KRZYSZTOF STROŃSKI
JOANNA TOKAJ
RAFAŁ JAWORSKI

DIACHRONY AND TYPOLOGY OF NON-FINITES IN INDO-ARYAN

Krzysztof Stroński
Joanna Tokaj
Rafał Jaworski

Diachrony and typology of non-finites in Indo-Aryan

Faculty of Modern Languages and Literatures
Adam Mickiewicz University in Poznań
2019

Reviewer
Liudmila Khokhlova
Moscow State University

This publication was supported by National Science Centre Poland grant
2013/10/M/HS2/00553

ISBN 978-83-954144-4-2
https://www.doi.org/10.14746/9788395414442

CONTENTS

FIGURES

TABLES

INTRODUCTION

The problem of non-finite forms in the Indo-Aryan (IA) languages is not a new one. There have been a number of studies by IA scholars devoted to synchronic and even diachronic descriptions of converbs. Old Indo-Aryan (OIA) received proper attention in the seminal study by Tikkanen (1985). Middle Indo-Aryan (MIA) was analysed by Hendriksen (1944) and more recently within a modern theoretical framework by Peterson (1998).

There have also been studies dealing with individual New Indo-Aryan (NIA) languages, such as a pragmatic account of the Hindi converb (cf. Davison 1981 and its refutation by Kachru 1981), the Bhojpuri converb (Lohar 2012), a diachronic account of Maithili converbs (Yadav 2004), a diachronic syntax of converbs in selected Hindi dialects (Dwarikesh 1971), and the most insightful synchronic study thus far, dealing thoroughly with the Nepali converb (Peterson 2002).

There have also been important typological and areal studies devoted to the problem of converbs in IA (cf. Masica 1976; Subbarao 2012).

However, despite the fact that the literature on IA non-finites is quite rich, there is a significant gap in historical research relating to the early NIA period (with the exception of Dwarikesh 1971, a work focusing predominantly on the morphology of the converb with something of a synchronic bias). This is a crucial period in the development of the IA languages. Within this period, there are drastic changes in the morphosyntactic alignment system, which seem to result from the reorganization of the case and verbal system inherited from MIA. In Early Hindi, for instance, we witness the rise of a new postpositional system as well as a reinterpretation of participial verbal forms.

What is more, other non-finite forms, such as adverbial participles, have very rarely been dealt with (cf. Pořízka 1950; 1952 on Hindi participles; Sigorski 2005 for a diachronic treatise on Hindi converbs and adverbial participles; Subbarao 2012 on the typology of adverbial participles and converbs in IA). Infinitives have also received little attention (cf. Davison 2008 and only recently Montaut 2018b).

The present work will not only provide a morphosyntactic analysis of different non-finite forms (converbs, infinitives and adverbial participles) appearing over the course of time; it will also attempt to approach the category of non-finite verbs from holistic and diachronic perspectives. We intend to combine the two viewpoints in analysing several IA dialectal groups, such as Early Rajasthani, Early Awadhi, Early Braj, Early Dakkhini and Early Pahari.

The structure of the work is as follows. After a short section 1 on the corpus utilized for our analysis, in section 2 we present basic theoretical prerequisites, giving a brief introduction to the problem of (non-)finiteness and then focusing on converbs, participles and infinitives from a typological perspective. Section 3 deals with morphosyntactic and control properties of converbs and participles in early NIA, and then in in section 4 we consider the problem of the scope of selected operators and the type of linking represented by converbal chain constructions. Concluding remarks on converbal chain constructions are presented in section 5. In section 6 morphosyntactic properties of infinitives are presented, and these are summarized in section 7. Section 8 contains a description of the tagging tool and artificial intelligence module used for enhancement of the linguistic analysis.

Here we would like to thank Saartje Verbeke from Ghent University for her help in Early Awadhi text annotation and her comments on several theoretical issues.

We are also grateful to Liudmila Khokhlova for reading the manuscript and giving a valuable empirical as well theoretical input.

1. CORPUS

The preparation of electronic corpora was a prerequisite to carrying out the research on non-finite forms. The texts include:

I. Excerpts from Early Rajasthani prose (Bhānāvat & Kamal 1997-1998):

1) Truṇprabhav Sūri – Dvitīya vrata satya para kathā (14c.) (henceforth RG.TS)
2) Somsuṃdar Sūri – Guru mahimā par kathā (15c.) (henceforth RG.SS)
3) Merusuṃdar – Amarsena-Vayarsena (15c.) (henceforth RG.M)
4) Gāḍaṇ Śiwdās – Vacanikā khīcī acaḷadāsa-rī (15c.) (henceforth RG.GŚ)
5) Vīsaḷadevarāsa (15c.) (henceforth VD)
6) Dalapata-vilāsa (16/17c.) (henceforth RG.DV)
7) Muhaṇot Naiṇasī – hadai surijamal-ri bata (17c.) (henceforth RG.MN)
8) Khiṛiyo Jaggo – vacanika rathod ratan ri (17c.) (henceforth RG.KJ)
9) Rathod duradavasa ro kagada (17/18c.) (henceforth RG.RDK)
10) Dhanush bhamg (18c.) (henceforth RG.DB)
11) Muhaṇot Saṃgrāmsiṃgh – Adalati nyay (18c.) (henceforth RG.MS)
12) Dokari ri bāt (18c.) (henceforth RG.DRB)

II. Excerpts from Early Awadhi:

1) Malik Muhammad Jāyasī's 'Padmāvat' AD 1540 (henceforth J)
2) Tulsīdās' 'Rāmacaritamānasa' AD 1574-1576 (henceforth T)

Early Awadhi texts were taken from https://wp.unil.ch/eniat/ and were accepted as standard, but where necessary compared with the following editions: for Tulsīdās' 'Rāmacaritamānasa', Prasad (1994); for Jāyasī's 'Padmāvat', Gautam (1954).

III. Excerpts from Early Braj texts:

1) prose – Indrajīt of Orchā 'Vivekadīpikā' from 1600 – edited by McGregor (1968: 1-21) (henceforth I)
2) poetry by Bhūṣaṇ Tripāṭhī 'Śivrājbhūṣaṇa' from 1673 – edited by Miśra (1994: 1-94) (henceforth Ś)
3) poetry by Hita Harivaṃśa 'Hita caurāsī / Caurāsī' from 16th c. – edited by Snell (1991a), verses 1.1-39.6 (henceforth HH)

4) poetry by Viṣṇudās 'Rāmāyan kathā' 1442 – edited by Dvivedī (1972: 50-59) (henceforth V)

IV. Excerpts from Early Dakkhini collection of prose (Śarma 1954):

1) Khwājā Bandā Navāz Gesūdrāz (1312/88-1422/37), prose (henceforth BN)
2) Goṁdā (1300-1351), poetry (henceforth G)
3) Eknāth (1548-99), poetry (henceforth E)
4) Qulī Qutub Śāh (1580-1612), poetry (henceforth QŚ)
5) Mullā Vajahī (1575/80-1660/71), prose (henceforth MV)
6) Fāyaz (1685), poetry (henceforth F)
7) Huseini (1641), poetry (henceforth H)
8) Saiyad Mīrāṁ Husenī (1623), poetry (henceforth SMH)
9) Valī Daknī/Dakkhinī (1682-1730), poetry (henceforth VD)

V. Early Pahari inscriptions (Joshi 2009; Pant 2009) and prose (Joshi 1983):

1) inscriptional data (Joshi 2009; Pant 2009)
2) the only edition of the first Kumaoni literary piece, Rajnitīśāstra (Joshi 1983) (henceforth RŚ)

The corpus amounts to 45000 words: excerpts I, III and IV – 10000 words each, II – 12000 words and V - 3000 words.

As regards the Early Pahari data, we have benefited greatly from the expertise of Indian and Nepali scholars, including Prof. Maheshwar Prasad Joshi from Kumaon University, and Prof. Madhav Pokharel and Dr. Chalise Bidurkumar from Tribhuvan University in Kathmandu.

Texts were stored and annotated by means of IATagger, a system designed for the optimization of multilayered analysis of early NIA data, which will be fully dealt with in section 8.

Optical recognition of texts available only in printed versions (Early Rajasthani, Early Dakkhini, Early Pahari, and Early Braj except Hita Harivaṃśa) was supported by a Hindi OCR program (HindiOCR 2013), for which we would like to thank Oliver Hellwig.

2. THEORETICAL PREREQUISITES

2.1 (NON-)FINITENESS

The term finiteness comes from the Latin *finitus*, the perfective participle of the verb *finio* 'finish, limit, set bounds to, determine' (Nikolaeva 2007: 1). As Nikolaeva (2007) explains, it was first used in defining personal pronouns, and later it was applied to verbs expressing person and number. Since then, grammars have employed the distinction between finite verbs, i.e. those determined by person and number, and non-finite verbs, i.e. the forms lacking person/number marking. "Although in Latin the finite/nonfinite distinction was initially motivated by the presence/absence of agreement (person and number), later other verbal categories were taken into account, most importantly tense" (Nikolaeva 2007: 1).

Infinitives, participles, converbs and action nominals are traditionally perceived as being non-finite. As van der Auwera (1998b: 275) and later Ylikoski (2003: 187) have noted, these forms are most typically used in more than one syntactic function. Ylikoski (2003: 187) "tentatively" presents the ideal system of non-finites, represented by Hungarian, with the infinitive occurring in (1a), the participle in (1b), the converb in (1c) and the action nominal in (1d).[1]

(1) Hungarian from Ylikoski (2003: 187)

a. *A lány sír-ni akar-t ~ kezd-ett.*
The girl cry-INF want-PST.3SG begin-PST.3SG
'The girl wanted ~ began to cry.'

b. *Egy sír-ó lány be-jö-tt a szobá-ba.*
A cry-PTCP.PRES girl in-come-PST.3SG the room-ILL
'A crying girl entered the room.'

c. *A lány sír-va jö-tt be a szobá-ba.*
The girl cry-CVB come-PST.3SG in the room-ILL
'The girl entered the room crying.'

[1] "It is important to note that of the four non-finites that illustrate the main types of non-finites, the last one, the action nominal in -ás/-és (1d), is not considered an inflectional verb form in traditional Hungarian grammar, but a derived deverbal noun instead." (Ylikoski 2003: 188). For more on the discussion of verb vs. non-verb forms cf. Ylikoski (2003: 188).

d. *A* *lány* *sír-ás-a* *ingerel* *engem.*
The girl cry-AN-3SG irritate.3SG I.ACC
'The girl's crying irritates me.'

The most popular criterion for defining non-finiteness is the lack of specification for tense (-aspect) and mood and for agreement with arguments (Haspelmath 1995: 4). Examples (2) and (3) illustrate the lack of agreement, although as we can see in example (2) below, tense or aspect may in the case of IE languages be somewhat problematic when participial forms are considered non-finite.

(2) Classical Greek from Goodwin (1889: 333)
taûta *eip-ὸ̀n* *apẽìei*
this say.PART.PRS.M.SG depart.IMPF.3SG
'Having said this, he departed.'

(3) Latin
lus-um *it* *Maecenas,* *dormi-tum* *ego* *Vergilius-que*
play-SUP go.3SG.PRS M.NOM sleep-SUP I.NOM V-and
'Maecenas goes to play, I and Vergilius go to sleep.' (Hor. Sat. 1.5.48)

Bickel (1998: 384) comments on Haspelmath's criterion of (non-)finiteness:

> "One challenge comes from the fact that, as Haspelmath acknowledges himself (pp. 5-6), finiteness is inherently a gradual concept. Traditionally, "finite" referred to verb forms that are delimited (*finitum*) with respect to person and number agreement, but with the recent inclusion of various degrees of tense, aspect, and mood specification, the concept now encompasses variable quantities of category specifications."

Even though the definition of non-finiteness is quite straightforward, considering the classical languages like Latin and Greek (with the exclusion of categories such as tense and aspect), Haspelmath points out that the notion of (non-)finiteness is problematic in a wider perspective:

> "[...] the traditional concepts of finiteness and nonfiniteness are just two extreme points on a *scale of desententialization* (cf. Lehmann 1988: 200), and other languages may show various intermediate points on this scale. Most notably, verb forms may lack tense and mood specifications, but still have subject agreement. [...]" (Haspelmath 1995: 5).

Lehmann's (1988) idea of *desententialization* is the concept of a continuum of subordinate clause reduction where the clause gradually loses its properties and becomes less finite. "Components of the clause which allow reference to a specific state of affairs are dropped; the state of affairs is 'typified'. At the same time, the subordinate clause increasingly acquires nominal properties, both internally and in its distribution. At the end of this

process of nominalization, the clause becomes a nominal or adverbial constituent of a matrix clause." Lehmann enumerates several consequences of *desententialization*: 1) lack of illocutionary force of the subordinate clause[2]; 2) constraints on the mood of the subordinate clause or its absence; 3) reduction of word order freedom inside the subordinate clause; 4) reduction and loss of tense and aspect; 5) affected relationality of the predicate verb[3]; 6) affected polarity (the subordinate verb can no longer be independently negated); 7) conversion of verbal to nominal government; 8) combinability with adpositions and case affixes. The continuum is shown below:

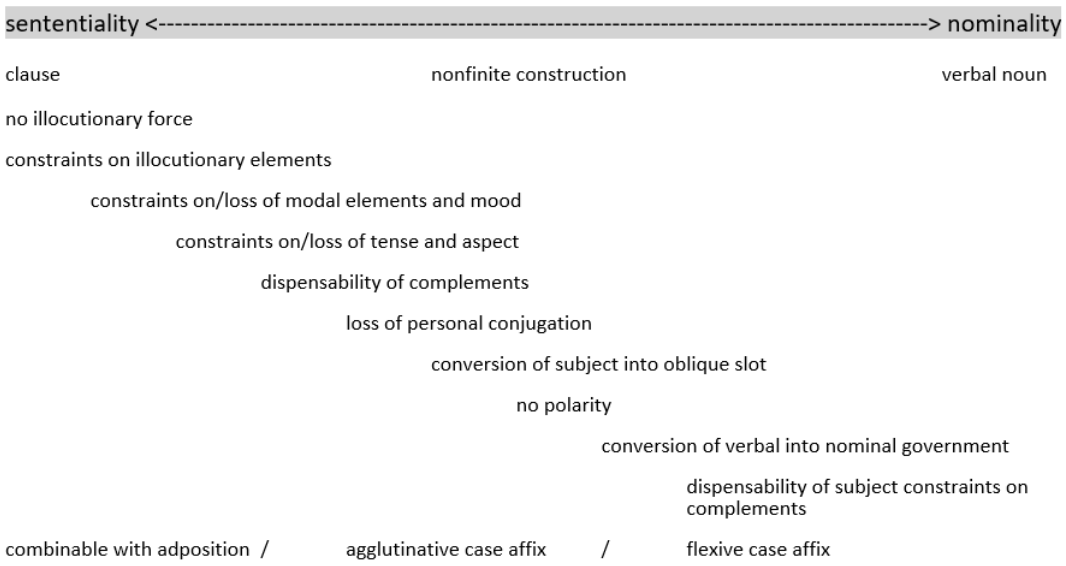

Figure 1. Desententialization, after Lehmann (1988).

As has already been pointed out, and can be seen again on the above continuum, the notion of non-finiteness should not be perceived as one point on a scale, but as a gradual process.

Beside TAM (tense, aspect, mood) and agreement with arguments, some scholars have considered dependency as a parameter forming the notion of non-finiteness. The idea acknowledges that only finite verbs can form independent clauses, and these have to contain only one finite verb. However, Nikolaeva (2007) argues that there are languages that allow non-finite forms with reduced tense and agreement to function as the only predicate in a clause. Using the findings of Johns and Smallwood[4] (1999) and Vincent (1998), she claims:

[2] However, there are a few classes of subordinate clauses that are allowed to have their own illocutionary force, e.g. non-restrictive relative clauses (cf. Lehmann 1988, Green 1976, Lakoff 1984).

[3] "The predication expressed in an independent clause [...] gets lost, and simultaneously the subject slot of the subordinate verb is either converted into an oblique slot or is entirely lost. The verb becomes nonfinite" (Lehmann 1988).

[4] "Johns and Smallwood (1999) show that among eight possible combinations of three finiteness features, main clausehood (MC), tense (T), and agreement (Agr) marking, at least four are associated with the term

> "So tense/mood/agreement morphology and dependent/independent status appear to be empirically independent parameters, although there may be implicational correlations between them: for all languages, if person and/or number and/or tense are marked on dependent forms, then they are also marked on independent forms (Vincent 1998)" (Nikolaeva 2007: 3).

Talmy Givón in his *Syntax: a functional-typological introduction*, similarly to Lehmann (1988), perceives non-finiteness as a feature of a gradual process of nominalization[5][6]:

> "*Finiteness* has been treated traditionally as a property of verbs, since many of its salient features (82a-d) [see footnote 6] indeed pertain to the verb. But the rest of the features (82e-g) [see footnote 6] pertain to other constituents of the clause. Finiteness is thus fundamentally an aggregate grammatical feature of clauses. Its converse, *non-finiteness*, is thus an aggregate grammatical feature of nominals, i.e. noun phrases" (Givon 2001ii: 24).

To illustrate the scale, he presents the following figure:

Least finite
a. Her good *knowledge* of math helped
b. Her *knowing* math well helped
c. For her *to know* math so well surely...
d. She wanted *to know* math well
e. *Having known* math well since high school, she...
f. She *should have known* math well
Most finite

Figure 2. Gradual process of nominalization (Givon 1990)

'nonfinite' in descriptive practice: -MC -T -Agr (English infinitives), -MC -T +Agr (European Portuguese infinitives), -MC +T -Agr (Tamil and Lezgian participles), and +MC -T -Agr (Russian and Middle Welsh infinitives)" (Nikolaeva 2007: 3).

[5] "Nominalization is the process via which a *finite verbal clause* — either a complete clause or a subject-less verb phrase — is converted into a *noun phrase*" (Givon 2001: 24).

[6] "Nominalization is best described in terms of the syntactic adjustments from the finite verbal-clause prototype to the nominal (NP) prototype (Hopper and Thompson 1984). The major components of such adjustment are:

Nominalization as the adjustment from the prototype finite verbal clause to the prototype noun phrase

a. verb becoming a *head noun*

b. verb acquiring *nominalizing* morphology

c. loss of *tense-aspect-modal* morphology

d. loss of *pronominal agreement* morphology

e. subject and/or object acquiring *genitive* case-marking

f. addition of *determiners*

g. conversion of adverbs into *adjectives*" (Givon 2001: 24-25).

At the very bottom of the scale is the most finite verb, marked for TAM. In (e), we have a form exhibiting aspectual marking (perfect). In (c) and (d), as Givon (1990: 26) indicates, "the subject is not marked as genitive and the verb is marked by the slightly more finite infinitive form", whilst in (b) the verb occurs in the already nominalized gerund form. Example (a) shows the most nominalized form.

Among syntactic features that determine the degree of finiteness of a given clause in comparison with a prototype transitive main clause, Givon (1990: 853) lists:

- Tense-aspect-modality (TAM)
- Pronominal ('grammatical') agreement
- Nominalizing affixes
- Case marking of the subject and object
- Articles, determiners

This short overview of (non-)finiteness leads us to a short examination of several non-finite forms, such as converbs, participles and infinitives.

2.2 CONVERBS

2.2.1 MAIN DEFINITIONAL ISSUES

The notion of converb has been analysed for quite some time. The term was introduced by John Ramstedt (1903: 55) for a dependent verb form found in Altaic languages. In time, scholars working on other languages realized that converb "is roughly synonymous with such terms as verbal adverb, adverbial/conjunctive participle, gerund (in the sense of the ablatives modi of the Latin gerund), deepričastie (Russian), gérondif (French), sentence equivalent (Finnish), and absolutive (Indo-Aryan)" (Tikkanen 2001: 1112). In this chapter, we will provide a thorough overview of the definitions and discussions of converbs offered by Haspelmath (1995), Nedjalkov (1995), van der Auwera (1998), Bickel (1998), Tikkanen (2001), and Coupe (2006).

There are two leading, widely accepted definitions of "converb", one proposed by Haspelmath (1995) and one by Nedjalkov (1995). Hasplemath (1995: 3) defines a converb as "[...] a nonfinite verb form[7] whose main function is to mark adverbial subordination.

[7] Haspelmath (1995: 4) claims that converb is not a separate word class. "Converbs never have the degree of autonomy that is associated with the status of lexemes, so they are clearly inflectional, not derivational forms", i.e., they carry grammatical information. So "converb is a verb form that is part of the inflectional paradigm of verbs", and because of that, it "cannot be easily analyzed as a verb plus a complementizer or subordinator". Haspelmath perceives converb as an inherently subordinate (embedded) verb form. And thus, he says: "It has been suggested that converbs should be understood as combinations of verb plus complementizer (which happen to be tightly bound), but this is just an attempt to fit an unfamiliar phenomenon into the procrustean

Another way of putting it is that converbs are verbal adverbs, just like participles are verbal adjectives." In this way, he tries to make a clear distinction between forms derived from verbs, as in the table below:

Word class:	Noun	Adjective	Adverb
Derived verb form:	masdar (= verbal noun)	participle (= verbal adjective)	converb (= verbal adverb)
Syntactic function:	argument	adnominal modifier	adverbial modifier

Table 1. Derived verb forms with different word class status (from Haspelmath 1995: 4).

Nedjalkov (1995: 97) suggests the following definition: "As a first approximation we can define a converb as a verb form which depends syntactically on another verb form, but is not its syntactic actant, i.e. does not realize its semantic valencies." Let us consider the following example:

(4) (Polish)

Przyszed-łszy	*do domu,*	*Ania zjadła*	*kolacj-ę.*
Come-PFV.CVB	to home	Ania eat-PST.PFV.F	dinner-ACC

'Having come home, Ania ate dinner.'

In the example above *przyszedłszy* 'having come' is an anterior converb which cannot stand alone as a main verb – it is dependent on the main clause verb *zjadła* 'ate'. The implicit subject of the converb is coreferential with the subject of the main verb – *Ania*. The converb is intransitive and thus has only one argument – the subject. The main verb is transitive and has two arguments: the subject *Ania* and direct object *kolację*. As regards valency, the converb is monovalent and assigns *Ania* the semantic role of the AGENT, while the main verb is bivalent, and here again *Ania* is the AGENT while *kolację* is the PATIENT. However, the PATIENT is assigned only to the main verb and not to the converb. Thus, the converb does not realize the semantic valency of the main verb.

The rest of Nedjalkov's definition makes a distinction between converbs and other verbal forms: "Thus, a *canonical* (i.e., noncombined) converb can occupy (1) the position of an adjunct, i.e., an adverbial, but cannot occupy the positions: (2) of the only predicate of a simple sentence (without additional auxiliary elements); (3) of nominal attributes; (4) of clausal actant (i.e., it cannot depend on verbs such as *begin*, *order*, etc.); (5) of nominal actant (i.e., it does not occur in subject and object position [...]" (Nedjalkov 1995: 97). He clarifies the last four positions, "the following verb forms occur canonically: in the second

bed of the European language type, which strongly prefers adverbial conjunctions to converbs (cf. Kortmann [in press])" (Haspelmath 1995: 4).

position – a finite form; in the third position – a participle; in the fourth position – an infinitive, in the fifth position – a gerund (i.e., a deverbal noun that is part of the system of verb forms)."

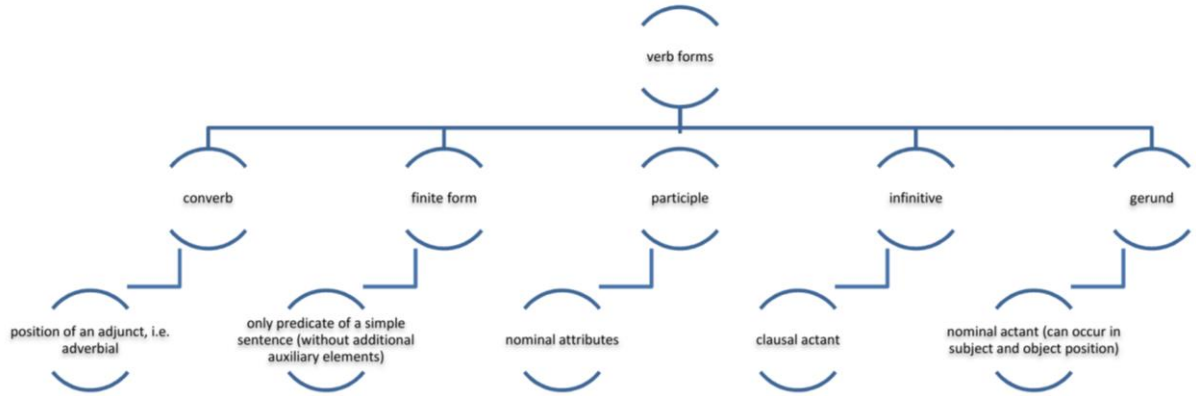

Figure 3. Nedjalkov's (1995) classification of verb forms.

Van der Auwera (1998) provides a systematic overview and comparison of both viewpoints. Firstly, he points out that these two definitions share the idea of a converb's "prototypical" (V. P. Nedjalkov and I. V. Nedjalkov) or "main" (Haspelmath) function. He approves of this view, giving the following examples:

(5) Russian (Weiss 1995: 251)

Ona	*menja*	*oskorbila,*	*priglasiv*	*moego*	*sopernika.*
She	me	offended.PST	inviting.CVB.PFV	my.ACC	rival.ACC

'She offended me by inviting my rival.'

(6) English (van der Auwera 1998: 275)

Inviting my rival was offensive to me.

Here the Russian *priglasiv* exhibits a prototypical function, while the English counterpart *inviting* does not, because it can function as a masdar as well (van der Auwera 1998: 274-275).

Nedjalkov (1995: 103) explains the problem of the delimitation of the main function of non-finite forms. He claims that "[...] it seems appropriate to consider the function with the highest text frequency as the main function" and to name the form according to this function. This is not straightforward, however, because of, among others, "nonequal statistical characteristics of different voice and aspect forms which are traditionally subsumed under one and the same nonfinite form" (Nedjalkov 1995: 103). As an example he gives the English participle, which can function as a nominal modifier or can be used in an adverbial function (especially in passive (also nonperfect) and perfect active forms).

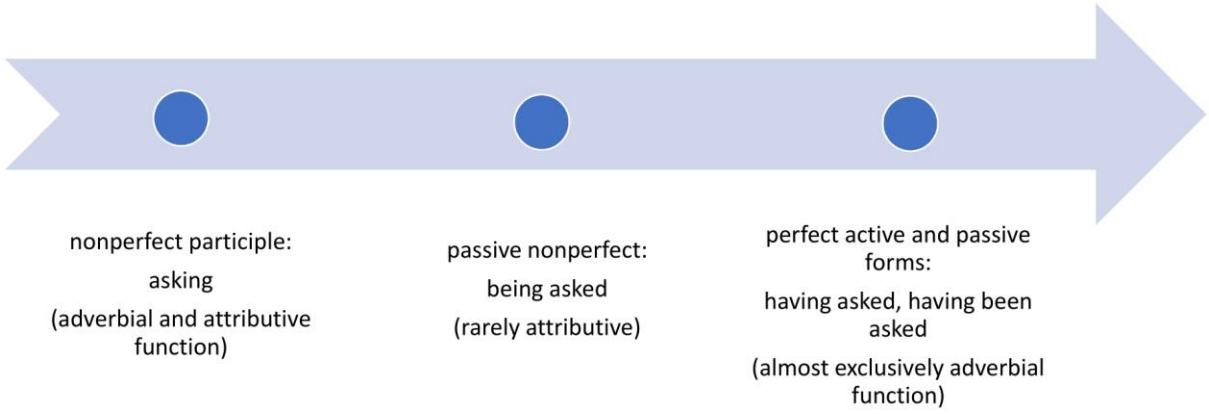

Figure 4. Functions of participles according to their tense, aspect and voice, based on Nedjalkov's (1995) analysis.

On the very left of the arrow we have a non-perfect participle, which gradually turns into passive and perfect forms. The functions of the participles are seen to become narrower as we move along the tense, aspect and voice scale. The non-perfect participle can function as both adverbial and noun modifier. The passive non-perfect is very rarely used attributively, and the perfect active and passive participles occur mainly in an adverbial function, "although they are perfectly grammatical in attributive function" (Nedjalkov 1995: 103).

What is more, Nedjalkov points out a possible twofold interpretation of the same non-finite form, as in:

(7) English (Nedjalkov 1995: 103)
The man, pacing the floor, said...
a. 'the man, who was pacing the floor, ...' – attributive function
b. 'the man, while pacing the floor, ...' – adverbial function

A striking difference between the two definitions is the presence of the notion of non-finiteness. Haspelmath (1995) includes this as one of the criteria defining a converb, while Nedjalkov (1995) does not.

Haspelmath comments on Nedjalkov's (1990, 1995) definition of converb where non-finiteness plays no role: "[...] finite verb forms which are used only in adverbial subordinate clauses also fall under the definition." However, he claims that "extending the term converb to finite subordination forms [...] seems an unjustified departure from traditional usage" (Haspelmath 1995: 4-5). He adds that only "a nonfinite adverbial subordination form could be said to be a 'verbal adverb', and the term converb seems ideally suited to fill the 'verbal adverb' position in Table 1."

Additionally, Haspelmath discusses a situation when the converb is marked for possessor agreement with its subject, which blurs the distinction between finite and non-finite. However, he observes that "in some languages it is not easy to tell whether person-number inflection is possessive or finite" (cf. Haspelmath 1995: 5-7).

Despite all these problems, he prefers to retain non-finiteness as a definitional criterion "because it restricts the notion *converb* in a way that is consistent with the traditional use of the term (and equivalent terms)" (Haspelmath 1995: 7).

Zúñiga (1998: 2), providing an overview of some approaches to converbs, mentions Ebert's work:

> "As to the desententialization continuum, Ebert (1993) presents evidence from several south Asian languages showing that functionally equivalent forms admit considerable formal diversity, ranging from nonfinite (verb stem + converbal suffix) in Tamil to finite (e.g. fully inflected verb form + case marker) in some Kiranti languages, among others. [...] Ebert (1993: 106) offers an intermediate position in allowing either TAM or PERS markers on the verb, but not both simultaneously, to consider a form converbal. She also suggests to restrict the use of *finite* to its functional usage with the meaning 'predicative' and to speak of *inflectedness* when referring to formal make-up."

Turning to Nedjalkov's definition, Tikkanen (2001) explains why it might be preferable to that proposed by Haspelmath:

> "But if non-finiteness is seen to exclude the possibility of subject agreement with regular subject markers, it may be that forms of the same paradigm must be assigned different finiteness status" (Tikkanen 2001: 1113). In other words, instead of trying to find some generalization for this particular dependency between two verb forms, we will end up with many different names for categories describing this dependency, as he says "depending upon where we draw the line between finite and non-finite". He says that "non-finites are sometimes allowed to take possessive markers, which show agreement with the (notional) subject."

He gives a very interesting example from Kurukh (North Dravidian):

(8) Kurukh (Dube 1983: 6)

sipaahi-r	*asan*	*bar-c-ar*	*kii*	*nerr-an*	*piṭi-y-ar*	*cic-c-ar.*
soldier-PL	thither	come-PST-3PL	CVB	snake-ACC	kill-PST-3PL	give-PST-3PL

'The soldiers came there and killed the snake.'

In this example we see a hybrid formation where the anterior converb is a finite, inflected verb form followed by a converb marker borrowed from Sadri/Sadani (Central Indo-Aryan).

Tikkanen (2001) suggests that: "At least in some cases it could then be expedient to introduce a distinction between conjugated and non-conjugated converbs. The label 'non-finite' could still be maintained, given that the forms in question lack expression or implication of mood and are hence unable to function as (prototypical) independent predicates. However, the restriction against expression of the absolute tense cannot be made criterial. Some languages, including Korean and some Papuan languages, have converbs that inflect for the absolute as well as relative tense."

Van der Auwera (1998) notices one more difference between these two definitions, namely adverbiality. Adverbial function is one of the components of Haspelmath's definition, but not of Nedjalkov's.

Haspelmath (1995) uses the defining criterion "adverbial (subordination)" to "exclude masdars/verbal nouns (nonfinite verb forms specialized for argument subordination, or complementation) and participles (nonfinite verb forms specialized for adnominal subordination)." He clarifies: "Converbal constructions are generally not arguments but modifiers, and they generally modify verbs, clauses or sentences, but not nouns or noun phrases" (Haspelmath 1995: 7).

He points out that there is a type of subordinate construction, called the *clause-chaining construction*, which is neither argumental nor adnominal, nor clearly adverbial. This construction indicates a sequence of successive events, as in the following example from Kumyk (Turkic):

(9) Kumyk (Džanmavov 1967: 234)
Bu-lar, köl-nü gör-üp, arba-syn toqtat-yp, čemodan-ny
this-PL lake-ACC see-CVB cart-3.POSS stop-CVB suitcase-ACC
Manaj-ğa da göter-t-ip, köl-nü jağa-syn-a
Manaj-DAT also take-CAUS-CVB lake-GEN bank-3.POSS-DAT
bar-yp, čemodan-ny ač-yp, šišla-ny čyğar-yp
go-CVB suitcase-ACC open-CVB bottle-ACC take.out-CVB
tiz-ip, suw-dan toltur-up, qajat-yp čemodan-ğa
put.in.row-CVB water-ABL fil-CVB return-CVB suitcase-DAT
sal-a.
put-PRES
'They see the lake, stop their cart, make Manaj bring the suitcase, go to the bank of the lake, open the suitcase, take out the bottles, put them in a row, fill them with water, and put them back into the suitcase.'

Haspelmath notices that this use of converbs is not central to his definition because it is not really adverbial. "However, it is not easy to make a clear-cut distinction between temporal adverbial subordination and clause-chaining. For example, one could use English adverbial subordinate constructions at least for a sequence of three events, e.g. *After they took out the bottles, putting them in a row, they filled them with water*. So it is not absurd to think of clause-chaining constructions such as [the one above] as successive adverbial subordination of a special type" (Haspelmath 1995: 8).

However, Tikkanen (2001: 1113) questions Haspelmath's concept of 'adverbial subordination' underlying the notion of clause-chaining. Similarly, Coupe (2006: 147) observes:

> "Some languages must rely on the contextual setting to distinguish between an adverbial modifying or clause-chaining function of a converb. In Hindi, for instance, alternative adverbial and sequential interpretations are sometimes apparent for the same converb clause." (see ex. 10)

(10) Hindi (Davison 1986: 1)

mujhe	*un*	*cīzõ-ko*	*dekh- kar*	*bahut*	*gussā*	*āyā.*
me.DAT	those.OBL	things-DAT	see-CVB	great	anger	come.PST

'When I saw those things, I became very angry.'
'Having seen those things, I became very angry.'

"Converbs are widely employed for clause chaining in languages of South, Central and East Asia, as well as in languages of Ethiopia (cf. Gasser: 1983: 101ff. and Lamberti & Sottile 1997: 222ff.)" (Coupe 2006: 147). As he goes on to say, European languages do not employ converbs to create sequences of events, but to encode purely adverbial propositions with respect to their matrix predicates (see examples below).

(11) Polish

Spojrza-wszy	*na niego,*	*za-śmia-ł-a*	*się.*
look at-CVB	at he.ACC	PERF-laugh-PST-F.SG	REFL

'Having looked at him she burst out laughing.'

(12) Polish

Śpiewa-ł,	*fałszuj-ąc*	*niemiłosiernie.*
sing.PST.M.SG	sing out of tune-CVB	mercilessly

'He sang (singing) out of tune mercilessly.'

On the other hand, V. P. Nedjalkov and I. V. Nedjalkov (1987) claim that converbs may, but need not, be adverbial. "If they are not adverbial, V.P. Nedjalkov and I. V. Nedjalkov (1987: 76) call them 'narrative'. [...] Besides the term 'narrative', one also finds 'coordinative' (V.P. Nedjalkov 1995: 95, 109-110; Bisang 1995: 154-155; Alpatov and Podlesskaya 1995: 467-474), 'syndetic', 'copulative', 'conjunctive' and 'propulsive' (all mentioned in Johanson 1995: 329-330), 'angeschlossen' (Müller-Bardey 1990: 14) and 'sequential' (Ebert 1993: 85)" (van der Auwera 1998: 276-277).

What is more, van der Auwera (1998) explains "the relation between Haspelmath's narrow concept of converb – 'converb *sensu stricto*' – and V.P. Nedjalkov and I.V. Nedjalkov's wider concept – 'converb *sensu latiore*'. Basically, converb in the narrow sense consists of the features [+ dependent, – argumental[8], – adnominal, + embedded, – finite], whereas converb in the wide sense requires only [+ dependent, – argumental, – adnominal]. After analysing these two approaches, he adds another notion that lies between the two, a converb that is [+ dependent, – argumental, – adnominal, – finite], whereby "converbs are either converbs

[8] Van der Auwera (1998: 278) uses the features "argumental" and "adnominal" to differentiate converbs *sensu stricto* from other [+ dependent, + embedded, – finite] verbal forms, i.e. masdars and participles. "Argumental" means here "can make up an argument by itself".

sensu stricto or medial verbs, but finite, subordinate mood forms are excluded" (van der Auwera 1998: 279). The following Table 2 from van der Auwera sums up the position:

<table>
<tr><th colspan="3">+ dependent, – argumental, – adnominal</th></tr>
<tr><th colspan="2">+ embedded
"subordinate"</th><th>– embedded
"cosubordinate"</th></tr>
<tr><th>+ finite</th><th colspan="2">– finite</th></tr>
<tr><th>subordinate mood</th><td>converb sensu stricto</td><th>medial verb</th></tr>
<tr><td colspan="3">converb sensu latiore</td></tr>
</table>

Table 2. Converb sensu stricto and converb sensu latiore (from van der Auwera 1998).

What is more, he considers whether other possibilities exist, i.e., verb forms that are [+ dependent, – argumental, – adnominal, – embedded, + finite]. Van der Auwera (1998) refers to Ebert (1993), who validates the existence of such forms in Kiranti languages, and adds that "Finite non-adverbial converbs are also found in Ethiopic and Cushitic languages."

