1 The Indo-Aryan Languages: a tour

- sub-branch of the Indo-European family, spoken mainly in India, Pakistan, Bangladesh, Nepal, Sri Lanka, and the Maldives by at least 640 million people (according to the 1981 census). (Masica (1991)).
- Together with the Iranian languages to the west (Persian, Kurdish, Dari, Pashto, Baluchi, Ormuri etc.) , the Indo-Aryan languages form the Indo-Iranian subgroup of the Indo-European family.
- Most of the subcontinent can be looked at as a dialect continuum. There seem to be no major geographical barriers to the movement of people in the subcontinent.

1.1 The Hindi Belt

According to the Ethnologue, in 1999, there were 491 million people who reported Hindi as their first language, and 58 million people who reported Urdu as their first language.

- Hindi-Urdu or Hindi and Urdu

The Hindi Belt involves a vast area that stretches across most of Northern India. Despite Hindi being the official language for this entire area, this area is home to many languages that are clearly distinct from Hindi.

Grierson (1883), Grierson (1969) divided the Hindi dialect area into the following subgroups:

- ‘Western Hindi’: Braj, Kannauji, Haryanvi
- ‘Eastern Hindi’: Awadhi, Bagheli, Chhattisgarhi
- ‘Bihari’: Bhojpuri, Magahi, Maithili
- ‘Central Pahari’: Kumauni, Garhwali
- ‘Rajasthani’: Marwari, Mewari, Harauti, Malvi

1.2 East of the Hindi Belt

The following languages are quite closely related:

- Assamese (Assam)
- Bengali (West Bengal, Tripura, Bangladesh)
- Oriya (Orissa)
- Bishnupriya Manipuri

This group of languages is also quite closely related to the ‘Bihari’ languages that are part of the Hindi belt: Bhojpuri, Magahi, Maithili.

1.3 Central Indo-Aryan

- Eastern Punjabi
- ‘Rajasthani’: Marwari, Mewari, Harauti, Malvi etc.
- Bhil Languages: Bili, Garasia, Rathawi, Wagdi etc.
- Gujarati, Saurashtra

The Bhil languages occupy an area that abuts ‘Rajasthani’, Gujarati, and Marathi. They have several properties in common with the surrounding languages. Central Indo-Aryan is also where Modern Standard Hindi fits in. Some central Indo-Aryan languages are spoken far from the subcontinent. These include the various forms of Romani, and Parya, spoken by about 1,000 people in Tajikistan and Uzbekistan.

1.4 Northern and Northwestern Indo-Aryan

The Pahari (Hill) languages:

- Eastern: Nepali (spoken in Nepal, adjoining parts of India (West Bengal, Sikkim), and Bhutan)
- Central: Garhwal, Kumauni
- Western: Dogri/Kangri, Jaunsari, Bilaspuri

Northwestern Indo-Aryan:

- Dardic
- Sindhi: Sindhi (Sindhi1, Pakistan), Kachchhi (Gujarat)

1Peccavi! (I have sinned!) British general Sir Charles James Napier sent this one word message to his commanding officer, Lord Ellenborough, after he had captured Sindh, in modern Pakistan.
1.4 Dardic Languages

- Lahnda: Hindko, Western Punjabi, Saraiki/Siraiki
- Dardic:
  - Chitral: Khowar, Kalasha
  - Kashmiri
  - Kohistani: Tirahi, Torwali
  - Kunar: Gawar-Bati, Shumashti
  - Shina: Brokskat/Brokpa, Domaki, Savi, Shina

Dardic languages used to be grouped together with the Nuristani (Kafiri) languages. But since Strand (1973), the Nuristani languages (Ashkun, Kati/Bashgali, Prasuni², Tregami, Waigali) have been analyzed as sisters of the Indo-Aryan and the Iranian language families.

