South Asian perspectives on relative-correlative constructions

Alexander R. Coupe
arcoupe@ntu.edu.sg
Nanyang Technological University Singapore

Abstract

This paper discusses relativization strategies used by South Asian languages, with the focus falling on relative-correlative constructions. The bi-clausal relative-correlative structure is believed to be native to Indo-Aryan languages and has been replicated by Austroasiatic, Dravidian and Tibeto-Burman languages of South Asia through contact and convergence, despite the non-Indic languages of the region already having a participial relativization strategy at their disposal. Various permutations of the relative-correlative construction are discussed and compared to the participial relativization strategies of South Asian languages, and functional reasons are proposed for its widespread diffusion and distribution.

1 Introduction

This paper surveys a sample of languages of South Asia that have bi-clausal structures closely resembling the relative-correlative clause construction native to Indic languages, which are thought to be the source of constructions having the same function and a very similar structure in many Austroasiatic, Tibeto-Burman and Dravidian languages. As the Indic languages employ a 'j-class' form of relative pronoun,[1] and many languages of South Asia are observed to either borrow such proforms or else press their own interrogative pronouns into the same function, it appears that relative-correlative constructions in non-Indic languages of the subcontinent have borrowed what might be identified as the South Asian Relative-Correlative Construction template as a result of language contact. Some prior work has been done on relative-correlative constructions in Indo-Aryan and Dravidian languages (e.g. Nadkarni 1975, Laksmi Bai 1985, Hock 1989, Hock 2016) but little if any investigation has been done on the diffusion of this construction into Tibeto-Burman languages, or for that matter Austroasiatic languages.

The paper is organized as follows. First, background information on the distinguishing typological features of the relative-correlative construction and the participial relative construction is presented in Section 2 and illustrated with data from a selection of old and new Indo-Aryan languages. In Section 3 our attention turns to bi-clausal constructions in Tibeto-Burman languages that express similar meanings and have very similar structures, albeit with slight modifications that account for the absence of a native relative pronoun word class in these languages. Section 4 discusses the related relativization strategies found in Dravidian and Munda languages, and presents arguments for why it is unlikely that relative-correlative constructions are indigenous to Dravidian. Finally, Section 5 concludes with a discussion of the findings

---

[1] These are so called because the relative pronouns of most Indic languages have a voiced postalveolar

and their relevance to our understanding of language contact and convergence in South Asian languages.
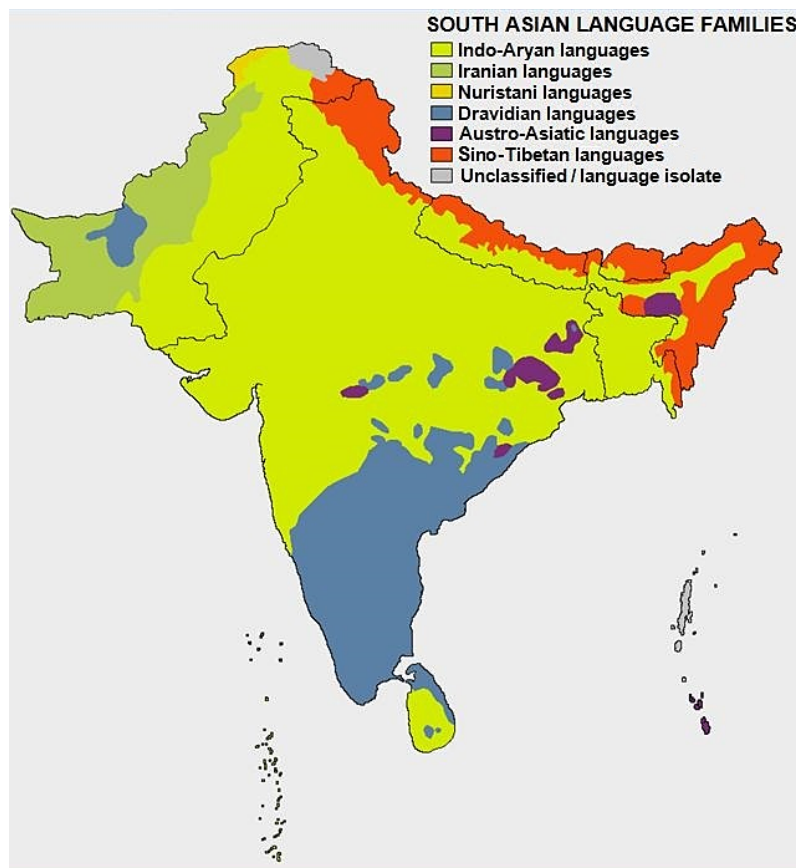
Map 1. Current distribution of South Asian languages[2]