Apart from the two mainstream definitions, we can find some different approaches. One of the alternatives is proposed by Bickel (1998).

Bickel disputes the notion of converb as a cross-linguistically applicable term. He postulates a differentiation between two types of converbs: "European converb" and "Asian converb". He states that the "European" converb follows Haspelmath's definition and "is generally confined to adverbial (verb-modifying) (13a) and adsentential (13b) subordination, with extensions to illocutionary force hedging (13c) and complement (13d) functions", giving the following examples from Russian:

(13) Russian (Bickel 1998: 394)

a. *On vyše-l* *posvistyvaj-a.*
He go out.PFV-PT whistle.IPFV-CVB
'He went out whistling.'

b. *Slušaj-a* *ego,* *ja* *čustvova-l* *sebja* *očen'* *ploxo.*
Listen-IPFV.CVB him I feel.IPFV-PST REFL very bad
'When listening to him, I felt very bad.'

c. *Otkrovenno* *govor'-a,* *èto* *sovsem* *nevozmožno.*
Frankly speaking.IPFV-CVB that at all impossible
'Frankly speaking, this is absolutely impossible.'

d. *My prove-l-i* *prazdniki* *kupaj-a-s'* *v more.*
We spend-PST-PL holidays bathe-CVB-REFL in sea
(vs. *My proveli Ø.)
'We spent the holidays bathing in the sea.'

As he claims, "this type does not include chaining function but rather stands in a binary relation to the main verb (cf. König, p. 72). In line with this, the European converb supports only what Tikkanen calls disjunctive scope integration" (Bickel 1998: 395). "Outside Europe, the Tungusic and Eskimo converbs seem to be similar to this type and many Kiranti languages feature a reduced version that supports only adverbial (verb-modifying) but not adsentential subordination" (Bickel 1998: 395).

The "Asian" converb (if we allow for regional "holes") merges adverbial modification and narrative chaining in a single (set of) dependent verb form(s). Bickel gives Nepali examples for this:

(14) Nepali (Bickel 1998: 395)

a. *hiḍ-era* *āun-chu*.
walk-CVB come-1SG.NPT
'I will come by foot.'

b. *Ek* *chin* *u* *saṅga* *kurā* *gar-era* *āun-chu*.
One moment he with talk do-CVB come-1SG.NPT
'I will talk to him for a moment and then I will come.'

"The systematic inclusion of modifying functions makes the Asian converb different from the Papuan *Satzinnenform* (medial verb) and may explain why Central and South Asian sequences only occasionally reach the length of the famous Fore chains or of Swahili ka-paragraphs" (Bickel 1998: 395).

The IA converb is a non-finite form, although Subbarao (2012: 281) claims that the Kashmiri converb has some features of finiteness, since it permits violation of the Subject Identity Constraint rule. What is more, due to its historical origin it still retains the +tensed feature. However, formally, IA converbs lack any features of finiteness and control properties, as will be discussed at length in the following sections, whereas in some Dravidian, Munda and Tibeto-Burman languages converbs have formal markers of finiteness such as person markers.

2.2.2 CONVERBS AS SUBORDINATING AND CHAINING DEVICES

In his definition of converb, Haspelmath uses the term *subordinate* "in the sense 'embedded', or 'incorporated into the superordinate clause', contrasting with *coordinate* clauses, which are not part of another superordinate clause" (Haspelmath 1995: 8). He observes that converbal constructions can often be paraphrased into coordinate constructions in languages that allow coordination of clauses. After analysing some examples, he adds that "one might therefore suspect that converbal constructions are also syntactically coordinate in some sense. However, converbal constructions consistently turn out to be subordinate by the most reliable criteria for subordination" (Haspelmath 1995: 8).

Haspelmath (1995) proposes the following criteria for subordination:

a. clause-internal word order
b. variable position
c. possibility of backwards pronominal anaphora (i.e. pronominal cataphora) and control
d. semantic restrictiveness, and hence focusability
e. possibility of extraction.

He observes that based on the proposed criteria, converbs are consistently subordinate (rather than coordinate). Even though not all converbs fulfil all of these criteria, they realize a subset of the criteria while non-subordinate clauses fulfil none of them.[9]

I. Clause-internal word order

A subordinate clause may appear inside its superordinate clause, making it discontinuous, as in the Japanese example:

(15) Japanese (Haspelmath 1995: 12)

John wa	*Mary ni*	*boosi*	*o*	*nui-de*	*aisatu*	*si-ta.*
John TOP	Mary DAT	hat	ACC	take off.CVB	greet	do-PST

'John took off his hat and greeted Mary.'

II. Variable position

Subordinate clauses may come after or before the superordinate verb without changing the meaning, but when coordinate clauses occur in different orders the meaning changes dramatically if the events are considered sequential (rather than simultaneous). "Because they are hierarchically equal, coordinate clauses show tense iconicity, i.e. the event reported in an earlier coordinate clause is interpreted as occurring earlier (Haiman 1985: 216). Meaning differences in converbs that are associated with different positions are also attested (e.g. Kortmann 1991), but they do not involve tense iconicity" (Haspelmath 1995: 13-14).

III. Backwards pronominal anaphora and control

Haspelmath (1995: 14) claims that: "Backwards pronominal anaphora is only possible in subordinate clauses [...]. The crucial point is, of course, that the pronoun must be

[9] Tikkanen (2001), commenting on Haspelmath's criterion of subordination, writes: "Finally, there is the question of subordination. Converbal clauses often do show characteristics of subordination, such as centre-embedding, extraposition, extraction, backwards pronominal anaphora and control, as well as focusing. Yet many converbs lack signs of subordination. For example, the Hindi-Urdu anterior-modal converb, which has all the other features of subordination, cannot be focused (asserted, negated or questioned) per se, i.e. on its interpropositional (inter-clausal semantic) relation, except in its reading as a manner adverbial:

Hindi (Davison 1981: 109)

??Kyaa	*vahdoost-õõ*	*see*	*mil-kar*	*deer*	*see*	*aa-y-aa?*
Q	s/he friend-OBL.PL	with	meet-CVB	lateness	with	come-PAST.MSG

'Did he come late [because of] having met his friends?'

c-commanded by its antecedent (postcedent), which is the case only in subordinate constructions." That is, the postcedent, an expression that gives a meaning to the pronoun, dominates the meaning of a pronoun.

(16) English (Haspelmath 1995: 14)
a. Talking to him, she solved all of Pedro's problems.
*b. *She talked to him$_i$ and she solved all of Pedro's problems.*

In (16a) *Pedro* is a postcedent of the pronoun *him* and c-commands it, in other words, the pronoun *him* of the subordinate clause 'Talking to him' refers to Pedro. In b. we have two coordinate clauses. There is no reference between Pedro and the pronoun *him*, and thus these two components are considered different entities.

Haspelmath adds that the referential control of the implicit converb subject also depends on c-command, and consequently is possible only with preceding subordinate clauses. He gives an example of backwards control of the implicit-subject clause and states that backwards control is possible only in subordinate clauses.

IV. Restrictiveness and focusability

Haspelmath (1995: 15) says:

> "Only subordinate clauses, but not coordinate clauses, may be interpreted restrictively (cf. Tikkanen 1987b), i.e., as modifying the main clause in such a way that its reference is narrowed. Since restrictiveness is a prerequisite for focusing, only subordinate clauses may be focused. Various types of focusing occur with converbs and show that they are indeed subordinate."

As he points out, converbal clauses may be focused by focus particles like *also* and *only*. Converbal clauses may be the focus of a polar question (yes-no question) and of focusing negation. What is more, converbal clauses may be the focus of cleft constructions.

V. Possibility of extraction

Haspelmath (1995) after Ross (1967) notices that coordinate structures severely restrict the possibility of extraction:

(17) English (Haspelmath 1995: 16)
a. Alexis sold his car and bought a bicycle.
*b. *What did Alexis sell his car and buy?*

whilst subordinate clauses do not:

(18) English (Haspelmath 1995: 17)
a. After he sold his car, Alexis bought a bicycle.
b. What did Alexis buy after he sold his car?

The same concerns converbs:

(19) English (Haspelmath 1995: 17)
a. What did Alexis buy, having sold his car?
b. What, having sold his car, did Alexis buy?

Focusing now on converbs in Indo-Aryan, we find that early as well as more modern analyses of IA converbs argue for their subordinate status, in line with the common definition of the converb (Davison 1981, Subbarao 2012). For example, according to Davison (1981) there is one underlying syntactic structure for the Hindi converb, which is [+ subordinate] combined with the aspectual value [+ perfective]. She lists several arguments for the subordinate status of the converb, which pertain to non-finiteness, the coreference of the subjects of the main and converbal clause, and word order.

Examples (20) and (21) show the role of word order in converbal constructions as opposed to the coordinate clause. In (20b) the object *cāy* 'tea' can be moved to the left of the converb, without a noticeable change in meaning; whereas in the equivalent coordinate structure (21b) such movement would produce an ungrammatical sentence (since *cāy* cannot be the argument of the verb *baiṭhā* 'sit').

(20) Hindi (Davison 1981: 107)
a. *akele* *baiṭh-kar* *cāy* *pī* *thī*
alone sit-CVB tea drink be.PST.F
(i) 'He used to sit alone and drink tea'
(ii) 'Having sat alone, he used to drink tea'
b. *cāy* *akele* *baiṭh-kar* *pī* *thī*
tea alone sit-CVB drink be.PST.F

(21) Hindi (Davison 1981: 107)
a. *akele* *baiṭhā* *aur* *cāy* *pī* *thī*
alone.OBL sit.PST and tea drink be.PST.F
b. **cāy* *akele* *baiṭhā* *aur* *pī* *thī*
tea alone sit-PST and drink be.PST.F
'He used to sit alone and drink tea.'

They also share many features with subordinate constructions and non-clausal adverbials. As a result, converbal constructions have been interpreted as lying between subordination and coordination (cf. Davison 1981: 105). As we can see in (22), at least two possible interpretations are acceptable. Even though Davison maintained the view that the IA converb is a subordinate device, she observed that the scope of question and negation in converbal constructions differs from that in subordinate and coordinate clauses, and that its blocking or extension is not based on any syntactic or semantic principles (Davison 1981: 108-116). That is

the reason why example (23) is acceptable, and the scope of the negation depends on pragmatic factors. According to Davison (1981: 116), it is exclusively world knowledge which blocks sequential reading and permits antithetical meaning, whereas Kachru (1981: 41) clearly denies the pragmatic motivation of antithetical meaning.

(22) Hindi (Davison 1981: 105)
us-ne *soc-kar* *kām* *kiyā*
s/he-ERG [think-CVB] work do.PST
'Having thought s/he did the work.'
'S/he did the work after thinking.' or
'S/he did the work carefully.'

(23) Hindi (Davison 1981: 115), (Kachru 1981: 41)
vah *patr* *na* *paṛh-kar* *bolā*
s/he letter not read-CVB speak.PST
'He spoke after not reading the letter./Instead of reading the letter, he spoke.'

We have already seen that the subordinating status of converbs is not the only possibility. One of the major functions of converbs is chaining, which from the syntactic perspective is closer to coordination, and as was mentioned above, is also a dominant feature of the so-called Asian converb. This brings us closer to the problem of the traditional dichotomy between coordination and subordination. The dichotomy has been challenged by several methodological schools. One of them is Role and Reference Grammar (henceforth RRG) as proposed by Foley and Van Valin (1984), Van Valin and LaPolla (1997), and Van Valin (2004).

RRG takes the notion of the layered structure of the clause as a point of departure. Thus, the *nucleus*, *core* and *peripheral* levels are distinguished. The *nucleus* is the predicate, the *core* consists of the nucleus and the arguments of the predicate, and the *periphery* consists of non-arguments (e.g. locative or temporal phrases). They all constitute components of the clause (see Fig. 5).

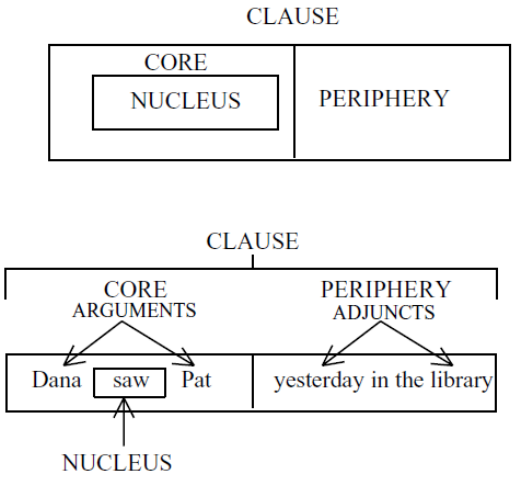

Figure 5. Clausal components (Van Valin 2005: 4).

Each level can be modified by one or more operators, which are grammatical categories such as aspect, tense, modality, etc. (see Table 3 for a summary).

Nuclear operators:
Aspect
Negation
Directionals (only those modifying orientation of action or event without reference to participants)
Core operators:
Directionals (only those expressing the orientation or motion of one participant with reference to another participant or to the speaker)
Event quantification
Modality (root modals, e.g. ability, permission, obligation)
Internal (narrow scope) negation
Clausal operators:
Status (epistemic modals, external negation)
Tense
Evidentials
Illocutionary force

Table 3. Nuclear, core and clausal operators (Van Valin 2005: 9).

Further, various levels can be joined, resulting in so called 'junctures' (nuclear, core and clausal junctures) (Foley and Van Valin 1984: 187-197; Van Valin and La Polla 1997: 442-448; Van Valin 2005: 188-198). A nuclear juncture assumes a single core with multiple nuclei – this can be best illustrated by a serial verb construction or a complex predicate construction[10] (Fig. 6). In a core juncture there is a single clause consisting of multiple cores (Fig. 7) each with its own nucleus. Finally, in a clausal juncture, clauses are joined.

[10] A complex predicate is a predicate that consists of a verb, a noun or an adjective (being the main predicational element) and a light verb, i.e. a verb which carries inflectional markers, is homophonous with a fully lexical verb, but does not convey the lexical meaning in the way the main verb does, but rather modifies the semantics of the main verb by such meanings as completion, inception, benefaction, etc. (cf. Butt and Lahiri 2013).

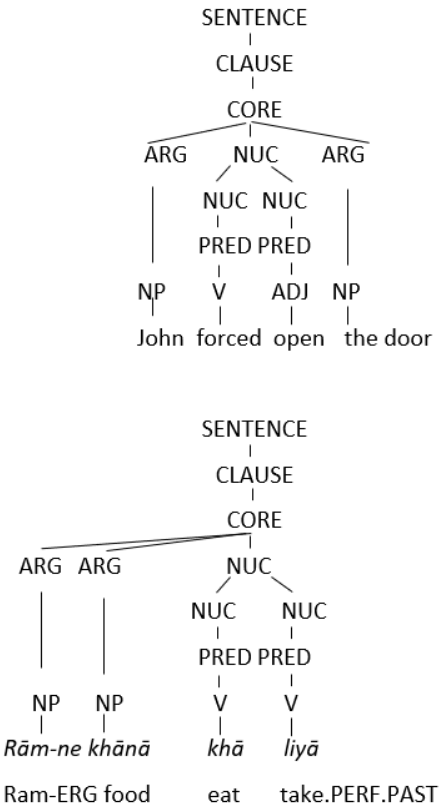

Figure 6. Nuclear juncture in English and Hindi.

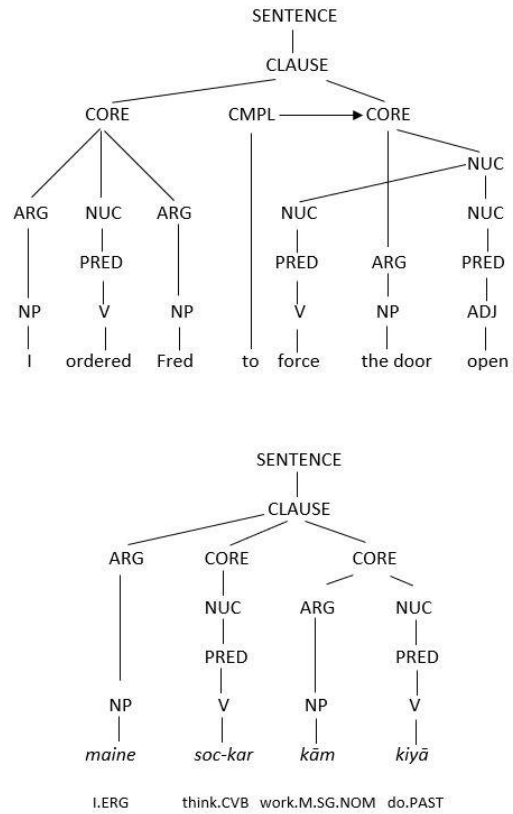

Figure 7. Core juncture in English and Hindi.

The type of syntactic relations between the units is defined in terms of 'nexus relation'. On this level, we find a division into three types of clause structure, i.e. coordination, subordination and cosubordination.

The notion of 'cosubordination' stems from Papuan linguistics (cf. Olson 1981), and occupies a crucial place in the RRG theory of clause linkage. It is characterized by the features of dependence and non-embeddedness. Switch reference constructions, which are dependent but non-embedded, can serve as an exemplification of 'operator sharing'. For instance, cosubordinated structures in the Papuan language Amele (24a) share the T (tense) operator – the tense marker is only on the main verb 'hit', but its scope is extended over the verb 'run out', whereas in the coordinate structures (24b) each conjunct is specified for tense. In other words, the main difference between cosubordination and coordination is that the former exhibits a type of dependency whereas the latter does not. The difference between cosubordination and subordination lies in the type of dependency. Subordination assumes structural dependency, i.e. the embedded clause functions as a main argument or a modifier (24c). On the other hand, cosubordination is based on the dependency of an operator (Van Valin and La Polla 1997: 454; Van Valin 2005: 183).

(24) Amele (from Van Valin 2005: 185-186)
a) switch reference = cosubordination
Ho busale-ce-b dana age qo-ig-a
pig run.out-DS-3SG man 3PL hit-3PL-TPST
'The pig ran out and the men killed it?'

b) coordination
Fred cum ho-i-an qa Bill uqadec h-ugi-an.
Fred yesterday come-3SG-TPST but Bill tomorrow come-3SG-FUT
'Fred came yesterday, but Bill will come tomorrow.'

c) subordination
I ja ja hud-ig-a eu nu, uqa sab mane-i-a.
1SG fire open-1SG-TPST that for 3SG food roast-3SG-TPST
'Because I lit the fire, she cooked the food.'

Cosubordination operates at all three levels of juncture, namely nuclear, core and clausal. At each juncture level there are operators of which at least one should be shared (Van Valin 2005: 201). Operator sharing is hierarchized, that is, operator sharing at a certain level of juncture assumes sharing of higher-level operators as well, but the reverse is only possible if it does not violate the semantics of the predicate (Van Valin and La Polla 1997: 455).

In (25) and (26) we see examples of nuclear and core cosubordination from Barai and Turkish respectively, with the operator projection given below.

(25) Barai nuclear cosubordination (Van Valin and LaPolla 1997: 456-457)
Fu *kai* *fu-one* *kume-fie* *va.*
3SG friend 3SG-GEN call-listen continue
'He continued calling and listening for his friend.'

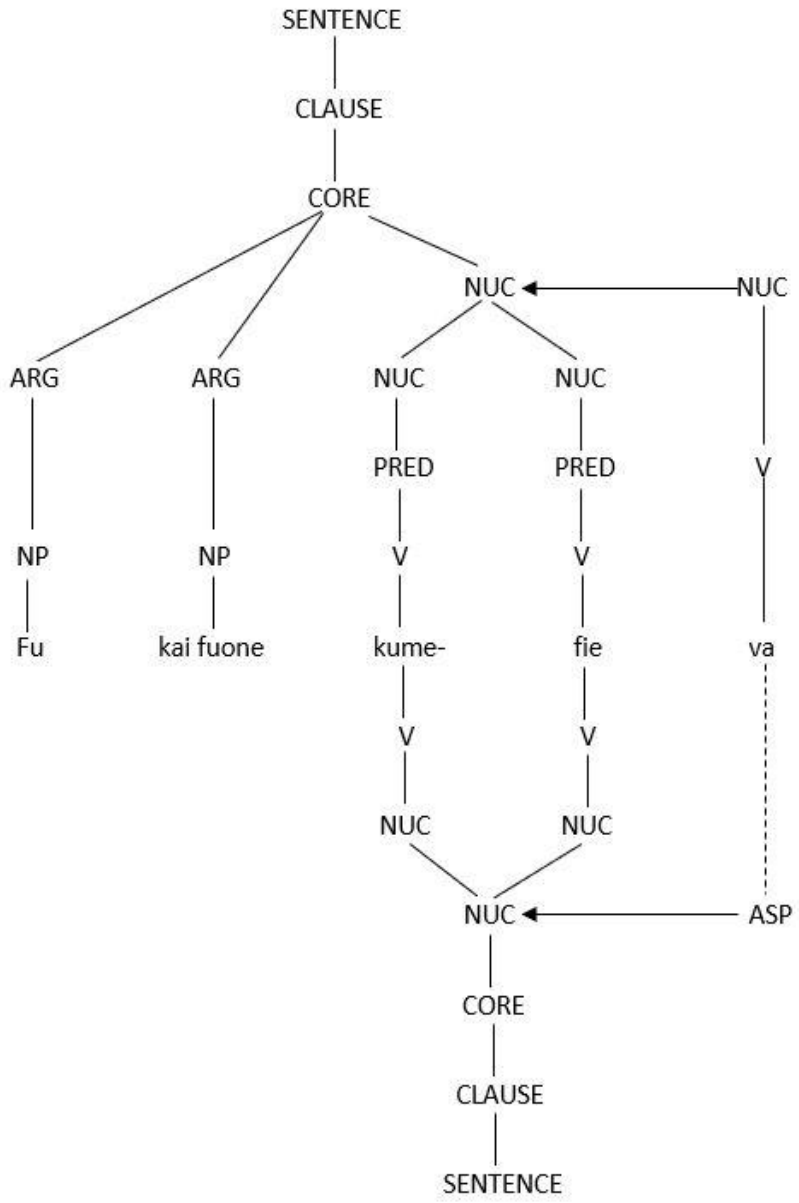

Figure 8. Operator sharing in Barai nuclear cosubordination

(26) Turkish core cosubordination (Van Valin and LaPolla 1997: 460-461)
Gid-ip *gör-meli-yiz.*
go-CMPL see-MODAL-PL
'We ought to go and see.'

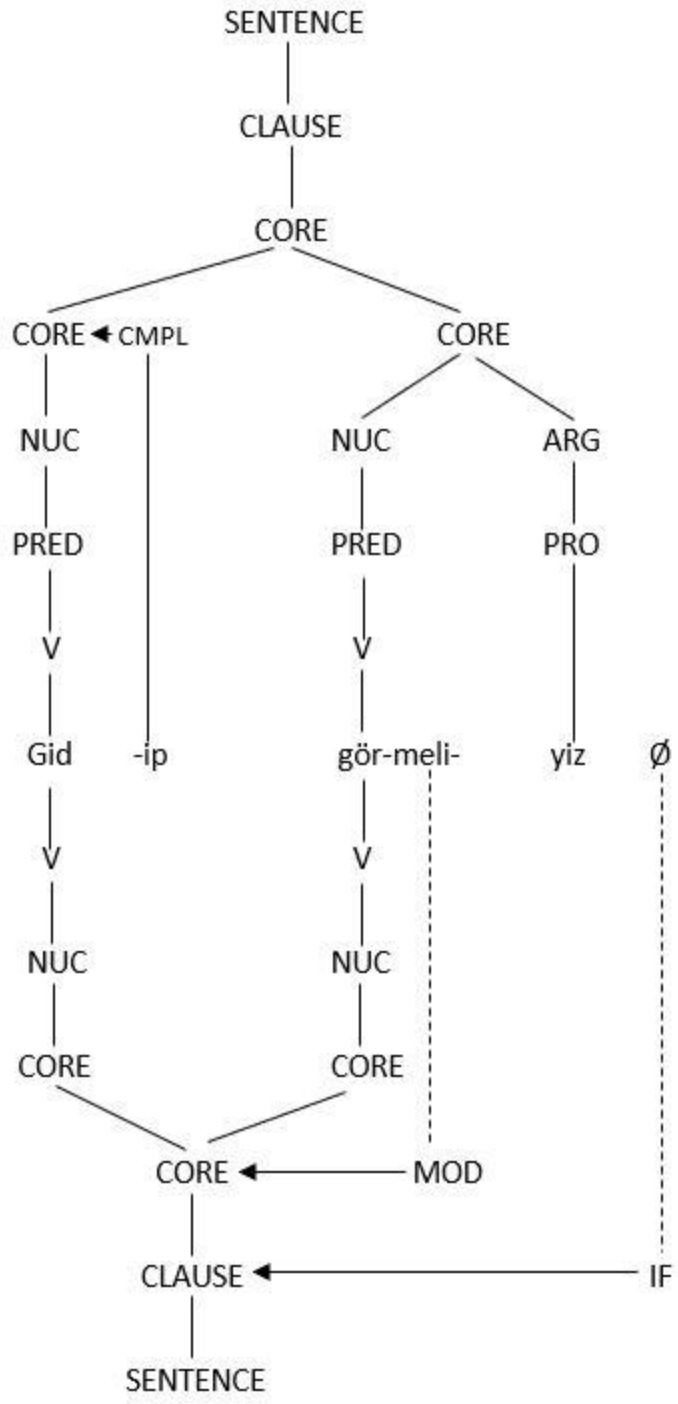

Figure 9. Operator sharing in Turkish core cosubordination.

From the RRG perspective, converbs can be a part of constructions which instantiate different types of cosubordination, namely nuclear, core and clausal.

We saw an example of a nuclear juncture in Fig. 6, which is an instantiation of nuclear cosubordination – the auxiliary verb carries the aspectual marker which is shared by the main verb. The bare stem, which is homophonous with the short converb, is a part of the light verb construction. In Figure 10 we give the operator projection.

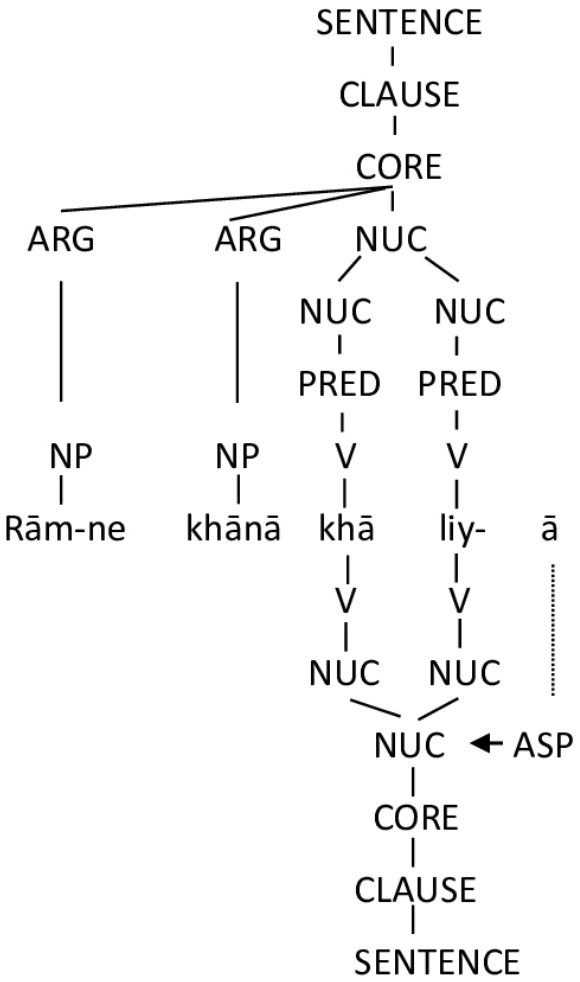

Figure 10. Operator sharing in Hindi nuclear cosubordination.

The IA converbal construction of the type represented by (27) is an instantiation of core cosubordination. There is a semantic difference between the converbal core of the sort exemplified in (27) and in (28), namely the adverbial character of the former and the clause-like character of the latter. Therefore only the latter is a converbal chain construction. Figures 11 and 12 illustrate the operator projection for (27) and (28) respectively. In (27) there are two cores sharing the core operator MOD; in (28) there are two clauses sharing the clausal operators IF and Tense.

(27) Hindi

niveśak-õ ko soc-samajh-kar bājār mẽ paisa lagānā cāhiye

Investor-PL.OBL DAT think-understand-CVB market in money invest should

'Investors should invest money in the market carefully (lit. having thought-understood).'

(28) Hindi

to tum-hī jā-kar dekho na?

CONJ you-EMPH go-CVB see.IMP.2PL not

'So you go and see, OK?' (Premcand, Kafan, 8)

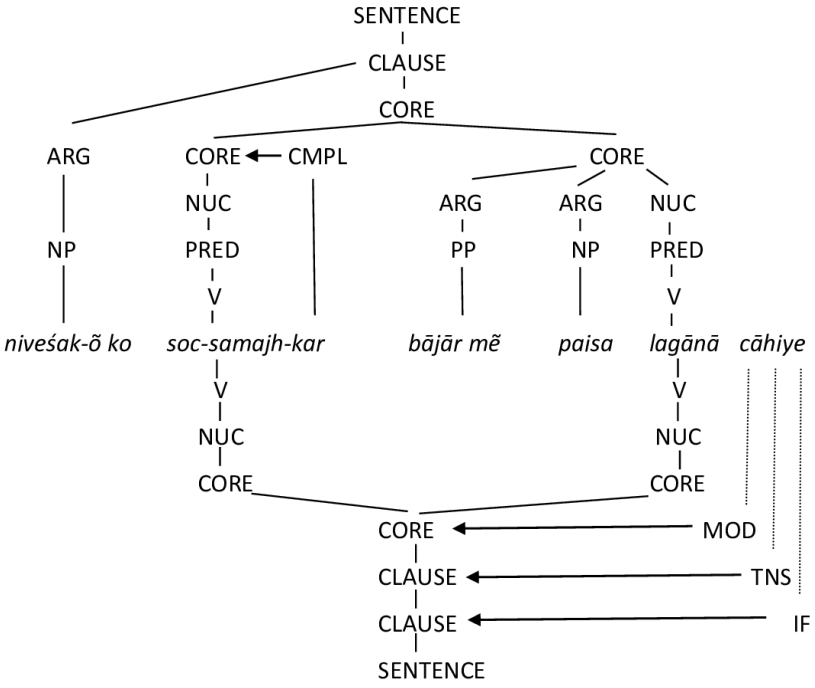

Figure 11. Operator sharing in Hindi core cosubordination.

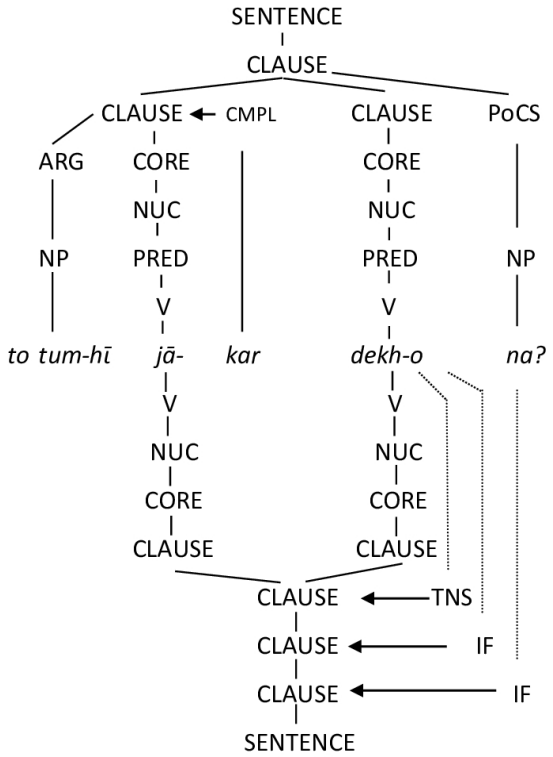

Figure 12. Operator sharing in Hindi clausal cosubordination.

However, there appear to be cases of cosubordination in which operator sharing is possible but not obligatory. For example, in (29) the first clause consists of a converbal core, and the second has a finite verb. The IF operator may have scope over the superordinate clause, over the converbal clause and over two clauses. This severely weakens the notion of

co-subordination, but Van Valin (2007: 80) sees no reason to treat such sentences as instantiations of all three nexus relations depending on the possibility of operator sharing.

(29) Nepali

phalphul	*tāch-era*	*nānī-lāī*	*di-ūṃ?*
fruit	peel-SEQ	little.girl-OBJ	give-INJ.1S

'After I peel the fruit, should I give it to the child?'
'Should I peel the fruit, before I give it to the child?'
'Should I peel the fruit and give it to the child?' (Peterson 2002: 105)

Tikkanen (1987), when discussing the syntactic properties of converbs, argues:

> "[...] it is not enough to define a non-finite form or construction in terms of its temporal, aspectual, interpropositional and coreferential features. Without knowing the inherent constraints on its elliptic and contrastive operational integratability we do not know the conditions under which it may paraphrase or translate finite subordinate and/or coordinate clauses" (Tikkanen 1987: 34).

His approach assumes an analysis of syntactic, semantic and pragmatic properties of converbs which is in line with RRG. However, according to Tikkanen, the concept of cosubordination is not a fully applicable notion for Indo-European languages, where, in contrast to Papuan languages, morphological marking of switch reference is lacking. He argues that the indeterminacy of non-finite constructions as to the parameters of embedding and dependency results from their discourse function (Tikkanen 1987: 28).

Since converbal chain constructions represent a type of linkage which shares features of subordination and coordination, a more fine-grained solution to this typological problem was proposed by Bickel (2010) in his multivariate analysis model. This model assumes that there is no discrete notion of cosubordination, and the type of non-embedded and dependent relation can be defined by a number of features. Bickel takes into consideration a number of features including Illocutionary scope (ILL scope – henceforth IF), Tense scope (T-scope), Finiteness, Illocutionary force marking (ILL-mark), Tense marking (T-mark), Symmetry, Wh, Extraction, FOC, Position, and Layer. However for the purpose of the present study we have selected only the scopes of IF, T and NEG, modifying slightly Bickel's proposal (Bickel 2010: 56-62; 81):

1. the scope of the IF operator can be: a) conjunct – extends to the main and the dependent clause; b) disjunct – extends to either the main or the dependent clause but never to both; c) local – limited to the main clause; d) extensible – extends to either the main clause alone or to both the main and the dependent clause but never to the dependent clause alone; e) constraint-free – not regulated by the linking type;
2. the scope of the Tense operator can be: a) conjunct – extends to the main clause and the dependent clause; b) local – limited to the main clause; c) extensible: extends to

either the main clause alone or to both the main clause and the dependent clause, but never to the dependent clause alone;

3. the scope of the Negation operator can be: a) conjunct – extends to the main and the dependent clause; b) local – limited to the main clause; c) variable – either limited to one of the clauses or conjunct; d) extensible – extends to the main or the dependent clause.[11]

A comment needs to be made with respect to the possibility of extension of NEG scope in the context of modern IA. The so-called 'NEG-transport' (Bickel 2010: 59) indicates that in a structure which looks like a cosubordinate one, NEG scope is local, but it has been 'transported' from the main to the converbal clause. Transport in the opposite direction seems also to be present in early NIA, and this will be discussed later in section 4.3. In example (30), clearly NEG is transported to the dependent clause. The operator projection for (30) is given in Figure 13.

(30) Hindi (Subbarao 2012: 292)

ye	*bacce*	*cal-kar*	*nahĩ*	*āye*
these	children	walk-CVB	not	come.PST.M.PL

'These children did not come by walking.'

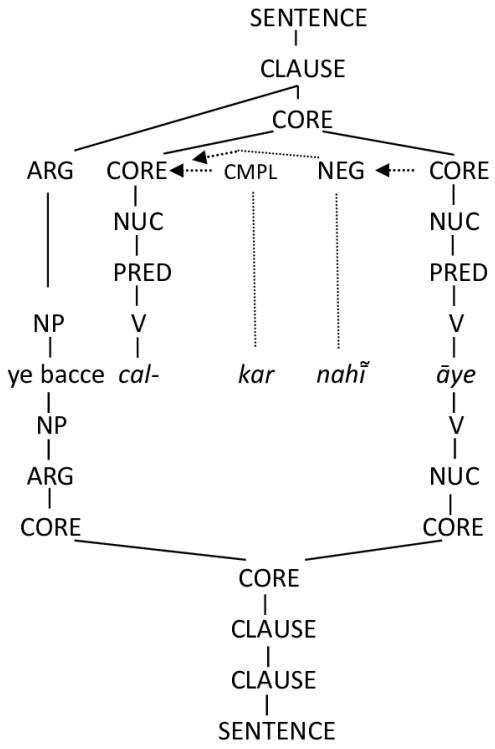

Figure 13. Operator projection in Hindi clauses with transportable scope.

[11] Bickel (2010: 58-60) actually discusses all options regarding the scope of negation, but the kind of NEG scope given here as extensible is restricted, in his view, to the option of extension to the dependent clause. Since we have come across examples with extension from the dependent to the main clause, we give that option here.

2.2.3 FORMAL MAKE-UP OF CONVERBS

Converbs are said to be usually marked by an affix on a verb stem (Coupe 2006: 148). Suffixes are generally preferred over prefixes, since converbs are found predominantly in verb-final languages, where suffixation is favoured. Nevertheless, prefixes (in Burushaski for example) and circumfixes (in Chukchi) are used as well. "A rare example of a non-affixal converb is provided by Ge'ez. Where converbs are formed by the vowel pattern *CaCiC* [...]" (Hasplemath 1995: 9).

Non-affixal particles may serve as converbal markers as well, "e.g., French *en* in the French *gérondif* (e.g., *en chantant* 'singing')" (Haspelmath 1995: 9). Converbal forms may be fully reduplicated, as in Turkish or Indo-Aryan.

(31) Turkish (Haspelmath 1995: 9)

Insan	*demir-i*	*döğ-e*	*döğ-e*	*demirci*	*ol-ur.*
person	iron-ACC	forge-CVB	forge-CVB	smith	become-AOR

'A person becomes a blacksmith by forging.'