1.5 Southern Indo-Aryan and Further

Southern Indo-Aryan

- Marathi (Maharashtra)
- Konkani (Goa, Mangalore, coastal areas of Maharashtra)

Further:

- Sinhala/Sinhalese (Sri Lanka)
- Veddah (Sri Lanka)
- Maldivian/Divehi (Maldives)

These languages have been separated from the rest of the Indo-Aryan languages since around the 5th century B.C. They are heavily influenced by the surrounding Dravidian languages.

1.6 The Non-Indo-Aryan Neighbourhood

The primary language family with which the Indo-Aryan languages came into contact with was Dravidian (Tamil, Kannada, Telugu, Malayalam). There are reasons to believe that Dravidian languages were spoken in parts of North India where Indo-Aryan languages are now spoken. Brahui, an isolated Dravidian language, is still spoken in parts of Pakistan. The influence of Dravidian languages is particularly clear on those Indo-Aryan that currently abut the Dravidian area (Marathi, Oriya) and on those which have been argued to abut the Dravidian area at some point in the relatively recent past (Gujarati).

The other language families in the area are:

- Iranian: Pashto, Dari, Persian, Baluchi, Ormuri
- Language Isolate: Burushaski
- Austroasiatic 1: Mon-Khmer: Amwi, Khasi, Pnar/Jaintia
- Austroasiatic 2: Munda: Mundari, Juang, Khariya, Gorum, Ho
- Tibeto-Burman: Kiranti, Newari, Meithei, Lepcha, Tangkhul, Hmar

2 Major Syntactic Phenomena

2.1 Case-Marking

Case-marking in Indo-Aryan languages is typically postpositional in nature. Subjects of finite intransitive clauses typically receive nominative case, which is usually unmarked.

(1) a. Leela kal aa-ii
   Leela.f yesterday come-Pfv.f
   ‘Leela came yesterday.’

b. Kalpna bahadur thii
   Kalpna.f brave be.Pst.F
   ‘Kalpna was brave.’

Nominative case (on subjects) seems to be licensed by finite Tense in many Indo-Aryan languages.

(2) [Shiraz*(-kaa) Ruta-se baat kar-naa] zaroori hai
   Shiraz-Gen Ruta-Instr talk do-Inf necessary be.Prs
   ‘It is necessary that Shiraz talk to Ruta.’

An nominative subject cannot appear in a non-finite clause in Hindi. Genitive marking is an option that is generally available. In certain ECM-like environments, an accusative ordative case can also be licensed. The conditions on the licensing of Nominative vary throughout Indo-Aryan. In particular, in Marathi, we find that nominative subjects can appear in what appear to be non-finite clauses.

2.1.1 Ergativity

The case on the subject of a finite transitive clause in the Western Indo-Aryan languages (Standard Hindi-Urdu, Marathi, Gujarati, Punjabi, Kashmiri, Sindhi but not Bengali, Oriya, Bhojpuri etc.) depends upon the aspect. In perfective tenses, the subject receives Ergative case. In all other tenses the case on the subject is nominative. Ergative case in Hindi-Urdu is marked by the case-clitic -ne.

(3) Perfective Tenses:
(4) Non-Perfective Tenses:  

- a. Habituals: imperfective/habitual participle plus tense auxiliary  
  Lataa-ji gaanaa gaa-e  
  Maybe Lataa-ji will sing (a song).  
- b. Future: subjunctive ending plus a participial ending  
  Lataa-ji gaanaa gaa-e-gii  
  Lataa-ji will sing a song.  

In Hindi-Urdu, the same perfective auxiliary appears in the simple past and the perfect. This is not the case in Kashmiri. However, the ergativity patterns are the same as in Hindi-Urdu.

(5) (Kashmiri from Wali and Koul (1997))  

- a. Simple Past:  
  Aslam-an por akhbaar  
  Aslam-erg read-MSg newspaper-MSg  
  ‘Aslam danced.’  
- b. Past Perfect (same facts obtain for the Present/Future Perfects):  
  me/tEm’ os akhbaar I-erg/he-erg be.Pst-MSg newspaper-MSg read-psp-Msg  
  ‘I danced.’