Where possible, naturalistic textual data or data collected from natural conversations have been used to ensure the authenticity of the presented examples. This data is supplemented by examples taken from various grammars and research papers by other authors, for which no claims can be made.

2   Indo-Aryan relativization patterns

For the purposes of preliminary exemplification, the Nagamese sentences of (1a-b) compare the Indo-Aryan relative-correlative construction (hereafter RCC) with examples of the Indo-Aryan participial relative clause (hereafter PRC) in (2a-b). Supplementary data from Vedic Sanskrit and Hindi are presented to demonstrate additional characteristics of Indo-Aryan relativization strategies, and this data will collectively serve as a baseline for investigating similar structures in unrelated languages of South Asia.

---

[2] Translated from image: Südasien Sprachfamilien.png, from *Language families and branches, languages and dialects* in *A Historical Atlas of South Asia*, Oxford University Press. New York 1992. Nihali, Kusunda, and Tai-Kadai languages are not shown. Author - User: BishkekRocks Translated by User: Kitkatcrazy (Photo credit: Wikipedia)

A RCC is a type of complex sentence consisting of two finite clauses. The dependent relative clause contains a relative pronoun, and its main clause may or may not contain a correlative pronoun or noun that is coreferential with the relativized argument. The relative pronoun *jun* of Nagamese most commonly occurs at the beginning of the relative clause and is morphologically invariant, due to the creolizing nature of Nagamese and the general attrition of inflectional categories in this language, but relative pronouns are typically inflected for gender, number and case in most Indo-Aryan languages (e.g. cf. the cognate relative pronoun *yo* of Vedic Sanskrit in [5]). Uncharacteristically for a Indo-Aryan language, a conditional marker may be used to delimit the boundary of the dependent relative clause in some Nagamese RCCs. As proposed in Coupe (2007a), this innovation appears to have resulted from the bidirectional influence of Mongsen Ao, which has replicated the structure of the Indo-Aryan RCC but additionally uses a native Mongsen Ao topic particle at the end of the relative clause, ostensibly for a similar boundary marking purpose (see Coupe [2007b: 234–236, 435–36] for discussion and examples, and Section 3 below).

The relative clause constituent is henceforth identified by square brackets [...]. Relative pronouns and their correlative (pro)nouns are bolded (where present), and the overtly mentioned heads of relative clauses are underlined. It is often observed that RCCs are used primarily to modify indefinite arguments, but indefiniteness is not necessarily an obligatory requirement for their use, as these examples clearly show. Also, note that the relative clause constituent can be embedded in a causal adverbial clause, as in (1b), thereby adding an additional layer of structural complexity.

(1) Nagamese RCC (Indo-Aryan, Nagaland)[3]

   a. kintu   [**jun**   jaga-te   mol   thik   na-hoi-le]
       CONJ   RP   place-LOC   manure   correct   NEG-be-COND

       to   ami-khan   bishi   na-pa-i   na?
       thus   1SG-PL   much   NEG-get-LINK   Q
       (Speaking of rice yields) 'But in places that don't contain the correct manure, then we don't get much, right?'   (Bhattacharya 2001: 334)[4]

---

[3] Transcription of Indic and Dravidian data follows established South Asian conventions (e.g. <j> = [dʒ]). Transcription of Austroasiatic and Tibeto-Burman data is consistent with IPA conventions with the exception of <j> and <y> in Watter's (2002) cited data, which follows South Asian conventions. As the phonology of Nagamese is not codified and typically varies according to the respective L1 influences of each speaker, the Nagamese transcription used in this article has been standardized for data that was not personally recorded by the author.
[4] Glossing and translations of Bhattacharya's Nagamese data have been checked with a native speaker and adjusted as deemed necessary. For example, Bhattacharya interprets (1a) as a having conditional meaning, even though it contains a relative pronoun and relativizes on *jaga* 'place'. See below for discussion on the source of the conditional suffix.