(32) Assamese from Subbarao (2012: 243)

rel-oloi	*ro-i*	*ro-i*	*ami*	*bhagori*	*pɔrilõ*
train-to	wait-CVB	wait-CVB	we	be tired	felt

'Waiting for the train we got very tired.'

As regards the place of occurrence of converbs in a sentence, they usually precede their matrix clauses (33). Nevertheless, converbal clauses can occur within the matrix clause (34) and can also follow the matrix clause (35).

(33) Korean (Haspelmath 1995: 2)

Achim	*mek-ko*	*hakkyo*	*ey*	*kassey*	*yo.*
Breakfast	eat-CVB	school	to	went	PT

'I ate breakfast and went to school.'

(34) Hindi (Subbarao 2012: 265)

ham-ne	*kamre*	*mẽ*	*baiṭh-kar*	*choṭe*	*baccõ-ko*	*dekh-ā*
we-ERG	room	in	sit-CVB	small	children-ACC	see-PST.M.SG

'We saw the small children while we were sitting (seated) in the room.'

(35) Russian (Haspelmath 1995: 13)

Xèvgun	*načal*	*novoju*	*žizn',*	*vernu-všis'*	*domoj.*
Khevgun	began	new	life	return-PERF.CVB	home

'Khevgun began a new life (after) returning home.'

In IA converbs have exclusively affixal (suffixal) markers of various origin. For example, the element *-ī/-i* has been ascribed different derivations by various scholars. The standard view, maintained even by some recent grammars, is that converbs are derived from the Sanskrit converb terminating in *-ya* via Prakrit *-ia* and Apabhramśa *-i*, e.g. OIA *a-karya* > MIA *a-karia* > Ap. *a-kari* 'not having done' (cf. Bloch 1965: 281-282; Oberlies 2005: 44). However, this view was challenged quite early by Tessitori (1915: 119), who claimed that in Early Rajasthani converbs are actually locative forms of the past passive participle, e.g. LOC *karī* < *kariu* 'done'. Similarly, other Early Rajasthani postpositional extensions which have also been employed as converbal markers such as *naï* and *karī* are themselves locative forms. Hindi *-kar* has its roots in the Apabhramśa converb *-kari* and presumably its shortened form *-ke* via apocope and vowel lengthening (in Early Braj or Early Awadhi *-kai*) as well (see Oberlies 2005 for a different opinion deriving shortened forms from the OIA locative form of the past passive participle *kṛte* < *kṛta* 'done'). The postposition *naï* could be a shortened form of *kanai* < *kanhaiṁ* < Apabramśa DAT/LOC *kaṇṇahiṁ* < OIA LOC **karṇasmin* (=*karṇe*) 'near' (Tessitori 1914: 226). Contemporary dialects have the following continuants: in Marwari *kənɛ*, in Mewari *kənɛ̃*, in Bagri *kəne* LOC, in Kumaoni *kəɳI, gəɳi, khəɳi* and in Garhwali *kʊɳi*, *khʊɳi* DAT/ACC, in Magahi *ne*, *kane* LOC, in Bhatri *āne* LOC, and in Halbi *ne* LOC, Nepali *kana* (cf. Tessitori 1914: 226; Telang 1966: 413, 493; Tripāthi 1993: 182). There are possibly also other sources of converbal markers, e.g. Marathi *-ūn*, derivable from the Prakritic rare form-*tūnaṃ* < OIA *-tum* or Kashmiri *-th* from OIA *-tvī* (cf. Chatterji 1927: 1009; Bloch 1965: 282).

Saksena (1937: 282) takes two possible sources for the converbal marker *-i* in Early Awadhi, namely the converbal marker in OIA *-ya* < MIA *-ia* or the OIA infinitive *-tum* < MIA *-iuṃ*.

Magadhan languages have the *-n-* feature, e.g. Maithili *-ne* from Old Maithili *-le*, whose origin is traced to the past passive participle in LOC (Jha 1958: 515-516). Bengali has *-e* and *-le,* the former having its roots in the OIA converbal affix *-iya* and the latter being the LOC of the past passive participle. Sinhala *-la* also has its origin in the past passive participle. There are also other suffixes serving as converbal markers, such as Nepali *-era* or Early Rajasthani *-ara*, of doubtful origin – possibly a shortening of the *-kara* form. Finally, the converb can be a bare root form.

2.2.4 REFERENTIAL CONTROL

According to previous research, converbal chain constructions exhibit different possibilities as regards the control properties of converbs (cf. Coupe 2006: 148). Many languages allow the subject of the converb to be implicit:

(36) Knowing the truth, she ran away.

When the subject of a converbal construction is implicit, then generally it is referentially controlled by the subject of the superordinate clause (*subject control*). In such a situation, we have a same-subject converb. Some languages distinguish between same-subject converbs

and different-subject converbs. Different-subject converbs imply the existence of same-subject converbs. In languages having a system contrasting same- and different-subject forms, we have the switch reference system (cf. 24a).

IA converbs are characterized by the Subject Identity Constraint (henceforth SIC), according to which the subject of the matrix clause must be identical with the subject of the converbal clause (see (37) with the main verb transitive and ERG subject, and (38) with the main verb intransitive). Therefore, examples of the type (39) are not acceptable (Davison 1981: 106; Subbarao 2012: 272-273).

(37) Hindi (Kachru et al. 1976: 90)

ghar	*ā-kar*	*mohan-ne*	*khānā*	*khā-yā*
home	come-CVB	Mohan-ERG	food	eat.PST

'Mohan ate (after) coming home.'

(38) Hindi (Kachru et al. 1976: 90)

patr	*paṛh-kar*	*k̲huś*	*ho*	*ga-yā*
letter	read-CVB	happy	be	go-PST

'He became happy (after) reading the letter.'

(39) Hindi

**tum*	*so-kar*	*us-ne*	*naśtā*	*ki-yā*
you	sleep-CVB	s/he-ERG	breakfast.M.SG	do.PST.M.SG

*'You having slept, s/he had breakfast.'

There are, however, exceptions to the SIC rule. Tikkanen (1995: 496) states that restrictions on coreference are predominantly of semantic or pragmatic but not morphosyntactic origin. Thus, even if there is no syntactic subject identity, there is a possessor or experiencer-like relation between the subject of the main clause and the subject of the converb.

Moreover, in case of violation of SIC, the converbal clause will be propositionally restrictive (temporal, causal, etc.) rather than non-restrictive (additive-sequential). This, of course, has direct implications for the scopal properties, which will be discussed at length in chapter 4.

According to Subbarao (2012: 274), in several IA languages as well as in all Dravidian, some Tibeto-Burman and the Munda languages, SIC can be violated when the subject of the converb is inanimate and the converb denotes a non-volitional act, as in example (40). In such cases the converbal subject can be overtly expressed.

(40) Assamese (from Subbarao 2012: 276)

bərəxuni	*pəry-i*	*xəisyəborj*	*baṛh-il*
rains	fall-CVB	crops	grow-PST

Literally: 'Rains having fallen, the crops grew (well).'

Subarao (2012: 277-279) notes also that in South Asian Languages (henceforth SAL), SIC can be violated even when the subject of the converb is animate and there is a relation of cause and effect or reason between the matrix and converbal clauses. This has been attested in some Dravidian, Tibeto-Burman and Munda languages. There are Indo-Aryan languages which seem to be quite strict as regards SIC, such as Hindi-Urdu and Punjabi, while others have relaxed the SIC rule, e.g. Nepali, Kashmiri and Sinhala (Gair and Paolillo 1997: 49; Henadreerage 2002: 82-85; Subbarao 2012: 279).

Examples (41-42) can be interpreted as cause and effect relations. However, example (43) from Kumaoni rather presents a chaining construction, which is quite anomalous in IA.

(41) Nepali (Peterson 2002: 12)

rām-le	*bhan-era*	*us-le*	*tyo*	*kitāb paḍh-yo*	*ki?*
Rām-ERG	say-SEQ	3S-ERG	that	book read-PST.3SG	Q

'Did he read that book because Rām told him to (and not because he wanted to)?'
'Since Rām told him to, did he read that book?'
'Did he read that book and did Rām tell him to?'

(42) Sinhala (Henadreerage 2002: 85)

Kalyani	*gedərə*	*gihil-la*	*mamə*	*kæmə*	*kæw*
Kalyani.NOM	house.ACC	go-CVB	I.NOM	food.ACC	eat.PST

'Kalyani went home and I ate./With Kalyani having gone home, I ate.'

(43) Kumaoni (Pant 2006: 58)

etuk	*nān*	*bhau-kaĩ*	*choṛi-ber*	*teri*	*sās-ul*	*tu-kaĩ*
so	small	infant-ACC	leave-CVB	your.F	mother-in-law-ERG	you-ACC
kilai	*lag-ā*					
why	send.PST					

'Why did you leave your infant and why did your mother send you here?'
Lit. (You$_i$) having left such a little infant, why did your mother-in-law$_j$ send you$_i$ (here)?

What is more, Haspelmath (1995) – and Coupe (2006) after him – claim that it is possible "for the implicit subject of the converb to be coreferential with a salient referent that is external to the sentence." In the French example (44), the subject of the converb in the second sentence is coreferential with the subject of the first sentence.

(44) French (Halmøy 1982: 179, cited in Haspelmath 1995: 35)

Il	*pensa*	*une*	*seconde*	*que c'*	*etait*	*sans*	*doute*	*cela*	*qui*
He	thought	a	second	that it	was	without	doubt	that	which
l'	*avait*	*sauvé,*	*lui,*	*trois*	*mois*	*plus*	*tôt,*	*mais*	*en*	*même*
him	had	saved	him	three	months	more	early	but	at	same

temps, il cherchait un moyen de lui prouver le contraire.
time he sought a means to him prove the opposite
En y réfléchissant, c' était elle qui dès le début
CVB about.it think.CVB it was her who from the beginning
de leur liaison avait pris toutes les initiatives...
of their relationship had taken all the initiatives
'He thought for a second that that was perhaps what had saved him three months earlier, but at the same time he was looking for a means to prove the opposite to him. Thinking about it, it was she who had taken all the initiatives from the beginning of their relationship...'

The fact that the converb control is not necessarily determined by a syntactically defined pivot has been noted previously by several scholars. Among the factors determining control of the converbal subject there are semantic as well as pragmatic ones, for example experiencer- or possessor-like relations between the matrix clause subject and converbal subject, and pragmatic salience of the converbal subject (see e.g. Coupe (2006: 149); see also Tikkanen (1995: 496), who uses the term *constructio ad sensum*).

2.2.5 CONVERBS VS. PARTICIPLES

It has been noted in the literature that some modern NIA languages have two syntactic devices which are functionally not much differentiated, namely the converbal chain construction and a construction based on the past perfect participle (PPP). If we compare (34) with (45), we see in (34) the clear coreference of the converbal and matrix clause subjects, whereas in (45) the participial clause can refer either to the matrix clause subject 'we' or to its object 'children'. Subbarao (2012: 268-272) tries to find a functional explanation for the coexistence of these two parallel constructions, stating simply that converbal chains are subject-oriented whereas participial constructions are not (and the referential control depends on the position of the participial clause).

(45) Hindi (Subbarao 2012: 271)
ham$_i$ ne choṭe baccõ$_j$ ko dekh-ā kamre mẽ baiṭhe hue
we erg small children ACC see.PST room LOC sit.PPP.OBL be.PPP.OBL
1. 'We saw the small children while we were sitting (seated) in the room.'
2. 'We saw the small children while they were sitting (seated) in the room.'

One can also notice that some of the above examples show the overt subject of the converb (41-42). This type of construction resembles an absolute construction (henceforth AC). A truly canonical AC can be defined as a subordinate temporal clause containing a head noun and a participle in an oblique case; the participle has a different subject from that of the main clause (cf. Bubenik 1998). However, Bauer (2000: 300) has convincingly demonstrated that classical languages such as Latin may also exhibit coreference of the

subject of the participle and the main verb. In the most recent monograph on AC's, Ruppel (2013: 29) does not take non-coreference of the subjects as a defining feature of AC's, stating that "absolute constructions are temporal expressions with non-temporal heads". Since AC's are attested in many branches of IE (e.g. in Sanskrit, Latin, Greek and Old Church Slavonic), some scholars have speculated, though with some caution, that they may be inherited from the protolanguage (Bauer 2000: 331-332; Ruppel 2013: 207-216).

Old Indo-Aryan (Sanskrit) attests two types of AC's: the locative AC (46) and the much rarer genitive AC (47). The locative has survived through the MIA period (48). Early NIA continued to use AC's with the subject in an oblique case and the predicate in the form of an inflected imperfective or perfective participle.

(46) Vedic Sanskrit from Ruppel (2013: 127)

yát	*adayá*	*sū́rye*	*udyatí*	*príyakṣatrāḥ*	*ṛtám*
when	today	sun.LOC.SG	up-going.LOC.SG	dear-ruler.VOC.PL	rite.ACC.SG
dadhá					
set.2PL.PERF					

'when you, beloved rulers, appoint the rite at sunrise today' (RV 8.27.19)

(47) Classical Sanskrit from Ruppel (2012: 135)

vasatas	*tasya*	*Rāmasya*	*vane*	*vanacaraiḥ*
live.PPP.GEN.SG	this.GEN	R. GEN	forest.LOC.SG	forest creatures.INS.PL

saha	*ṛṣayo*	*'bhyāgaman*	*sarve*	*vadhāyāsurarakṣasām*
with	seer.NOM.PL	come.IMPF.3PL	all	for the killing of the asuras and rakṣasas

'While Rāma lived in the forest together with the forest creatures, all the seers came for the killing of the asuras and rakṣasas.' (Rāmāyaṇa 1.1.35)

(48) Pali from Peterson (1998: 144)

tehi	*ce*	*bhikkhave*	*bhikkhūhi*	*kate*
DEM.INS.PL	if	monk.VOC.PL	monk.INS.PL	make.PPP.LOC.SG

pavāraṇāsaṃgahe	*aññataro*	*bhikkhu*	*evaṃ*	*vadeyya:*
an abridged pavāraṇā.LOC.SG	certain.NOM	monk.NOM	thus	speak.3SG.OPT

'And if, oh monks, after monks have conducted an abridged pavāraṇā a certain monk should speak thus [...]'

Modern NIA also has a construction based on the past passive participle which can be regarded as an AC. A temporal expression in Hindi employing the head noun in the DAT and PPP in the oblique case is presented in (49). An interesting phenomenon has been observed in Bengali, where apart from the converb terminating in *-e* there is a conditional participle form in *-le*, which is clearly of participial origin and which may be functionally equivalent to an absolute construction, i.e., it allows the same subject (50a) as well as a non-same subject (50b) of the matrix clause and the converbal clause (see Chaterji 1927: 1003-1004).

(49) Hindi

mere	*dost*	*ko*	*mare*	*hue*	*sirf*	*ek*	*haftā*
my.OBL	friend	DAT	die.PPP.M.OBL	be.PPP.M.OBL	only	one	week
huā	*hai*						
be.PPP.M	be.3SG.PRS						

'It has been only one week since my friend died.'

(50) Bengali (Radice 2003: 170)

a. *porīkṣai pas kor-le (ami) notun saikel pabo*
exam pass do-CVB I new cycle get.1SG.FUT
'If I pass the exam, I shall get a new bicycle.'

b. *se e-le ami yabo*
s/he go-CVB I go.SG.FUT
'I shall go if s/he comes.'

In section 3.4 we will discuss in detail how the converbal chain constructions with overt converbal non-coreferential subjects, which are actually AC's, coexisted in the early NIA period with those AC's based on the PPP.

2.3 INFINITIVES

2.3.1 MAIN PROPERTIES

As in the case of converbs, infinitives have also been defined in terms of functional syntactic criteria. Infinitives are thus 'a nonfinite verb form used in the object function in complement clauses' (Nedjalkov 1998: 422). This view has been maintained in the typological literature since the late '80s (Haspelmath 1989; Koptjevskaja-Tamm 1993; Nedjakov 1995, to name a few), but apart from their primary complement functions they can also be used as adverbial modifiers in purposive constructions (see e.g. Ylikoski 2003). For example, (51a) shows an infinitive as object complement, whereas in (51b) the infinitive is not a complement but rather an adverbial modifier.

(51) from Noonan (2007: 52-53)
a) Zeke remembered to leave.
b) Roscoe hit Floyd to cause trouble.

In (51) the subjects of both infinitive clauses have been equi-deleted; that is, the subject of the infinitive clause, which is coreferential with the subject of the matrix clause, is not explicitly mentioned. There are also other options with infinitives, namely equi-deletion of an argument being coreferential with other arguments of the matrix clause, such as objects (52).

(52) from Noonan (2007: 76)
The woman forced the man to winnow the millet.

However, even IE languages may not apply the equi-deletion rule to objects, as in the case of Vedic (53):

(53) Vedic (V. 74.3d) from Disterheft (1980: 77)

vayáṁ	*vāṁ*	*uśmasi*	*iṣṭáye*
we	you.ENCL	wish.1PL.PRS	come quickly.INF

'We wish that you come quickly.'

Other syntactic peculiarities of infinitives include raising or matrix coding constructions. Traditionally they are of two types: raising to subject and raising to object. In Latin grammar these have been labelled Nominativus cum infinitivo and Accusativus cum infinitivo respectively. In RRG the subject raising construction receives the label 'matrix coding as PSA construction' and object raising the label 'matrix coding as non-PSA construction', where PSA stands for Privileged Syntactic Argument, which can be considered equal to 'syntactic subject' (cf. Van Valin and LaPolla 1997: 176; 562-576).

For purposes of exemplification we may consider (54), in which *ille* 'he' has been raised from the subject of the embedded clause to the subject of the matrix clause. In (55) *te* 'you' has been raised from the subject of the embedded clause to the object of the matrix clause.

(54) Latin (Catullus 51)

Ille	*mi*	*par*	*esse*	*deo*	*videtur*
He.NOM	me.DAT	equal	be.INF	god.DAT	seem.MED-PASS

'He seems to me to be equal to a god.' < It seems to me that he is equal to gods.

(55) Latin (Cicero Att. 1.18.6)

volō	*tē*	*hoc*	*scīre*
I want	you.ACC	it	know.INF.PRS

'I want you to know this.' < I want that you know this.

Despite the fact that infinitives are classified as non-finites in some IE languages, they can inflect for almost all verbal categories such as tense, aspect, voice and object agreement. However, infinitives are usually sensitive to a smaller number of categories than finite verbs. In the IE context infinitives can be inflected for tense and voice (e.g. Ancient Greek, Latin), as in Greek *paideúein/paideúsein* 'bring up.PRS.ACT/FUT.ACT', '*paideúesthai/paideúsesthai*' bring up.PRS.MED-PASS/FUT.MED-PASS', and for aspect (e.g. Slavonic), as in Polish *czytać/przeczytać* 'read.INF.IMPERF/PERF', but in some IE languages they seem to retain nominal features without having significant verbal characteristics (e.g. Sanskrit).

The common assumption is that the complement clauses of which the infinitive is a part are subordinate. However, in RRG, constructions with control verbs are not interpreted as subordinate, but as either cosubordinate or coordinate. Sentences like (56) are not subordinate, because in (56a) there is coreference of the subjects of the infinitive and the matrix clause and there is a core-level shared modality operator, whereas in (56b) there is no coreference and no sharing of the operator. Moreover, subordinate units should serve as core arguments, and this is not the case with both constructions – they neither cleft nor passivize (see Van Valin and LaPolla 1997: 459-462). The operator projection in Figure 15 shows that (56a) is an instance of core cosubordination, and (56b) of core coordination.

(56) English
  a) John must try to wash the car
  b) John must tell Bill to wash the car

Purposive clauses of the type (51b) also represent core cosubordinative nexus. As regards both types of the matrix coding construction (PSA and non-PSA), RRG considers them core coordinate nexus (Van Valin and LaPolla 1997: 561-576).

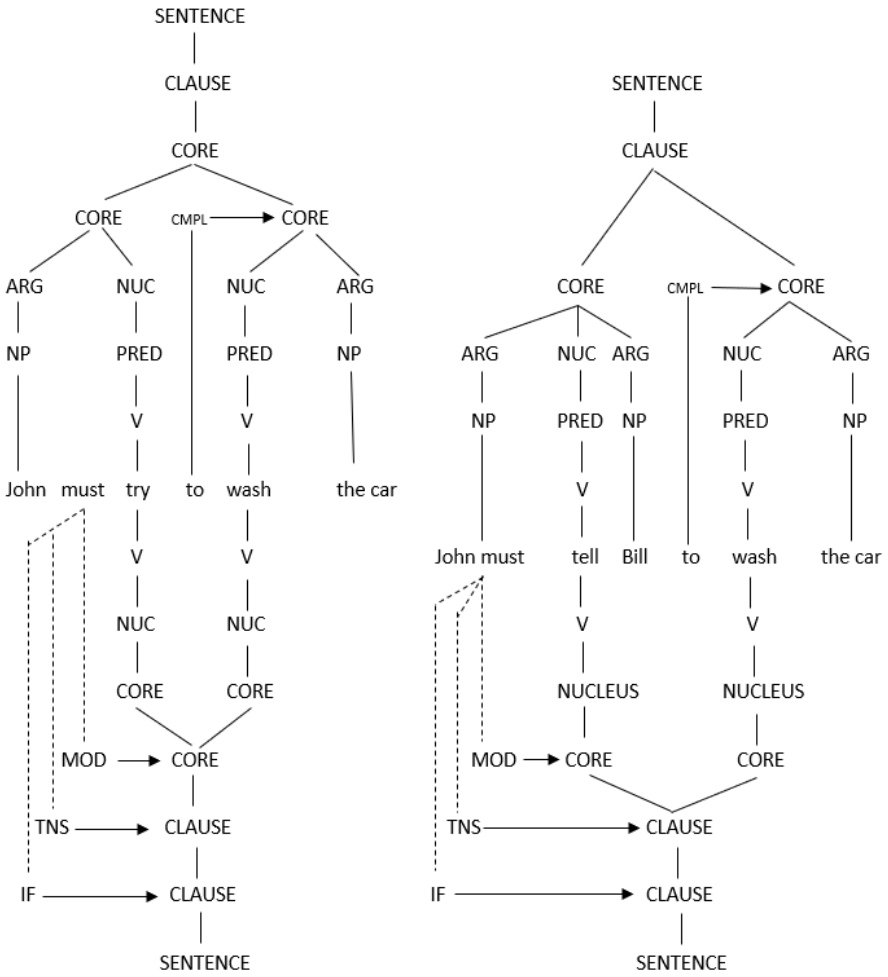

Figure 14. Operator projection in core cosubordination and core coordination.

In a Hindi-Urdu complement clause (57) semantically equivalent to English (56a) with the verb *try* and modal 'must',[12] we have an entirely different construction type based on the non-nominative pattern with DAT subject, but also have core cosubordination – the MOD operator has a scope over two cores; see the operator projection in Figure 15.

(57) Hindi

hamẽ	*is*	*paiṭarn*	*ko*	*samajhne*	*kī*	*pūrī*	*kośiś*	*karnī*
we.OBL	this.OBL	pattern	ACC	understand.INF.OBL	GEN.F	full	try	do.INF.F

cāhiye
should
'We should try to understand fully this pattern.'

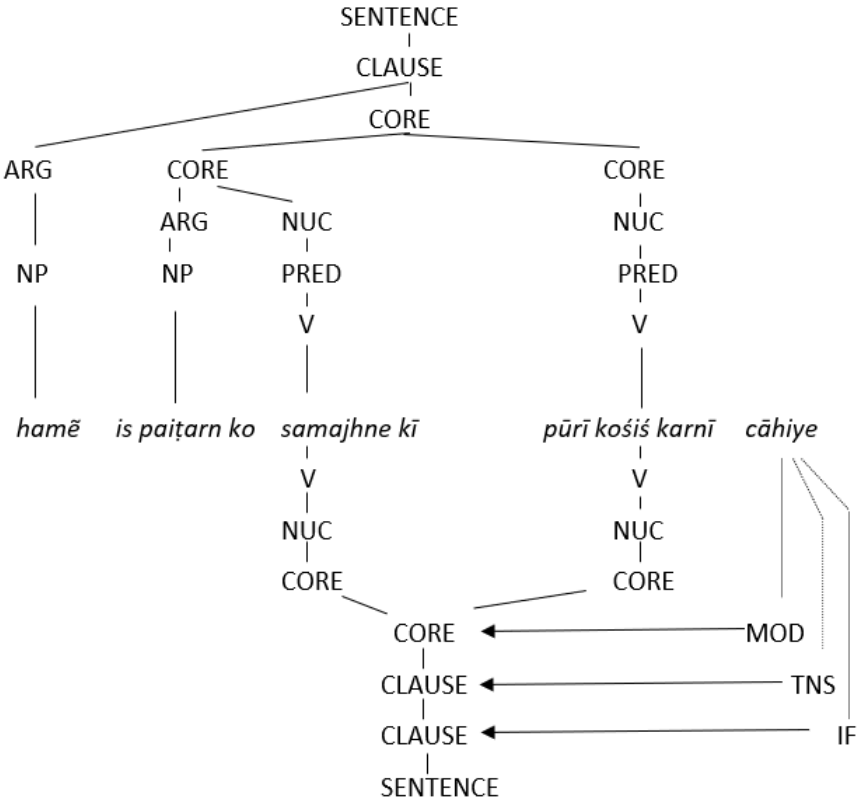

Figure 15. Hindi core cosubordination (obligative pattern).

A construction of the type (56b) with verbs like *tell* or *force* (*John told/forced Bill to wash the car*) representing core coordinate nexus is also available in Hindi-Urdu (58) with the operator projection given in Figure 16. However, it does not occur with modals which require DAT subjects – in a modal context the fact that the subject of the infinitive clause is different from that of the matrix clause would result in a construction representing clausal subordinative nexus.

[12] In Hindi-Urdu, modals such as *must* or *should* require DAT subjects, while *want* occurs with NOM subjects or with ERG in perfective tenses.

In purposive clauses in Hindi-Urdu we have core coordination – in (59) the MOD operator does not have scope over the infinitive clause, but it is limited only to the matrix clause, as the operator projection in Figure 17 demonstrates. The matrix coding as PSA construction is in fact not available in Hindi-Urdu – the only type comparable to it is that of (60), with an existential covert verb 'to be'. The matrix coding as non-PSA construction cannot be realized by the construction with an infinitival embedded clause, the only option being similar to the matrix coding as PSA construction, with a covert existential 'to be' infinitive as in (61) (see also Montaut 2004: 215-216). Otherwise, the matrix coding as non-PSA construction can be realized with a participial embedded clause (62), which is quite a widespread construction in IA (cf. Bickel and Yadava 2000: 360-361).

(58) Hindi

mere	*bās-ne*	*mujhe*	*sārā*	*kām*	*turant*	*karne-ko*
my.M.OBL	boss-ERG	I-OBL	all	work	immediately	do.INF.OBL-ACC
majbūr	*kiyā*					
force	do.PST.M					

'My boss forced me to do this work immediately.'

(59) Hindi

ve	*naukrī*	*ḍhũḍhne*	*mumbaī*	*āe*	*hõge*
they	job	search.INF.OBL	Mumbai	come.PST	be.FUT.PL

'They must have come to Mumbai in order to search for a job.'

(60) Hindi

mujhe	*vah*	*pagal*	*lagtā*	*hai*
I.OBL	he	crazy	seem.M.SG.PRS	be.3SG.PRS

'He seems (to be) crazy to me.'

(61) Hindi

maĩ	*to*	*use*	*bevkūf*	*samajhtā*	*hũ̃*
I	PART	s/he.OBL	stupid	understand	be.3SG.PRS

'I considered him/her (to be) a fool.'

(62) Hindi

mīnā-ne	*bhaũkte*	*hue*	*kutte-ko*	*dekhā*
mīnā-ERG	bark.PTCP.PRS.OBL	be.PTCP.PST	dog.M.OBL-ACC	see.PST.M.SG

'Mina saw a barking dog.'

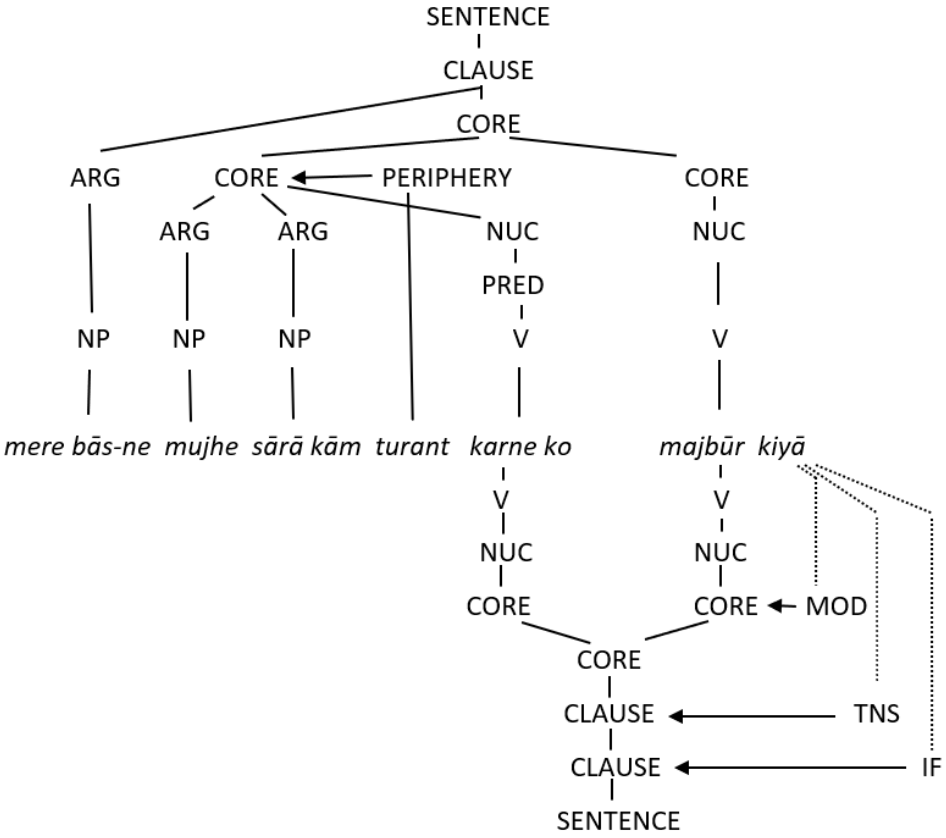

Figure 16. Hindi-Urdu core coordination (control construction).

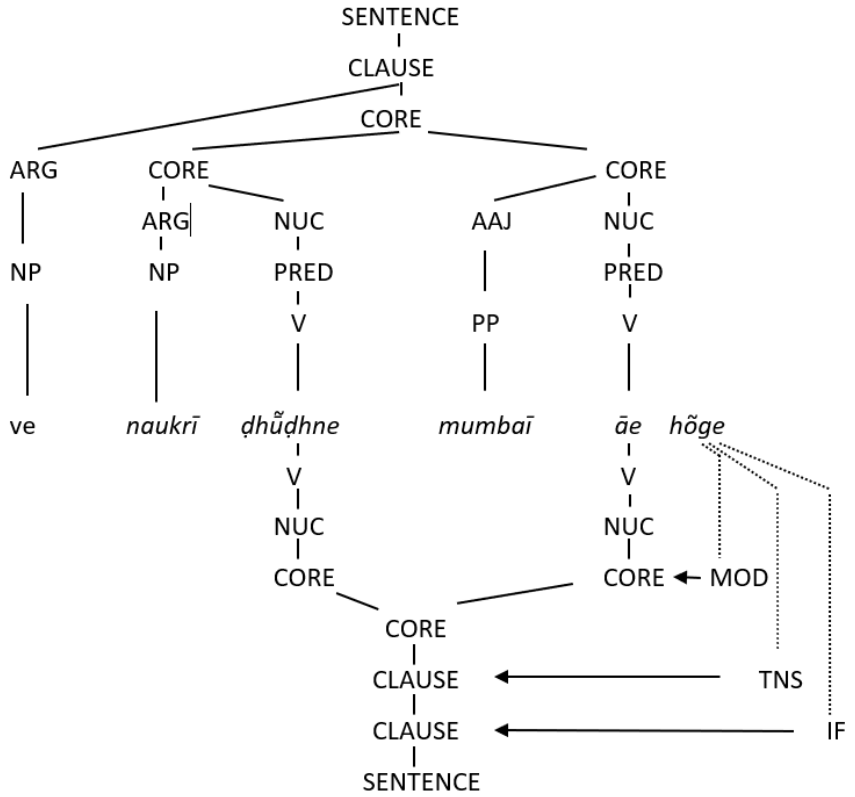

Figure 17. Hindi core coordination (purposive clause).

2.3.2 INFINITIVES VS. ACTION NOMINALS VS. GERUNDIVES

The historical origin of infinitives in Indo-European has been traced back to case forms of verbal nouns or action nominals (Disterheft 1980; Haspelmath 1989, among many others).

For example "the Latin *-tum*, the Old Church Slavonic (OCS) supine in *-tъ* probably represents the former accusative form of a verbal noun; the Latin *-tū* derives from the ablative case and the infinitive endings seem to stem from the locative (Latin *-re*) and from the dative (OCS *-ti*) forms of earlier verbal nouns" (Ylikoski 2003: 211).

A possible path of development is from a purposive action nominal to infinitive, which is well attested in several IE languages. Such purposive forms in the course of time have usually lost their original purposive meaning and required further reinforcement. For example, as Haspelmath (1989) convincingly demonstrated for German, there is a quite well documented path of the development of the infinitive from the purposive, during which the bare infinitive gradually lost its original directional-purposive modality and had to be reinforced first by the element *zu* and later on by *um*. Even more interestingly, such developments have been attested outside of IE, for example in Finno-Ugric (e.g. Ylikoski 2003: 207-208).

The nominal character of the infinitive is present at all stages of IE, and its reanalysis leads to the completion of the circle 'noun-to-infinitive-to-noun' (Disterheft 1980: 198).

The nominal origin of infinitives is not exclusive to IE languages – it is quite a common phenomenon observed in many other language families. Therefore, what we can observe now cross-linguistically is a rather vague boundary between infinitives and action nominals (see e.g. Koptjevskaja-Tamm 1993: 33-34; Ylikoski 2003: 224), although in some languages this difference is more evident. For example, in Slavonic, infinitives have subjects equi-deleted (63), whereas action nominals can have their subject overtly expressed (64).

(63) Polish
Paweł lubi pracować w nocy.
Paweł like.3SG.PRS work.INF at night
'Paweł likes to work at night.'

(64) Polish
Słyszę jej głośne chrapanie.
hear.1SG.PRS her noisy snoring.AN
'I hear her noisy snoring.'

A summary of the properties of infinitives and verbal nouns is given in Table 4 (after Noonan 2007: 75). But we have already seen that the syntactic relation of the subject to the predicate in the case of infinitives, i.e. the fact that they cannot form a constituent with the subject, does not always hold (see example (53) from Vedic).

Complement type	Part of speech of predicate	Syntactic relation of subject to predicate	Range of inflectional categories	Other characteristics
infinitive	Verb	predicate cannot form constituent with subject	reduced; cannot take subject–verb agreement	relations with object same as indicative
verbal noun	Noun	genitive relation between subject and predicate	reduced; may take nominal categories such as case and number	may have internal structure of NP; frequent gradation between nominalizations and infinitives

Table 4. Properties of infinitives and verbal nouns.

Moreover, the category of the future passive participle, i.e. the gerundive, is also formally very close to an infinitive or verbal noun. The Latin infinitive and gerund have the same formal markers (*leg-ere* 'to read / reading'), but it is the gerund which takes inflectional endings: *leg-end-i* 'reading-GER-GEN'. The feature *-nd-* is shared by the gerund and the gerundive, the former having a more abstract noun-like character (lacking plural forms) and the latter being of adjectival nature: *liber legendus* 'a book to be read'.

It is therefore rather difficult, or in some languages even impossible, to delineate infinitives from action nominals and action nominals from gerundives. NIA languages are of this type. In modern descriptions of NIA languages there is very often no differentiation between infinitives and action nominals. In the Hindi-oriented grammatical tradition there is usually one term for both categories, namely *kriyārthak sangyā* 'verbal noun' (for example Kāmtāprasād: 2060: 99-401; Pāṁḍe 2012: 153-154; Tivārī 2005: 237). In some descriptions another category is introduced, namely *kriyārthak/kriyāvācak kṛdant* 'verbal participle' or *vidhyarthak kṛdant* 'modal participle', which in fact refers to the category known under the label of gerundive (Pāṁḍe 2012: 164; Tivārī 2005: 238).

The lack of differentiation between the two or even three forms (infinitive, verbal noun and gerundive) is visible at the morphological level. Western languages such as Punjabi, Hindi, Kashmiri and Pahari have forms terminating in *-n-* or *-ṇ-*, which is derived from OIA form *-anīya* (cf. Southworth 2005: 137), although other options are also considered as possible antecedents, such as an enlarged verbal noun *°(a)nā(ṁ)*, *°(a)nāu(ṁ)* < Ap. *°anahaṁ* (:*°*anakaṁ*) (cf. Oberlies 2005: 44 and most recently Montaut 2018: 119-120). Rajasthani has two affixes forming infinitives/verbal nouns, namely those terminating in *-n-* and *-b-*, whereas Gujarati has forms terminating in *-v-*. Marathi has three forms which are called infinitives or gerunds (*-ṇ-*. *-l-*, and *-u*) and an obligative construction based on a subjunctive terminating in *-v-* (cf. Katenina 1963: 180-185; Wali 2005: 34-35; Montaut 2018: 120-122). The *-b-*/*-v-* forms come from the OIA future passive participle in *-(i)tavya*. Interestingly, the

-b- forms have remained in use in western languages such as Early Braj and Bundeli, but they are no longer attested in Bangaru (cf. e.g. Varmā 1954: 111-113; Singh 1970; Prakāś 1986: 71-72 for Early Braj; Śarmā 2006: 85 for Bangaru; and Agravāl 1963: 128-133 for Bundeli). Early Awadhi from its earliest sources has suffixes containing *-n-* and *-b-* features (Saksena 1937: 283-285). Magadhan tongues such as Bengali or Maithili have a variety of endings marking what is usually labelled a verbal noun; for example Maithili has suffixes based on *-n-* and *-b-* features. Jha (1958: 584) analyses under the heading of verbal nouns, nominalized past passive participles terminating in *-ā,* forms terminating in the *-l-* feature (the same as for the past passive participle), forms terminating in the *-t-* feature (going back to the OIA infinitive *-tum* or *-nt-* participial feature), and forms terminating in vowels, e.g. *-ae*, *-ai*, of rather obscure origin (cf. Chatterji 1926: 1012-1019; Jha 1958: 521).