This optionality has been related to notions of volitionality, with the version with ergative marking being more volitional. Hindi and Kashmiri make a surface distinction between Ergative and Nominative case in all person-number combinations. Many of the other Ergative Indo-Aryan languages collapse the distinction between Ergative and Nominative in several person-number combinations (typically 1st and 2nd person, and plurals) (e.g. Gujarati (cf. Cardona (1965))),
Marathi (cf. Pandharipande (1997)), Punjabi (cf. Bhatia (1993)). This apparently surface fact has interesting implications for questions pertaining to markedness and the directionality of syntactic change as discussed in Deo and Sharma (2002).

It was noted in the discussion on the licensing of nominative on the subject that in many Indo-Aryan languages, nominative subjects cannot appear in non-finite environments. Somewhat curiously, in these languages this also seems to be true of Ergative subjects. Further the Indo-Aryan languages where nominative subjects can appear in non-finite environments also seem to allow Ergative subjects in certain non-finite environments.

2.1.2 Specificity Marking on Direct Objects

Direct Objects in most Indo-Aryan languages (with exception of Kashmiri and Sinhalese) use the following strategy:

1. Animate proper names must be marked by the postposition ko:
   (9) a. Animate Proper Name:
      Madhukar-ne Tara*(ko) dekh-aa
      Madhukar-Erg Tara-Acc see-Pfv.3MSg
      ‘Madhukar saw Tara.’
   b. Inanimate Proper Name:
      Madhukar-ne ‘Titanic’ dekh-ii hai
      Madhukar-Erg Titanic.f see-Pfv.F be.Prs
      ‘Madhukar has seen ‘Titanic’.

If the ko is omitted, Tara cannot be interpreted as an animate object, only as ‘the blob ‘Tara’.

2. With most other potentially referential objects, ko-marking is an option that correlates with a ‘specific’ reading (cf. Butt (1993), Singh (1994), among many others).

    (10) a. Rahul akhbaar-ko phaar raha hai
        Rahul.m newspaper-Acc tear Prog.MSg be.Prs.Sg
        ‘Rahul is tearing the newspaper.’
    b. Rahul akhbaar phaar raha hai
        Rahul.m newspaper tear Prog.MSg be.Prs.Sg
        ‘Rahul is tearing newspapers.’ / ‘Rahul is tearing the newspaper.’

The ‘Rahul is tearing newspapers’ reading can be given a plausible paraphrase as ‘Rahul is newspaper-tearing’. For this and other reasons, object incorporation of a non-standard sort has been proposed for them in Mohanan (1995b). There is also much work on this topic by Veneeta Dayal (cf. Dayal (1992), Dayal (1999), Dayal (2002a), Dayal (2002b)).

3. Nominals that are clearly non-referential cannot take ko.

   (11) Atul mehnat(-ko) kar raha hai
       Atul.m hardwork-Acc do Prog.MSg be.Prs.Sg
       ‘Atul is working hard.’

   • Agreement and Specificity
   • An issue of terminology: the ko that appears on some Direct Objects also appears on Indirect Objects and certain Experiencer Subjects. The unmarked option is also used by Subjects. For this reason, occasionally certain authors will refer to the unmarked case on the object as ‘Nominative’ and the ko-marked option as ‘Dative’.
   
   I will be using the following labeling strategy:
   1. If ko cannot appear: nominative
   2. If ko may (but not must) appear: accusative
   3. If ko must appear: dative

   This strategy is based on the source for case as opposed to the surface form of the case. Nominative is licensed higher in the tree, above vP. Accusative and Dative are licensed lower.

2.1.3 Non-nominative Subjects

In addition to Ergative subjects, the Indo-Aryan languages display a wide-range of constructions where what seems to be the subject receives a non-nominative case (cf. Mohanan (1995a)). The relevant argument has been called a subject because it meets a subset of subjecthood tests.

