66:261-296.
- Harris, Alice C. and Lyle Campbell. 1995. *Historical Syntax in Cross-Linguistic Perspective*. Cambridge: Cambridge University Press.
- Hock, Hans Henrich. 1986. *P-oriented Constructions in Sanskrit*. In Bh. Krishnamurti (ed.) *South Asian Languages: Structure, Convergence and Diglossia*. Delhi: Motilal Banarsidass.

- Kellogg, S. H. 1893. Grammar of the Hindi Language. Delhi: Munshiram Manoharlal Publishers Pvt. Ltd.. Second Edition, reprinted 1990.
- Kiparsky, Paul. 1996. The Shift to Head-Initial VP in Germanic. In Ans van Kemenade and Nigel Vincent *Studies in Comparative Germanic Syntax*, 140-179. Kluwer Academic Publishers.
- Kiparsky, Paul. 2000. From Passive to Ergative Shift in Indic. Handout, West Coast Conference of Linguistics, Fresno.
- Klaiman, M. H. 1978. Arguments Against a Passive Origin of the IA Ergative The Proceedings of the 14th Meeting of the Chicago Linguistic Society, 204-216.
- Mahajan, Anoop . 1990. The A/A-Bar Distinction and Movement Theory. PhD Thesis, MIT.
- Mohanar, Tara. 1994. Argument Structure in Hindi. Stanford, California: CSLI Publications.
- Plank, Frans. 1979. Ergativity, Syntactic Typology and Universal Grammar: Some past and present viewpoints. In Frans Plank (ed.) *Ergativity: Towards a Theory of Grammatical Relations*, 3-36. New York, New York: Academic Press.
- Pott, A.F. 1873. Unterschied eines transitiven und intransitiven nominativs. Beiträge zur vergleichenden Sprachforschung auf dem Gebiete der arischen, celtischen und slawischen Sprachen 7:71-94.
- Sen, Subhadra Kumar. 1973. Proto-New Indo-Aryan. Calcutta: Eastern Publishers.
- Sharma, Devyani. 1999. Nominal Clitics and Constructive Morphology in Hindi. In editor = Miriam Butt and Tracy Holloway King (eds.) *Online Proceedings of the LFG99 Conference*. Stanford, California: CSLI Publications. <http://csli-publications.stanford.edu/LFG/4/lfg99.html>.
- Speijer, J. S. 1886. Sanskrit Syntax. Republished 1973. Delhi: Motilal Banarsidas.
- Trask, R. L. 1979. On the Origins of Ergativity. In Frans Plank (ed.) *Ergativity: Towards a Theory of Grammatical Relations*, 385-404. New York, New York: Academic Press.
- Zakharyin, Boris. 1979. On the Formation of Ergativity in Indo-Aryan and Dardic. *Osmania Papers in Linguistics* 5:50-71.