In the recent history of IA languages we observe an evolution of the construction based on the *-b/v-* infinitive from the obligative to the future tense. The OIA construction based on the *-tavya* future passive participle with the A argument marked by the instrumental (65) continued to be used through MIA (66) up to NIA (67) and (68). We can see at all stages a construction with non-nominative subject and with object verb agreement. This construction finally paved the way for the future tense in Magadhan tongues.

(65) Sanskrit from Montaut (2017: 111)
mayā *tat* *kartavyam*
1SG.INS DEM.NOM.N.SG do.GER.NOM.N.SG
'I have to/should do that.' (lit. 'by me this to-be-done')

(66) Aśokan Prakrit from Montaut (2017: 112)
iyam *sāsane* *vĩnapayitavye*
this.DEM.NOM.M.SG principle.M.SG make known.GER.NOM.M.SG
'This principle should/will be made known'. (made to be known)

(67) Old Bengali from Montaut (2017: 112)
tabẽ *to-ka* *rakhiba* *kona* *jāne*
then you.ACC protect.GER which person.OBL
'Then who will protect you?'

(68) Old Maithili from Jha (1958: 495)
garimā *ghabi* *kañone*
gravity.F hold.GER.F who.OBL
'Who would hold gravity?'

Modern eastern NIA languages have undergone reorganization of the main argument marking in the obligational construction, the final result being a fully nominative pattern of the future. Such a development is widely attested in languages of the world. The Romance

languages, for example, replaced the former gerundive + *esse* construction with obligative meaning with an infinitive + *habere* construction which acquired a futurative meaning (Kuryłowicz 1960 [1931]; Montaut 2017: 115). Similarly, in the history of Greek we observe during the early medieval Greek period an increase in the use of the construction *ekhō* 'have' + infinitive referring to the future (cf. Markoppoulos 2009: 94-104).

The evolution of the construction based on the infinitive/verbal noun in *-b/v-* did not always result in the emergence of a future tense. In Oriya and Assamese, as well as in western languages (Early Rajasthani, Early Braj, Gujarati and Marathi), the *-b/v-* forms are still preserved and have not been reinterpreted as future tense, instead retaining their original obligative meaning (cf. Masica 1991; Khokhlova 2013; Montaut 2017).

The majority of western NIA languages developed the obligative pattern based on the *-n/ṇ-*feature with several options regarding main argument marking. Contemporary Hindi-Urdu has DAT marking, although the encroachment of ERG marking under the influence of Punjabi has been noted in the last twenty years (see (69), and for an extensive discussion Bashir 1999). Eastern Pahari languages such as Nepali allow NOM, ERG and DAT marking with intransitives (70a) and ERG or DAT marking with transitives (70b), the last being certainly an innovation (cf. Masica 1991: 336). Kumaoni has ERG(INS) marking and DAT as well (71), the last having been introduced only recently under the influence of the dominant language of the region, Hindi (cf. Stroński 2010: 93).

(69) Hindi-Urdu
mujhe / maĩ-ne *jānā* *hai*
I-DAT / I-ERG go.INF be.3SG.PRS
'I have to go.'

(70) Nepali from Masica (1991: 336)
a. *mai / mai-le / ma-lāī* *jānu* *cha*
I.NOM / I.ERG / I-DAT go.INF be.3SG.PRS
'I have to go.'

b. *mai-le / ma-lāī* *yo* *kitāp* *paṛnu* *parcha*
I.ERG / I-DAT this book read.INF have to.3SG.PRS
'I have to read this book.'

(71) Kumaoni
mi-l / mi-kai *ya* *kām* *karaṇ* *chu*
I-ERG / I-DAT this work do.INF be.3SG.PRS
'I have to do this work.'

Interestingly, some Western Pahari languages have only recently reinterpreted the construction based on the *-n-* infinitive with ergatively marked subjects as future tense (72).

An interesting development can be traced in Bundeli, in which both *-b-* and *-n-* forms are preserved. Both forms allow obligative and future readings, the latter being restricted only to some dialectal areas and occurring only in 1 pers. plural. The *-b-* form requires a subject in the NOM case (73), whereas *-n-* requires a subject in DAT (74).

(72) Kului from Ṭhākur (1975: 305)
śohrū-ai *būṭe* *kāṭṇe*
son.ERG tree.PL cut.INF.PL
'The son will cut the trees.'

(73) Bundeli from Agravāl (1963: 129)
ham *kām* *karabī*
we.NOM work do.INF
'We will do the work.'

(74) Bundeli from Agravāl (1963: 131)
hamẽ *jānaĩ*
we.DAT go.INF
'We will go.'

We would thus suggest that in the western languages, despite the fact that *-b-* forms are not dominant, the construction based on the *-n-* form is currently undergoing a change leading to a stadial reorganization of the verbal system in the domain of the future tense by means of the obligative non-nominative pattern.

3. MORPHOSYNTAX OF CONVERBAL CHAIN CONSTRUCTIONS IN EARLY NIA

	i	*ī*	*ī+ karī*	*i+ karī*	*i+nai*	*i+ara*	*i+ari*	*i+kāri*	*ya*	*ya + nai*	root	root + *nai*	root + *karī*	root + *ara*	root + *kara*
dvitīya vrata satya para kathā (XIV)	11	15	19	1									2		
guru mahimā par kathā (XV)	1	10													
amarasena-vayarasena kathā (XV)		55													
vacanikā khīcī acaḷadāsa-rī (XV)	8														
vīsaḷadevarāsa (XVI)	32	5													
daḷapata-vilāsa (XVI/XVII)	15				2	9	1	1	3	1				3	
haḍai sūrijamala-rī vāta (XVII)		1							3	2	22	5			
vacanikā rāṭhoṛa ratana-rī (XVII)	9				1				1						
rāṭhoṛa durādavāsa-ro kāgada (XVII/XVIII)											1	1			
dhanuṣa-bhaṅga (XVIII)	17				6				8	6	1	7			1
adālatī nyāya (XVIII)									12	1	28	3			
ḍokarī rī vāta (XVIII)										2					

Table 5. Morphology of converbal forms in Early Rajasthani, after Stroński et al. 2019.

3.1 MORPHOLOGY OF CONVERBS IN EARLY NIA

Early varieties of NIA show considerable variation of converbal forms. In general the attested forms can be divided into simple and extended. Simple forms are either root forms or forms with an ending which is a continuant of the earlier Apabhramśa converbal form (either terminating in *-i, -ī* or in *-ia, -ya*). Extended forms may take the suffix *-kara*/*-kari*, but also other forms such as *-ara* or *-nai*.

There is a basic functional difference between simple and extended forms, namely both can serve as converbs, but only the former can be a part of a complex predicate together with light verbs (see footnote 10).

The most complex situation is found in Early Rajasthani, with a rich variety of simple and extended forms – see Table 5, which shows the distribution of the attested forms in our corpus. We could not find any substantial functional difference between the extended forms. Apart from the adverbial functions, they have all been used for chaining.

Far less complex is the system in Early Awadhi, presented in Table 6. Since the Early Awadhi corpus is quite homogeneous, it may be difficult to discover any greater variety of forms. However, we know also from Saksena (1937: 280-281) that in later Early Awadhi works such as Nur Muhammad (18th c.) there can be found only three more forms, namely those terminating in *-a* and *-i+kara* and *-i+ke*.

Early Braj and Early Dakkhini also exhibit some variety of converbal forms, although as is the case in Early Awadhi and to a lesser degree in Early Rajasthani, there is a tendency to use unextended forms (see Table 7 and Table 8). Early Dakkhini also has a preponderance of unextended, i.e. root, forms.

There are also additional data from Early Pahari showing that before early NIA we do not in fact witness extended forms. In Early Pahari only *-i* forms (Kumaoni) or both *-i* and *-ya* forms (Nepali) were in use; extended forms such as *-i+bera* in Kumaoni appeared around the 18th century, and forms such as *-ya+ra* in Nepali around the 19th century (cf. Pokharel 2008: 209-210).

Contemporary Rajasthani has two converbal forms, terminating in *-ara* or *-nai* or a bare root form (Lāḷas 1996: 120).

Contemporary Braj shows considerable complexity, with simple *-i*/*-y* as well as extended forms in *-kaĩ*/*-kẽ*/*-kai*/*-ke*. But there are also forms which must have emerged due to long-lasting contacts with Rajasthani, formed by the addition of a postposition having a variety of forms: *nãi*/*naiyā*/*naiyã̄*/*naĩyã̄* (Liperovskij 1987: 130).

Modern Awadhi has a simple form terminating in -i which can further be reinforced by a postposition *kai* (with a few variants, *kā, kaïhã̄, kehã̄, ke*; Liperovskij 1997: 161).

Dakkhini presents the most interesting case. It has a continuant *-ko* which is a regular converbal marker, and it has also developed a form in *-ndeko* which is labelled ‘post action’ and which allows violation of SIC (see Mustafa 2000: 149-150).

Kumaoni has only one form, in *-ber*, whereas Nepali has *-era* and two less frequently used forms in *-ī* and *-īkana* (Mathews 1992: 115-116).

	-i	-i+kai	i-kari
Padamāvat 1540	282	6	
Rāmcaritmānas 1574-1576	246	2	1

Table 6. Morphology of converbs in Early Awadhi.

	-i	-root	Ī	-ya	-i+kai	i-kari	-ya +kari
Vishnudas 15th c.	112	6	2		1		
Hitaharivamśa 16th c.	68	1					
Indrajit of Orcha 1600	47			1	10	2	
Śivarajabhushana 1673	92	1		8	10	3	1

Table 7. Morphology of converbs in Early Braj.

	root	-ya	-root+kara	-root+ke
Khwājā Bandā Navāz 1312/18-1422/37	1		3	2
Gorndā 1300-1351	20		6	11
Eknāth 1548-99	7		1	10
Qulī Qutub Śāh 1580-1612	9		2	
Valī Daknī/Dakkhinī 1682-1730	8		5	3
Saiyad Mīrāṁ Husenī 1623	10	1	3	
Mullā Vajhī 1636	22		7	
Huseini 1641	10		1	3
Fāyaz, 1685	7	4	1	

Table 8. Morphology of converbs in Early Dakkhini.

3.2 A NOTE ON ALIGNMENT IN MODERN NIA

IA languages are a textbook example of split ergative languages. The discussion as to whether the ergative alignment resulted from reanalysis of the passive or was a genuine ergative from the very beginning has not reached any final conclusion (for recent discussion from various perspectives see Dahl and Stroński 2016).

The type of alignment attested in several IA languages has been labelled "split ergativity based on aspect".[13] In IA, ergative alignment is manifested in perfective tenses (with the exception of Shina, where it is present in all tenses) by case marking and agreement. Both can (but do not necessarily) co-occur, as in (75-76). Example (75) shows an ergatively marked A and an unmarked O which agrees with the verb form in gender and number. In (76) we have unmarked S also showing agreement with the verb in gender and number. This morphological pattern is not the only one available in the perfective domain, since O can also be marked (with a default verb form) as is the case in (77):

[13] For the sake of brevity we will consistently use semantic-syntactic primitives such as S (subject of an intransitive verb), A (subject of a transitive verb) and O (object of a transitive verb), which have been widespread in the typological literature since Dixon (1979 and later 1994).

(75) Hindi-Urdu
rām-ne *railgāṛī* *dekhī*
Ram-ERG train.ABS.F.SG see.PPP.F.SG
'Ram saw a train.'

(76) Hindi-Urdu
railgāṛī *calī*
train.ABS.F.SG go.PPP.F.SG
'A train departed.'

(77) Hindi-Urdu
Rām-ne *sītā-ko* *dekhā*
Ram-ERG Sita-ACC see.PPP.M.SG
'Ram saw Sita.'

Rich variation is also observed in the attested patterns of split systems in the context of Indo-Aryan constructions with object marking, mainly along the lines of animacy and definiteness and various types of object or subject agreement. A summary of the attested main argument marking patterns and agreement patterns is given in Table 9. The introduction of O marking in the perfective domain or the introduction of subject–verb agreement has been perceived as a drift towards nominative morphology (see for example Stump 1983; for an extensive discussion on variation of ergative and nominative patterns see also Verbeke 2013a).

Main argument marking	Agreement	Features	Remarks
S=O≠A	OV	Gender, Number	
S=O≠A	OV	Gender	
S=O≠A	AV	Person	
S=O≠A	OV	Gender Number	With a main verb
	OV	Person Number	With an auxiliary
S≠O≠A	NO	-	
S≠O≠A	OV	Gender, Number	
S≠O≠A	AV	Person	
A=S≠O	OV	Gender, Number	With a main verb
	AV	Person Number	With an auxiliary
A=S=O	OV	Gender, Number	With a main verb
	AV	Person Number	With an auxiliary
S=O≠A	OV	Person, Number	
A=O≠S	NO	-	

Table 9. Attested agreement patterns and main argument markings in modern NIA.

In some IA languages, along with aspect-based split ergativity, NP split ergativity is also attested. NP splits operate along the lines of the animacy hierarchy (or more recently 'referential hierarchy') first introduced by Silverstein (1976), which correlates main argument marking with the referential features of these arguments.

first/second person > third person > proper nouns > human > animate > inanimate

In languages showing NP split ergativity one would expect arguments to the right of the hierarchy to be more likely marked by ergative case markers than those to the left. This is indeed the case in many Australian languages, which for example do not mark 1st and 2nd person pronouns while marking other pronominal and nominal arguments. In IA, one such example is provided by Punjabi, where 1st and 2nd pers. pronominal forms remain unmarked while the 3rd person and nouns receive ergative marking (78). The same pattern has been noted for Marathi (cf. Wali 2005: 45). To a certain extent, similar phenomena have also been attested in Rajasthani. For example, Marwari, Shekhavati and supposedly Mewari and Jaipuri display an opposition (although not consistently) of A and S forms in the 3rd person pronouns but not in the 1st and 2nd person pronouns, whereas Bagri maintains this opposition also in the 2nd person pronouns (but only in the singular). Interestingly, only those nouns terminating in *-o* inconsistently mark the opposition between A and S. Other classes of nouns do not distinguish between A and S (cf. Magier 1983: 311-312; Khokhlova 1995: 20-21; Stroński 2011: 71).

(78) Punjabi

a. *maĩ* *te-nữ* *saṛak-te* *vekhya*
I.**NOM** you-ACC street-LOC see.PST.M.3SG
'I saw you in the street.'

b. *tữ* *maĩ-nữ* *saṛak-te* *vekhya*
you.**NOM** I-ACC street-LOC see.PST.M.3SG
'You saw me in the street.'

c. *o-ne* *te-nữ* *saṛak-te* *vekhya*
s/he-**ERG** you-ACC street-LOC see.PST.M.3SG
'S/he saw you in the street.'

Ergative morphology has also been studied in the context of subjecthood at least since the 1970s. Several studies on IA clause linking have led to the conclusion that ERG arguments are endowed with all behavioural subject properties, and what is more in some languages (e.g. Nepali) they can even trigger agreement. One of the classical tests for subjecthood in IA is a construction based on a non-finite verbal form, i.e. a converb. We have already seen that arguments with ergative marking may control and undergo reduction in converbal chain

constructions (examples (37-38)). Such examples, along with conjunction reduction (79), demonstrate that IA languages are ergative exclusively at the level of morphology, whereas syntactically they are nominative.

(79) Hindi

Sītā$_i$-ne	*ek*	*larkī$_j$*	*dekhī*	*aur*	*∅$_i$*	*bhag*	*gaī*
Sita-ERG	one	girl	see.F.SG.PST	and		run away.PST	go.AUX.PST

'Sita$_i$ saw one girl$_j$ and ∅$_{i/*j}$ ran away.'

3.3 MAIN ARGUMENT MARKING IN CONVERBAL CHAIN CONSTRUCTIONS

In this section we analyse the case marking of the arguments of the main verb in converbal chain constructions.

In modern Hindi and in a number of other Indo-Aryan languages, one finds split ergativity as well as differential object marking, two phenomena which pertain to the marking of A and O respectively. Both phenomena already occur in Early NIA, though in a different form. First, we will discuss the marking of the A-argument.

3.3.1 MARKING OF THE A-ARGUMENT

A marking in early NIA shows variability to a certain extent, with an optional ergative marking which may vary according to parameters such as animacy or volitionality (see for example Drocco 2008; Butt 2001). In the modern literature this phenomenon has been labelled Optional Ergative Marking (henceforth OEM). OEM is a subset of phenomena placed under the more general umbrella of Optional Case Marking (OCM). McGregor (2009: 1610) defines OCM as follows: 'the situation in which, in specifiable lexical or grammatical environments, a case marking morpheme (inflectional affix, clitic, or adposition) may be either present or absent from an NP of a specifiable type without affecting the grammatical role borne by that NP.'

OEM is present, for example, in Lhasa Tibetan. The semantic content of the sentences given below is exactly the same, i.e. they refer to an event in which somebody is preparing meals. However, (80a) is neutral, (80b) focuses on the person preparing and contrasts him with somebody else, and (80c) emphasizes the fact that he is the one who is cooking (i.e. an agent).

(80) Lhasa Tibetan from McGregor (2010: 1610-1611)

a. *khōng khāla' sǫ̲kiyo:re'*
   he food make-IPFV.GNOM
   'He prepares the meals.'

b. *khōng-ki' khāla' sọ̲kiyo:re'*
he-ERG food make-IPFV.GNOM
'He prepares the meals.'

c. *khāla' khōng-ki' sọ̲kiyo:re'*
food he-ERG make-IPFV.GNOM
'He is the one who prepares the meals.'

OCM has been noted for contemporary IA and Iranian with so-called anti-impersonal verbs (see Lazard 1983 for a more detailed typological overview and the semantic areas covered by such verbs). In Indo-Iranian, verbs of bodily emissions as well as unergative verbs can display OEM. The following examples show the variation between marked and unmarked S with an intransitive verb. There have been attempts to connect this optional use with volitionality (Butt 2001; Butt and King 2005), i.e. the construction with unmarked S denotes a non-volitional act whereas the construction with marked S denotes a volitional one (see (81)), but they do not seem to be very convincing:

(81) Hindi (Butt 2001: 122)
a. *Ram khɑ̃̄sā*
Ram cough.PST
'Ram coughed.'

b. *Ram-ne khɑ̃̄sā*
Ram-ERG cough.PST
'Ram coughed (purposefully).'

OCM has only recently been discussed in the context of the contemporary IA languages. De Hoop and Narasimhan (2009) discuss OEM in Hindi, which they attribute to syntactic and semantic factors such as coding of prototypical subject properties by the ergative marker. Verbeke has put forward an interesting proposal as regards OEM in Nepali, connecting it with the perfective meaning of the marker itself (2013b; see also Verbeke and De Cuypere 2015).

The OEM in early NIA finite constructions, as found in our data, is quite common. As may be expected, it can occur only in perfective constructions with transitive verbs. We see no semantic difference between constructions with marked and unmarked forms. In the following examples we observe marked (82a–86a) and unmarked (82b–86b) A forms in perfective clauses.

(82) Early Awadhi
a. *su-aiṁ asīsa dīnha*
parrot.OBL blessing.M.SG give.PST.M.SG
'The parrot gave blessing (to the king).' (J.81.1) AD 1575

b. *cāri* *mīta* *kabi* *muhamada* *pāe*
four friend.M.PL poet Muhammad.NOM.M.SG get.PST.M.PL
'Poet Muhammad got four friends.' (J.22.1) AD 1540

(83) Early Rajasthani

a. *ghaṇī vāri* *mātaṅgiṁ* *vidyā* *kah-ī*
manytimes outcaste.INS knowledge.NOM.F.SG say.PST.F.SG
'Many times the outcaste (A) taught (him).' (RG.SS.9) 15th c.

b. *kumār* *lakuṭ-ai* *te* *tim* *haṇy-ā* *jima...*
prince.NOM wood-INS they.DIR so beat.PST.M.PL that...
'The prince beat them with a wooden stick so that...' (RG.M.98) 15th c.

(84) Early Braj

a. *jihiṃ* *aneka* *vāra* *samara* *viṣai* *jītyau* *hai*
who.OBL many time fight.NOM.M.SG in win.PST.M.SG be.3SG.PRS
'Who many times has won in fights' (I.) AD 1600

b. *mokahaṃ* *dasaratha* *āisu* *diyau*
I.OBL.DAT Daśaratha.NOM permission.NOM.M.SG give.PST.M.SG
'Daśaratha gave me permission.' (V.) AD 1442

(85) Early Dakkhini

a. from Šamatov (1974: 130)
bahutāṁ-ne *apnā* *sar* *bhāye*
many-ERG own head.M throw.PST.PL.M
'Many rested their heads.'

b. from Šamatov (1974: 131)
khudā *tujhe* *fursat* *diyā* *hai*
God.NOM you.OBL opportunity give.PST be.AUX.3SG
'God has given you opportunity.'

(86) Early Pahari

a. Kumaoni from Joshi (2009: 344)
rājā-lai *datta* *dinhi* *guṇākara* *pāṇḍe-lai*
king-ERG gift.F.SG give.PPP.F.SG Guṇākar Paṇdey-ERG
data *pāī*
gift.F.SG get.PPP.F.SG
'The king gave the donation, Guṇākar Paṇdey received the donation.' AD 1374

b. Kumaoni from Joshi (2009: 360)

śrī jāgeśvara	*kī*	*jātrā*	*karī*	*sutradhāra*	*rajau*
J.	GEN	pilgrimage.F	do.PST.F.SG	Sutradhara	Rajau.NOM

putra	*rāmu*
son	Ramu

'Sutrahāra Rajau son of Ramu undertook (lit. did) the pilgrimage of Jāgeśvara.' 14th c.

The above examples demonstrate that ergative A marking is not obligatory with the subjects of transitive verbs in early NIA.

Early Awadhi and eastern branches such as Maithili seem to show a preference for a fully nominative pattern, i.e. there is a preponderance of unmarked A forms and regular AV agreement from a very early period.

(87) Early Awadhi

ahā	*bāṁdi*	*kīnhesi*	*niti*	*sevā*
stay.M.SG.PST	cage.CVB	do.3SG.PST	always	service.F.NOM.SG

'When he (the bird) stayed caged, he served you always.' (J.68.2) AD 1540

(88) Maithili

bihi	*chalalihu*	*mohi*
fate	cheat.PPP.3SG	I.OBL

'Fate has cheated me.' (V.P. 47.4) 15th c.

Early Awadhi from the so-called Eastern Hindi branch, as well as Maithili from the Magadhan group, have reorganized their verbal systems, ousting the forms based on the *-ta* participle and developing finite forms showing subject–verb agreement. This had taken place at a very early stage, although certainly it was maintained longer in Early Awadhi than Maithili (cf. Saksena 1971 [1937]: 234; Jha 1958: 472, 503-505; Stroński 2010: 142-150).

In Early Rajasthani, one can observe reorganization of the case marking in the nominal system, i.e. around the 18th century the scope of the old instrumental case became limited to masculine nouns terminating in *-o* (both sg. and pl.) and masculine nouns terminating in a consonant (but only pl.) (for detailed study see Khokhlova 1992; 2000; 2001).

Early Braj and Early Dakkhini, both from the so-called Western Hindi group, developed postpositional A markers in perfective tenses. In both languages the marker was a postposition *-ne*, attested in Early Braj only from the 17th century onwards, and in Early Dakkhini as early as the 14th century. However, in the two languages the use of the agentive postposition was quite divergent. In Early Dakkhini at least up to the 16th century we observe OEM, with the postposition used in the obligative construction; later A marking disappeared due to language contact with fully nominative Dravidian languages. In Early Braj its use became quite regular over the centuries (Drocco 2017).

Early Pahari (both early Nepali and Kumaoni, for which there is an ample corpus of inscriptions dating to the 13th and 14th centuries) introduced the ergative postposition *-le* in the 14th century, but until the end of the 16th century its use was rather irregular (see Stroński 2014).

OEM has previously been explained in terms of the early decay of ergativity in dialects belonging to the eastern branches of IA and to the Early Rajasthani group (cf. Khokhlova 2000; 2005; Stroński et al. 2019).

Poudel (2008) tried to ascribe recipient semantics to the DAT marker, but we have clear counter-evidence to this claim in Early Kumaoni (Stroński 2014: 284).

Another example of what can be labelled DEM is present in contemporary Shina, which has a native ERG marker used exclusively in the perfective and past tenses, and a borrowed one (from adjacent Tibetan) which has encroached on the imperfective and non-past tenses as well (Hook 1996).

In converbal constructions, A marking depends exclusively on the transitivity of the main verb. The subject is only marked when the main verb is transitive, irrespective of the converb's transitivity. The examples below show converbal chain constructions with the main verb transitive and the converb either transitive (89a) or intransitive (89b), as well as a transitive converb and intransitive main verb (89c).

(89) Early Awadhi

a. *rājaiṃ* *sun-i* *biyoga* *tasa* *mānā*
king.INS.M.SG hear-CVB separation.NOM.M.SG like consider.PST.M.SG
'The king having heard [this] felt desolated.' (J.88.1) AD 1540

b. *taba* *lagi* *rānī* *suā* *chapāvā* *jaba* *lagi* *ā-i*
then ABL queen parrot hide.PPP.M.SG when for come-CVB
maṁjāri-nha *pāvā*
cat-PL.OBL get.PST.M.SG
'Then the queen hid the parrot until the cats, having come, could find (it).' (J.56.4) AD 1540

c. *guru* *sikha* *dēi* *rāya* *pahiṃ* *gay-au*
master.NOM instruction give.CVB king next to go.PST.M.SG
'After giving this teaching, the guru went to the king.' (T.2.10) AD 1574-1576

However, OEM is widespread – examples (90a) with marked A and (90b, c) with unmarked A forms show that transitivity of the main verb does not always trigger A marking.

(90) Early Rajasthani

a. *iyāṃ* *bulāya* *turatībega* *nūṃ* *kahi-yo*
they.OBL call.CVB Turtibeg ACC say-PST.M.SG
'They, having called Turtibeg, said.' (RG.DV.51) 16th/17th c.

b. *mātaṃga* *tihāṃ* *āwī* *rājā* *nu* *ādeśa* *kahiu*
outcast.NOM there come.CVB king GEN order.NOM.M.SG tell.PST.M.SG
'Having come there, the outcast explained the king's order.' (RG.M.11) 15th c.

c. *isiuṃ* *sāṃbhaḷī* *bewai* *bāṃdhawa* *deśāṃtara* *bhaṇī*
such hear.CVB two brothers.NOM abroad for
cālyā
go.PST.M.PL
'After having heard such a thing both brothers went abroad.' (RG.M.16) 15th c.

In our corpus, in converbal constructions with a transitive main verb, in dialects such as Early Awadhi, Early Braj and Early Dakkhini the majority of A forms are unmarked, whereas in Early Rajasthani marked and unmarked forms are evenly distributed.

Table 10 shows the distribution of marked and unmarked A forms in four dialects in converbal constructions with the main verb transitive. No S forms are marked, except in two examples from Early Awadhi (in fact the second in our main corpus) showing OEM with the verb 'laugh', which can be interpreted as one of the anti-impersonals (see Lazard 1983).

(91) Early Awadhi
a. *haṁsā* *suā*
laugh.PST.M.SG parrot.M.SG.**NOM**
'The parrot laughed .' (J.84.1) AD 1540

b. *jeĩ* *mukha* *dekhā* *teĩ* *haṁsā*
who.OBL face see.PST.M.SG s/he.**OBL** laugh.PST.M.SG
'The one who saw (his) face, laughed.' (J.23.8) AD 1540

The data from Early Pahari, although rather limited, show that this particular dialectal group has a preference for the ergative pattern, lacking any attestation of unmarked forms in converbal chain constructions.

	CVB + marked A	CVB + unmarked A
Early Awadhi	28%	72%
Early Rajasthani	48.5%	51.5%
Early Braj	29%	71%
Early Dakkhini	33.3%	66.7%

Table 10. Distribution of A marking in converbal constructions with transitive main verb.

An opposite phenomenon, namely A marking dependent on the transitivity of the converb, has been noted for some modern NIA languages. This is the case, for example, in Nepali.[14]

[14] Similar phenomena have been observed in Shina (Hook 1996).

In (92a) both the CVB and the main verb are intransitive, and the A remains unmarked, whereas in (92b) the main verb is intransitive and the converb transitive, and it is thus the transitivity of the converb which triggers A marking. We consider this phenomenon an innovation; similar attestations in other Early Pahari languages (e.g. Kumaoni or Garhwali; Stroński 2014) let alone other early NIA tongues are not found.[15]

(92) Nepali from Wallace (1982: 168)

a. *u* *bimar* *bha-era* *mar-yo*
he sick be-CVB die-PST.3SG
'He became sick and died.'

b. *us-le* *bikh* *khā-era* *mar-yo*
he-ERG poison eat-CVB die-PST.3SG
'He died after eating poison.'

3.3.2 DIFFERENTIAL OBJECT MARKING (DOM)

The second case marking phenomenon which is relevant to the discussion of the control properties of arguments in sentences with converbs is Differential Object Marking (henceforth DOM). Since Bossong (1985) the term has gained popularity, although similarly to Differential Subject Marking (henceforth DSM) it has recently been criticized for its rather constrained focus on a single grammatical relation (cf. McGregor 2010: 1614). The phenomenon of DOM has already been dealt with from various perspectives – theoretical and typological (e.g. Aissen 2003; Iemmolo 2010; Iemmolo and Klumpp 2014, to name a few) and recently also typological-diachronic (Seržants and Witzlack-Makarevich 2018). DOM refers to the phenomenon of marking or lack of marking of the object argument depending on semantic (animacy, definiteness and specificity) or pragmatic (topicality) factors. Aissen (2003) reanalysed Silverstein's model (1976) and proposed two scales based on animacy and definiteness. The O-arguments that lie further to the left of both scales are more likely to be marked:

a. Animacy scale:

Human > Animate > Inanimate

[15] Interestingly we find examples of ergative A with transitive converbs and intranistive main verbs in early Braj such as:

Old Braj *Caurāsi vaiṣṇavana kī vārtā* 1-2 from Snell (1991b: 71)

Tānasena *neṃ* *eka* *pada* *sūradasa* *kau* *sīkhi-kai*
Tanasena ERG one lyric hymn Surdas GEN.M.SG learn-CVB
akabara *bādasāha* *ke āgai* *gāyau*
Akbar emperor in front go.PST.M.SG
'Tanasena having learnt one lyric hymn of Surdas went in front of the emperor Akbar'

b. Definiteness scale:
Personal pronoun > Proper name > Definite NP > Indefinite specific NP > Non-specific NP (Aissen 2003: 436-437)

More recently, Witzlack-Makarevich and Seržants (2018: 5-11) summarized the previously discussed referential hierarchies, making distinctions between those based on inherent lexical and morphological properties of arguments on the one hand, and those based on non-inherent, i.e. discourse-based, properties on the other hand. Lexical properties are given in Table 11, morphological properties in Table 12, and finally non-inherent properties in Table 13.

Dimension	Example
Person	First & Second person > Third person > (Obviative / Fourth person) (cf. Dixon 1979: 85; Croft 2003: 130)
Animacy	Humans > Animate non-humans (animals) > Inanimate (cf. Bossong 1991: 159; Silverstein 1976; Aissen 2003)
Uniqueness	Proper nouns > Common nouns (e.g. as part of Croft 2003: 130)
Discreteness	Count nouns > Mass nouns (cf. Bossong 1991: 159)
Number	Singular vs. Plural vs. Dual

Table 11. Inherent semantic argument properties (after Witzlack-Makarevich and Seržants 2018: 6).

part-of-speech distinction – pronoun vs. noun
gender/inflectional classes distinction

Table 12. Inherent morphological argument properties (after Witzlack-Makarevich and Seržants 2018: 6).

definite > (indefinite) specific > (indefinite) non-specific
topicality vs. focality

Table 13. Non-inherent discourse-based argument properties (after Witzlack-Makarevich and Seržants 2018: 10).

Languages differ as regards the choice of scales, or more precisely the cut-off point of a preferred scale at which O marking starts to operate. The question of a certain symmetry between A and O marking has also been raised by several scholars, some of whom put forward the hypothesis of 'markedness reversal', whereby features which make O-arguments marked make A-arguments unmarked, and vice-versa (cf. Aissen 2003: 437; Fauconnier and Verstraete 2014). However, more recent research demonstrates that the two phenomena are far from being symmetrical (see Verbeke 2013: 28-39 for an overview). There is of course a general tendency for ergative languages to exhibit DSM and accusative languages DOM, and in this context the IA languages, which are morphologically ergative only in the perfective tenses and accusative outside the perfective domain, lie somewhere in between.

The main parameters that determine object marking in IA languages are definiteness and animacy, but DOM has been present in IA only for the last few centuries (cf. Montaut 2018).

Early varieties of NIA seem to show that the parameters of animacy and definiteness began to shape DOM, leading to behaviour similar to that found in modern NIA but it was presumably quite a long process. In Early NIA finite verb constructions still even animate definite Os could remain unmarked (e.g. Awadhi (82b) or Rajasthani (83b)), let alone definite inanimate Os (cf. Dakkhini (85a)) or Pahari (86b). In addition to that, DOM had also a secondary motivation, namely aspectual. For example, Early Rajasthani introduced O marking first in imperfective tenses and only later in perfective ones (cf. Khokhlova 2000; 2006). Early Pahari dialects show a similar path of development of DOM (see ex. (106) from early Nepali and 107) from early Kumaoni) and, what is more, some Early Pahari varieties are not consistent as regards the implication of definiteness and animacy even today (cf. Wallace 1981; Stroński 2011; 2014).

O marking in converbal constructions is attested only for definite or animate nouns. The introduction of O marking occurred in different periods in the NIA languages. In Early Awadhi, from the mid-16th century onwards we observe the first instances of animate as well as inanimate and definite O marking, but this is only a tendency, rather than a strict rule. In 'Padmāvat' by Jayasī from the mid-16th century we find both marked forms (as in (93) and (94)) and unmarked forms (as in (95) and (96)), for animate and definite nouns respectively.

(93) Early Awadhi

beṃca-i	*lāga*	*hāṭa*	*la-i*	*ohīṃ*	*mola*	*ratana*
sell-CVB	start.PPP.M.SG	market	take-CVB	he.OBL	price	gemstone
mānika	*jahaṁ*	*hohīṃ*				
ruby	where	be.3PL.SBJ				

'He started selling (scil. [the parrot)], having brought him to the market where the price of gemstone and ruby was settled.' (J.76.2) AD 1540

(94) Early Awadhi

pātī	*likhī*	*saṁvari*	*tumha*
letter.NOM.F.SG	write.PPP.F.SG	remember.CVB	you.OBL
nāmāṁ			
name.O.OBL.M.SG			

'I have written a letter having remembered your name.' (J.225.6) AD 1540

(95) Early Awadhi

dhāi	*suā*	*lai*	*mārai*
wet-nurse.NOM.F.SG	parrot.O.NOM.M.SG	take.CVB	kill.INF.OBL
gaī			
go.PPP.F.SG			

'The wet-nurse having taken the parrot went to kill it.' (J.86.1) AD 1540

(96) Early Awadhi

saṁvari	*rūpa*	*padumāvati*	*kerā*	*haṁsā*
remind.CVB	beauty.O.NOM.M.SG	Padmavati	GEN	smile M.PPP.SG
suā				
parrot.NOM.M.SG				

'Having recalled the beauty of Padmavati, the parrot smiled.' (J.84.1) AD 1540

The earliest occurrences of O marking in Early Rajasthani are for both animate definite and animate indefinite arguments, as exemplified in (97) and (98) respectively. Only much later can O marking be found with inanimate definite arguments, as in (99), becoming consistent around the 18th century. Although we find very early occurrences of marked O arguments, the earliest Early Rajasthani texts show considerable variability of O marking (cf. (100) with unmarked animate and definite argument):

(97) Early Rajasthani

mṛga-rahaiṃ	*mati*	*prayog-i*
deer.M.SG-ACC	intelligence.F.NOM.SG	use.M.SG-INS
choḍawī	*karī*	
release.CAUS.CVB	do.CVB	

'Having saved the deer by the use of intelligence...' (RG.TS.8) 14th c.

(98) Early Rajasthani

muṃḍa	*pākhaṃḍika*	*eka-rahaiṃ*	*dekhī*	*karī*
shaven.NOM.M.SG	ascetic.NOM.M.SG	one-ACC	see.CVB	do.CVB

'Having seen one shaven ascetic...' (RG.TS.9) 14th c.

(99) Early Rajasthani

tina	*sahanāna-nūṃ*	*dekha*	*mo-nūṃ*	*khabara*
that	sign-ACC	see.CVB	I-ACC/DAT	information
paṛasai				
be found.PRS.SBJV.3SG				

'Having seen the sign, I would get the information.' (R.G.MS.79) 18th c.

(100) Early Rajasthani

yakṣ-i	*arjuna*	*ripu*	*bāṃdhī-karī*	*page*
Yaksha.M.SG-INS	Arjuna	enemy.NOM.SG	bind-CVB	foot[M]LOC.PL
āṇi	*ghātiu*			
come.CVB	throw.PST.M.SG			

'Yaksha, having bound the enemy named Arjuna, threw him on his feet.' (RG.TS.20) 14th c.