One of the concerns that we will be concerned with is the very notion of ‘subject’. Is a unitary notion of subject necessary/desirable? Does every sentence have to have a subject?

The most well-discussed of the non-nominative subjects is the Dative/Experiencer subject construction (cf. Verma and Mohanan (1990)).

   (12) a. Gajaanan-ko yeh tathya maaluum the
       Gajaanan-Dat this.f fact.M known be.Pst.Pl
       ‘Lit. to Gajaanan, these facts were known.’
   b. Naim-ko Rina pasand hai
       Naim-Dat Rina pleasing be.Prs
       ‘Naim likes Rina’. (Lit. To Naim, Rina is pleasing.)

   The others are:

   (13) a. Instrumental Subject:
       Ram-se per nahi: kat-aa
       Ram-Instr tree.m Neg cut-MFSg
       ‘Ram was not able to cut the tree.’ (Lit. By Ram, the tree did not cut.)
   b. Locative Subject:
       Ravi-par bahut bojh hai
       Ravi-on much burden.f be.Prs.Sg
       ‘Ravi has a big burden.’ (Lit. On Ravi is a big burden.)
2.2 Agreement

Many Indo-Aryan languages display object agreement and default agreement. One common pattern is the one displayed by Hindi. Most prominent non-overtly case-marked argument triggers agreement:

(14) a. Nominative subject, Accusative object, both non-overtly case-marked
   Rahul kitaab pāṛh-ṭaa thaa
   Rahul.M book.F read-Hab.MSG be.Pst.MSG
   ‘Rahul used to read (a/the) book.’

b. Ergative subject, Accusative object, only object is non-overtly case-marked
   Rahul-ne kitaab pāṛh-ii thii
   Rahul-Erg book.F read-Pfv.F be.Pst.FSg
   ‘Rahul had read the book.’

c. Ergative Subject, Overtly marked accusative object
   Rahul-ne kitaab-ko pāṛh-aa thaa
   Rahul-Erg book-Acc read-Pfv.MSg be.Pst.MSG
   ‘Rahul had read the book.’

Long Distance Agreement (LDA):

(15) Vivek-ne [kitaab pāṛh-nii] chaah-ii
    Vivek-Erg book.F read-Inf.f want-Pfv.f
    ‘Vivek wanted to read the book.’

LDA (as well as the Hindi-Urdu case system) is analyzed in Mahajan (1989), Butt (1995), and Bhatt (2003) among others. There is much variation with respect to the particulars of agreement in the Indo-Aryan languages and some of this is addressed in Subbarao (2001) and Deo and Sharma (2002).

2.3 Passives

Passives in Modern Indo-Aryan tend to be analytical and are composed of the following elements:

(i) Ablative or Locative form of the infinitive + the verb come (Marathi, Gujarati, Kashmiri)
(ii) Infinitive + the verb receive (Sinhalese)
(iii) Perfective Participle + the verb go (Punjabi, Hindi, Assamese, Bengali, Oriya)

Gujarati, Kumaoni, Nepali, Lahnda, Marwari, and Sindhi have a morphological passive.

2.3.1 Exceptions to Burzio’s Generalization

Passives in several Indo-Aryan languages present a potential counterexample for Burzio’s generalization. They seem to involve suppression of the external argument without promotion of an internal argument. (cf. Pandharipande (1982)).

(16) a. Active:
   Rashmi-ne Nupur-ko bazaar-mē dekh-aa
   Rashmi-Erg Nupur-Acc market-in see-Pfv
   ‘Rashmi saw Nupur in the market.’

b. Passive, without promotion:
   Nupur-ko (Rashmi-dwaaraa) bazaar-mē dekh-aa gayaa
   Nupur-Acc Rashmi-by market-in see-Pfv Pass-Pfv
   ‘Nupur was seen in the market by Rashmi.’

c. Passive, with promotion:
   Nupur (Rashmi-dwaaraa) bazaar-mē dekh-ii gayii
   Nupur Rashmi-by market-in see-Pfv.f Pass-Pfv.f
   ‘Nupur was seen in the market by Rashmi.’