b. hoilebi itu   <u>elektor</u>   <u>khan</u>   [**kun**   itu   elekshan   nimite   chuti
   CONJ   this elector   PL   **RP**   this election   BEN   holiday

   mangi-kena   ja-i-jo-a   para]
   ask-CVB   go-LINK-ECHO-NMLZ   INS

   itu   kam   kur-a   jaga   te   diktar   ho-bole   ase
   this   work   do-NMLZ   place LOC   difficult be-INF   be.PRS
   But because of these electors who have asked for and gone on holidays for this
   election, it is going to be difficult in this work place.
   http://nagamesekhobor.com/election-din-chuti/, accessed 8/7/2018

RCCs can also be headless. Such structures may employ an interrogative pronoun instead of the relative pronoun in Nagamese. Despite this, they still express the illocutionary force of declarative mood. The recruitment of an interrogative pronoun as a relative pronoun is common in the replicated RCC pattern found in Tibeto-Burman, Dravidian and Munda languages of South Asia (see Sections 3 and 4 below for further discussion and examples).

(2) Nagamese RCC (Indo-Aryan, Nagaland)

[isor   **kun**   ase]   apuni   no-jan-e
god   **who**   is   2SG   NEG-know-HAB.
'You don't realize who God is.'   (Bhattacharya 2001: 357)

Turning now to the structure of the Nagamese PRC, the relative clause constituent contains a verb stem nominalized by the participial suffix -*a*, plus the nominalized verb stem's notional argument(s), if any. The relative clause precedes the head it modifies if the reference is restricted to a subset of entities denoted by the head (see 3a). Alternatively, the relative clause may follow the head, in which case the reference is non-restrictive and the PRC is appositional in nature, merely adding further elaboration of the properties of the referent (see 3b).

(3) Nagamese PRC (Indo-Aryan, Nagaland)

a. pura   india-te   bishi   jaga,   mane   [eku   gas   hi
   whole   India-LOC   much   place,   that is any   tree   also

   ula-bo   na   pare   thak-i-a]   <u>jaga</u>   bi   ase…
   grow-IRR   NEG   can   remain-LINK-NMLZ   place   also   exist.PRES
   'All over India, there are a lot of places where no trees can grow…
                                                          (Bhattacharya 2001: 357)

b. to   itu   <u>mān</u>   [kori   thak-a]
   PTCL   DEM   worship   do   stay-PTCP
   'Thus, this worship that was being done …'   (Coupe 2007a: 353)

An observation of related relevance is that some Nagamese PRCs qualify as 'clausal noun-modifying constructions' (e.g. Matsumoto 1988, Matsumoto, Comrie & Sells 2017 and papers therein). Note that the head noun *famasi* in (4) below is not a notional

4

argument of its dependent clause and does not seem amenable to a gapped NP interpretation. Despite this, it is still the case that the nominalized clause *itu pani lo-a* serves to restrict the reference of its phrasal head.

(4) Nagamese PRC

[itu	pani	lo-a]	*famasi*	te	aro	ek	bar	pani
this	water	take-NMLZ	**pharmacy**	LOC	CONJ	one	time	water

lo-bo
take-IRR
'(We) will take water once again at the pharmacy where (we previously] bought water.'   (author's field notes)

In Sanskrit and the New Indo-Aryan languages, the structural order of relative clause to correlative clause is similarly exploited for encoding a restrictive versus non-restrictive meaning. Hock (2016: 570) proposes that the variable ordering of clauses in Indo-Aryan RCCs is a feature inherited from Proto-Indo-European. The following Vedic Sanskrit examples of (5) demonstrate the possibility of varying the position of the relative clause vis-à-vis the correlative clause.