In Early Braj, the earliest occurrences of converbal constructions with marked O's can already be found in the 15th century. Initially O marking occurs with pronouns (101), and then it extends to nouns (see the postpositional marking in (102)). Inanimate definite arguments may also be marked very early (103), but DOM is variable until the 16/17th century, and only from the 17th century onwards do inanimate definite O arguments appear to be regularly marked (104). This also coincides with the period when postpositional accusative marking became more common in the perfective tenses (e.g. Drocco 2017; Stroński 2011).

(101) Early Braj
tihi *baiṃṭhāri* *chatru* *siru* *diyo*
this.OBL make seated.CVB umbrella head give.PPP.M
'Making him seated, [he] gave an umbrella [to protect his] head.' (V.1.73) 15th c.

(102) Early Braj
kurukha *cikattā-kauṃ* *nirakhi* *kīnau* *sarajā*
having ugly face Chagatai-ACC see.CVB do.PPP.M.SG lion
sāhasa
courage.NOM.M.SG
'Having seen ugly-faced Chagatai (i.e. Aurangzeb), the Lion (i.e. Shivaji) did a courageous act.' (Ś.77) 1673

(103) Early Braj
tāhi *dekhi* *kopaṃtanu* *bhayau*
it.OBL see.CVB angry.M.SG be.PPP.M.SG
'...having seen it [he] became angry.' (V.2.114) 15th c.

(104) Early Braj
avasthā-hi *pāi-kai* *vastu* *ghaṭai*
state.F.SG-OBL obtain-CVB wealth decrease.3PRS.SG.SBJ
'Having obtained this state, the wealth decreases.' (I.a.105) 1600

In the Early Dakkhini corpus, only a few attestations of marked O's are found with converbs, all occurring with the main verb in imperfective tenses.

(105) Early Dakkhini
dono *kūṃ* *aisā* *kasa-ke* *[...]* *khelo*
both ACC such tighten-CVB play-2.IMP
'Having bound both [in] such [way], play.' (E.1.40) 16th c.

The Early Pahari data are rather ambiguous – marked O's are attested in early inscriptional Nepali corpora, but entirely absent in Kumaoni or Chambyali (cf. Stroński 2014). With finite verb constructions they first appear around the 14th century in Nepali inscriptions (106), but in Kumaoni only in the 18thcentury (cf. (107) from the first Kumaoni literary work), and in

converbal constructions in Nepali in the 17th century (108). Interestingly, in Kumaoni, unmarked definite and animate O's with converbs can be found even as late as the 19th century (see (109)).

(106) Old Nepali (Wallace 1981: 112)
deva-ṁ *ghale*
God-OBL destroy.FUT
'You will destroy the god.' AD 1356

(107) Old Kumaoni (Joshi 1983: 65)
kārya *akārya* *kana* *jāṇaṁ* *cha*
act non-act ACC understand be.3.SG.PRS
'He understands the acts to be followed and to be avoided.' (RŚ) AD 1728-1729

(108) Old Nepali (Wallace 1981: 112)
rāvaṇa-kana *māri*
R-ACC kill-CVB
'having killed Ravana' AD 1773

(109) Old Kumaoni (Grierson 1916: 172)
unan *maiṃ* *dekhi* *baṛi* *rīs* *ai*
he.OBL.PL I.NOM.SG see.CVB big anger.F come.PST.F.SG
'Having looked at me they became very angry.' 19th c.

3.3.3 THE SUBJECT IDENTITY CONSTRAINT (SIC)

As mentioned in section 2.2.4, IA converbs share their subject with main verbs, but there are several exceptions to the coreference constraints.

In our data we find a few instances of the violation of SIC which seem to have a semantic motivation. Thus, in (110a, b), the implicit subject of the converb is in an experiencer-like relation to the subject of the main clause. In (111a, b, c), the implicit subject of the converb is in a possessor-like relation to the covert subject of the main clause, which appears in an earlier portion of the text.

(110)

a. Early Awadhi
suni *jogī* *kai* *ammara* *karanī* *nevarī* *biraha*
hear.CVB ascetic GEN immortal deed.F.SG end.PPP.F.SG separation
bithā *kai* *maranī*
pain GEN death.NOM.F.SG
'After [the person separated from the lover] had heard some immortal deeds of the ascetic (lit.: 'Having heard about some immortal deeds of the ascetic...'), the death caused by the pain of separation was gone.' (J.259.1) AD 1540

b. Early Braj
sasī *dekhi* *rāma* *mana* *sītā* *basī*
moon see.CVB Ram mind Sita settle down.PPP.F.SG
'Having seen the moon [...], Ram's mind/heart became inhabited by Sita.' (V.1.102) 15th c.

c. Early Dakkhini
tirā *laba* *dekha* *haivāṁ* *yāda* *āve*
your lips see.CVB beast memory come.3SG.PRS
'Having looked [at] your lips, a beast comes to memory.' (VD.1) 18th c.

(111)
a. Early Awadhi
kathā *kahānī* *suni* *suṭhi* *jarā* *jānahuṁ* *ghīu* *baisaṃdara*
story story hear.CVB very much burn as if ghee fire
parā
fall.PPP.M.SG
'Having heard the stories, (the body) burnt as if ghee fell into fire.' (J.226.7) AD 1540

b. Early Rajasthani
ara *hemū* *pāṇīpaṃtha* *āi* *ḍerā* *paṛiyā*
and Hemu Panipat come.CVB camp.M.NOM.PL fall.PST.M.PL
'And after Hemu had come to Panipat, the camps were established.' (RG.DV.58) 16/17th c.

c. Early Braj
avasthāhi *pāikai* *vastu* *ghaṭai*
state.F.OBL obtain.CVB wealth decrease.3SG.PRS
aru *vaḍhai*
and increase.3SG.PRS
'Having obtained the state, the thing/wealth (may) decrease or increase.' (I.a105) AD 1600

The instances of SIC violation contain converbs denoting both non-volitional acts and volitional ones (cf. (114) vs. (112-113)) and their subjects can be both animate and inanimate (cf. (110-113) vs. (114-115)). This goes against the rule formulated by Subbarao (2012: 274) according to which the subject of the converb should be inanimate and the converb denote a non-volitional act (as in (40)).

(112) Early Awadhi
sakami *haṁkāri* *phāṁdi* *giyaṁ* *melā*
power call.CVB noose neck.OBL.F.SG put.PPP.M.SG
'[Birds] having called [one another] with power [loudly], the noose was put on their neck.' (J.72.3) AD 1540

(113) Early Rajasthani

ti	*puruṣa*	*raja-nai*	*vacani*	*karī*
these	man.M.NOM.PL	king-GEN	speech	do.CVB
saṃgha	*māhi*	*gayā*		
community	in	go.PST.M.PL		

'These men on hearing the king's speech (lit. of the king having spoken) went happy to their community.' (RG.TS.13) 15th c.

(114) Early Awadhi

biraha	*agini*	*tana*	*jari*	*bana*	*jare*
separation	fire	body	burn.CVB	forest	burn.PPP.M.PL

'From the fire of separation my body burnt and forests burnt.'
Lit. 'From the fire of separation, my body having burnt, forests burnt.' (J.225.5) AD 1540

(115) Early Dakkhini

to	*begānagī*	*jā-kara*	*tamāma*	*egānagī*	*āve*
then	strangeness	go-CVB	whole	excellence	come.3SG.PRS

'Then the strangeness would go [away] [and] the whole excellence would come [in its place].' (MV.4.65) 1636

3.4. ABSOLUTE CONSTRUCTIONS

Some of the above examples contain the overt subject of the converb (113-115). This type of construction resembles an absolute construction (AC), the basic properties of which we have already discussed in section 2.2.5.

Early NIA continued to use AC's with the subject in an oblique case and the predicate in the form of an inflected imperfective or perfective participle.

In Early Rajasthani both present and past participles can be a part of the AC. Present participles are attested in the oblique plural (genitive plural) with ending *-ā̃* (116). The past participle also has the ending *-ā̃* or *-aï*, which may be either locative or instrumental (cf. Tessitori 1915: 103-104). The head noun is not always in the oblique case, but some cases are evident, as in (117).

In our corpus Early Awadhi AC's have imperfective participles terminating in *-ta* (118) and perfective participles terminating in *-e(ṁ)* (119) (see also Śukla 2022: 138-143). Quite similar is the situation in Early Braj, where imperfective forms in *-ta* (120) and perfective forms in *-e* (121) also occur in AC's, although in our Early Braj examples, in both types of AC's – with imperfective as well as with perfective participles – the head noun may appear with a postposition GEN (120) or DAT/ACC (122). For Dakkhini we have just one example with a perfective participle in an oblique case (123).

(116) Early Rajasthani

tai	*narasiṃghadāsa*	*kā*	*kaṭaka-baṃdha*	*cālitāṃ*
this.OBL	Narasimhadasa	GEN.OBL.SG	army.M.SG	go.PRS.PTCP.OBL.PL
sāṃtar-i	*āgal-ai*	*daḷ-i*		
best.M.SG-LOC	front.M.SG-LOC	army.M.SG-LOC		

'When the army of Narasinghadasa moved, in the frontal part of (his) best army (there was water).' (RG.GŚ.9) AD 1428

(117) Early Rajasthani

teṇi	*pātisāh-i*	*āyāṁ*	*sāṁtari*	*kuṇa*	*sahaï*
this.INS	king-INS	come.PPP.OBL.PL	burden.F.SG	who	bear.3SG.PRS

'When the king came, who bears the burden?' (RG.GŚ.43) AD 1428

(118) Early Awadhi

bhora	*hota*	*bāsahiṃ*	*cuhacuhī*
morning	be.PTCP.PRS	sing.3PL.PRS	cuhcuhi bird

'When the morning comes cuhcuhi birds sing.' (J.29.2) AD 1540

(119) Early Awadhi

khela	*gaeṁ*	*kata*	*khelai*	*koī*
play.NOM.M.SG	go.PPP.M.OBL	how	play.3SG.PRS	someone

'If a play is over, how can anyone play?' (J.63.6) AD 1540

(120) Early Braj

udita	*hota*	*sibarāja*	*ke*	*mudita*	*bhae*	*dvija*	*deva*
arrived.PRS.PTCP	be.PRS.PTCP	king Shivaji	GEN	happy	be.PST.PL	twice-born	god

'On king Shivaji's coming (i.e. birth) gods and twice-born became happy.' (Ś.29) 1673

(121) Early Braj

aise	*vacana*	*kahe*	*lachimanā*	*tajyau*	*soga*	*rāghava*
like	speech	say.PPP.M.OBL	Lakshmana	abandon.PPP.M.SG	sadness	Rama

tatachinā
immediately

'When Lakshmana said such words, Rama abandoned [his] sorrow immediately.' (V.1.121) 15th c.

(122) Early Braj

su uni	*kahuṃ*	*kachū*	*sumeru*	*pāyeṃ*	*hī*	*ānaṃdu*	*na*	*hvaihai*
so this.OBL.PL	DAT.ACC	some	Meru	obtain.PPP.OBL	very	joy	not	be.3SG.FUT

'So there won't be any joy after some obtain [the mountain] Meru.' (I.b113) AD 1600

(123) Early Dakkhini

bhaī	*use*	*dekhe*	*piche*	*pachānata*	*kā*	*eka*	*qatarā*
be	DEM.3SG.OBL	see.PPP.M.OBL	after	recognize	of	one	drop
ātā	*hai*						
come.PRP.M.SG	be.3SG.PRS						

'After seeing it (lit. after it [sc. the light] has appeared), a drop [out] of the recognized one comes/appears.' (BN.124) 14th c.

The coexistence of the two AC's, namely that based on the converb and that based on the participle, raises several questions as regards their functional differences. As we have already seen in section 2.2.5, contemporary NIA also has two such constructions, but the converbal one is clearly subject-oriented, whereas the participial one is not.

From the morphological point of view, the participles which formed part of AC's up to the MIA period showed their inflectional characteristics – they agreed with the head noun in gender, number and case. This was never the case with converbs, which had an uninflected form. The situation changed in early NIA, because participles did not show any agreement with the head nouns, and in AC's they often occur in a petrified form, e.g. genitive plural as in the case of Early Rajasthani, or oblique sg. as in the case of Early Awadhi, Dakkhini or Early Braj.

In Early NIA AC's both constructions, i.e. (i) 'different-subject' converbal constructions and (ii) participial AC's, are rare, and as in modern NIA languages, subject-orientation of converbs is dominant. In our corpus there are only a few instances of the type represented by (41-43) compared with a few hundred instances of same-subject converbal chains. AC's based on adverbial participles are extremely rare, and there are also a few examples of quasi AC's with coreferential subjects, as in (124):

(124) Early Dakkhini

khela	*khelate*	*avidye*	*ke*	*khalīte*	*meṃ*	*ghusā*
play	to play.ADV.PTCP	ignorance	of	sack	in	enter.PPP.M.SG

'While playing, [it, i.e. a pigeon] entered into the sack of ignorance.' (E.II.35) AD 1548-1599

The interpretation of constructions in which the converb can receive a passive interpretation poses certain difficulties. When the main verb is in the passive voice, the converb receives a passive reading as well (125).

Our texts contain isolated instances of constructions in which the main verb is intransitive and the converb may have a passive interpretation. Such constructions naturally have S of the main verb coreferential with the subject of the passive converb, cf. *bisari* 'forget' in (126). The main problem with such rare constructions is that converbs usually do not take any passive morphology. Only isolated instances of passive converbs with passive morphology can be found (127). Such converbs with passive morphology are also attested in modern NIA, but they are extremely rare (cf. Kāmtāprasād 2060: 406).

A possible passive interpretation of converbs in the ergative construction is presumably the most intricate. In (128) the main verb, being in the participial form, may receive either an active or a passive interpretation, and the interpretation of the converbal form may be analysed in a similar manner. But in (129) the interpretation is more difficult, since the converb, if analysed actively, cannot have a coreferential subject with the subject of the main clause (because it was not the beneficiary of the grant who exempted himself from the taxes) and we are left with an AC *per se*. The passive interpretation of the converb gives another logical and possibly more preferred option.

(125) Early Rajasthani

pātisāhāṃ	*rī*	*gaja-ghaṛā*	*paṛījai*
Shah.M.PL.OBL	F.GEN	elephant's army.F.SG.NOM	be attacked.3SG.PRS.PASS
aujhaṛāṁ		*māri*	*ṭhelījai*
slanting blow.M.PL.OBL		beat.CVB	chase away.3.PRS.PASS

'The shahs' army of elephants is being attacked and having been beaten with slanting blows is chased away.' (RG.KJ.12) 18thc.

(126) Early Awadhi

pāī	*bhuguti*	*sukkha*	*mana*	*bhaeū*
get.PPP.F.SG	eating.F.SG	happiness	heart	become.3SG.PST
ahā	*jo*	*dukkha*	*bisari*	*saba*	*gaeū*
be.3SG.PST	which	sorrow	forget.CVB	all	go.3SG.PST

'He ate and the happiness filled [his] heart, the sorrow which was [there], having been forgotten, was all gone.'

'After eating, his hunger disappeared and he was happy and, having forgotten the sorrow, all is gone.' (J.66.5)

(127) Old Awadhi

niti	*gaṛha*	*bāṁci*	*calai*	*sasi*	*surū*
always	fort.NOM.M.SG	avoid.CVB	go.3SG.PRS	moon.NOM.M.SG	sun.NOM.M.SG
nāhi	*ta*	*bāji*	*hoi*	*ratha*	*cūrū*
not	this	strike.CVB	be.CVB	cart.NOM.M.SG	smash.PPP.M.SG

'Moon and sun always go avoiding the fort, if not, their carts being struck are smashed' (J.41.1) AD 1540

(128) Early Awadhi

sakala	*dīpa*	*mahaṁ*	*cuni*	*cuni*	*ānī*
all	island.M.PL.OBL	in.LOC	choose.CVB	choose.CVB	bring.F.SG.PPP

'Having been chosen (the queens) have been brought from all islands.'

'Somebody having chosen the queens brought them from all islands.' (J.49.7) AD 1540

(129) Early Pahari (Kumaoni)

sarvva	*kara*	*akarī*	*sarvva*	*dvaṁda*	*viśuddha*	*kari*
all	tax.M.SG.NOM	undo.CVB	all	dispute.M.SG.NOM	purified	do.CVB

pāi	*Viru*	*le*
get.F.SG.PPP	Viru.M.SG.OBL	ERG

‘(The king) has exempted (lit. having exempted) all taxes and purified all disputes and Viru got (this grant).’

‘After all taxes have been exempted, all disputes purified, Viru got (this grant).’ AD 1418

The overall status of different-subject constructions based on both participles and converbs is rather marginal, but in comparison to contemporary NIA, SIC violations and the frequency of AC’s seem to be higher in early NIA. This seems to be a result of several morphosyntactic phenomena observed during this period, among which the following can be mentioned: a) two opposite tendencies – the emergence of the new ergative pattern based on a new postpositional system and the attrition of A marking – both resulting in temporal OEM; and b) the establishment of DOM.

4. SCOPE OF SELECTED OPERATORS

At the end of section 2.2.2 we defined a number of features and the values according to which we plan to demonstrate the type of linking represented by converbal chain constructions. We have taken three operators, namely T (Tense), IF (Illocutionary Force) and NEG (Negation), investigating their scope in four corpora, namely Early Rajasthani, Early Awadhi, Early Braj and Early Dakkhini. In this section we also aim to put forward some hypotheses regarding a) the syntactic status of the converb, which cannot be easily defined in terms of a subordinating, coordinating or even co-subordinating device, and b) the perfect meaning of the converb. We expect these hypotheses to find substantial support in the diachronic data that we have collected.

4.1 IF-SCOPE

Illocutionary force (IF) is a very important operator of universal nature. In every language it is possible to express questions, commands or statements, and there are various means of doing this, e.g. question words (Wh-words), particles, clitics, word order, etc. In IA, intonation, grammatical markers of mood and Wh-words are regularly used to mark IF.

Theoretically IF-scope can be: a) conjunct; b) disjunct; c) local; d) extensible; e) constraint-free. In fact, in Early New Indo-Aryan we find numerous examples of converbal chains in which two readings may be available, either with conjunct or with local scope. This is the case with converbal chains with the main verb in the imperative mood. In all instances below, the converbal clause is preposed and it may lie within the scope of the imperative.

(130) Early Rajasthani

paṇi	*tumhẽ*	*mayā*	*karī*	*deśāntari*	*pahucaü*
but	you	mercy	do.CNV	abroad.OBL	reach.IMP

'But you, having shown mercy, go abroad.'
'But you show mercy and go abroad.' (RG.M.15) 15th c.

(131) Early Awadhi

kai	*cali*	*hohu*	*suā*	*sanga*	*satī*
or	go.CNV	be.IMP	parrot.M.SG	with	satī.F.SG

'... or having gone become satī with the parrot.'
'... or go and become satī with the parrot.' (J.88.8) AD 1540

(132) Early Braj

baṃdi	*mokhū*	*devani*	*kauṃ*	*dehu*
pay homage.CVB	liberation	god	DAT	give.2SG.IMP

'Give liberation, having paid homage to gods.'
'Give liberation and pay homage to gods.' (V.1.120) 15th c.

(133) Early Dakkhini

pāṁca	*mila-ke*	*insāpha*	*karanā*
five	meet-CVB	decision	do.2IMP/INF

'Make decision/verdict/justice [in] five! (lit. having met [in] five).' (E.III.44) 16th c.

From the collected data we conclude that the conjunct scope in commands is not available when the main verb is preposed, as in (134) and (135). This certainly requires further investigation, but the fact that the imperative IF-scope can depend on the constituent order is quite evident in our corpus.

(134) Early Braj

sukha	*sovahi*	*kāyara*	*kī*	*nāri*	*hāranu*	*jitanu*
happiness	sleep.2SG.IMP	coward	GEN.F	woman	lose.INF	win.INF
virogu	*nivāri*					
separation	dismiss.CVB					

'Sleep [in] happiness, cowardly woman, having dismissed the separation [from] losing and winning.' (V.1.6) 15th c.

(135) Early Braj

deṣi	*saṁbhāri*	*pīta*	*paṭa*	*ūpara*	*kahāṁ*	*cūnarī*	*rātī*
see.2SG.IMP	put on.CVB	yellow	cloth	above	where	(partly dyed) cloth	red

'See [how] you have put on the yellow cloth, where [is] the red garment?' (HH.20.3) 16th c.

In assertive sentences the IF-scope is exclusively conjunct, but the negation of the main verb blocks extension of the IF-scope (cf. (136) with a local IF-scope). This is in line with Tikkanen's observations regarding earlier stages of IA, that negation of the main verb causes backgrounding and a restrictive reading of the converbal clause (Tikkanen 1987: 161-162).

(136) Early Awadhi

rasa	*taji*	*risi*	*kabahuṁ*	*na*	*kījai*
love	leave.CVB	anger	when	not	do.3SG.PRS/IMP.PASS

'Having abandoned love one should never be angry.' (J.90.5) AD 1540

The position of Wh-words determines the IF-scope in questions. There are several interdependencies found in the corpus. If the Wh-word is in the leftmost position in the

clause, then its scope is preferably conjunct, as in (137). If the Wh-word occurs in front of the main verb with the main clause postposed, the scope is preferably local (138). The same tendency pertains to AC's, as can be seen in (117) and (119). We may state here that different subject constructions *per se* block IF extension, as seems to be the case with AC's, but if converbal constructions obeying SIC also show local interrogative IF-scope in cases where the Wh-word is in the postposed main clause, then we suggest that the position of the Wh-word is a decisive factor here.

(137) Early Awadhi
kata cirihāra ḍhukata lai lāsā
how fowler hide.PRS.PTCP.M take.CVB bird-lime
'Why would the fowler hide himself and use an adhesive to catch birds?' (J.70.4) AD 1540

(138) Early Braj
haraṣita iṃdu tajata jaisai jaladhara so bhrama ḍhūṃḍhi
delighted moon abandon.IMPRF.PTCP like cloud so error search.CVB
kahāṁ hauṃ pāūṁ
where I.NOM obtain.1SG.PRS.SBJV
'As the delightful moon abandons the cloud, so having searched, where would I find out the error (delusion)?' (HH.14.4) 16th c.

The IF-operator cannot be taken as a decisive factor as regards linking types in early NIA. It is very often not too clear whether IF-scope should be conjunct or local. An analysis of Imperative IF-scope shows that both options can be considered, with a slight preference for conjunct scope. It is thus not obvious here whether we are dealing with constraint-free scope, but this option cannot be excluded. What we also see is that with the left position of the matrix clause the imperative IF-scope is preferably local. Interrogative IF-scope can be conjunct or local, and this depends on the position of the Wh-word. Since the default postion of the Wh-word in IA is leftmost, it is not surprising that there is a preference for conjunct scope when the Wh-word occurs in its default position and the matrix clause is preposed. Moving the matrix clause with the Wh-word to the right results in blocking of the possibility of interrogative IF-scope extension. Thus, IF-scope precludes the applicability of the discrete notion of cosubordination for early NIA. We are able to check the IF-scopal properties and to capture the main properties of clausal junctures involving converbs.

4.2 T-SCOPE

Tense is a clause-level operator. As we have seen in Figure 12, in structures labelled tentatively as cosubordinative, the T-operator is supposed to be shared. However, in early NIA this is not always the case, and we assume that there may be other grammatical factors which influence the T-scope. One of them, at which we shall take a closer look, is the tense of the main verb.

According to our analysis, in constructions with the main verb in the past tense T-scope is preferably conjunct, whereas in constructions with the main verb in non-past tenses it is expected to be local (compare (139) with conjunct scope to (140) with local scope).

(139)

a) Early Awadhi
paṃkhi-nha *dekh-i* *saba-nhi* *ḍara* *khāvā*
bird-OBL.M.PL see-CVB all.OBL.PL fear.M.SG eat.M.SG.PST
'... birds saw all of that and got scared.' (J.69.2) AD 1540

b) Early Rajasthani
āmbā *nī* *ḍāḷa* *namāṛī* *āmbā* *leī*
mango.OBL.M.SG GEN.F branch.F bend.CVB mango.M.PL take.CVB
ḍohalu *pūriu.*
craving.M.SG fill.PST.M.SG
'(The outcast) bent the branches of mango tree, took mangos, fulfilled (her) craving.' (RG.SS.5) 15th c.

c) Early Braj
rac-i *rac-i* *apane* *hātha* *saṁvāryau*
make-CVB make-CVB own.OBL.M.SG hand arrange.PST.M.SG
nikuṃja *bhavana*
bower.M.SG house.M.SG
'[He] made [it with his] own hands, [he] arranged the bower house.' (HH.39.2) 16th c.

d) Early Dakkhini
hora *suka* *duka* *dono* *jāna-kara* *khudā* *kī* *yāda* *meṃ*
and happiness sorrow both know-CVB god GEN.F memory LOC
rahyā
remain.PST.M.SG
'And [he] experienced both happiness and sorrow [he] remained in God's memory.' (BN.73) 15th c.

(140)

a) Early Awadhi
siddha *ḍarahiṃ* *nahiṃ* *apane* *jīvāṁ*
holyman fear.3PL.PRS not self.M.OBL.SG life.M.OBL.SG
kharaga *dekh-i* *kai* *nāvahiṃ* *gīvāṁ*
sword see-CVB CVB bow.3PL.PRS neck.F.OBL.SG
'Holy men do not fear for their own life but they bow their necks having seen a sword.' (J.240.3) AD 1540

b) Early Rajasthani
phūladhārā *vica* *uḍi* *paṛāṃ*
stream of flowers middle fly.CVB fall.1PL.PRS
'Having flown in the middle of the stream of flowers we fall.' (RG.KJ.30) 18th c.

c) Early Braj
kapi *nāyaka* *ke* *jūtha* *bulāi* *rājahiṃ* *preta*
monkey leader GEN.M.PL troop call.CVB shine3SG.SBJ spirit
'The spirit would shine, having summoned troops of monkey leaders.' (V.1.50) 15th c.

d) Early Dakkhini
to *bīja* *hora* *jhāṛa* *dono* *mila-kara* *qālabiyata* *ku* *apaṛe*
then seed and heavy rain both meet-CVB model ACC attain.3SG.SBJ
yāne *isa* *khākī* *vajūda* *kūṁ*
that is this earthy manifestation ACC
'Then both the seed and the rain, having met, deliver/attain the body/model, i.e. the material manifestation.' (BN.28) 14th c.

In general, in Early NIA the converb is usually attested with past tenses. In his work on the Sanskrit converb, Tikkanen (1987: 129-132) assumes the perfectivity of this formation to be a direct consequence of its semantics of relative past tense. According to the Sanskrit data collected by him, the incompatibility of the perfective aspect and non-past tenses is quite evident. Similarly, Davison (1981) argued that Hindi converbs are perfective. If we look at (141) we can see that this incompatibility of perfective aspect and non-past tense exists also in contemporary NIA languages. The T-scope is local – the tense of the matrix clause is future, whereas the tense of the converbal form is present perfect.

(141) Hindi (Davison 1981: 123, fn.7)
ham *sṭāks* *āj* *kharīd-kar* *do* *sāl* *ke bād* *bec* *dẽge*
we stocks today buy-CVB two year after sell give.AUX.FUT.1PL
'We have bought stocks today and we will sell them in two years.'

The question should be posed here whether we have any typological support for a claim that perfective converbs are not fully compatible with non-past tenses. We are not aware of any large-scale studies dealing with this problem, but, for example, some recent research on Polish converbs shows that anterior converbs (which are semantically the closest equivalents to IA converbs) are rarely used in non-past tenses. According to Bojałkowska (2010: 206, 213-215), in Polish, anterior converbs co-occur with main verbs along the following hierarchy:

past tense > present tense > future tense > imperative, subjunctive

In her database, 82% of the main verbs are in the past tense, 12% in the present tense and only 3% in the future tense. The remaining forms are subjunctives and imperatives.

T-scope in anterior converb constructions can be conjunct when the main verb is in the past or future tense (Bojałkowska 2010: 208-211). This results from the aspectual structure of the Polish verb, which maintains the imperfective/perfective opposition in the future and past tenses, but not in the present tense, where the only available aspectual value is imperfective. Anterior converbs are considered to be perfective, and therefore, if they denote an event completed before the one expressed by the future or past tense, they can refer to an event in the future or past. For an event completed before the one expressed by the verb of the main clause in the present tense, there is no present perfective option, and for that reason converbal constructions with the main verb in the present tense have exclusively local T-scope. If we compare converbal chains (142), (143) and their finite verb equivalents in (142'), (143') we see that the T-scope in the converbal chains is clearly conjunct – the finite verb equivalent of the converbal clause has the same tense as the superordinate clause. However, in a converbal chain with the main verb in the present tense (144) the T-scope is local, because as the finite verb construction (144') shows, the tenses of the subordinate and superordinate clauses are different. The equivalent (144') also shows that the temporal interpretation of the converb can to a certain extent be ambiguous in terms of tense (examples from Bojałkowska 2010: 208-211).

(142) Polish

Dalej	*będziemy*	*działać,*	*zapoznawszy*	*się*	*z*	*tym*	*materiałem.*
further	be.FUT.1PL	act.INF,	acquaint.CVB	REF	with	this.INS	material.M.INS.SG

'We will act further, having acquainted ourselves with this material.'

(142') Polish

Dalej	*będziemy*	*działać,*	*kiedy*	*zapoznamy*	*się*	*z*	*tym*
further	be.FUT.1PL	act.INF,	when	acquaint onself.FUT.1PL	REF	with	this.INS

materiałem.
material.M.INS.SG
'We will act further, when we acquaint ourselves with this material.'

(143) Polish

Przeszedłszy	*do*	*drugiego*	*pokoju*	*kucnęła*
move on.CVB	to	second.GEN.SG.M	room.GEN.SG.M	squat.PST.3SG.F

przy *książkach.*
near book.INS.PL.F
'Having moved on to the second room she squatted near the books.'

(143') Polish
Kiedy *przeszła* *do* *drugiego* *pokoju*,
when move on.PST.3SG.F to second.GEN.M.SG room.GEN.M.SG
kucnęła *przy* *książkach.*
squat.PST.3SG.F near book.INS.F.PL
'When she moved on to the second room she squatted near the books.'

(144) Polish
[...] *wyrwawszy* *się* *z* *korka* *chętnie* *stosujemy*
break out.CVB REF from traffic jam.GEN.M.SG willingly apply.1PL.PRS
'ciężką *nogę'.* [. . .]
heavy.ACC.F.SG leg.ACC.F.SG
'Having broken out from the traffic jam we speed up.'

(144') Polish
[...] *po tym,* *jak* *wyrwiemy* / *wyrwaliśmy* *się* *z*
after this.INS how break out.1.PL.FUT / break out.1.PL.PST REF from
korka, *chętnie* *stosujemy* *ciężką* *nogę* [. . .]
traffic jam.GEN.SG willingly apply.1PL.PRS heavy.ACC.F.S leg.ACC.F.SG
'After we break out/have broken out from the traffic jam we speed up.'

Even though Polish anterior converbs occur with non-past tenses only rarely, such constructions are perfectly grammatical. This tendency clearly shows the interdependency between the meaning of the converb and its incompatibility with the non-past tenses. This in turn is an additional piece of evidence for the perfectivity of the converb. We assume that in early NIA the situation may be to a certain extent similar to Slavonic, and therefore our diachronic data confirm the synchronic analysis of the IA converb as perfective (e.g. Davison 1981). There is at least one more important theoretical implication, namely that if the variable scope of the T-operator (i.e. conjunct or local) depends on the tense of the main verb, it will require two different linking types. As in the case of the IF-operator, we cannot posit a discrete notion of clausal cosubordination for clausal junctures involving converbs, because it would require quite stable T-operator sharing.

4.3 NEG-SCOPE

Negation is an operator which can be nuclear, core and clausal. We are concerned here with internal and external negation, i.e. core and clause-level negation. According to the metrics proposed in section 2.2.2, NEG-scope can be a) conjunct; b) local; c) variable; and d) extensible. Kachru (1981: 42) states that in modern New Indo-Aryan languages NEG-scope can be variable at least in sequential constructions (chaining). Thus, (145) can have three

interpretations. There are also less ambiguous examples where apart from the conjunct NEG-scope no other interpretation is available – these include converbal constructions with adverbial readings, as in (146), as well as, interestingly, some sequential constructions (147).

(145) Hindi (Kachru 1981: 42)
us-ne *nahā-kar* *khānā* *nahĩ̃* *khāyā*
s/he bathe-CVB meal.M.SG not eat.PST.M.SG
'He did not eat after bathing.'
1. s/he ate before bathing
2. s/he went directly to the office without eating
3. s/he only had tea after washing her/his face

(146) Hindi (Kachru 1981: 42)
tum *man* *lagā-kar* *nahĩ̃* *paṛhte*
you mind apply-CVB not study.IMPF.M.PL
'You don't study diligently.'

(147) Hindi (Tikkanen 1987: 31; from Premcand's 'Godān')
ab *god* *se* *utar-kar* *pā̃v-pā̃v* *kyõ* *nahĩ̃* *caltī*
now lap ABL descend-CVB foot-foot why not walk.IMPF.F.SG
'Now why don't you get down from (daddy's) lap and walk on your feet.'

In early NIA NEG-scope can also be conjunct in sequential constructions (148) with the NEG marker in the main clause. The position of the NEG marker can determine the NEG-scope, but it is at the same time closely correlated with the position of the subject. (148) and (149) differ in NEG-scope because they have different subject positions – in (148) we have a converbal clause *in situ* (cf. Subbarao 2012: 265), whereas in (149) the subject is fronted, which implies local scope. We assume that in (150) the covert subject will also be fronted, and as a result the NEG-scope is also local.

In adverbial constructions such as (151) the scope of NEG is extended to the converbal clause. This phenomenon is labelled NEG-transport, and it has already been discussed in the typological literature (cf. Bickel 2010: 59-60). This seems to be an exclusive property of the NEG operator – we have already given examples from contemporary NIA in section 2.2.2 (24) – and it seems to be mostly reserved for core junctures.

One more option is represented by (152), where the frontal position of the NEG marker in the preposed converbal clause may result in the extension of the NEG-scope to the adjacent main clause. This is the opposite kind of NEG-transport to that attested in (151).

(148) Early Braj

rājakāja	*dekhi*	*koū*	*pāvata*	*na*	*bheu*	*hai*
royal act	see.CVB	someone	obtain.IMPF	not	secret	be.3SG.PRS

The one who has not seen your royal act does not reach (your) secret.
[lit. Having seen the royal act, no one reaches (your) secret.] (Ś.170) 17th c.

(149) Early Awadhi

jauṃ	*tivāīṁ*	*kai*	*kaja*	*na*	*jānā*
if	woman.INS.F.SG	do.CVB	work	not	know.PPP.M.SG

'If the woman didn't realize what she has done' (J.86.4) AD 1540

(150) Early Awadhi

sūjhai	*bhuguti*	*na*	*sūjha*	*biādhū*
see.3SG.PRS	eating	not	see.CVB	fowler

'[We] see eating not having seen the fowler.' (J.72.5) AD 1540

(151) Early Dakkhini

moammā	*nahīṃ*	*kahe*	*isa*	*bāta*	*kūṃ*	*khola*
enigma	not	say.3SG.SBJ	this	matter	ACC	open.CVB

'One does not say this enigmatic thing openly (lit: having opened).'
= 'One says this thing not being open.' (H.39) AD 1641

(152) Early Awadhi

na	*āi*	*sahī*	*yaha*	*dhūpā*
not	come.CVB	suffer.PPP.F.SG	this.DEM	sunshine.NOM.F.SG

'even if having come again, did not suffer this sunshine.' (J.27.7) AD 1540

5. CONVERBAL CHAIN CONSTRUCTIONS – CONCLUDING REMARKS

Contemporary NIA appears to maintain the morphosyntactic and scopal properties of clausal operators in converbal chain constructions.

SIC is regularly preserved, although the absolute use of converbs is also attested (see (153) with a volitional subject). SIC violation is also possible when there is an experiencer- or possessor-like relation between the subject of the main verb and the subject of the converb (see the DAT experiencer in (154) and (155)).

(153) Braj from Liperovskij (1987: 132)

nagar	*mẽ*	*āi-kẽ*	*khūb*	*dhūm-dhām*	*bhaī*
city	LOC	come-CVB	good	pomp and show	be.PST

'Having come to the town, there was a great pomp and show.'

(154) Braj from Liperovskij (1987: 132)

bvā	*totā*	*kī*	*bāt*	*suni-kẽ*	*bā-kũ̃*	*tarasu*	*āi*	*gayau*
this.OBL	parrot	GEN.F	speech	listen-CVB	she-DAT	pity	come	go.PST.M.SG

'Having heard the speech of this parrot she felt pity.'

(155) Awadhi from Liperovskij (1997: 163)

yū dekhi-kai	*un-kā*	*baṛā*	*acaraju*	*bhā*
he see-CVB	he.OBL-DAT	big	surprise	be.PST

'Having seen him he became very surprised.'

Scopal properties show similar regularities as in early NIA. The scope of the IF-operator does not have to be conjunct. The local scope of an imperative IF-operator is exemplified in (156).

(156) Awadhi from Lieprovskij (1997: 162)

tum	*baniyā*	*kere*	*mehariyā-kā*	*phusalāy-kai*	*lai*	*āwo*
you	merchant	GEN	wife-ACC	seduce-CVB	take	come.IMP

'Having seduced the wife of the merchant bring (her).'

The scope of NEG can thus be conjunct when the NEG marker is on the main verb with the converbal clause preposed (157). The NEG-scope is not necessarily transported to the main clause when the preposed converbal clause is negated (158).

(157) Braj from Liperovskij (1987: 132)
ab *hanumān* *ke* *mandir* *mẽ* *jāy-kai* *nə̃y* *soungo*
now Hanuman GEN.M.OBL.SG temple LOC go-CVB not sleep.FUT.1SG
'Now I will not go to Hanuman's temple and I will not sleep there.'

(158) Braj from Liperovskij (1987: 132)
sīdī sādī *na* *kah-kai* *bā-nai* *apanī* *bāt* *mīṭhī* *rīt* *te*
straight not say-CVB he-ERG own.F speech sweet way ABL
kah *dīnī* *ī*
say give.PST.F be.PST.F
'Not having spoken straightforwardly, he expressed himself in a hypocritical manner.'