Passives in the Modern Indo-Aryan languages are distinctive in that they can apply quite freely to (non-unaccusative) intransitives as well as transitives.

2.3.2 Inabilitative Passives

Passive constructions with the demoted external argument realized by a -se (instrumental) phrase behave like polarity items. For most speakers, they can only appear in affective environments. They have a special modal meaning indicating (in)ability.

(17) a. Vikram-se sīrf ek per kaat-aa gayaa
    Vikram-Instr only one tree cut-Pfv Pass.Pfv
    ‘Vikram could only cut one tree.’

b. Saira-se per ukhaar-e naḥī: jaa-te
    Saira-Instr tree.m uproot.Pfv.MPl Neg Pass-Hab.MPl
    ‘Saira is unable (to bring herself) to uproot trees.’

c. mujh-se Dilli naḥī: jaa-yaa gayaa
    I-Instr Delhi Neg go-Pfv Pass.Pfv
    ‘I couldn’t (bring myself to) go to Delhi.’

2.4 Causatives

The Indo-Aryan languages have a complex system of causative formation where we can distinguish at least three distinct processes.
2.5 ‘Intransitivization’

In this class of verbs, there is no overt causative affix. The phonological form of the intransitive is derived from the phonological form of the transitive via shortening.

(18) a. Jaayzaad bāt rahii hai.
    property divide PROG-FEM be-PRES
    ‘The property is dividing.’

b. Ram-ne jaayzad bāt dii.
    Ram-ERG property divide GIVE-PERF
    ‘Ram divided the property.’

(19) a. Madhu per kaṭ rahii hai
    Madhu.f tree.m cut Prog.F be.Prs.Sg
    ‘Madhu is cutting a/the tree(s).’

b. per kaṭ rahe hē
    tree.m cut, Prog.MPl be.Prs.Pl
    ‘The trees are cutting.’

These intransitives differ from passives in that they do not involve any agentivity in their semantics.

2.6 Direct Causativess

In this class, an intransitive with no overt affix is paired with a transitive showing the suffix -aa.

(20) a. Makan jal raha hai.
    house.M burn PROG.M be.Prs
    ‘The house is burning.’

b. Dakaitō-ne makaan jal-aa diyya.
    bandits-ERG house.M burn-CAUS GIVE-PERF
    ‘Dacoits burned the house.’

2.7 Indirect Causatives

In addition to these two types of derivation, which involve a lower or so-called ‘lexical’ causativization, there are causatives with the affix -vaa, which have an indirect causative interpretation.

'Dacoit', a term used in India for a robber belonging to an armed gang. The word is derived from the Hindustani dakaīt, and being current in Bengali got into the Indian penal code. By law, to constitute dacoity, there must be five or more in the gang committing the crime. (from http://21.1911encyclopedia.org/D/DA/DACOIT.htm)

(21) zamindaar-ne (dakaitō-se) makaan jal-vaa diyya.
    landlord-Erg bandits-Instr house.M burn-CAUS GIVE-PERF
    ‘The landlord had the house burned (by the dacoits).’

In some Indo-Aryan languages, the same exponent is used to mark direct and indirect causation. Whether we get direct or indirect causation depends upon the predicate the causative exponent appears on.

(22) (Kashmiri, from Hook and Koul (1984), pg. 102)
    a. do + Caus = Indirect Causation:
        su chu no:kra̱ a̱ thi kE:m kar-inav-a:n
        he is servant.Dat by work do-Caus-Impfv
        ‘He is having the work done by the servant.’

b. laugh + Caus = Direct Causation:
    mohnI chu asḻ am-as as-inav-a:n
    Mohan is Aslam-Dat laugh-Caus-Impfv
    ‘Mohan is making Aslam laugh.’