(5) Vedic Sanskrit (**Hock 2016: 570**)

a. [**yo**   yatnena   kāryaṁ   karoti]
   RP.NOM.SG.M effort.INS.SG.M   work.ACC.SG.N   do.PRS.3SG

   sa   saṁsāre   puṣyati
   CP.NOM.SG.M   world.LOC.SG.M   thrive.PRS.3SG
   'Who works hard progresses in life.'

b. **sa**   saṁsāre   puṣyati
   CP.NOM.SG.M   world.LOC.SG.M   thrive.PRS.3SG

   [**yo**   yatnena   kāryaṁ   karoti]
   RP.NOM.SG.M effort.INS.SG.M   work.ACC.SG.N   do.PRS.3SG

As we saw in (3a-b) above, this variable ordering pattern is mirrored by Nagamese PRCs, and it is also attested in the PRCs of many Tibeto-Burman languages, resulting in the same kind of semantic contrast (e.g. see Coupe [2017: 218–221] for examples and discussion).

According to Hock (2016: 571), the post-nominal head type of RCC demonstrated by the Nagamese example of (1b) above is a recent "bookish" innovation in Hindi and is of marginal acceptability for some speakers, but it seems to be widespread in Nagamese, and possibly it is in Hindi as well. For example, Barz & Yadav (1993: 210–211) present Hindi RCC examples having a near-identical structure to the post-nominal RCC of Nagamese in their pedagogical grammar. Many speakers of Hindi and Nagamese are also familiar with English, which could have served as a model for the innovative structure of these RCCs in Indic languages.

(6) Hindi (Indo-Aryan) (Barz & Yadav 1993: 210-11)[5]

merī	bhai	[jis-kī	ummr	battīs	sāl	kī	hai]
1SG.POSS.M	brother	RP-GEN.F	AGE	thirty-two	year	GEN.F	be

vah	abhī	tak	nā-kār-ā	baiṭ-ā	hai
3sg	still	until	NEG-work-PFV.SG.M	sit-PFV.SG.M	be.SG

'My brother, who is thirty two, is still sitting around doing nothing.'

It appears that the more complex the sentence construction, the more likely it is for a speaker to resort to the RCC strategy for forming a relativization. This is especially likely to occur when an argument is in an oblique function and access to relativization using the PRC strategy would require a much greater expenditure of cognitive effort – this may be motivated by Zipf's (1949) principle of least effort. Or perhaps an argument may not be accessible to relativization at all using the PRC strategy due to language-specific constraints, in which case the RCC strategy becomes the only option. We will return to a discussion of motivations for the existence of two relativization patterns in the languages of South Asia in the final section.

3   Relativization patterns in Tibeto-Burman languages

Like the Indo-Aryan languages, many Tibeto-Burman (TB) languages of South Asia have RCCs and PRCs in their syntactic inventories, and they are respectively structurally similar to those found in the Indic languages. Firstly, the TB PRC is also a type of nominalization in which a head noun is modified by a nominalized verb stem, which may or may not include its notional arguments. The head may be external, internal or omitted. Overall, internally-headed PRCs are less commonly encountered in narrative texts of the South Asian TB languages surveyed.

To illustrate, the data of (7) demonstrates the structural possibilities of (a) post-headed, (b) pre-headed, (c) headless, and internally-headed PRCs in Chang, a TB language of central east Nagaland.

(7) Chang (Bodo-Konyak-Jinghpaw, Nagaland)

a. [tʃáŋ-éj   màŋ-á   lé-pʉ̀]   **hélé-èj-pʉ̀**   ɲên
   Chang-ERG   body-LOC   wear-NMLZ   **dress-INS-GEN**   name
   'The names of the adornments that the Chang wear on their bodies …'

b. hò   **káwtàk**   hò   pántów   [mìʃnəlì   á-sɤ́ɯ-pʉ̀]
   this   **world**   this   all   missionary   NEG-reach-NMLZ
   '... all this world that the missionaries didn't reach'

c. ŋèj[[ní   hì-pʉ̀]   tʃè   kù-lápʉ́.
   1SG:ERG   2SG:ERG   ask-NMLZ   TOP   give-IRR
   'I will give you what you ask for.'

---

[5] Interlinear glosses have been added to this example.

d.