Early NIA languages present an interesting example of transitional languages which display a rather unstable system of main argument marking in perfective tenses. There is an important areal differentiation in A marking – all analysed languages mark A's in converbal chain constructions inconsistently and only with the main verbs transitive. Languages belonging to the eastern branch, such as Early Awadhi, and to the western branch, such as Early Rajasthani, lost A marking at quite an early stage. Languages of the West such as Braj and Pahari maintained A marking, and Dakkhini lost it due to long-lasting contact with Dravidian. We have no instances of marking triggered by the transitivity of the converb; therefore we conclude that this was a later development.

The phenomenon of DOM shows considerable variability. Even though the first instances of O marking are attested at quite early stages in all dialectal groups, the full establishment of the DOM rule (i.e. animacy and definiteness) is quite a late phenomenon. Moreover, as we can see from the presented data, some modern NIA languages, such as Early Pahari, still show unmarked O's even in contexts which require regular marking.

SIC was violated in instances where there was a possessor- or experiencer-like relation between the subjects of the main and converbal clauses, and unlike in modern NIA converbs could denote both non-volitional and volitional acts. Violation of the SIC rule led to the rise of the absolute construction based on the converb, which coexisted with the absolute construction based on the past participle. As we have seen, converbal constructions predominantly conformed to SIC, whereas those based on participles did not. This differentiation seems also to exist in modern NIA.

Finally, the scope of selected operators, such as T-scope, IF-scope and NEG-scope, can be interpreted as a diagnostic tool for the syntactic status of the converbal chain construction.

Converbal constructions in early NIA do not clearly show operator dependency. Davison (1981) is certainly right to state that the scope of the IF and NEG operators is often dependent on pragmatic factors such as the discourse prominence of the clauses being a part of the converbal construction, discourse context, speaker's knowledge, etc. However, we have also seen that in early NIA, IF- and NEG-scope may often result from the position of the markers

and the position of the subject as well. By contrast, T-scope has a structural motivation, being dependent on the tense of the main verb. Evidence for this feature is found in several branches of early NIA and in modern NIA (in modern Hindi in particular) as well as in other branches of IE.

What is more, scopal properties are dependent on the juncture type, as is quite obvious. Although we have been chiefly concerned with clausal junctures, sometimes (given the fact that we are dealing with written texts belonging to various genres) the boundary between core and clausal junctures may be fuzzy, and this causes additional difficulties in the analysis of scopal properties.

As a result of applying the limited and modified set of variables taken from the Multivariate Analysis Model, we believe that it has been demonstrated that the third type of linking proposed within the RRG model, namely cosubordination, is not fully applicable to IA. As in the case of modern NIA, early NIA cosubordinate-like structures, i.e. coverbal chains, do not always show operator dependency, a defining feature for cosubordination. Even though Van Valin (2007: 80) considered the optionality of operator sharing, we believe that Bickel's (2010) model, in which terms such as 'cosubordination' are decomposed into sets of variables, captures in a more precise way the major properties of converbal constructions in IA, and in turn the main tendencies in their development. In our opinion IA exhibits quite a stable system of clause linking with respect to non-embedded structures between early and modern NIA, with early NIA being more labile due to the transitional phase in alignment reorganization.

6. MORPHOSYNTAX OF INFINITIVES IN EARLY NIA

In section 2.3.1 we have already given the major properties of infinitives from a more general perspective, and in section 2.3.2, while discussing the relation between infinitives and other non-finite categories such as gerundives and action nominals, we have also referred to the morphological properties of NIA infinitives.

In this section we shall deal with the formal make-up of infinitives in our corpora and the distribution of various forms.

	-n/ṇ-	*-v/w-*
dvitīya vrata satya para kathā (XIV)	4	5
guru mahimā par kathā (XV)	-	2
amarasena-vayarasena kathā (XV)	1	19
vacanikā khīcī acaḷadāsa-rī (XV)	-	-
vīsaḷadevarāsa (XVI)	4	-
daḷapata-vilāsa (XVI/XVII)	1	-
haḍai sūrijamala-rī vāta (XVII)	9	
vacanikā rāṭhoṛa ratana-rī (XVII)	-	-
rāṭhoṛa durādavāsa-ro kāgada (XVII/XVIII)	-	-
dhanuṣa-bhaṅga (XVIII)	11	3
adālatī nyāya (XVIII)	16	-
ḍokarī rī vāta (XVIII)	2	-

Table 14. Distribution of infinitive forms in Early Rajasthani.

In Rajasthani there are two forms attested, in *-v/w-* and *-n/ṇ-*, and their distribution shows that the former was more frequent at least up to the 15th century. Their distribution in our corpus is given in Table 14.

In early texts the *-v/w-* form is used in purposive clauses (159). We have a bare infinitive in its oblique form, but in some cases the infinitive form had already undergone extension, as in (160) or (161). This is in line with the hypothesis put forward by Haspelmath (1989), already discussed in section 2.3.1, according to which the infinitive in purposive clauses, in losing its purposive modality, often needs reinforcement to maintain it.

(159) Early Rajasthani

sīmāḷau	*purī*	*levā*	*āwiu*
standing on the border.M.SG.PPP	town.F.SG.ACC	take.M.SG.INF.OBL	come.M.SG.PPP

'Standing at the border (he) came to take the town.' (RG.TS.3) 14th c.

(160) Early Rajasthani

teha	*mārāwiwā*	*kāraṇ-i*	*amhe*	*mokaḷiyā.*
he.SG.OBL	kill.INF.SG.OBL	reason.M.SG.INS	we.1PL.NOM	send.M.PL.PPP

'He sent us here to kill (him).' (RG.TS.9) 14th c.

(161) Early Rajasthani

amarasena-vayarasena	*āścarya*	*jovā*	*nai*	*kāji*	*deśāṃtara*
Amarsena-Vayarsena.NOM	wonder.M.PL	see.INF.OBL	M.GEN.OBL	for	abroad.OBL
bhaṇī	*cālyā.*				
for	go.3PL.PPP				

'Amarsen and Vayarsen (...) went abroad to see the unknown/wonders.' (RG.M.19) 15thc.

From the earliest sources, the *-v/w-* form is used in an inceptive construction with the verb *-lag* 'begin' (162) and the infinitive in the oblique case. This usage is quite frequent. Similarly there are permissive constructions attested with the verb *de-* 'give' (163). Interestingly, an obligative construction is also attested only with the *-v/w-* form (164) – in the obligative pattern A, if present, is marked by INS and there is object–verb agreement (for details of the development of the obligative pattern in Early Rajasthani see Khohklova 2013).

(162) Early Rajasthani

pachai	*bewai*	*putra*	*kahivā*	*lāgā.*
after	two	son.M.SG.NOM	say.INF	start.M.PL.PST.PTCP

'Then two sons started to say.' (RG.M.12) 15th c.

(163) Early Rajasthani

sūrajamala	*āvaṇa*	*dai*	*nahīṃ*
Surajamal.NOM	come.INF.OBL	give.PRS.3SG	not

'Surajmal does not allow (us) to come.' (R.G.MN.45) 17th c.

(164) Early Rajasthani

taiṃ (...)	*eka*	*yakṣa*	*nī*	*pūjā*	*variwī.*
you.2SG.INS	one	yaksha.M.SG.OBL	F.SG.GEN	worship.F.SG	perform.INF.F

'You (...) should perform worship of one Yaksha.' (RG.M.69) 15thc.

The *-n/ṇ-* form is used basically in the complement function (165) and in the subject function (166) as well, although the *-v/w-* form is also used in the complement function (167)

at least up to the 15th c. We have also come across an example of the matrix coding as PSA construction (168), which is interesting as this possibility is not available in contemporary NIA.

(165) Early Rajasthani
bheṭaṇau *mānau* *chai*
make a gift.INF accept.3.PRS be.3.AUX.PRS
'He accepts my offerings.' (RG.M.62) 15thc.

(166) Early Rajasthani
pāchauṃ *vaḷanu* *yuktau* *nahīṃ*
back return.M.SG.INF proper.M.SG not
'Returning is not proper.' (RG.TS.4) 14thc.

(167) Early Rajasthani
vyāpāra *viṇa* *ewaṛau* *kharicivau* *kima* *pahucai*
activity/trade.M.SG.OBL without that much spend money.INF how obtain.3SG.PRS
chai.
be.3SG.AUX.PRS
'How can you afford to spend so much without (having) a job?' (RG.M.38) 15thc.

(168) Early Rajasthani
kūṛai *bhaṇii* *saṃgha-lūṃṭaṇu*
lie.M.SG.LOC say.SG.INS.PPP plundering the community.M.SG.NOM.INF
dūsaṇu *lāgai*
sin.M.SG.NOM seem.3SG.PRS
'If I tell a lie, the sin seems to be like plundering the Jain community.' (RG.TS.13) 14thc.

From the 16th century onwards all functions were in fact taken over by the *-n/ṇ-* form. Similarly to the case of the *-v*/w- form, initially in purposive clauses the bare infinitive in the oblique case (169) is used, whereas later on it is reinforced by a postposition (170). Despite the dominance of the *-n/ṇ-* form, still in 18th-century texts we occasionally find the *-v/w-* form in obligative constructions, but already with A marked by a DAT postposition (171) (cf. Khokhlova 2013: 102).

(169) Early Rajasthani
tathā *pahalī* *sūrajamala* *sināna* *karaṇa*
and before Surajamal.NOM bathing.M.SG.NOM do.M.SG.INF.OBL
gayo *tho*
go.M.SG.PPP be.M.SG.PST
'But Surajmal went first to take a bath.' (RG.MN.134) 17th c.

(170) Early Rajasthani

rājā	*janaka*	*ghaṇī*	*bhuṃya*	*tāṃī*
king.M.SG.NOM	Janaka.M.SG.NOM	much.F.PL.NOM	land.F.SG.NOM	up to/until.LOC

puhacāwaṇa	*nūṃ*	*āyo*
cause to reach.M.INF.OBL	DAT	come.M.SG.PPP

'King Janaka came to accompany covering a very long distance.' (RG.DB.77) 18th c.

(171) Early Rajasthani

au	*dhanuṣ*	*mo-nūṃ*	*cāṛhṇo*	*sītā*	*parṇawī*
this	bow.M.SG	me-DAT	draw.INF.M.SG	Sita.F	marry.INF.F

'I am to draw this bow and marry Sita.' (RG.DB.48) 18th c.

The *-n-* form infinitives are also seldom used in the imperative function. When used as imperatives, infinitives in Early and Middle Early Rajasthani occur in unmarked form (172).

(172) Early Rajasthani

raiyata	*vastai*	*bhalo*	*cāhaṇo*
people.F.SG.OBL	for	good	want.INF

'Want good for people!' (RG.MS.113) 18th c.

The lack of formal difference between infinitive and gerund is evident. As infinitives of purpose may receive a postpositional marker (170), gerunds also do so (cf. (173) with the bare form and (174) with the genitive postposition).

(173) Early Rajasthani

iṇa-nūṃ	*māraṇa*	*matai*	*chai*
s/he.OBL-DAT	kill.INF.OBL	intention.M.SG	be.PRS.3SG

'[She] thinks about killing him.' (RG.MN.138) 17th c.

(174) Early Rajasthani

mhārai	*rowaṇai-ro*	*kāraṇa*
my	cry.INF.OBL-GEN	reason

'The reason for my sadness (crying).' (R.G.MS.1.76) 18th c.

In Early Awadhi there is a variety of forms, namely the forms having *-n-* and *-b-* features as well as forms which are formally very close to converbs terminating in *-ai* (the forms are summarized in Table 15). In Early Rajasthani, thanks to our rather diverse corpus, we can trace the process of the disappearance of the *-b-* infinitive, while in Early Awadhi, despite the fact that the corpus is much less diverse, one can still see that the formation of infinitives is again quite complex. In the earlier corpus (Padmāvat 1540) we find only a few instances of *-b-* forms, whereas they are still widely attested in the 18th-century text

'Indrāvatī' by Nūr Muhammad, composed in 1757. In fact Saksena (1937: 283) does not mention *-b-* forms for Padmāvat, but they are attested there (although not in our portion of the text) and they are used as complements or in the obligative pattern (cf. Śukla 2022: 152).

	-n-	-b-	-ai
Padamāvat 1540	12	-	13
Rāmcaritmānas 1574-1576	25	3	-

Table 15. Distribution of infinitives in Early Awadhi.

The two forms attested in Padamāvat are not in fully complementary distribution as regards their functions, but there are some interesting regularities. The form in *-ai* occurs in purposive clauses (175) and as a complement of modal verbs such as *cāh-* 'want' (176) and *pā-* 'get' (177). Moreover, it is the form used in inceptive constructions with *lag-* (178).

(175) Early Awadhi
pāni *bharai* *āvahiṃ* *panihārī*
water pour.INF.OBL come.3PL.PRS water carrier.F.PL
'The water carriers come to pour the water.' (J.32.1) AD 1540

(176) Early Awadhi
cahauṃ *bikāi* *bhūli* *gā* *paḍhā*
want.1PRS.SBJ sell.INF.OBL be lost.CVB go.PST read.PST
'What was learned is lost and now I want to be sold.' (J.77.3) AD 1540

(177) Early Awadhi
pāya *chuai* *maku* *pāvauṃ*
leg touch.INF.OBL perhaps get.1PRS.SBJ
'...perhaps I could touch [her] legs.' (J.61d) AD 1540

(178) Early Awadhi
lāgīṃ *keli* *karai*
start.PPP.F.PL amorous play do.INF.OBL
'(Friends) started playing.' (J.63.1) AD 1540

An interesting example of functional differentiation of the two infinitives is attested in (179), which actually consists of the example (149) and its subsequent text. The infinitive *parai* 'fall' serves the same purpose as the converbal clause, and the other part with the infinitive in *-n-* occurs in the obligative construction. In fact the obligative function is reserved only for forms in *-n-* and occasionally the *-b-* form. The complement function is partly shared by the two infinitival forms (cf. (176) with (180)), but with the reservation that *-n-* is never used with any modals (e.g. 'want'). Unlike in Early Rajasthani, in the Early Awadhi corpus there are no reinforced forms in purposive clauses.

The *-n-* form is formally and functionally closer to the nominal forms and can therefore also be used as a gerund in a subject (181) or modifier position (182). There is also an isolated example of a matrix coding as non-PSA construction (183).

(179) Early Awadhi
jauṃ *tivāīṁ* *kai* *kaja* *na* *jānā* *parai* *dhokha* *pācheṃ*
if woman.INS.F.SG do.CVB work not know.PPP.M.SG fall.INF deceit afterwards
pachitānā
regret.INF
'If the woman didn't realize what she has done [she will] fall into a mistake and regret afterwards [on falling into a mistake she will have to regret afterwards].' (J.86.4) AD 1540

(180) Early Awadhi
puna *bisarā* *bhā* *saṁvaranā*
but forget.PST.M.SG be.PST.M.SG remember.INF
'But (they) forget to remember.' (J.66d) AD 1540

(181) Early Awadhi
kita *āvana* *puni* *apane* *hāthāṁ*
where come.INF again self.M.PL.OBL hand.M.PL.OBL
'Then it will not be in our hands to come here again.' (J.60.6) AD 1540

(182) Early Awadhi
oiṁ *uṛāna* *phara* *tahiai* *khāe*
that.3.OBL fly.INF fruit.M.PL.NOM that.3.INS eat.M.PL.PPP
'He tasted the fruit of flying.' (J.68.4) AD 1540

(183) Early Awadhi
pāvā *sakhinha* *caṃda* *bihaṁsānā*
get.M.SG.PPP friend.F.PL.OBL moon.M.SG.NOM smile.INF
'Friends got the moon [Padmavati] smiling.' (J.65.5) AD 1540

Both in 'Padmāvat' and in 'Rāmcaritmānas' there are instances of the future tense based on the *-b-* feature. Yet, in both authors in some contexts it is not entirely clear whether it is an instance of an obligative construction or a future tense. Looking at examples (184) and (185), we tend towards the futurative interpretation of the former and the obligative interpretation of the latter. Additional support for the obligative reading may come from the A and S marking – there is a tendency (although it is certainly not a strict rule) to prefer agentive forms in some *-b-* based constructions which may be close to an obligative reading. In (185) A is used in a form which can be interpreted as oblique, and in (186) we have S in the agentive form *maiṃ*, in contrast to *hauṃ* in (187), which is a PRS/SBJ form. Thus in Early

Awadhi we may witness the final phase of transition from obligative to future. Interestingly, Early Awadhi has reduced the use of *-b-* forms to 1st PL, and only eastern dialects allow 2nd SG and 1st and 2nd PL with the *-b-* feature (Saksena 1937: 264; Ahmad 1986: 146; for the most recent discussion on the syntactic patterning observed in obligative/future formation see Khokhlova 2013 and Montaut 2017).

(184) Early Awadhi

kita	*milikai*	*khelaba*	*eka*	*sāthāṁ*
where	meet.CVB	play.1PL.FUT	one.DET	together

'How having met we will play together.' (J.60.6) AD 1540

(185) Early Awadhi

tinha	*nija*	*ōra*	*na*	*lāuba*
this.PL.OBL	own.M.NOM	end.F.SG.NOM	not.NEG	bring.INF

bhōrā
hallucination.M.SG.NOM
'They should/will not depart from their ways.' (T.1.5.1) AD 1574-1576

(186) Early Awadhi

ghara	*kaiseṃ*	*paiṭhaba*	*maiṃ*	*chūṁchai*
home.M.SG.NOM	how	enter.1SG.FUT	I.1SG.NOM	empty-handed

'How will/should I enter home empty-handed and what will I answer if someone asks?' (J.75.7) AD 1540

(187) Early Awadhi

āū	*hauṃ*	*aba*	*banobāsa*	*kahaṁ*	*jāūṁ*
and	I.1SG.NOM	now	life in jungle	for.DAT	go.1SG.SBJV

'I am going to live in the jungle now.' (J.57.2) AD 1540

In 'Rāmcaritmānas', which is a part of our corpus, there is great preponderance of *-n-* forms which occur as complements, e.g. (188). It is also the form used in purposive clauses (189), where it can be reinforced, for example by the grammaticalized noun *hetu* 'in order to < purpose' (190). The *-n-* form is also attested with modals such as *cāh-* 'want' (191) and *sak-* 'can, be able' (192) and in inceptive constructions with *lag-* 'begin' (193).

(188) Early Awadhi

dēi	*asīsa*	*sikhāvanu*	*dēhīṃ*
give.CVB	blessing.F.SG.NOM	teach.INF	give.3PL.PRS

'They, having given blessing, give (her) teaching.' (T. 1.334.2) AD 1574-1576

(189) Early Awadhi
Rāmadhāma *sikha* *dēna* *paṭhāē*
Rama's place instruction/advice.F.SG give.INF send.M.SG.PST
'(The king) sent (him) to Rama's place to give teachings.' (T.2.9.1) AD 1574-1576

(190) Early Awadhi
calē *janaka-maṃdira* *mudita* *bidā*
move.M.PL.PPP Janaka's palace.M.SG.NOM happy.M.SG.NOM farewell.SG
karāvana *hētu*
do.INF in order to
'Joyfully he came to Janaka's palace to bid him farewell.' (T.1.334) AD 1574-1576

(191) Early Awadhi
avadhanāthu *cāhata* *calana*
lord of Avadh.M.SG.NOM want.M.SG.NOM.PRS.PTCP go.INF
'(...) Avadh's lord wishes to depart.' (T.1.332) AD 1574-1576

(192) Early Awadhi
ehi *bidhi* *rāma-biāha* *uchāhū* *sakai*
this.OBL way.OBL Rama's wedding.M.SG.NOM joyous.M.NOM be able to.3PL.PRS
na *barani* *sahasa* *mukha* *jāhū*
not.NEG describe.INF thousand mouth.M.SG.NOM of which.REL
'Such were the rejoicings at Rama's wedding that not even Shesha could describe with his thousand tongues.' (T.1.331.4) AD 1574-1576

(193) Early Awadhi
lagē *sumaṃgala* *sajana* *saba* *bidhi* *anukūla* *bicāri*
be attached happy.M.NOM prepare.INF all god.M.SG.NOM favourable think.CVB
'And thinking that the God is favourable, [they] began to make happy preparations.' (T.2.8) AD 1574-1576

The distribution of the infinitival forms in Early Braj is quite interesting because, contrary to our earlier assumptions that *-v/b-* forms are recessive, we do not see them in early texts from the 15th and 16th centuries. Excerpts from texts from the 17th century show either exclusive employment of *-b-* forms or the use of both forms (see Table 16).[16] Other 17th-century texts also contain both forms (cf. Snell 1991b: 16).

[16] In the text of Indrajit *-n-* forms are also attested, but they are not present in our portion of the text (see McGregor 1968: 215).

	-n-	-v/b-
Viṣnudās 15th c.	15	-
Hitaharivaṃśa 16th c.	11	-
Indrajit of Orchā 1600	-	10
Śivarājabhūṣaṇa 1673	12	8

Table 16. Distribution of infinitives in Early Braj.

The *-n-* form is used in purposive clauses in all texts (194) excluding Indrajit. This function is also attested in the 17th century in 'Śivarājabhūṣaṇa', where the infinitive can occur in its oblique form (195) or be additionally reinforced by the dative postposition *kauṃ* (196).

(194) Early Braj

baṃdara	*paṭhae*	*caṃdana*	*laina*
monkey.M.PL.NOM	send.3PL.M.PPP	sandal tree.M.SG.NOM	to take.INF

'[He] sent the monkeys to bring the sandal tree.' (V.14) 15th c.

(195) Early Braj

sāhitanai	*sarajā*	*ke*			
son of Shahji.M.SG.OBL	lion.M.SG.OBL	GEN.M.SG.OBL			
bhaya	*sauṃ*	*bhagāne*	*bhūpa*	*meru*	
fear.M.SG.OBL	to	escape.INF.OBL	king.M.PL.NOM	Meru.M.SG.OBL	
ke	*lukāne*	*te*	*lahata*	*jāiṁ*	
GEN.M.SG.OBL	hide.INF.OBL	DEM.M.PL.NOM	obtain.3PL.PRS	go.CVB	
ota	*haiṃ*				
relief.F.SG.NOM	be.3PL.PRS.AUX				

'From fear of Shahji's son the Lion, kings having gone to run away to hide in Meru obtain relief.' (Ś.83) AD 1673

(196) Early Braj

sāhitanai	*sarajā*	*ke pāsa*	*āibe*	*kauṃ*
son of Shahji.M.SG.OBL	lion.M.SG.OBL	near	come.INF.OBL	to
baḍhīṃ	*ura*	*hauṃsana*		*kau*
increase.F.PL.PFV.PTCP	breast.M.PL.NOM	strong desire.F.PL.OBL		GEN.M.SG
aila	*haiṃ*			
flood.M.SG.NOM	be.3SG.PRS.AUX			

'[There] is a flood of strong desire [that] increased [in the] hearts to come to Shahji's son, the Lion.' (Ś.62) AD 1673

Both forms are also used as complements (197-198). In the latter example *-n-* is presumably a part of a compound verb. Modals such as *cāh-* 'want' (199) or inceptives with *lag-* 'begin' (200) employ exclusively *-n-* forms at all stages of Early Braj (the latter also in Indrajit's text – see McGregor 1968: 215).

(197) Early Braj

siya	*saṃmhari*	*dukha*	*jāina*	*sahyau*
Sita.F.SG.NOM	destroy.CVB	sorrow.M.SG.NOM	go.INF.OBL	endure.1SG.M.PST

'I have withstood Sita's going having destroyed the sorrow.' (V.2.61) 15th c.

(198) Early Braj

kiṃtu	*ve*	*to*	*saṃtoṣasaṃbaṃdhī*	*suṣu*	*kari*
but	DEM.PL.NOM	then	related to satisfaction.M.PL.NOM	joy.M.SG.OBL	from.INS
pūrana	*rahivoī*	*karatu*	*hai*		
full.M.PL.NOM	remain.INF.NOM	do.M.PL.IPFV.PTCP	be.3PL.PRS.AUX		

'But they usually remain full of happiness and content.' (I.114) 1600

(199) Early Braj

cāhatu	*tina*	*ke*	*sahana*
want.3SG.M.PRS	that.SG.OBL	GEN.M.PL.OBL	endure.INF.NOM
prahāra			
wound.M.PL.NOM			

'[He] wants to withstand its wounds (or: wounds caused by it).' (V.2.93) 15th c.

(200) Early Braj

kaṃta	*kaṃta*	*kari*	*rovana*	*lāgīṃ*
lord.M.SG.NOM	lord.M.SG.NOM	do.CVB	cry.INF.OBL	start.3PL.F.PST

'[They] started to cry calling: [oh] lord, [oh] lord!' (V.1.56) 15th c.

Both forms are also attested in the obligative construction (201-202). They show object–verb agreement, but in the corpus we have not encountered any marked A's.

(201) Early Braj

dinabharu	*chāyā*	*megha*	*samaggu*
whole day	shadow.F.SG.NOM	cloud.M.SG.NOM	entire.M.SG.NOM
lachimana	*taiso*	*dīsana*	*laggu*
Lakshmana.M.SG.NOM	such.M.SG.NOM	be seen.INF.OBL	near

'The whole day Lakshman is almost to be seen in such a way [as under] all clouds' shadow.' (V.2.11) 15thc.

(202) Early Braj

teī	*niṃdya*	*haiṃ*	*aru*	*teī*
this.NOM.PL	wrong.F.PL.NOM	be.3PL.PRS.AUX	and	this.DEM.PL.NOM
kutsita	*parīkṣaka*	*jānive*		
spoiled.F.PL.NOM	tester.M.SG.NOM	know.INF.OBL		

'Just as, [when] the value of good jewels is unknowingly lessened, the tester should know the value.' (I.125) 1600

There are isolated attestations of the matrix coding as non-PSA construction with the infinitive terminating in *-n-* (203).

(203) Early Braj

dāvedāra	*kau*	*risānau*	*dekhi*
making claim.M.SG.OBL	ACC	anger.INF.NOM	see.CVB

'Having seen the one making claims being angry...'
('= Having seen that the one making claims is angry') (Ś.33) AD 1673

As was the case in Early Rajasthani and Early Awadhi, the formal proximity of the infinitive and the action nominal often makes it impossible to distinguish the two categories. As the examples below show, both forms in *-n-* and those in *-b/v-* retain some nominal features, occurring with the GEN (204) or DAT (205-206) postpositions, or without them (207), occupying a modifier or a subject position.

(204) Early Braj

tau	*kīje*	*rākhana*	*ko*	*dāu*
then	do.2.IMP	protect.INF.OBL	of.GEN.M.SG	turn.M.SG.NOM

'Since you do not help now, then take your turn and protect!' (V.2.87) 15th c.

(205) Early Braj

uttama	*kathā*	*sunive*	*kahuṃ*	*visanu*
best.F.SG.NOM	story.F.SG.NOM	listen.INF.OBL	to	diligence.M.SG.NOM

'The diligence to listen to the best story.' (I.133) 1600

(206) Early Braj

nābhi	*gaṁbhīra*	*mīna*	*mohana*
navel.F.SG.NOM	profound.F.SG.NOM	fish.F.SG.NOM	Mohan (Krishna).M.SG.NOM
mana	*ṣelana*	*kauṃ*	*hṛdanī*
mind.M.SG.NOM	play.INF.OBL	DAT	river.F.SG.NOM

'Deep navel [is] a river for playing (to play) [of] the fish [of] Mohan['s] mind.' (HH.29.13) 16th c.

(207) Early Braj

nāṃtaru	*vrata*	*kīvau*	*vaḍoī*	*gunu*	*hai*
otherwise	task.M.SG.NOM	do.INF.NOM	big.M.SG.NOM	quality.M.SG.NOM	be.3SG.PRS

'Otherwise, accomplishing (lit. to do) the task is indeed a big advantage.' (I.161) AD 1600

It is very important to note that, since Early Dakkhini is a variety of language based on the dialect spoken around Delhi, i.e. Khariboli, which is geographically very close to Early Braj and Early Rajasthani, one might expect more variation in infinitival forms, at least at the initial stage. However, Early Dakkhini texts show exclusively one form, that in *-n-*. This is consistent with the very small amount of variation of the infinitival forms in the earliest Braj texts (15th–16th c.) but not later Braj texts, let alone Rajasthani texts.

From the earliest texts, the form was used in the obligative construction. The obligative construction is predominantly agentless. S arguments, if they occur, can remain unmarked (208). Marking of the A argument is attested in an oblique form (209), but in later texts we also witness the DAT (210) and ERG (211) postpositions. The ERG postposition is only attested from the 17th century, which at present is quite difficult to explain.

(208) Early Dakkhini

jo	*terī*	*ibādata*	*bagaira*	*uṭhenā*
who.NOM	your.F.SG	divine worship.F.SG.OBL	without	to rise.INF.NOM

'Those who should rise without your worship.' (BN.99) 14th c.

(209) Early Dakkhini

so	*uneṃ*	*karanā*
that.DEM.3.NOM	those.PL.DAT	to do.INF.NOM

'They should do/act.' (BN.36) 14thc.

(210) Early Dakkhini

naīṁ	*usa*	*ko*	*ānā*	*jānā*
no.NEG	he.M.SG.OBL	to.DAT	come.INF.NOM	go.INF.NOM

'He does not [have] to come [and] go [away].' (MH.38) AD 1623

(211) Early Dakkhini (from Šamatov 1974: 169)

āśik-ne	*kyā*	*karnā*
lover-ERG	what	do.INF

'What should the lover do?' (MV) AD 1636

The infinitival form in Early Dakkhini was regularly used as an imperative form. This modality is quite close to the obligative, and in such form it has survived in Modern NIA. In the examples we have found it is often ambiguous whether the infinitival form is an instance of an obligative or imperative (212).

(212) Early Dakkhini

murīda	*yaha*	*irśāda*	*yāda*
disciple.M.SG.NOM	he.M.SG.NOM	instruction.M.SG.NOM	memory.F.SG.NOM

rakhanā
keep.INF.NOM
'Disciple! Keep this instruction in [your] mind...' (BN.114) 14th c.

There are two more interesting examples of deontic modality in Early Dakkhini, one of which can also be found in Early Braj. It is based on the copula and the oblique infinitive form with the DAT postposition (cf. (213) from Dakkhini and (214) from Early Braj). There is also a construction with a similar meaning based on the oblique infinitive and the GEN postposition (215).

(213) Early Dakkhini

disane	*kūṁ*	*kanta*	*naiṁ*	*hai*
appear.INF.OBL	to.DAT	beloved one.M.SG.NOM	no.NEG	be.3SG.PRS

'The beloved one (i.e. God) does not appear (lit. is not to appear).' (SMH.45) AD 1623

(214) Early Braj

kahata	*dharesa*	*dharā*	*dharibe*	*kauṃ*	*sesa*
say.3PL.M.PRS	king.M.PL.NOM	earth.F.SG.NOM	hold.INF.OBL	to.DAT	Sesha.M.SG.NOM

'Kings say [that he is] Shesha holding the earth ...' (Ś.169) AD 1673

(215) Old Dakkhini

agara	*dekhane*	*kā*	*nā*	*hotā*	*hora*	*nā*
if	see.INF.OBL	of.GEN.M.SG	not.NEG	be.SG.PRS.PTCP	and	not.NEG

dekhā	*jātā*
see.M.SG.PPP	go.3SG.M.PRS.PASS.COND

'If [he (i.e. the God)] were not to be seen and [if he] were not seen...' (MV.79) 1636

From the earliest sources the infinitival form was used in the purposive construction. In earlier texts we have a bare infinitive in the oblique form (216), and in later texts the infinitive may receive reinforcement in a form of an adposition; this may be a dative postposition (217), which is quite regular in IA, or other types of postpositions, possibly borrowed from Arabic (218).

(216) Early Dakkhini

pīra	*ājamata*	*bulā*	*lāva*	*isa*
saint.M.PL.NOM	respectable.M.PL.NOM	call.CVB	bring.3SG.PRS.SBJV	this.M.SG.OBL

vakata	*karāmāta*	*dekhaṇeṃ*
time.M.SG.OBL	miracles.F.PL.NOM	see.INF.OBL

'[The emperor] calls the respectable saints [only] in order to see the miracles.' (G.34) 14th c.

(217) Early Dakkhini

use	*duda*	*pīne*	*ko*	*dāī*	*ḍhuṁḍhāya*
he.DAT	milk.M.SG.NOM	drink.INF.OBL	to.DAT	wet-nurse.F.SG.NOM	to look for.CVB

‘(...) looking for a wet-nurse for him to drink the milk.’ (F.43) AD 1685

(218) Early Dakkhini

iśqa	*baṛhāne*	*khātira*	[...]	*mūna*	*khola*
love.M.SG.NOM	increase.INF.OBL	for the sake		mouth.M.SG.NOM	open.CVB
dikhlāe					
show.3PRS.SBJV					

‘[She] would appear having uncovered her mouth for the sake of increasing the love/passion.’ (MV.4.13) AD 1636

The infinitival form is used as a complement (219). It is also attested with various modals – for instance, as was the case in other dialectal groups, with the inceptive *lag-* ‘begin’ (220), and in permissive constructions with *de-* ‘give’ with and without the DAT postpositon (221). There are also two types of constructions expressing abilitative meanings with *ā-* ‘come’ (222) and *pā-* ‘be able’ (223), both having continuants in modern IA.

(219) Early Dakkhini

nita	*basanā*	*yāda*	*e*	*allā*	*hara*
always	dwell.INF.NOM	memory.F.SG.NOM	of	Allah.M.SG.NOM	each.M.SG.NOM
dama	*pāūṁ*				
breath.M.SG.NOM	obtain.1SG.PRS.SBJV				

‘Let me obtain a perpetual possession (lit. dwelling) [of] the memory of Allah [in] every breath.’ (BN.49) 14th c.

(220) Early Dakkhini

chū	*na*	*na*	*na*	*na*	*kahane*	*lagā*
touch.2.IMP	not.NEG	not.NEG	not.NEG	not.NEG	say.INF.OBL	to start.3SG.M.PPP

‘[He] started to say: do not touch [me]! no! no! no!’ (E.1.16) 16th c.

(221) Early Dakkhini

kahe	*hukuma*	*āne*	*deva*	*phera*
say.3.PRS.SBJV	order.M.SG.NOM	come.INF.OBL	give.2.IMP	turn away.CVB
deva	*jāne*	*kū*		
give.2.IMP	go.INF.OBL	to.DAT		

‘[He/they] say the order: let [us/me] come in, turn away to let [us/me] go.’ (G. 27) 14th c.

(222) Early Dakkhini

jo	*koī*	*vo*	*nūra*
who.REL.NOM	someone.SG.NOM	that.DEM.M.SG.NOM	light.M.SG.NOM

pāyā	*phira*	*bolne*	*na*	*āyā*
obtain.3SG.M.PPP	then	speak.INF.OBL	not.NEG	come.3SG.M.PPP

'Whoever attained this light, then he was not able to speak [up].' (SMH.13) AD 1623

(223) Early Dakkhini

so kabī	*āvane*	*na*	*pāve*
so sometime	come.INF.OBL	not.NEG	get.3SG.PRS.SBJV

'So [that], [it/she, i.e. the female serpent] would never be able to come.' (E.1.36) 16th c.

As was the case in other dialectal groups, the infinitival form in Dakkhini retains many nominal features, also occurring in a subject (224) or modifier position (225). A very early attestation of the double possessive construction – a native one with the GEN postposition and a borrowed one based on *ezafe* – gives additional support for the nominal character of the infinitival form in Early Dakkhini (226). A single attestation of the matrix coding as non-PSA construction based on the infinitival form has also been found in the corpus (227).

(224) Early Dakkhini

usa	*kū*	*pachānanā*	*vājiba*	*hora*
that.SG.OBL	ACC	recognize.INF.NOM	necessary.M.SG.NOM	and

farza	*huā*
religious obligation.M.SG.NOM	be.3SG.M.PPP

'Recognizing it became necessary and [it became] a religious obligation.' (BN.120) 14th c.

(225) Early Dakkhini

rahane	*ke*	*tanta*	*naīṁ*	*hai*
to stay.INF.OBL	of.GEN.M.PL	rule.M.PL.NOM	no.NEG	be.3PL.PRS

'[Waves] have no [...] no rule (i.e. no way) to stay [still].' (SMH.44) AD 1623

(226) Early Dakkhini

yāne	*marane*	*ke*	*aṃga*	*e*	*maranā*
that is	die.INF.OBL	of.GEN.M.PL	limb.M.PL.NOM	of	die.INF.NOM

'It means: the death (dying) of the limbs [which are supposed] to die.' (BN.70) 14th c.

(227) Early Dakkhini

khudā	*kā*	*hone*	*maṃgtā*	*hai*
lord.M.SG	GEN.M.SG	be.INF.OBL	request.2IND.IMPF.M	be.2SG.PRS

'[Since] [you] request God's presence, ...' (MV.61) AD 1636

The data from Early Pahari show that this dialectal group had exclusively one form, in *-n-*, occurring in at least three major functions, namely as an action nominal in the subject

position or as a modifier, and in the obligative construction. In Early Kumaoni the former two can be traced only from the first literary texts (228-229), whereas the latter is amply attested from the earliest inscriptional sources, and therefore major developments in main argument marking as well as possible transformation of the obligative pattern into the future tense can be traced at least from about the beginning of the 14th century.

Early Pahari exhibited OEM and DEM. In Early Eastern Pahari we have A-arguments unmarked (230a), marked by nasalization (230b), marked by the ERG postposition *le* (230c), and marked by a postposition homophonous with the contemporary DAT postposition in Nepali, *lai* (224d). The reading of these constructions is purely obligative. Also, contemporary data provide no evidence for the 'obligative-to-future' transformation, although contemporary grammarians tend to label the constructions based on infinitives 'obligative future' (Hin. 'karaṇīy bhavishyat'; Juyāl 1973: 146-147), which carries implications of futurity.