3 Selected ‘Higher in the tree’ Phenomena

3.1 Scrambling and wh-movement

All Indo-Aryan have scrambling. Scrambling in Hindi-Urdu has been analyzed in some detail in Mahajan (1990), Mahajan (1994), and Kidwai (2000). There seems to be some variation in the degree to which long scrambling (i.e. out of finite clauses) is deemed acceptable. Most Indo-Aryan languages seem to be wh-in-situ. (but see Bhattacharya and Simpson (2000) who argue that Bengali should be treated as involving overt wh-movement despite apparent wh-in-situ behavior).

(23) a. Yunus-ne kyaa parh-a:
    Yunus-ne what read-Pfv
    ‘What did Yunus read?’

b. Fronting is dispreferred:
    ??Kyaa Yunus-ne parh-a:
    what Yunus-Erg read-Pfv
    ‘What did Yunus read?’

c. (?)kis-ne Mona-se baat kii thii
    who-Erg Mona-with talk.f do-Pfv.F be.Pst.F
    ‘Who had talked to Mona?’
With the exception of Kashmiri, they have been all claimed to be \textit{wh}-in-situ.

(24) Kashmiri

a. \textit{yi kitaab kem' che} \textit{pArmlIts}
   this book \ who \ is \ read
   'Who has read this book?'

b. 'yi kitaab \textit{ kem' che} \textit{pArmlIts}
   this book \ who \ read
   'Who has read this books?'

c. \textit{kem' che yi kitaab pArmlIts}
   who \ this book \ read
   'Who has read this books?'

d. 'kem' \textit{yi kitaab che pArmlIts}
   who \ this book \ is \ read

The \textit{wh}-in-situ nature disappears once we consider extraction out of finite clauses. Then one of two strategies needs to be used:

(25) a. Long Movement:
   \textit{kis-ko Ram soch-taa hai [ki Sita t pasand kar-tii}
   who-Acc Ram.m think-Hab.MSg be.Prs.Sg that Sita.f like do-Hab.f
   hai] be.Prs.Sg
   'Who does Ram think that Sita likes?'

b. Scope Marking:
   \textit{Ram kyaa soch-taa hai [ki Sita kis-ko pasand kar-tii}
   Ram,m what think-Hab.MSg be.Prs.Sg that Sita who-Acc like do-Hab.f
   hai] be.Prs.Sg
   'What does Ram think who does Sita like?'

It is reported that question formation via long movement is unavailable in Kashmiri and Punjabi.

3.2 Correlatives

Correlative clauses are one of the most distinctive features of the Modern Indo-Aryan languages. (cf. Srivastav (1991), Dayal (1996)).

(26) a. [jo laarki kharii hai] [yo lambii hai]
   Rel girl.f standing.f be.Prs.Sg Dem tall.f be.Prs.Sg
   'The girl who is standing is tall.' (Lit. which girl is standing, she is tall.)

b. [jo CD sale-par hai] [muji vo CD chaahiye]
   Rel CD sale-on be.Prs.Sg me.Dat Dem CD want
   'I want the CD which is on sale.' (Lit. which CD is on sale, I want that CD.)

In addition to relativization, correlatives are also used to form conditionals, \textit{when}-clauses, \textit{until}-clauses, and comparatives.

(27) conditional

a. If he studies, he will pass.

b. [dzar tyane abhyas kelaa] [tar to pas hoil]
   if be-Erg studying do-Pst-3MSg then he pass be-Fut-3Sg
   'If he studies, then he will pass.' Marathi

(28) \textit{when}-clauses

a. When Harry met Sally, she was living in Montreal.

b. [jab Harry Sally-se mil-aa] [tab vo Montreal-me rah rahi thii]
   when Harry Sally-with met then she Montreal-in live Prog was
   'When Harry met Sally, she was living in Montreal.'