[ŋèj	ŋ̂ə̂ɯ	làw-pów]	khò	ʃàtân	kí-Ø-kèj.
1SG:ERG	**word**	say-ANMLZ	that	incorrect	be-PST-DECL

'The word that I said was incorrect.'   (Coupe 2017: 219-220)

In common with the Nagamese data of Section 2, locating the head after or before the nominalized verb stem of the dependent clause determines whether the reference is restrictive or non-restrictive, and similar structural and semantic contrasts are reported in many other Tibeto-Burman languages of South Asia (e.g. Kham [Watters 2004], Mongsen Ao [Coupe 2007], Tibetan [Beyer 1992], and Yakka [Schackow 2015]; variable orders have also been attested in Sangtam, Lotha and Khiamniungan textual data [Coupe 2017]).

The structure of the RCC in TB languages that have borrowed the RCC pattern deviates slightly from the Indo-Aryan template, because these languages do not have an native relative pronoun word class. Languages replicating this structure may instead used interrogative pronouns, or they can resort to borrowing members of the j-class relative pronoun paradigm from an Indo-Aryan language in contact, as the Bodic language Kham does. According to Watters (2002: 165) Nepali is the source of at least one relative-correlative pair in this language (*jo* is a relative pronoun in Nepali) and, given the widespread bilingualism in Nepali, this *lingua franca* is likely to have served as a model for the creation of other RCCs in Kham and other TB languages of the Himalayas. Only three examples are provided in Watter's grammar. All are of the headless indefinite type of RCC in which the relative clause precedes the main clause, and the first word of the relative clause is a relative pronoun.

(8) Kham (Bodic, Nepal)

jo	nə-pə̃ĩ-zya,	**ho**	zə	ŋa-jã:
whatever	2SG-want-CONT	**that**	EMP	1sgA-give.2sgO

'Whatever you desire, that I will give you.   (Watters 2002: 166)

Breugel (2014: 172) provides a list of "indefinite proforms" in Atong, one of which is *je*, and expresses his suspicion that this may have an Indic origin. This is almost certainly the case, and it parallels the borrowing of j-class relative pronouns in the South Munda language Kharia (Peterson 2011: 408ff. – also see Section 4. Although Breugel does not recognize it as such, the following Atong example fits the Indic model of the RCC in employing a j-class relative pronoun to represent the relative clause argument, and this has presumably been borrowed from Assamese or Bangla along with the structure of the RCC. This has the characteristics of a headless RCC in which the relative clause precedes the main clause.

(9) Atong (Bodo-Garo, Meghalaya)

[**je**-səkən	naŋʔ=ci	gananŋ]	cən=ari=bo,	kamal=na
RP-QUANT	2SG=LOC	exist	offer=SIMP=IMP	priest=GOAL

'However much you have, just offer it to the priest.   (Breugel 2014: 174)

According to Joseph (2007: 486) the relative pronoun of the Bodo language Rabha is also borrowed from Assamese or Bangla, and he reports that the RCC is used

frequently. This may reflect the high degree of convergence resulting from contact with Indic languages in the Assam valley. Joseph states that the same relativized meaning can be conveyed with less emphasis if the relative pronoun is omitted, but note that (10c-d) actually use the non-finite PRC structure with a nominalized verb stem, whereas (10a-b) have bi-clausal finite structures that we can identify as the RCC. The nominalizing suffixes are *-ba* and *-kai*. These are used in (10c) to form a headless PRC, and a post-head PRC in (10d).

(10) Rabha (Joseph 2007: 486)

a. [**ja**-si   náh   to-eta]   **u**-si   ah-ba   toh-o
   RP-LOC   you   stay-CONT   **there**-LOC   I-also   stay-FUT
   'I too will stay where you (are) stay(ing).'

b. [**ja**   pan   aŋ   kái-nata]   **okai**   pan-be   thé-jo
   RP   tree   I   plant-PST   **that**   tree-DEF   bear.fruit-PST
   'The tree which I planted bore fruit.'

c. [náŋ   toŋ-ba-i]   aŋ-ba   toh-o
   you   stay-NMLZ-LOC   I-also   stay-FUT
   'Where you stay I too will stay.'

d [aŋ   kái-kai]   pan-be   thé-jo
   I   plant-attr   tree-DEF   bear fruit-PST
   'The tree I planted bore fruit.'