(228) Early Kumaoni

śilpina	*ko*	*kāma*	*ālakasa*	*varjjaṇo*	*paṃḍitāī*	*mitrana*	*ko*	
craftsman.OBL.M.PL	GEN.M	work	laziness	abandon	wisdom	friend.OBL.PL	GEN.M	
saṃgraha	*karaṇī*	*jai*	*kana*	*cora*	*lagai*	*nai*	*haranā*	*eśī*
collection	do.INF.OBL	that	ACC	thief	up to	not	remove.INF.M.PL	such.F
vidyā	*in*	*pāṃca*	*akṣayanidhi*		*chana*			
knowledge.F	this.OBL.PL	five	undecaying treasure.F		be.3.PL.PRS			

'Such knowledges are those five undecaying treasures: craftsman's work, abandoning laziness, wisdom, gathering friends.' (RŚ.3.4) AD 1728-1729

(229) Early Eastern Pahari (Kumaoni)

jaśo	*kasauṭi*	*meṃ*	*ghasi-bera*	*kāṭanā*	*tapauṇā*	*le*
like	touchstone.F.SG	in	rub-CVB	cut.INF.OBL	heat.INF.OBL	ERG

tāḍana-le	*cautira-le*	*sunā*	*kī*	*parīkṣā*	
beat.INF.OBL-ERG	fourfold way-ERG	gold	GEN.F	examination.F.SG	
karī	*cha*				
do	be.3PL.PRS				

'Just as [one] examines the gold by fourfold way having rubbed [it] against the touchstone, by cutting, heating, beating.' (RŚ.4.4) AD 1728-1729

(230) Early Eastern Pahari

a) Early Kumaoni

vāsūdeī-ki	*noṭha naṭhyālī*	*gaṃḍilī pe(ṭi)l*	*ī*
V.-GEN	property owned by a person without progeny.F	military tax.F.	O

rajā	*na*	*pāuṇī*
king.NOM	not	get.INF.F

'The king should not resume the property and the military tax belonging to Vāsūdeī.' AD 1337 (Joshi 2009: 338)

b) Early Kumaoni (Joshi 2009: 338)
jo *āi* *upa(tt)a* *mara* *tas-(k)i*
which come.CVB sonless die.PPP.M.SG he.OBL.GEN.F
naṭhyālī *rajā-ṁ* *pāuṇī*
property owned by a person without progeny.F king.OBL get.INF.F
'The king should receive the property of the one who died without a son.' AD 1337

c) Early Kumaoni (Joshi 2009: 344-345)
jiulo *virsighaṁ* *ki* *saṁtati-le* *bhuchaṇu*
jiulo V. GEN progeny.F-ERG enjoy.INF.MASC
'The progeny of Virsinha should enjoy a jiulo (of land).' AD 1380

d) Early Kumaoni (Joshi 2009: 340)
te *rita* *bhada bhāṭa-lai* *nirvahaṇu*
this.NOM custom B.B.-DAT carry on.INF.MASC
'Bhada Bhāṭa should carry on that custom.' AD 1395

Similarly to Early Eastern Pahari, Western Pahari also has traces of OEM and DEM, as the examples below show. In (225) we have A-argument unmarked (a), marked by an OBL marker (b), and finally marked by a GEN postposition (c). The GEN postposition in (225c) does not seem to be a Pahari feature, but strongly points towards Punjabi influence.

(231) Early Western Pahari
a) Chambyali (Chhabra 1957: 35)
eha *śrīrā-e-ke* *putra potra* *pālaṇa*
this.DEM king-OBL-GEN.PL sons protect.INF.M
'The descendants of the king should protect this (gift).' AD 1446

b) Chambyali (Chhabra 1957: 31)
dāpa-paṭ-e *4 śāghas-trīi* *rā-e* *leṇe*
elephant rug-piece-PL 4 rhino's horn-3 king-OBL take.INF.M.PL
'The king should receive 4 pieces of elephant rugs and 3 rhino's horns.' Mid -5th c.

c) Chambyali (Chhabra 1957: 160)
eha *dharma* *ihnā-kī* *mahārāj-e-de* *vaś-e-de* *pālāṇā*
this.DEM pious gift they.OBL.-DAT king-OBL-GEN descendent.PL protect.INF
'The descendants of the king should protect this pious gift for themselves.' AD 1664

7. MORPHOSYNTAX OF INFINITIVES – SUMMARY

Contemporary NIA languages have preserved most of the features of infinitives which are present in Early NIA. Infinitives are used as complements, in purposive constructions and in the obligative construction.

As regards the morphological make-up of converbs in contemporary NIA, Rajasthani shows dialectal variation: namely Marwari has only one form in *-ṇ-*, whereas eastern Rajasthani dialects may have two forms, in *-ṇ-* or *-b-* (e.g. Harauti), or only one form in *-b-* (other eastern Rajasthani dialects).

Languages such as Awadhi still use three forms, namely in *-ai*, *-b-* and *-n-*. The first two forms are used in all functions attested in Early Awadhi, whereas the third is constrained to certain modals (Liperovskij 1997: 140-151).

Braj has *-b-* and *-n-* forms, although some scholars label the former as gerunds and the latter as infinitives (see for example Liperovskij 1987: 113-119). Only the *-n-* forms are used in obligative constructions. The infinitive in *-n-* has also undergone phonological reduction, and with modals – inceptives, permissives and in the abilitative construction – a form without any inflectional marker, i.e. terminating in *-n*, can be used (Liperovskij 1987: 118-119).

In the obligative pattern, Rajasthani, Braj and even Awadhi (which has lost the ergative pattern) follow a non-nominative pattern, marking A and S with the DAT postposition (see Liperovskij 1987: 118; 1997: 142).

Dakkhini has one form, in *-n-*, which is functionally equal to an infinitive.[17] In purposive clauses it is used in the OBL form *-ne*. As was the case with the ergative construction, Dakkhini completely lost A marking in the obligative construction (Mustafa 2000: 147).

The types of matrix coding constructions found in Early NIA are not available in contemporary NIA. The matrix coding as PSA construction attested in Early Rajasthani (162) is surely exceptional. The matrix coding as non-PSA construction occurs sporadically in Early Awadi (177), Braj (197) and Dakkhini (221).

As we have seen in section 2.3.2, the basic obligative pattern for contemporary Eastern Pahari (Nepali Kumaoni) may be one with the A or S arguments marked by the ERG case being replaced by the DAT case marker. Some Western Pahari languages, such as Kului, have

[17] Mustafa (2000: 185, ref. 26) rejects the existence of an infinitive in Dakkhini, but does not give any evidence as to why the *-n-* forms should not be interpreted as infinitives. His examples (both obligative and purposive constructions) in fact support the interpretation of *-n-* forms as infinitives.

developed a future tense based on the *-n-* infinitive and A marked by a synthetic ERG case. Other dialects can also have GEN marking in obligative constructions, which may evolve into the future tense, as in Kothgari (Hendriksen 1986: 164), or retain their obligative value, as in Bangani (Zoller 2008: 297). It has been noted in the literature that Western Pahari developed a range of genitively marked A's (for an exhaustive discussion see Zoller 2008, and more recently Renkovskaja 2018).

Montaut (2018: 118) is certainly correct when she states that 'this evolution from modality to future is not pan-Indian', but there is visible extension of this evolution to the west, and Western Pahari appears to be the best proof of this. Clearly the infinitive in *-n-* was the basis here for the formation of a new future tense with main arguments (both S and A) marked by agentive suffixes (or in the case of personal pronouns occurring in the agentive form), exactly the same as those used in the ergative construction (cf. Ṭhākur 1975: 305).

8. APPENDIX – DESCRIPTION OF TOOLS

The following sections describe the techniques, computer algorithms and software used to facilitate and enhance the process of tagging texts in early New Indo-Aryan languages. The techniques described are used in two separate components: a tagging tool and an artificial intelligence module.

The first component, the IA Tagger tool, is a browser-based database system which facilitates manual tagging of texts. The tagging can be performed by multiple users at once, and all the annotations are stored in a central database. Such a solution, though standard in modern IT for business, is not always applied in linguistic research due to the relatively high costs of implementation of the software. In our case, however, the software was largely based on well-known publicly available frameworks, which made it possible to reduce the time and costs of its development.

The benefits of using a central database for linguistic information include:

- data security – protection of sensitive texts
- data backups – protection against data loss
- instant access to the annotations provided by co-researchers (consulting, reviewing)
- ease of exporting the data in standardized formats

Moreover, the IA Tagger tool itself provides functionalities which speed up the process of tagging, such as tag suggestions. A detailed description of the tool is presented in section 8.1.

The second component – the artificial intelligence module – was used after a considerable amount of data had been collected by the IA Tagger tool. The data were used to train multiple machine learning algorithms, with the aim of developing an automatic tagger. More specifically, the main requirement for the automatic tagger was to identify converbal forms in unannotated texts.

8.1. THE IA TAGGER TOOL

8.1.1 TOOL OVERVIEW

The process of collecting linguistic data was facilitated by means of the aforementioned computer software – the manual tagging system IA Tagger, based on the open-source Tagger framework, available on the GitHub platform https://github.com/rjawor/tagging.

More information about systems of the Tagger family can be obtained at http://rjawor.vm.wmi.amu.edu.pl/wiki.

Similar systems are commonly used in contemporary linguistic research; see for instance Haig and Schnell (2015). Features of IA Tagger include:

- import of texts in textual format (.txt or Microsoft Word) into documents;
- storing of information about the language of the document and the epoch in which it was written;
- automatic splitting of the text into sentences and individual words;
- manual assignment of tags to individual words.

IA Tagger is a tool for text annotation specialized for Indo-Aryan languages. It supports, for instance, the tagging of postpositions as separate tokens. The key functionality of the tool is multi-level annotation of words and sentences of early NIA texts. IA Tagger provides several features that improve the efficiency of use. For most annotation levels the system displays a context-sensitive list of prompts of available annotation tags. For a word under annotation the system displays a "prompt cloud", which consists of a set of tag suggestions.

IA Tagger minimizes the cost of usage errors or system failure. Each annotation decision is saved automatically in a periodically backed-up database. There is no save button. This solution ensures protection against the loss of valuable annotations.

The wide variety of configuration settings ensures the flexibility of the tagger, allowing it to be used in various scenarios.

On request, IA Tagger generates statistics concerning occurrences of specific classes of words and word collocations – in a specified document or collection of documents.

The system is intended for open access. It is accessible using any popular Internet browser at http://rjawor.vm.wmi.amu.edu.pl/tagging. Access credentials can be obtained on request from rafal.jaworski@amu.edu.pl.

8.1.2 MULTI-LEVEL TAGGING

To start the tagging process, the user uploads a text document into the system. Upon upload, the document is automatically split into sentences .

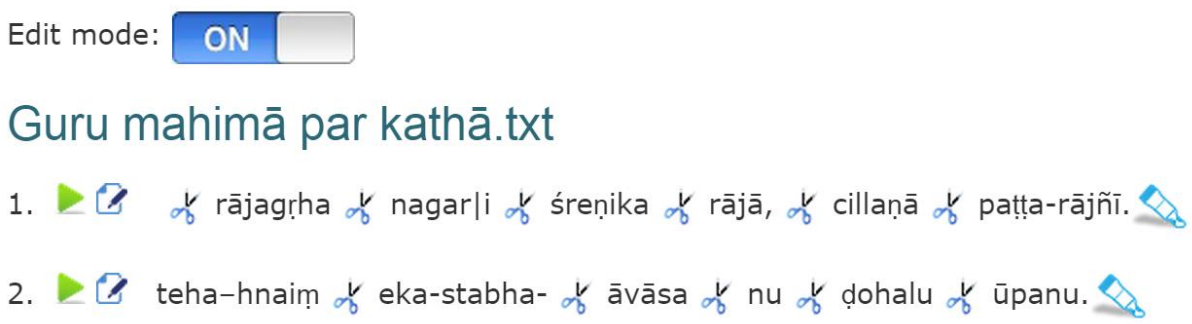

Figure 18. Sentence splitting.

The user can easily override the automatic sentence split (using "scissors" or "glue"; Fig. 18). The document is then annotated in sentence-by-sentence mode.

1.	rājagṛha	nagar|i	śreṇika	rājā,	cillaṇā	paṭṭa-rājñī.
lexeme	rājagṛha	town	śreṇika	king	cillaṇā	queen
grammar	M NOM SG	LOC	M NOM SG	M NOM SG	F NOM SG	F NOM SG
POS	NOUN	NOUN	NOUN	NOUN	NOUN	NOUN
syntax			S	S	S	S
semantics			TH	TH	TH	TH
pragmatic						
add info						
english	King Shrenika and queen Chillana (ruled) in the city of Rajgrah.					

Figure 19. An annotated sentence.

Each sentence is automatically split into words. A split sentence is shown in Fig. 19. The user may also override the word split, for example to divide a word into a stem and a suffix. Words are annotated at six levels: Lexeme (where the closest English lexical equivalent is given), Grammar (annotated using the Leipzig Glossing Rules), POS (Parts of Speech), Syntax (exploring the basic Dixonian (Dixon 1994) scheme based on the three primitive terms: A, S and O (see footnote 13), Semantics (where we distinguish six basic thematic roles: Agent, Patient, Experiencer, Recipient, Stimulus and Theme, based on the RRG approach, e.g. Van Valin and LaPolla 1997), and Pragmatics. Figure 19 shows also an annotated sentence from Early Rajasthani.

8.1.3 AUTOMATICALLY GENERATED SUGGESTIONS

To improve tagging efficiency, the system suggests hints whenever possible, i.e. when a word has already been tagged or when the tagging could be deduced automatically. Tag suggestions appear in a "cloud" above the word (Fig. 20).

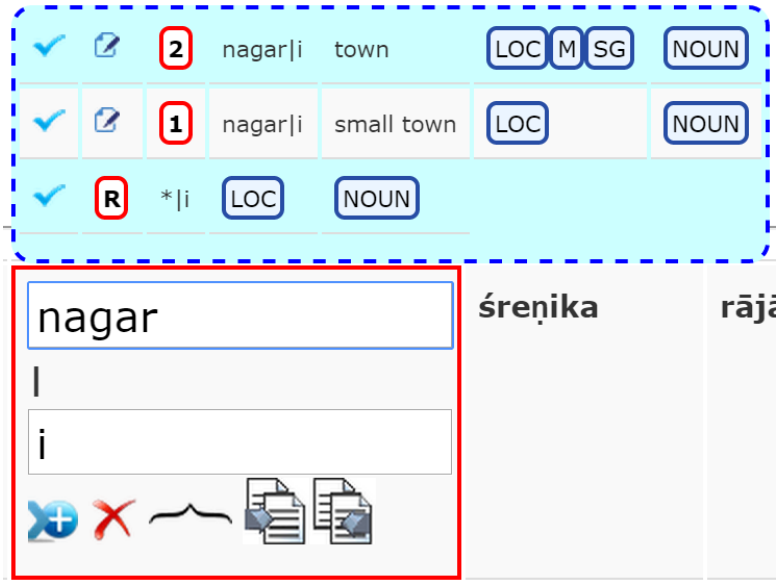

Figure 20. Automatically generated suggestions.

Figure 20 shows tag suggestions for the word 'nagar—i' (the pipe indicates that the word 'nagari' has been split into a stem and a suffix). The first two lines come from previous annotations, whereas the third line is the set of suggestions deduced automatically. The user can accept the set of suggestions by clicking the 'check' symbol in the leftmost column. The annotation shown in Fig. 20 was obtained by applying the set of tags from the third line.

8.1.4 CONFIGURATION

IA Tagger may be configured to serve a variety of annotation tasks. The "configuration" option allows one to manage the languages of tagged documents as well as to configure annotation levels. Annotation levels may be freely ordered, added, deleted or edited. Editing of an annotation level consists in defining admissible values of respective tags.

8.1.5 COLLECTED DATA

The Tagger system for New Indo-Aryan languages has been in active use since September 2014. The counts of annotated words in four languages from the database are presented below:

Language	Word count
Early Awadhi	12 281
Early Braj	10 016
Dakkhini	10 055
Early Rajasthani	10 157
Total	42 512

The collected data have been used to develop a mechanism that automatically assigns linguistic information to words.

8.1.6 STATISTICS GENERATOR

The IA Tagger tool has a Statistics Module, which is used to compute – on demand – statistics based on the data collected in the system. The module has two key functionalities: searching for words or phrases which meet specified criteria, and computing various statistics regarding those words or phrases. There are several functions that can be invoked from the Statistics Module.

Search for specific words

When searching for specific words, the user inputs the literal form of the word and the system presents results in the following form:

No.	Document	Language	Epoque	Word	Context
1	amarsen-vayarsen.txt	RJ		ghara	pachai vayarasena ni-dravya jāṇī **ghara** hūṃtu kāḍhiu.
2	vacanika rathod ratan ri_18th century.txt	RJ		gharā	au to kahai jaḷāboḷa riṇa-samaṃda māhe asi-jihāja dharāṃ , kilaṃbāṃ **gharā** māri pārī karāṃ ;
3	jaayasii padmavaat sinvala dviip varnan khand.txt	AW		ghara	JayP36.3 rāu rāṁka saba **ghara** ghara sukhī jo dekhia so haṁsatā mukhī |
4	jaayasii padmavaat sinvala dviip varnan khand.txt	AW		ghara	JayP36.3 rāu rāṁka saba ghara **ghara** sukhī jo dekhia so haṁsatā mukhī |

Figure 21. Search for specific words with contexts.

Each word found is presented in the context of the sentence that contains it. Information about the document, language and epoch (if available) is also shown alongside the word. By clicking on the "Edit" icon, the user is redirected to sentence edit mode, where the word in question is automatically highlighted.

The search results can be refined using the filter panel:

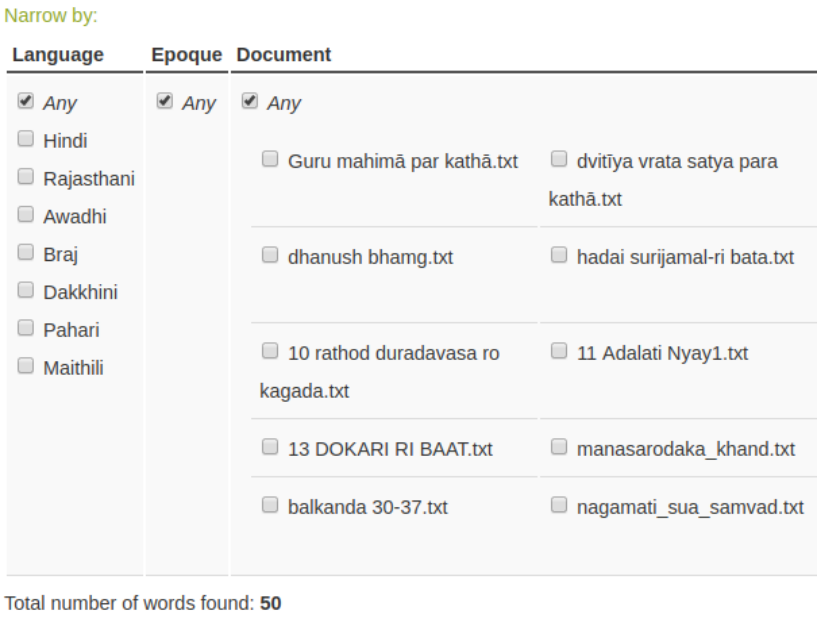

Figure 22. Application of filters for searching for specific words.

The panel allows narrowing of the search results by language, epoch or specific document.

Search for specific word forms

Another option in the Statistics Module is to search for specific word forms. The user can specify criteria for word search using the following search window:

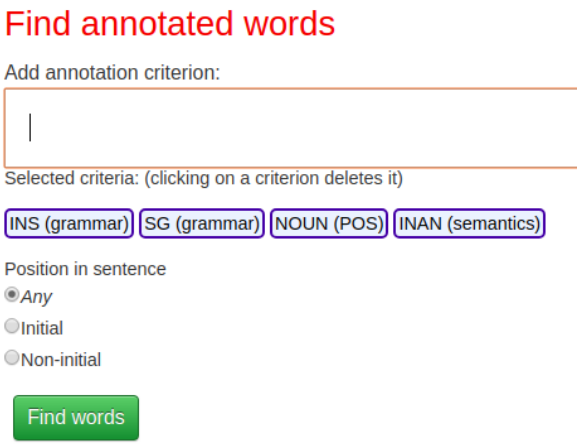

Figure 23. Search for specific word forms – window.

In this case the system will look for singular nouns in the instrumental case, tagged as inanimate on the semantic level. The presentation of results and filter are analogous to the case of searching for specific words.

Collocations search

The user can also search for multiple words which appear together in a single sentence. These are referred to in the system as collocations. An example of a collocation search is the following:

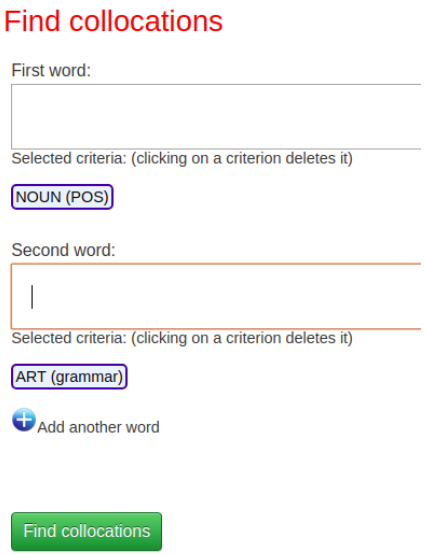

Figure 24. Search for collocations – window.

In this case the user searches for sentences which contain both a noun and an article. It is possible to search for collocations of a maximum of five words. The number of criteria for each word is unlimited (i.e. limited only by the number of tags available in the system). Results of collocations search are presented in the same way as in specific word search and word forms search.

Proportional statistics

The last option is the proportional statistics search. This kind of search is based on the word form search, although it takes two sets of criteria: main search criteria and additional specific criteria. The system first takes the main search criteria and counts the number of words that meet those criteria. It then appends the additional specific criteria to the main criteria list and counts how many words meet all these criteria. Lastly, it computes the ratio of the number of words meeting all the criteria to the number of words meeting only the main criteria.

For example, if the main search criteria specify noun, singular, and the additional search criteria specify instrumental case, the result is the following:

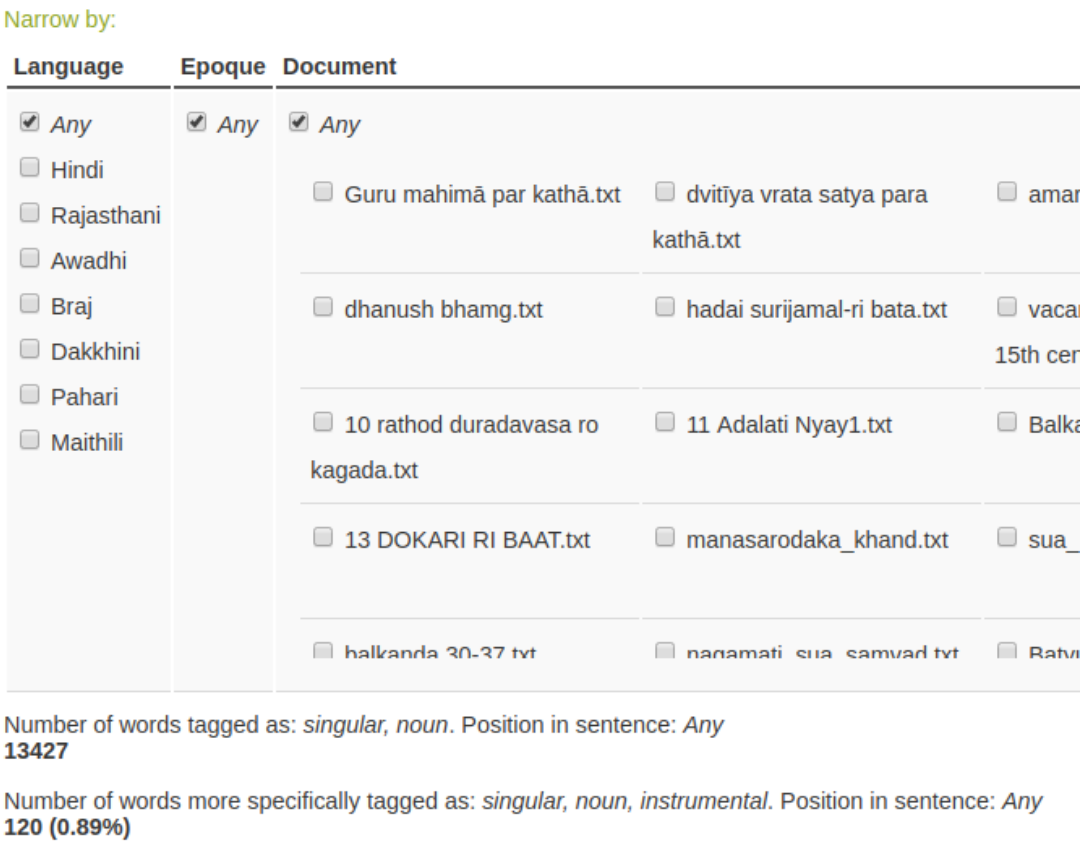

Figure 25. Proportional statistics search.

Note that the results can be refined with the use of filtering to a specified language, epoch and documents.

Predefined statistics

Frequently used statistics searches can be stored in the system for convenient use. The current list of such predefined statistics includes, among others:

- Converbs (as word form search)
- Collocations of converbs and transitive subjects (in various noun cases)
- Collocations of past participles with transitive subjects

8.2 AUTOMATIC POS-TAGGING

8.2.1 SIMILAR EXPERIMENTS

Experiments with automatic POS-tagging of less-resourced languages have already been conducted in recent years. This subsection briefly describes the techniques used and the outcome of two projects: an automatic tagger for Urdu, developed by Hardie (2005), and Sanskrittagger (Hellwig 2008).

The tagger for Urdu was developed by Andrew Hardie (Hardie 2005). The main difficulty in tagging Urdu texts identified by the author was word sense disambiguation. Two techniques were implemented in order to resolve this problem. One was based on hand-crafted rules prepared by a linguist, whereas the other relied on the statistical analysis of manually annotated Urdu texts. The author reports the low effectiveness of the latter method, attributing it to the relatively small quantity of training data. Hence the author decided to use the tagger based on hand-crafted rules. It must be pointed out, however, that the statistical model used was HMM (Hidden Markov Models; Baum and Petrie 1966), which was considered state-of-the-art in the early 2000s, but was replaced in the following years by several other methods, such as Conditional Random Fields and Maximum Entropy.

The resulting rule-based tagger used a tagset of approximately 80 tags and achieved an accuracy of 88-90%. The author admitted that these results were lower than those of taggers for well-resourced languages, such as English. Such taggers score at least 95% accuracy. This, however, should not be considered the main flaw of this system. A more important drawback of the approach presented by Hardie is the heavy reliance on manually designed rules, which account for most of the positive results of the system. These rules were specially designed to work with Urdu, and even more specifically – with the Urdu texts that were at the author's disposal. In a different scenario the same rules may prove to be inapplicable, thus impairing the performance of the system significantly.

Sanskrit tagger, described in Hellwig (2008), is an automatic tokenizer and tagger for Sanskrit. Like Hardie's Urdu tagger, it uses HMM to perform the tagging. Interestingly, the same model is also applied to the task of tokenization, which is a non-standard solution. The system uses a tagset of 136 tags. Unfortunately, accuracy figures are not known, as the evaluation of the system was performed on only five short passages of text. However, it is revealed that the system is purely statistical.

Among suggested methods of improvement, one seems particularly interesting – integrating tokenization and POS-tagging into one mechanism. The author argues that this might be a good approach for Sanskrit, even though it is not commonly used for other languages.

8.2.2 FULL POS-TAGGER

The work on annotation of NIA texts resulted in the creation of a manually tagged corpus. These data were used for preparing statistical models with the use of artificial intelligence algorithms during training. Annotated words became the training set. The aim of the preparation of the statistical models was to create an automatic POS-tagger, which was tested using a 10-fold cross-validation scheme. Experiments were run to determine whether it is possible to create a usable POS tagger for early NIA. Firstly, a full POS tagging system was developed. It uses 22 tags to annotate the text. The tags are hierarchical, e.g. there is a NOUN tag and its child – NOUN-SINGULAR. The task of annotation with 22 tags was seen as a multi-class classification problem. To implement such a tagger, the well-known Maximum Entropy (Jaynes 1957) tagging mechanism was used. This idea was first proposed by Ratnaparkhi (1996) and later used to implement the Stanford Part-Of-Speech Tagger (Toutanova and Manning 2000; Toutanova et al. 2003). The automatic tagger for early NIA is based on the Stanford software.

The main difficulty in training automatic taggers using the Maximum Entropy principle is the identification of the feature set. Possible features may include: suffix(*n*) of the word (i.e. last *n* letters), length of the word, whether the word starts with a capital letter (boolean feature) and many others. It is crucial, however, that all these features should be computable on unannotated text. Thus, features like "is located between a noun and a verb" are not acceptable.

The described automatic tagger for early NIA texts uses the following set of features: Suffix(6), Previous word (i.e. the literal text form of the previous word), Next word and Distributional similarity class.

Distributional similarity (often abbreviated distsim) is a method for categorizing words in a large corpus based on their contexts. Each word falls into a category with other words that appeared in similar contexts. The id of such a category can be used as a word feature. To compute distributional similarity classes, an unannotated modern Early Rajasthani corpus of 81 843 words was used. It was processed with the help of word2vec software (Mikolov 2013). The words were categorized into 209 classes, each containing between 1 and 66 words. For example, one of the classes contained the following words: *te* 'this', *teha* 's/he', which are pronouns.

Unfortunately, the overall results of the multi-class classification were not satisfactory (see Tables 17 and 18). The system achieved an overall accuracy of only 57.9% counting only

exact matches (guessing the precise tag), and 64.1% including partial matches (guessing a tag from the same hierarchy).

8.2.3 CONVERB DETECTOR

The second approach involved the training of a separate tagger, focused solely on identifying words of special interest – converbs. This is a case of binary classification. Two such binary converb detectors were implemented – one based on the Maximum Entropy algorithm, and another using the Vowpal Wabbit library (Langford 2009).

The implementation of the converb detector using the Maximum Entropy (ME) algorithm is based on the Python NLTK library (Loper and Bird 2002) using additional optimization techniques. This makes it possible to create a robust binary classifier. This converb detector was trained on the same data as the multi-class tagger described in previous section. The features used by this detector are presented in Table 17. Note that the features cvbEnding and firstOrLast use linguistic knowledge about converbs. Firstly, Early Rajasthani converbs typically terminate in /i/ and /a/, although from the earliest texts onwards other suffixes are also attested. Secondly, converbs would never appear as the first or last word in the sentence. This approach recalls the hand-crafted rules as seen in Hardie (2005). However, the features are never strict. The decision on whether or not to use a specific feature is made by the statistical model.

Feature name	Parameters	Description
word	None	Literal text of the word
wordContext	N	*n* words to the left and to the right of the word
Suffix	N	*n* last characters of the word
wordClass	None	Distributional similarity class
classContext	N	Classes of *n* words to the left and to the right of the word
cvbEnding	None	Whether or not the word has a typical converb ending
firstOrLast	None	Whether or not the word is first or last in the sentence

Table 17. Features used by the ME converb detector.

In a separate experiment, a second converb detector was built with the help of the Vowpal Wabbit (VW) software and was trained on a set of 5596 Early Awadhi words. All of these words came from one text, 'Padmāvat' by Malik Muhammad Jāyasī. Because of the homogeneity of the texts, we expected better evaluation results than in the previous experiments.

On the other hand, the classification algorithm used by the Vowpal Wabbit software is based on classic regression and features numerous improvements, described thoroughly by Langford (2009). Importantly, the software features a tool for assessing the importance of individual features in the process of prediction. The most informative features identified

with the help of this tool were used in the process of classification and are presented in Table 18.

Feature name	Description
Suffix(3)	Last three letters of the word
wordContext(1)	Literal forms of the previous and next word
classContext(1)	Distributional similarity classes of the previous and next word

Table 18. Features used by the VW converb detector.

8.2.4 CONVERB DETECTOR TESTS

A series of experiments was conducted to measure the performance of the converb detector. In the first scenario the whole set of 5596 words was used as both training and test set in a 10-fold cross-validation scheme. This cross-validation scheme is divided into 10 steps. In the first step, the first 10% of the text constitutes the test set and the rest becomes the training set. In the second step, the next 10% of the text is treated as the test set, while the rest is used for training, etc. This technique makes it possible to maximize the magnitude of training and test data and optimize the experimental setup. Two experiments were conducted in the cross-validation scenario. First, a baseline system was developed in order to assess the complexity of the task itself. The baseline system simply creates a dictionary of converbs from the training set (this operation is trivial as every word in the training set is annotated) and then predicts a word from the test set as converb if and only if it is found in the converb dictionary. This approach yields a precision of 46.7% and a recall of 57.0%. The value of the precision parameter in this case is interpreted in the following statement: out of all the words in the test set that were found in the converb dictionary, only 46.7% are indeed converbs. On the one hand this leads to the conclusion that the same word (in terms of spelling) can serve as a converb in one context but not in another. On the other hand, the recall value leads to the conclusion that dictionary-based search for converbs is only capable of finding 57% of all actual converbs, while the remaining 43%, from outside of the (too narrow) dictionary, are missed. Based on these findings, it can be stated that the task of converb detection in the researched texts is non-trivial.

The second experiment measured the performance of the statistical converb detector based on Vowpal Wabbit software. The recorded precision score was 80.2%, with a recall of 64.4%. These results can be viewed as a success.

Apart from the cross-validation experiments, the VW converb detector was additionally tested in another scenario. The experiment consisted in preparing separate annotated data, based on texts from outside of the IA Tagger system. For the needs of the experiment, an excerpt from the above-mentioned 'Padmāvat' not tagged in the IA Tagger, consisting of 11 501 words, was manually annotated with converb tags. The magnitude of the test data complies with the standards for human evaluation experiments in the field of natural language processing (see for instance Seljan et al. 2015). The converb detector in this

scenario was trained on the whole set of 5596 words from IA Tagger and tested on the excerpt. The results achieved were the following: precision 74.8% and recall 66.4%. These results further confirmed the success achieved in the cross-validation experiment.

8.2.5 EXPERIMENTAL RESULTS

This section presents the results of the experiment conducted using both of the automatic POS taggers. In both cases the tagged corpus (13 022 words) was used to perform 10-fold cross-validation. The magnitude of the test data complies with the standards for human evaluation experiments in the field of natural language processing (see for instance (20)).

Table 19 presents results for the multi-class tagger. It assigned tags to 10 730 out of 13 022 words (82.4%), leaving the remaining words untagged. Exact tag matching counts a tag as correct only if it matches exactly the tag in the golden standard. Partial tag matching allows, for example, the tagging of a NOUN-SINGULAR with the tag NOUN.

Metric	Number of correct tags	Accuracy
Exact	6210	57.9%
Partial	6874	64.1%

Table 19. Overall results of the multi-class tagger.

Some specific word forms were investigated more thoroughly. Table 20 presents precision, recall and F-measure scores (as proposed in Makhoul 1999) for identification of these forms. All results assume the partial tag matching metric.

Word form	Precision	Recall	F-measure
Verb	0.61	0.70	0.65
Noun	0.41	0.52	0.46
past participle	0.70	0.60	0.64
converb	0.33	0.07	0.11

Table 20. Detailed performance of the multi-class tagger.

The accuracy of the multi-class tagger, which was as low as 64%, was not a satisfactory result. However, the results in Table 20 reveal that even though the overall accuracy of the system is low, some word classes can be detected more accurately, such as verbs. However, converbs, the forms of our special interest, were detected poorly by the multi-class tagger. This inspired further study using the specialized converb detector.

The detector was expected to attain higher precision and recall scores in finding converbs than the multi-class tagger. The scores of the Maximum Entropy detector are presented in Table 21. These indeed show a considerable improvement over the multi-class tagger (see Table 20). This justifies the decision to implement a separate detector solely for word forms of particular interest.

Metric	Value
Precision	0.83
Recall	0.39
F-measure	0.53

Table 21. ME converb detector scores.

The best results, however, were achieved by the Vowpal Wabbit converb detector. A summary of these results is presented in Table 22.

Scenario	Precision	Recall	F-score
Baseline (converb dictionary)	46.7%	57%	51.3%
Cross-validation on training data	80.2%	64.4%	71.4%
Testing on golden standard	74.8%	66.4%	70.4%

Table 22. VW converb detector scores.

8.2.6 OPTIMIZING LINGUISTIC RESEARCH

Based on the quality of results, the VW converb detector was selected to perform the following procedure. The detector was again trained with the manually tagged Early Awadhi words and run on the remaining untagged parts of 'Padmāvat'. Next, all sentences containing automatically annotated converbs were extracted and input to the IA Tagger system for full manual annotation at all annotation levels. Thus, the annotators worked with sentences that had a high probability of containing converbs. Based on the results of the experiment with the excerpt, this probability can be assessed as 74.8%. Furthermore, the annotators will identify approximately two-thirds (66.4%) of all converbs in 'Padmāvat' (as follows from the recall score). In a normal scenario, this would require manual tagging of at least 66% of the whole of 'Padmāvat', which would require an excessive amount of work and time.

8.2.7 CONCLUSIONS

The described experiments have demonstrated the usefulness of enhancing linguistic analysis with the help of modern computer science. The tedious work of performing statistical computations can easily be transferred to the software, provided that data are stored in a digital database in a standardized format.

Moreover, recent advances in natural language processing, involving machine learning algorithms, provide a powerful set of tools for data analysis. These tools can be utilized to perform linguistic analysis even for less-resourced New Indo-Aryan languages.

REFERENCES

Agravāl, Rāmeśvar Prasād. 1963. *Bundelī kā bhāṣāśāstrīy adhyāyan* [A linguistic study of Bundeli]. Lakhnau: Lakhnau Viśvavidyālay.

Ahmad, Suhail. 1986. *A descriptive analysis of Eastern Awadhi dialect of Jaunpur.* Aligarh: Aligarh Muslim University, Department of Linguistics. (PhD Thesis).