(29) \textit{until} clauses

a. I will stay here until John arrives.

b. [jab tak John nahii aa jaa-taa] [tab tak mE yahi: rahugaa]
   when till John Neg come Hab then till I here stay-will
   'I will stay here until John arrives.'
   (Literally: [Till when John hasn’t come], [I will stay here till \textit{then}])

(30) Comparatives

a. Michael Jordan has more scoring titles than Dennis Rodman has tattoos. \textit{(Chicago Tribune, 7/17/98, Kennedy (2000))}

b. [Rodman ke jitne tattoo hE] [Jordan ke-paas us-se jyaadaa Rodman Gen how-many tattoo are Jordan near that-than more khitaab hE]
   title are
   'Michael Jordan has more scoring titles than Dennis Rodman has tattoos.' (Literally: [How many tattoos Dennis Rodman has],
   [Michael Jordan has more scoring titles than \textit{that}])
3.3 Peculiar Extrapositions

In addition to the usual kinds of finite complement clause and relative clause extraposition, the Indo-Aryan languages permit systematic violations of the Right Root/Upward Bounded Constraint of Ross (1967). This constraint is illustrated in (31).

(31) a. [That the girl [who John likes] is tall] is obvious.
b. [That the girl is tall [who John likes]] is obvious.
c. *[That the girl is tall] is obvious [who John likes].

The fact that the Indo-Aryan (and Dravidian) languages violate this constraint was noted by Subbarao (1984).

(32) a. [un jhuuthō-ko [jo Ram-ne mujhe bataa-ye the] dohraa-naa] those lies-Acc Rel Ram-Erg me.Dat tell-Plv.MPI be.Pst.MPI repeat-Inf galat hai wrong be.Prs.Sg
   ‘[To repeat the lies that Ram had told me] is wrong.’
b. [un jhuuthō-ko t dohraa-naa] galat hai [jo Ram-ne mujhe those lies-Acc repeat-Inf wrong be.Prs.Sg Rel Ram-Erg me.Dat bataa-ye the].
tell-Plv.MPI be.Pst.MPI
   ‘[To repeat the lies that Ram had told me] is wrong.’
   (Lit. ‘*[To repeat the lies] is wrong [that John had told me].’)

3.4 Compound Verbs

Compound Verbs are not as much an Indo-Aryan feature as they are an areal feature of the South Asian sprachbund (cf. Masica (1976)). Given locutions like aa jaa (Lit. come go, actually: Come in!), they are also initially quite puzzling. Compound verbs are drawn from a small class of verbs such as jaa ‘go’, le ‘take’, de ‘give’, baith ‘sit’ and a few others.

(33) (from Hook (1979), pg. 63)

a. jaa ‘go’:
   ham steshan pahutch gaye we station reach GO-Plv.MPI
   ‘We got to the station.’
b. le ‘take’:
   mē kabab khaa lugaa I kabab eat TAKE-Fut.1MSg
   ‘I’ll eat up the kababs.’
c. de ‘give’:

   is-ne sabkuchh bataa di-yaa
   s/he-Erg everything tell GIVE-Pfv
   ‘S/he told all.’

When used in the compound verb construction, the above verbs do not contribute their lexical meaning. Instead the semantic contribution concerns aspect, manner, and for le ‘take’, modality. Complex verb construction behave like positive polarity items. They cannot co-occur with a surface negation, unless that negation is in some sense (that needs to be made precise) cancelled.

(34) (from Hook (1974), pg. 221)

a. lagaan ghataa di-yaa gayaa land-tax.m reduce GIVE-Pfv.MSg Pass-Pfv.MSg
   ‘The land tax was reduced.’
b. # lagaan ghataa nahī: di-yaa gayaa land-tax.m reduce Neg GIVE-Pfv.Ms Pass-Pfv.MSg
   c. ‘Double Negation’:
   koi vajah nahī: ki lagaan ghataa nahī: di-yaa jaa-e some reason Neg that land-tax.m reduce Neg GIVE-Pfv.Ms Pass-Sbjv.MSg
   ‘There is no reason that the land tax should not be reduced.’

References

Dayal, V. (2002b) “Number Marking and (Im)definiteness in Kind Terms,” manuscript, Rutgers, on website.