A more common response of the replicating language is to recruit its interrogative pronouns to serve as relative pronouns if it does not borrow these along with the RCC structure. Such a pattern is found in the Patsho dialect of Khiamniungan, a Konyak language of extreme eastern Nagaland. This is another language which can use its RCC for expressing restrictive reference.

My work on the grammar of this language is at a preliminary stage, and there are insufficient examples to make definitive claims. Nevertheless, it may be the case that the RCC structure is employed in Khiamniungan when a speaker needs to relativize on a genitive, as oblique arguments are generally much less accessible to relativization cross-linguistically (Keenan & Comrie 1977). This is demonstrated in (11), in which the head is a genitival noun. The sentence has the syntactic structure of a post-head relative clause, in common with the Nagamese example of (1b) and the Hindi example of (6), suggesting that this pattern may be widespread in South Asia. The structure also bears strong similarities to Dravidian replicants of the Indic RCC (discussed below in Section 4), as it employs an interrogative pronoun in the role of the Indic relative pronoun, and it marks the dependent clause with an interrogative particle. This is interesting for the fact that contact between Dravidian and Konyak languages of the Bodo-Konyak-Jinghpaw branch of Sino-Tibetan is highly unlikely, even in antiquity when Dravidian languages were assumed to have a much wider and more northerly distribution in the subcontinent.

(11) Patsho Khiamniungan (Konyak, Nagaland)

noŋ³³ni⁵⁵   [ʃaʊ⁵⁵   khiʊ¹¹u?¹¹   khɔ³³]
this.one   **who**   hair   Q

tʃə¹¹-miɛ³³ɲu³¹-a³³   thi-ɛ   tə¹¹   ɲu-n¹¹
1SG:POSS-wife-?   be-IRR   thus   say-PST
'"The one whose hair this is will be my wife", [he] said.'

Tshangla is a Bodic language of Bhutan with a type of complex sentence closely resembling the RCC of Khiamniungan. Andvik (2010: 267–71) describes a structure that he refers to as an 'embedded question complement with *mo*'. Despite the presence of the interrogative pronoun and interrogative particle, these dependent clauses do not express the illocutionary force of a question, in common with Patsho Khiamniungan.

(12) Tsangla (Bodic, Bhutan)

a. unyu   chas   khepa   [**hang**=ga   korgai   gila   mo]
   DEM   talk   TOP   **what**-LOC   about   COP   Q

   lok   yek-co
   return   speak-IMP
   'Repeat back what this talk is about.'   (Andvik 2010: 268)

b. Za   [chilu   apa-gi   **hang**   a-n-ca   mo]
   son   great   father-AGT   what   do-SE-COP   Q

   thamcen   se-le.
   all   know-INF
   'The oldest son knows everything that the father is doing.' (Andvik 2010: 268)

c. [Jang   **ibi**   ngan   phi-wa-la   **mo**]   lok   ngan   phi
   1SG   **who**   curse   do-NMLZ-COP   **Q**   return   curse   do.IMP
    'Whoever put a curse on me, put a curse on them as well!'
                                                          (Andvik 2010: 271)

The Mongsen dialect of Ao, spoken in west central Nagaland, follows the familiar pattern of using an interrogative pronoun as a relative pronoun in the RCC, but as previously mentioned, it deviates by using a topic particle at the end of the relative clause. This plausibly has developed through the two-way convergence of the Indic and the native TB structure. The topic marker *la* of this language carries a fairly heavy functional load and also appears as a grammaticalized formative in some converb suffixes (Coupe 2007b: Ch 11), in addition to marking topicalized constituents. The language appears to have adopted the RCC pattern, but retains the topic marking function of *la* to encode the dependency of the relative clause constituent. Frequency of the RCC in Mongsen Ao texts varies from speaker to speaker; some use it a lot in their narratives, while other speakers use it more sparingly.