Aissen, Judith. 2003. Differential object marking: Iconicity vs. economy. *Natural Language and Linguistic Theory* 21: 435-483. https://doi.org/10.1023/A:1024109008573.

Alpatov, Vladimir M. and Podlesskaya, Vera. 1995. Converbs in Japanese. In Martin Haspelmath and Ekkehard König (eds.), *Converbs in Cross-Linguistic Perspective*, 465-485. Berlin: Mouton de Gruyter.

Azeb, Amha and Dimmendaal, Gerrit Jan. 2006. Converbs in an African perspective. In Felix K. Ameka, Alan Dench and Nicholas Evans (eds.), *Catching language: the standing challenge of grammar writing,* 393-440. Berlin: Mouton de Gruyter.

Bashir, Elena. 1999. The Urdu and Hindi ergative postposition ne: Its changing role in the grammar. In Rajendra Singh (ed.), *The Yearbook of South Asian Languages and Linguistics*. 11-36. New Delhi: Sage.

Bauer, Brigitte. 2000. *Archaic Syntax in Indo-European*. Berlin–New York: Mouton de Gruyter.

Baum, Leonard and Petrie, Ted. 1966. Statistical inference for probabilistic functions of finite state markov chains. *The Annals of Mathematical Statistics*. 37 (6): 1554-1563.

Bickel, Balthasar and Yādava, Yogendra P. 2000. A fresh look at grammatical relations in Indo-Aryan. *Lingua* 110: 342-373. https://doi.org/10.1016/S0024-3841(99)00048-0.

Bickel, Balthasar. 1998. Review article of 'Converbs in cross-linguistic perspective'. *Linguistic Typology* 2: 381-397.

Bickel, Balthasar. 2010. Capturing particulars and universals in clause linkage: a multivariate analysis. In Isabelle Brill (ed.), *Clause Linking and Clause Hierarchy*: Syntax and Pragmatics. 51-101. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Bisang, Walter. 1995. Verb serialization and converbs – differences and similarities. In Martin Haspelmath and Ekkehard König (eds.), *Converbs in cross-linguistic perspective*. 137-188. Berlin: Mouton de Gruyter.

Bloch, Jules. 1965. *Indo-Aryan. From the Vedas to modern times*. Paris: Adrien Maisonneuve.

Bojałkowska, Krystyna. 2010. *Opis składniowy imiesłowów przysłówkowych we współczesnym języku polskim*. [Syntactic description of adverbial participles in the contemporary Polish language]. Toruń: Wydawnictwo Naukowe Uniwersytetu Mikołaja Kopernika.

Bossong, Georg. 1985. *Differentielle Objektmarkierung in den neuiranischen Sprachen*. Tübingen: Narr.

Bossong, Georg. 1991. Differential object marking in Romance and beyond. In Dieter Wanner and Douglas A. Kibbee (eds.), *New analyses in Romance linguistics. Selected papers from the XVIII Linguistic Symposium on Romance Languages, Urbana-Champaign, April 7-9, 1988*, 143-170. Amsterdam: John Benjamins.

Bubenik, Vit. 1998. *A Historical Syntax of Late Middle Indo-Aryan (Apabhramśa)* [Current Issues in Linguistic Theory 165]. Amsterdam: John Benjamins. https://doi.org/10.1075/cilt.165.

Butt, Miriam. 2001. A reexamination of the accusative to ergative shift in Indo-Aryan. In Miriam Butt & Tracy Holloway King (eds.), *Time Over Matter. Diachronic Perspectives on Morphosyntax*. 105-141. Stanford CA: CSLI.

Butt, Miriam, and King, Tracy Holloway. 2005. The Status of Case. In Veneeta Dayal and Anoop Mahajan (eds.), *Clause Structure in South Asian Languages*. 153-198. Berlin: Kluwer Academic Publishers.

Butt, Miriam and Lahiri, Aditi. 2013. Diachronic pertinacity of light verbs. Lingua 135, 7-29. https://doi.org/10.1016/j.lingua.2012.11.006.

Chatterji, Suniti Kumar. 1926 [1970]. *The Origin and Development of the Bengali Language*. Delhi: Motilal Banarsidass.

Coupe, Alexander R. 2006. Converbs. In Brown, Keith (ed.), *Encyclopedia of languages and linguistics*, 2nd Edition, Vol. 3, 145-152. Oxford: Elsevier.

Croft, William. 2003. *Typology and universals*. Cambridge: Cambridge University Press.

Dahl, Eystein and Stroński, Krzysztof. 2016. Ergativity in Indo-Aryan and beyond. In Eystein Dahl and Krzysztof Stroński (eds.), Indo-Aryan Ergativity in Typological and Diachronic Perspective, 1-38. Amsterdam/Philadelphia: John Benjamins Publishing Company.

Davison, Alice. 1981. Syntactic and semantic indeterminacy resolved: A mostly pragmatic analysis for the Hindi conjunctive participle. In Peter Cole (ed.), *Radical pragmatics*. 101-128. New York: Academic Press.

Davison, Alice. 1986. Hindi -kar: the problem of multiple syntactic interpretation. In Bhadriraju Krishnamurti, Collin P. Masica and Anjani K. Sinha (eds.), *South Asian languages. Structure, convergence, diglossia.* 1-14. Delhi: Motilal Banarsidass.

Davison, Alice. 2008. Case restriction on control: Implications for movement. *Journal of South Asian Linguistics* 1(1): 29-54.

de Hoop, Helen and Narasimhan, Bhuvana. 2009. Ergative case-marking in Hindi. In Helen de Hoop and Peter de Swart (eds.), *Differential Subject Marking, Studies in Natural Language and Linguistic Theory*. 63-78. Dordrecht: Kluwer.

Disterheft, Dorothy. 1980. *The syntactic development of the infinitive in Indo-European*. Columbus, Ohio: Slavica Publishers, Inc.

Dixon, Robert M. W. 1979. Ergativity. *Language* 55(1): 59-138. https://doi.org/10.2307/412519.

Dixon, Robert M. W. 1994. *Ergativity*. Cambridge Studies in Linguistics. Cambridge University Press.

Drocco, Andrea. 2008. *L'ergativita in hindī. Studio diacronico del processo di diffusione della posposizione* 'ne'. Alessandria: Edizioni dell'Orso.

Drocco, Andrea. 2017. Rājasthānī Features in Medieval Early Braj Prose Texts: The Case of Differential Object Marking and Verbal Agreement in Perfective Clauses. *Annali di Ca' Foscari. Serie orientale* Vol. 53. 205-234. https://doi.org/10.14277/2385-3042/AnnOr-53-17-7.

Dube, K. C. (ed.). 1983. *Ādibhāratī Bhāg 2. Kuṛukẖ-Hindī.* Bhopāl: Ādimjāti Anusandhān evam Vikās Saṁsthā.

Dwarikesh, Dwarika Prasad Sharma. 1971. *Historical syntax of the conjunctive participle phrase in New Indo-Aryan dialects of Madhyadesa (Midland) of northern India*. (University of Chicago Ph.D. dissertation).

Džanmavov, Jusup Dž. 1967. *Deepričastija v kumyskom literaturnom jazyke (srvanitiel'no s drugimi tjurkiskimi jazykami)* [Converbs in Kumyk literary language (in comparison with other Turkic languages)]. Moskva: Nauka.

Ebert. Karen H. 1993. Kiranti subordination in the South Asian areal context. In Karen H. Ebert (ed.), *Studies in clause linkage. Papers from the first Köln-Zürich workshop*, 85-110. Zurich: ASAS.

Fauconnier, Stefanie and Verstraete, Jean-Christophe. 2014. A and O as each other's mirror image? Problems with markedness reversal. *Linguistic Typology* 18(1). 3-49.

Foley, William A., and Van Valin, Robert D. 1984. *Functional syntax and universal grammar*. Cambridge: Cambridge University Press.

Gair, James W. and Paolillo, John C. 1997. *Sinhala*. München: Lincom Europa.

Gasser, Michael. (1983). Topic continuity in written Amharic discourse. In Talmy Givón (ed.), *Topic continuity in discourse*. 95-139. Amsterdam & Philadelphia: John Benjamins.

Givón, Talmy. 1990. *Syntax: a Functional-Typological Introduction 2.* Amsterdam; Philadelphia: John Benjamins.

Givón, Talmy. 2001a. *Syntax,* Vol. I. Amsterdam: John Benjamins. https://doi.org/10.1075/z.syn1 z.syn1.

Givón, Talmy. 2001b. *Syntax,* Vol. II. Amsterdam: John Benjamins. https://doi.org/10.1075/ z.syn2.

Goodwin, William Watson. 1889. *Syntax of the moods and tenses of the Greek verb*. Boston: Ginn and Company.

Green, Georgia M. 1976. *Main clause phenomena in subordinate clauses*. Language 52: 382-397.

Grierson, George A. 1916. *Linguistic Survey of India*, Vol. IX, Part IV: *Specimens of the Pahari Languages and Gujuri*. Calcutta.

Haig, Geoffrey and Stefan Schnell. 2015. *Annotations using graid (grammatical relations and animacy in discourse); manual version 7.0.* Otto-Friedrich-Universität, Bamberg.

Haiman, John and Thompson, Sandra A. 1984, *'Subordination' in universal grammar*. *BLS* 10: 510-523.

Haiman, John. 1985. *Natural Syntax*. Cambridge: Cambridge University Press.

Halmøy, Jane-Odile. 1982. *Le gérondif, éléments pour une description syntaxique et sémantique* (thèse). Trondheim: Tapir.

Hardie, Andrew. 2005. Automated part-of-speech analysis of Urdu: conceptual and technical issues. *Contemporary issues in Nepalese linguistics*. 48-72. Kathmandu.

Haspelmath Martin and König Ekkehard (eds.). 1995. *Converbs in cross-linguistic perspective: structure and meaning of adverbial verb forms – adverbial participles, gerunds*. (Empirical approaches to language typology, 13). Berlin: Mouton de Gruyter.

Haspelmath, Martin. 1989. *From purposive to infinitive – a universal path of grammaticization. Folia Linguistica Historica* 10.1-2: 287-310.

Haspelmath, Martin. 1995. The converb as a cross-linguistically valid category. In Martin Haspelmath and Ekkehard König (eds.), *Converbs in cross-linguistic perspective: structure and meaning of adverbial verb forms – adverbial participles, gerunds*. (Empirical approaches to language typology, 13). 1-55. Berlin: Mouton de Gruyter.

Hellwig, Oliver. 2008. Sanskrittagger: A stochastic lexical and POS tagger for Sanskrit. In Gérard P. Huet, Amba P. Kulkarni and Peter M. Scharf (eds.), *Sanskrit Computational Linguistics*, Lecture Notes in Computer Science 5402. Berlin: Springer. 106-115.

Henadreerage, Deepthi Kumar. 2002. *Topics in Sinhala syntax*. PhD thesis. Canberra: Australian National University.

Hendriksen, Hans. 1986. *Himachali Studies,* III: *Grammar* [Det Kongelige Danske Videnskabernes Selskab Historisk-filosofiske Meddelelser 48(3)]. Kobenhavn: Munksgaard.

Hendriksen, Hans (1944). *Syntax of the Infinite Verb-forms of Pāli*, Copenhagen: Einar Munksgaard.

Hook, Peter Edwin. 1996. Kesar of Layul: A Central Asian Epic in the Shina of Gultari. In William Hanaway and Wilma Heston (eds.), *Studies in Pakistani Popular Culture*. 121-183. Lahore: Sang-e-Meel and Lok Virsa.

Iemmolo, Giorgio and Klumpp, Gerson (eds.). 2014. Special thematic issue on *Differential Object Marking. Linguistics* 52(2).

Iemmolo, Giorgio. 2010. Topicality and differential object marking. Evidence from Romance and beyond. *Studies in Language* 34(2): 239-272.

Jaworski, Rafał. 2015. IATagger. http://rjawor.vm.wmi.amu.edu.pl/tagging/.

Jaynes, Edwin. 1957. Information theory and statistical mechanics. *Physical Review.* 106 (4): 620-630.

Jha, Subhadra. 1958 [1985]. *The Formation of the Maithili Language*. New Delhi: Munshiram Manoharlal.

Johanson, Lars. 1995. On Turkic converb clauses. In Martin Haspelmath and Ekkehard König (eds.), *Converbs in cross-linguistic perspective: structure and meaning of adverbial verb forms – adverbial participles, gerunds*. (Empirical approaches to language typology, 13). 313-347. Berlin & New York: Mouton de Gruyter.

Johns, Alana. and Smallwood, Carolyn R. 1999. On (non-)finiteness in Inuktitut. Toronto Working Papers in Linguistics 17. The Hague: Thesus – Holland Academic Graphics.

Juyal, Govind. 1973. *Madhya Pahari Bhasha (Garhvali Kumaoni) ka anushilan aur uska hindi se sambandh* [The study of Middle Pahari language (Garhwali and Kumaoni) and its connection with Hindi]. Lucknow: Navyug Granthagar.

Kachru, Yamuna. 1981. On the Syntax, Semantics and Pragmatics of the Conjunctive Participle in Hindi-Urdu. *Studies in Linguistic Sciences.* 11(2): 35-49.

Kachru, Yamuna, Kachru, B. and Bhatia, K. Tej, 1976. 'Subject'. A note on Hindi-Urdu. Kashmiri and Punjabi. In Manindra K. Verma (ed.), 1976. *The Notion of Subject in South Asian Languages*. Madison: University of Wisconsin. (South Asian Studies, Publication Series 2). 79-108.

Kāmtāprasād, Guru. 2060. *Hindī vyākaraṇ* (Hindi grammar). Kāśī: Nāgarīpracāriṇī sabhā.

Katenina T. E. 1963. *Očerk grammatiki jazyka maratxi*. Moskva: Izdatel'stvo literatury na inostrannych jazykach.

Khokhlova, Liudmila V. 1992. Trends in the development of ergativity in New Indo-Aryan. *Osmania Papers in Linguistics* 18: 71-98.

Khokhlova, Ludmila V. 2000. Typological evolution of Western New Indo-Aryan languages. *Berliner Indologische Studien* 13-14: 117-142.

Khokhlova, Ludmila V. 2001. Ergativity attrition in the history of western New Indo-Aryan languages (Punjabi, Gujarati and Rajastahani). In Peri Bhaskararao and Karumuri Venkata Subbarao (eds.), *The yearbook of South Asian languages and linguistics*. 159-184. New Delhi/London: Sage Publication.

Khokhlova, Ludmila V. 2006. Sintaktičeskaja evolucija zapadnych novoindijskich jazykov v 15-20 v. [Syntactic evolution of Western New Indo-Aryan languages in 15-20 c.]. In Dybo et al. (eds.), Aspekty komparativistiki [Aspects of comparative studies]. 151-186. Moskva: Rosijskij Gosudarstvennyj Gumanitarnyj Universitet (Orientalia et Classica: Trudy Instituta Vostočnych Kultur i Antičnosti: Vypusk VIII).

Khokhlova, Liudmila V. 2013. Obligational Constructions in New Indo-Aryan Languages of Western India. *Lingua Posnaniensis* 55(2): 91-107.

König, Ekkehard. 1995. The meaning of converb constructions. In Martin Haspelmath and Ekkehard König (eds.), *Converbs in cross-linguistic perspective*. 57-96. Berlin: Mouton de Gruyter.

Koptjevskaja-Tamm, M. 1994. Finiteness. In Asher, R. E. & Simpson, J. M. Y. (eds.), *The encyclopedia of languages and linguistics* (1st edn.). Oxford & New York: Permagon Press. 1245-1248.

Koptjevskaja-Tamm, Maria. 1993. *Nominalizations.* London: Routledge.

Kortmann, Bernd. 1991. *Free Adjuncts and Absolutes in English: Problems of Control and Interpretation*. London/New York: Routledge.

Kuryłowicz, Jerzy. 1960 [1931]. Les Temps composés du roman. In Esquisses linguistiques. 104-108. Kraków: Polska Akademia Nauk.

Kuryłowicz, Jerzy. 1965. The evolution of grammatical categories. *Diogene* 51: 51-71.

Lakoff, George. 1984. *Performative subordinate clauses*. *BLS* 10: 472-480.

Lamberti, Marcello and Sottile, Roberto. 1997. *The Wolaytta language*. Köln: Rüdiger Köppe Verlag.

Lāḷas, Padmaśri Sītārām. 1997. *Rājasthānī vyākaraṇ* [A grammar of Rajasthani]. Jodhpur: Rājasthānī Śodh Saṃsthān, Ćaupāsanī.

Langford, John, Lihong Li, and Tong Zhang. 2009. *Sparse online learning via truncated gradient*. In The Journal of Machine Learning Research vol. 10: 777-801, University of Pennsylvania, USA.

Lazard Gilbert. 1983. Anti-impersonal Verbs, Transitivity Continuum and the Notion of Transitivity. In Hansjakob Seiler and Gunter Brettschneider (eds.), *Language Invariants and Mental Operations*. 115-123. International Interdisciplinary Conference held at Gummersbach/Cologne, Germany September 18-23.

Lehmann, Christian. 1988. Towards a typology of clause linkage. In John Haiman and Sandra. A. Thompson (eds.), *Clause combining in grammar and discourse*. 181-226. Amsterdam: Benjamins.

Liperovskij, Vladimir P. 1988. *Očerk grammatiki sovremennogo bradža*. Moskva: Nauka.

Liperovskij, Vladimir P. 1997. *Očerk grammatiki sovremennogo avadhi*. Moskva: Izdatel'skaja Firma 'Vostočnaja Literatura' RAN.

Lohar, Gopal Thakur. 2012. Converbal constructions in Bhojpuri. *Nepalese Linguistics* 27: 217-222.

Loper, E. and Bird, S. 2002. NLTK: The natural language toolkit. In Proceedings of the ACL-02 Workshop on Effective Tools and Methodologies for Teaching Natural Language Processing and Computational Linguistics. ETMTNLP '02, Association for Computational Linguistics, Stroudsburg, PA, USA. 1: 63-70.

Makhoul, J., Kubala, F., Schwartz, R., Weischedel, R. 1999. Performance measures for information extraction. In Proceedings of DARPA Broadcast News Workshop. 249-252, Herndon, Virginia, USA.

Markoppoulos, Theodore. 2009. *The Future in Greek. From Ancient to Medieval*. Oxford: Oxford University Press.

Masica, Colin P. 1991. *The Indo-Aryan Languages.* Cambridge: CUP.

Masica, Colin P. 1976. *Defining a linguistic area. South Asia*. Chicago and London: Chicago University Press.

Matthews, David. 1992. *A course in Nepali*. London: SOAS.

McGregor, William B. 2009. Typology of ergativity. *Language and Linguistics Compass* 3(1): 480-508. https://doi.org/10.1111/j.1749-818X.2008.00118.x.

McGregor, William B. 2010. Optional ergative case marking systems in a typological-semiotic perspective. *Lingua* 120: 1610-1636. https://doi.org/10.1016/j.lingua.2009.05.010.

Mikolov, Tomas, Kai Chen, Greg Corrado, and Jeffrey Dean. 2013. *Efficient estimation of word representations in vector space*. In Proceedings of the International Conference on Learning Representations (ICLR 2013), Scottsdale, AZ, USA.

Montaut, Annie. 2004a. Oblique main arguments in Hindi as localizing predications: Questioning the category of subject. In Peri Bhaskararao & Karumuri Venkata Subbarao (eds.),

Non-nominative Subjects, Vol. 2 [Typological Studies in Language 61]. 33-56. Amsterdam: John Benjamins. https://doi.org/10.1075/tsl.61.04mon.

Montaut, Annie. 2004b. *A Grammar of Hindi*. Munich: Lincom.

Montaut, Annie. 2017. Grammaticalization of participles and gerunds in Indo-Aryan; preterite, future, infinitive. In Andrej Malchukov and Bisang Walter (eds.), *Unity and diversity in grammaticalization scenario*. 95-134. Language Science Press (Studies in Diversity Linguistics 99).

Montaut, Annie. 2018a. The rise of differential object marking in Hindi and related languages. In Ilja A. Seržant and Alena Witzlack-Makarevich (eds.), Diachrony of differential argument marking. [Studies in Diversity Linguistics]. 281-313. Berlin: Language Science Press.

Montaut, Annie. 2018b. On the nature of the Hindi infinitive: History as an answer to its syntactic behavior? In Ghyan Sharma and Rajesh Bhatt (eds.), *Trends in Hindi linguistics* (Trends in Linguistics, Studies and Monographs, Volume 325). 115-146. Berlin/Boston: De Gruyter Mouton.

Müller-Bardey, T. 1990. Koprädikation als grundlegende syntaktische Operation. In Hansjakob Seiler (ed.), *Internationales Interdisziplinäres Kolloquium "Sprache und Denken, Variation und Invarianz in Linguistik und Nachbardisziplinen": Lenzburg/Schweiz, 16.-19. Mai 1989*. 1-21. Cologne: University of Cologne.

Mustafa, Khateeb S. 2000. *A descriptive grammar of Dakkhini*. Delhi: Mushiram Manoharlal Publishers Pvt. Ltd.

Nedjalkov, Igor V. 1998. Converbs in the languages of Europe. In Johan van der Auwera and Dónall Ó Baoill (eds.), *Adverbial Constructions in the Languages of Europe*. 421-456. Berlin: Walter de Gruyter.

Nedjalkov, Vladimir P. and Nedjalkov, Igor V. 1987. On the typological characteristics of converbs. In Toomas Help and Sirje Murumets (eds.), *Symposium on Language Universals*. 75-79. Tallinn: Academy of Sciences of the Estonian SSR.

Nedjalkov, Vladimir P. 1995. Some typological parameters of converbs. In Martin Haspelmath and Ekkehard König (eds.), *Converbs in cross-linguistic perspective: structure and meaning of adverbial verb forms – adverbial participles, gerunds*. (Empirical approaches to language typology, 13). 97-136. Berlin: Mouton de Gruyter.

Nedjalkov, Vladimir P. 1990. Osnovnye tipy deepričastij. [Basic types of converbs]. In Viktor Xrakovskij (ed.), Tipologija i grammatika: 36-59. Moscow: Nauka.

Nikolaeva, Irina. 2007. Introduction. In Irina Nikolaeva (ed.), *Finiteness: Theoretical and empirical foundations*. 1-19. Oxford: Oxford University Press.

Noonan, Michael. 2007. Complementation. In Timothy Shopen (ed.), *Language Typology and Syntactic Description, Second edition, Volume II: Complex Constructions*. 52-150. Cambridge University Press.

Oberlies, Thomas. 2005. *A Historical Grammar of Hindi*. Grazer Vergleichende Arbeiten, Band 19. Graz: Leykam.

Olson, Michael L. 1981. *Barai clause junctures: toward a functional theory of interclausal relations*. Unpublished Ph.D. dissertation, Australian National University.

Pāṁḍe, Hemcandra. 2012. *Samasāmayik hindī vyākaraṇ* [Grammar of contemporary Hindi]. Dillī: Granthlok.

Peterson, John. 1998. *Grammatical Relations in Pali and the Emergence of Ergativity in Indo-Aryan*. Munich: Lincom.

Peterson, John. 2002. The Nepali converbs: a holistic approach. In Rajendra Singh (ed.), *Yearbook of South Asian Languages and Linguistics*. New Delhi/Thousand Oaks/London: Sage Publications. 93-133.

Pokharel, Madhav P. 2008. Grammaticalization of clause combining in the inscriptional Nepali. *Nepalese Linguistics.* 28: 208-220.

Pořízka, Vincent. 1950. Hindi participles used as substantives. Marginal notes on the syntax of the imperfective and perfective participles in Hindi prose. *Archiv Orientalni* 18: 166-187.

Pořízka, Vincent. 1953. The adjectival and adverbial participles in Hind syntax. *Archiv Orientalni*, 20: 524-538.

Prakāś, Śyām.1986. *Braj bhāṣā kī kriyāpad-saṁracnā* [Verbal structure of the Braj language]. Lakhnaū: Sneha prakāśan.

Pray, Bruce R. 1976. From passive to ergative in Indo-Aryan. In *The Notion of Subject in South Asian Languages*, Manindra K. Verma (ed.), 195-211. Madison WI: Department of South Asian Studies, University of Wisconsin.

Radice, William. 2003. *Teach yourself Bengali*. London: Hodder Headline Ltd.

Ramstedt, John G. 1903. Über die Konjugation des Khalkha-Mongolischen. (Mémoires de la Société finno-ougrienne 19). Helsingfors: Société finno-ougrienne.

Ratnaparkhi, Adwait. 1996. A maximum entropy model for part-of-speech tagging. In Proceedings of the Conference on Empirical Methods in Natural Language Processing, April.

Renkovskaya, Evgeniya. 2018. Genitive marking of arguments in Kullui (Indo-Aryan). Paper presented at 51 Societas Linguistica Europea Meeting 29 August – 1 September in Tallinn.

Ross, John Robert. 1967. Constraints on Variables in Syntax. Ph.D. dissertation, MIT.

Ruppel, Antonia. 2012. *The Development of the Genitive Absolute in Sanskrit*. In Jared Klein and Kazuhiko Yoshida (eds.), *Indic across the Millennia: from the Rigveda to Modern Indo-Aryan,* 14th World Sanskrit Conference Kyoto, Japan. 135-143. Bremen: Hempen Verlag.

Ruppel, Antonia. 2013. *Absolute Constructions in Early Indo-European*. Cambridge: Cambridge University Press.

Saksena Baburam. 1971 [1937]. *Evolution of Early Awadhi*. Delhi-Patna-Varanasi: Motilal Banarsidass.

Šamatov, Azad N. 1974. Klassičeskij dakhini (Južnyj hindustani XVII v.). [Classical Dakkhini (Southern Hindustani XVII)]. Moskva: Izdetel'stvo 'Nauka'.

Śarmā, Viśvanabdhu. 2006. *Hariyaṇvī bhāṣā, svarūp aur pahcān* [The Hariyaṇvī language essential properties and characteristics]. Dillī: anang prakāśan.

Schumacher, Rolf. 1977. *Untersuchungen zum Absolutiv in modernen Hindi*. Frankfurt a.M.: Lang.

Sigorskij, Aleksandr. 2005. Deepričastija v jazyke xindi [Converbs in the Hindi language]. *Sbornik naučnychtrudov/MGIMO(U) MID Rossii*, 21 (36), 54-63.

Silverstein, Michael. 1976. Hierarchy of features and ergativity. In *Grammatical Categories in Australian Languages*, Robert M. W. Dixon (ed.), 112-171. Canberra: Australian Institute of Aboriginal Studies & Atlantic Highlands NJ: Humanities Press.

Singh Jag Deva. 1970. *A Descriptive Grammar of Bangru*. Kurukshetra: Kurukshetra University.

Southworth, Franklin C. 2005. *Linguistic Archeology of South Asia.* London: RoutledgeCurzon. https://doi.org/10.4324/9780203412916.

Stroński, Krzysztof. 2010. Non-nominative subjects in Rajasthani and Central Pahari. The status of the ergative and obligative constructions. *Lingua Posnaniensis* LII(1): 81-97.

Stroński, Krzysztof. 2010. Variation of ergativity patterns in Indo-Aryan. *Poznań Studies in Contemporary Linguistics* 46(2): 237-253. https://doi.org/10.2478/v10010-010-0012-8.

Stroński, Krzysztof. 2011. Synchronic and diachronic aspects of ergativity in Indo-Aryan. Poznań: Wydawnictwo Naukowe UAM.

Stroński, Krzysztof. 2014. On Syntax and Semantics of the Past Perfect Participle and Gerundive in Early NIA – Evidence from Eastern Early Pahari. *Folia Linguistica Historica* 35: 275-305.

Stroński, Krzysztof, Tokaj, Joanna and Verbeke, Saartje. 2019. A diachronic account of converbal constructions in Early Rajasthani. In Michela Cennamo and Claudia Fabrizio (eds.), *Historical Linguistics 2015. Selected papers from the 22nd International Conference on Historical Linguistics, Naples, 27-31 July 2015*. Current Issues in Linguistic Theory. 423-441. Amsterdam – Philadelphia: John Benjamins. https://doi.org/10.1075/cilt.348.20str.

Stump, Gregory T. 1983. The elimination of ergative patterns of case-marking and verbal agreement in modern Indic languages. OSU *Working Papers in Linguistics* 27: 140-164.

Subbarao, Kurumuri V. 2012. *South Asian languages: A syntactic typology*. Cambridge: Cambridge University Press. https://doi.org/10.1017/CBO9781139003575.

Śukla, Prabhākar. 2022 (V.). *Jāyasī kī bhāṣā* [The language of Jayasi]. Lakhnaū: Lakhnaū Viśvavidyālay.

Telang, Bhālchandra Rāo. 1966. *Chattīsgarhī, halbī, bhatrī boliyõ kā bhāṣāvaigyānik adhyayan* [A linguistic study of Chattisgarhi, Halbi and Bhatri]. Bombai: Hindī Grantha Ratnākar Private Limited.

Tessitori Luigi. 1914-16. Notes on the grammar of the Old Western Rajasthani, with special reference to Apabhramsha and to Gujarati and Marwari. *Indian Antiquary* 43-44: 1-106.

Tessitori, Luigi. 1914. On the origin of the perfect participles in the Neo-Indian vernaculars. *Zeitschrift der Deutschen Morgenlandischen Gesellschaft* 68: 571-578.

Ṭhākur, Maulūrām. 1975. *Pahāṛī bhāṣā kuluī ke viśeṣ sadarbh mẽ* [The Pahari language with special reference to Kului]. Dillī: Sanmārg Prakāśan.

Tikkanen, Bertil. 1995. Burushaski converbs in their South and Central Asian areal context. In Haspelmath Martin and König Ekkehard (eds.), *Converbs in cross-linguistic perspective:*

structure and meaning of adverbial verb forms – adverbial participles, gerunds. (Empirical approaches to language typology, 13). Berlin: Mouton de Gruyter. 487-528.

Tikkanen, Bertil. 1987. *The Sanskrit Gerund: A Synchronic. Diachronic and Typological Analysis*. Studia Orientalia, 62, Helsinki.

Tikkanen, Bertil. 2001. Converbs In Martin Haspelmath, Ekkehard König and Wulf Oesterreicher and Wolfgang Raible (eds.), *Language typology and language universals: an international handbook Vol 2*. 1112-1123. Berlin: Walter de Gruyter.

Tivāri Bholanāth. 2005. *Hindī bhāṣā* [The Hindi language]. Ilāhābād: Kitāb Mahal.

Toutanova, K., Klein, D., Manning, C.D., Singer, Y. 2003: *Feature-rich part-of-speech tagging with a cyclic dependency network* In NAACL '03 Proceedings of the 2003 Conference of the North American Chapter of the Association for Computational Linguistics on Human Language Technology – Volume 1, 173-180, Edmonton, Canada.

Toutanova, K., Manning, C.D. 2000: *Enriching the knowledge sources used in a maximum entropy part-of-speech tagger*. In EMNLP '00 Proceedings of the 2000 Joint SIGDAT conference on Empirical methods in natural language processing and very large corpora: held in conjunction with the 38th Annual Meeting of the Association for Computational Linguistics. 13: 63-70. Hong Kong.

Tripāthi, Satroj Kumār. 1993. *Hindī aur magahī vyākaraṇik sãracnā* [Grammatical structure of Hindi and Magahi]. Paṭnā: Candr Prakāśan.

Van der Auwera, Johan. 1998. Defining converbs. In Kulikov, Leonid and Vater, Heinz (eds.), *Typology of verbal categories: Papers presented to Vladimir Nedjalkov on the occasion of his 70th birthday*. Tübingen: Niemeyer, 273-282.

Van Valin Jr., Robert. 2004. *An Introduction to Syntax.* Cambridge: Cambridge University Press.

Van Valin Jr., Robert. 2005. *Exploring the Syntax-semantics Interface*. Cambridge: Cambridge University Press.

Van Valin, Robert J. and LaPolla, Randy. 1997. *Syntax*. Cambridge: Cambridge University Press.

Varmā, Dhīrendra. 1954. *Brajbhāṣā* [The language Braj]. Ilāhābād: Hindustāni Akedemī.

Verbeke, Saartje. 2013a. Ergativity and alignment in New Indo-Aryan languages. Berlin-New York: Mouton de Gruyter. https://doi.org/10.1515/9783110292671.

Verbeke, Saartje. 2013b. Differential subject marking in Nepali: The agent marker le in imperfective constructions. *Linguistics* 51(3): 585-610. https://doi.org/10.1515/ling-2013-0021.

Verbeke, Saartje and De Cuypere, Ludovic. 2015. Differential subject marking in Nepali imperfective constructions: a probabilistic grammar approach. *Studies in Language* 39(1): 1-23.

Vincent, Nigel. 1998. On the grammar of inflected nonfinite forms (with special reference to Old Neapolitan). In Iørn Korzen and Michael Herslund (eds.), *Clause Combining and Text Structure*. 135-158 Copenhagen: Samfundsliteratur.

Wali, Kashi and Koul Omkar N. 2006. *Modern Kashmiri grammar*. Springfield: Dunwoody Press.

Wallace, William D. 1981. Object-marking in the history of Nepali: A case of syntactic diffusion. *Studies in the Linguistic Sciences* 11(2): 107-128.

Wallace, William D. 1982. The evolution of ergative syntax in Nepali. *Studies in the Linguistic Sciences* 12(2): 147-211.

Weiss, Daniel. 1995. Russian converbs: a typological outline. In Martin Haspelmath and Ekkehard König (eds.), *Converbs in cross-linguistic perspective: structure and meaning of adverbial verb forms – adverbial participles, gerunds*. (Empirical approaches to language typology, 13). 239-282. Berlin & New York: Mouton de Gruyter.

Witzlack-Makarevich, Alena & Ilja A. Seržant 2017. Differential argument marking: Patterns of variation. In Ilja A. Seržant and Alena Witzlack-Makarevich (eds.), Diachrony of differential argument marking. [Studies in Diversity Linguistics]. 1-40. Berlin: Language Science Press. http://langsci-press.org/.

Yadav, Ramawatar. 2004. On diachronic origins of converbs in Maithili. *Contributions to Nepalese Studies*, 31(2): 215-241.

Ylikoski, Jussi. 2003. Defining non-finites: Action nominals, converbs, and infinitives. SKY Journal of Linguistics 16: 185-237.

Zoller, Claus Peter. 2008. Genitive marking of subjects in West Pahāṛī. *Acta Orientalia* 69: 121-151.

Zúñiga, Fernando. 1998. Nomina sunt odiosa: A critique of the converb as a cross-linguistically valid category. Ms. University of Zurich.

PRIMARY SOURCES:

Bhānāvat, Narendra and Kamal, Lakshmi (eds.). 1997-1998. Rājasthānī gadya: vikās aur prakāś [Rajasthani prose: development and publicity]. Āgrā: Śrīrām Mehrā end Kampanī. (R.G.)

Chhabra, Bahadur Chand. 1957. *Antiquities of Chamba state*. Part II. New Delhi: ASI.

Dvivedī Silākārī, Loknāth. 1972 (ed.), *Rāmāyan kathā* [The story of Ramyana]. Ilāhābād: Sāhitya bhavan limiṭeḍ.

Gautam Manmohan. 1954. *Jāyasī granthavalī* [Books of Jayasi]. Delhi: Rigal Buk Ḍipo. (J.)

Joshi, Maheshwar P. 1983. *Rājanītiśāstra of Chāṇakya* (Text and translation). Almora: Śri Malika Publications.

Joshi, Maheshwar P. 2009. Advent of polities in Uttarkhand (Kumaon and Garhwal). In Marie Lecomte-Tilouine (ed.), *Bards and mediums: History, culture, and politics in the central Himalayan kingdoms*. 327-371. Almora: Shri Almora Book Depot.

McGregor, Ronald S. 1968. *The Language of Indrajit of Orchā. A Study of early Braj Bhāsā prose*. Cambridge: CUP.

Miśra, Viśvanāth P. 1994 (ed.), *Bhūṣaṇa granthāvalī* [Collected works of Bhushan]. Naī Dillī: Vāṇī prakāśan.

Pant Prabhā. 2006. *Kumāunī lok kathā* [Kumaoni folk tales]. Almoṛā: Maniś Prakāśan.

Pant, Mahes Raj. 2009. Towards a history of the Khasa empire. In Marie Lecomte-Tilouine (ed.), *Bards and mediums: History, culture, and politics in the central Himalayan kingdoms.* 293-326. Almora: Shri Almora Book Depot.

Prasad Ram Chandra (ed.). 1994. *Tulsidasa's shriramacaritamanasa*. Delhi: Motilal Banarsidass Publishers Private Limited.

Snell, Rupert. 1991a. *The Eighty-four Hymns of Hita Harivaṃśa*. Delhi: Motilal Banarsidass.

Snell, Rupert. 1991b. *The Hindi Classical Tradition. A Braj bhāṣā Reader*. London: School of Oriental and African Studies, University of London.

Śarmā Śrīrām. 1954. *Dakkhinī kā padya aur gadya* [Poetry and prose of Dakkhini]. Haidrābād: Hindī Pracār Sabhā.

ELECTRONIC RESOURCES:

www.kavitakosh.org.
https://wp.unil.ch/eniat/.

ABBREVIATIONS

Glossing follows the Leipzig Glossing Rules (http://www.eva.mpg.de/lingua/resources/glossing-rules.php). Aditional abbreviations employed in this book are:

AN	action nominal
ARG	argument
ASP	aspect
CMPL	complementizer
DCM	differential case marking
DEM	differential ergative case marking
DIR	direct case
DOM	differential object marking
DSM	differential subject marking
ENCL	enclitic
GNOM	gnomic (nonevidential)
IA	Indo-Aryan
IF	illocutionary force
MIA	Middle Indo-Aryan
MOD	modality
NIA	New Indo-Aryan
NUC	nucleus
O	object of a transitive verb
OCM	optional case marking
OCS	Old Church Slavonic
OEM	optional ergative case marking
OIA	Old Indo-Aryan
PPP	past passive participle
SCM	split case marking
SIC	subject identity constraint
TAM	tense, aspect, mood
TNS	tense
TPST	today's past.

INDEX OF LANGUAGES

ISBN 978-83-954144-4-2