(13) Mongsen Ao (Indo-Burmic, Nagaland)

a. pa   thak   ku   [**tʃə́pá?**   tʃhà-mì-ə̀ɹ   la]   tʃhà-aŋ."
   3SG   PLACE   LOC   **what**   do-DESID-PRS   TOP   do-IMP
   'To it [i.e. the lightning, he said] "Whatever you want to do, do it!"' [in order
   to placate it].   (Coupe 2007: 355)

b. [sə́pá? nə̄   ì   tʃə̄lāj   ā-tshə̄   phāŋā   tsə̄ŋ-ī?-ɹū   lā]
   who AGT   1PL.INCL   daughter   NRL-mithun   five   attach-CAU-IMM   TOP

   ājī   tʃə̄lāj   **pā**   tsə̄-ì-ù?   tè   sā-Ø
   1PL.POSS   daughter   **3SG**   take-IRR-DECL   thus   say-PST
   '"Whoever ties five mithuns (*Bos frontalis*) [as a bride price for] our daughter,
   he can take our daughter", [he] said.' (Coupe 2018: 2015)

To summarize thus far, it has been shown that the RCC is found in different
branches of a number of TB languages of South Asia that are in contact with Indo-
Aryan languages. Some of these not only borrow the RCC template, but also borrow
one or more j-class relative pronouns from Indo-Aryan; others instead use their
interrogative pronouns innovatively for this function in the relative clause constituent.
Some languages also use either a topic particle or an interrogative particle at the end of
the dependent relative clause. As we shall see in the next section, these patterns are also
found in other non-Indic languages of South Asia.

4   Relativization patterns in Munda and Dravidian

Relativization patterns in the Munda languages of South Asia have a great deal in
common with TB languages, and show even stronger evidence of contact with Indo-
Aryan and structural borrowing. In addition to the PRC structure, RCCs are also
attested.

Peterson (2011: 408–10) describes two different RCC templates in the South Munda
language Kharia. One of these uses a set of j-class relative pronouns, which are
borrowed from Indo-Aryan (see 14a). The other structure uses forms that are identical
to the interrogative pronouns (see 14b).

(14) Kharia (South Munda, eastern India)

a. …aɖi   [**je**   <u>bhere</u>   eŋ=ki]   **se**   <u>bhere</u>   aɖi=ya?
   ANAPH   **RP**   time   return=MID.PST **that**   time   ANAPH=GEN

   poʈom=te   soreŋ   kui=ki
   bundle-OBL   stone   find=MID.PST
   '…when he returned, he found a stone in the bundle (*which time he returned,*
   that time)   (Peterson 2011:409)

b. … [a = boʔ = te pujapaʈh   karay = na] aw = ki   **ho** boʔ = te
    Q=place=OBL sacrifice   do=INF   QUAL-MID.PST **that** place=OBL

ɖɑm = ke   ho = ki   ho   ɖoli = te   mɑɽɑy = oʔ = mɑy
arrive = SEQ that = PL that palanquin = OBL   put.down = ACT.PST = 3PL
'… having arrived at the place where the sacrifice was to be done, they put the palanquin down.'   (Peterson 2011:409)

The Kharia RCC can also be headless. Both the je-class and Q-class interrogative types of RCC are used to relativize on the same kinds of oblique arguments – instrument, locative, temporal, genitive, and comparative (Peterson 2011:410). This versatility may explain the adoption of the RCC structure in the language.

Steever (1998: 35) notes that Dravidian languages lack relative pronouns, but employ a RCC structure by using interrogative pronouns as relatives paired with distal deictic pronouns. He proposes that that the RCC can be reconstructed to Proto-Dravidian and cites the use of clitics at the end of the relative clause in daughter languages as evidence of the antiquity of the construction. The following Konda example is illustrative.

(15) Konda (Dravidian)

[embe   nī   iṣṭam   kinid] = o   bān   idʔa
 where   you   desire   do-FUT-2SG = or there   put-IMP
Put it wherever you want to.'   (Steever 1998:35)

Similarly, Laksmi Bai (1985) argues that the direction of borrowing must be from Dravidian to Indo-Aryan. She presents of variety of evidence, but much of this is based on translations of English sentences into Tamil, and the data is just as likely to reflect a species of translationese, rather than Tamil. My own experience of attempting to elicit relative clauses in Mongsen Ao using Dahl's (1980) TMA questionnaire resulted in a perplexing preponderance of RCC structures, even though these were fairly rarely encountered in texts, and the PRC strategy was clearly the dominant pattern in naturalistic textual data. It turns out that the speaker, who spoke English fluently, was attempting to accommodate me by giving structurally similar sentences to the English relative clauses. Laksmi Bai also underestimates the impact of convergence and model replication in the subcontinent. It is likely that there has been a millennium and a half of language contact between Indo-Aryan and Dravidian before the earliest attested examples of RCCs in old Tamil sources (early centuries AD), which is more than enough time for such a structure to diffuse, especially if it fills a functional gap.

5   Conclusion

This paper has investigated relativization strategies in Indo-Aryan, Tibeto-Burman, Munda and Dravidian languages of South Asia, focusing on the RCC structure. It finds that the RCC pattern is widespread and has diffused into TB languages as a result of replication of the Indo-Aryan pattern. While it is common in South Asia, it does not appear to be attested in TB languages that are not in contact with Indo-Aryan languages. The likely functional motivation for its spread into Munda, Dravidian and TB languages

can be explained by accessibility to relativization. Speakers will resort to the RCC strategy when when language-specific constraints prevent access to relativization using the PRC strategy.

References

Andik, Erik. 2010. *A grammar of Tshangla*. Leiden: Brill.

Barz, Richard &Yogendra Yadav. 1993. *An introduction to Hindi and Urdu*, 5th ed. New Delhi: Munshiram Manoharlal.

Bhattacharya, Dwijen. 2001. The genesis and development of Nagamese: Its social history and linguistic structure. Unpublished PhD dissertation, City University of New York.

Breugel, Seino van. 2014. *A grammar of Atong*. Leiden: Brill.

Beyer, Stephan V. 1992. *The classical Tibetan language*. Albany: State University of New York Press.

Coupe, Alexander R. 2007a. Converging patterns of clause linkage in Nagaland. In Miestamo, Matti and Wälchli, Bernhard (eds), *New challenges in typology: Broadening the horizons and redefining the foundations.* Trends in Linguistics, Studies and Monographs 189, 339–361. Berlin and New York: Mouton de Gruyter.

Coupe, Alexander R. 2007b. *A grammar of Mongsen Ao.* Berlin: Mouton de Gruyter.

Coupe, Alexander R. In press. Grammaticalization processes in South Asian languages. In Narrog, Heiko & Bernd Heine (eds). *Grammaticalization from a typological perspective*, 189–218. Oxford: Oxford University Press.

Emeneau, Murray B. 1956. India as a linguistic area. *Language* 32(1): 3–16.

Hock, Hans Heinrich 1989. Conjoined we stand: Theoretical implications of Sanskrit relative structures. *Studies in the Linguistic Sciences* 19(1): 93–126.

Hock, Hans Heinrich. 2016. Finite and non-finite subordination. In Bashir, Elena & Hans Heinrich Hock (eds.) *The languages and linguistics of South Asia: A comprehensive guide*. Berlin/Boston: De Gruyter Mouton.

Keenan, Edward L. & Bernard Comrie. 1977. Noun phrase accessibility and universal grammar. *Linguistic Inquiry* 8(1): 63-99.

Koul, Omar N. & Kashi Wali. 2015. *Kashmiri*. Muenchen: LINCOM.

Laksmi Bai, B. 1985. Some notes on correlative constructions in Dravidian. *Oceanic Linguistics Special Publications* (For Gordon H. Fairbanks) 20: 181–190.

Masica, Colin P. 1976. *Defining a linguistic area*. Chicago, London: University of Chicago Press.

Matsumoto, Yoshiko. 1988. Semantics and pragmatics of noun-modifying constructions in Japanese.

Matsumoto, Yoshiko, Bernard Comrie and Peter Sells (eds.) 2017. *Noun-modifying clause constructions in the languages of Eurasia: Rethinking theoretical and geographical boundaries*. Amsterdam: John Benjamins.

Nadkarni, Mangesh V. 1975. Bilingualism and syntactic change in Konkani. *Language* 51 (3): 672–683

Peterson, John. 2011. *A grammar of Kharia*. Leiden: Brill.

Steever. Stanford B. 1998. Introduction to the Dravidian languages. In Steever (ed.) *The Dravidian languages*. London: Routledge.

Watters, David E. 2002. *A grammar of Kham*. Cambridge: Cambridge University Press.

Zipf, George Kingsley. 1949. *Human behavior and the principle of least effort*. Cambridge, Mass.: Addison-Wesley Press.