URDU CORRELATIVES: THEORETICAL AND IMPLEMENTATIONAL ISSUES

Miriam Butt, Tracy Holloway King and Sebastian Roth
Universität Konstanz, PARC, Universität Konstanz

Proceedings of the LFG07 Conference

Miriam Butt and Tracy Holloway King (Editors)

2007

CSLI Publications

http://csli-publications.stanford.edu/

Abstract

The inclusion of South Asian languages in multilingual grammar development projects that were initially based on European languages has resulted in a number of interesting extensions to those projects. Butt and King (2002) report on the inclusion of Urdu in the Parallel Grammar Project (ParGram; Butt et al. (1999, 2002)) with respect to case and complex predicates. In this paper, we focus on a possible integration of *correlatives* into the computational analysis. Hindi/Urdu correlative clauses have received various analyses in the past that treat them as distinct from other strategies of relativization. We follow Bhatt (1997), who argues that the syntax and semantics of single-headed correlative clauses strongly resemble those of free relative clauses in European languages, but we analyze these as specifiers of a DP, rather than as adjuncts.

1 Introduction

This paper aims at introducing the discussion of so-called *correlative* constructions, a special strategy of relativization commonly found in a number of Indo-European, especially Indo-Aryan, languages, into LFG analyses. In particular, we look at correlatives from within the context of creating broad-coverage grammars as part of the Parallel Grammar project (ParGram; Butt et al. 1999, 2002). Among the aims of the ParGram project is to test the LFG formalism for its universality and coverage limitations and to see how far crosslinguistic parallelism at f-structure can be maintained. Where possible, the analyses produced by the grammars for similar sentences in each language are parallel. The standardization of the analyses has the computational advantage that the grammars can be used in similar applications and it can simplify cross-language applications such as machine translation. Parallelism, however, is not maintained at the cost of misrepresenting the language. Given this context, the phenomenon of correlatives is particularly interesting as it is a puzzling construction from the perspective of most European languages.

The pattern in correlatives is that a demonstrative pronoun, which also functions as determiner in Urdu/Hindi[1], in this case *vo*, always occurs in correlation with a relative pronoun, *jo*. In fact, the language employs a series of such pronouns: e.g., *jɪs/ʊs* 'which/that' (oblique), *jahã/vahã* 'where/there' (distal), *jɪdər/ɪdər* 'where/-there' (proximal).

We base our analysis in large part on Srivastav (1991), who argues convincingly that correlative constructions in Hindi fall into two classes: one in which the relative *jo* clause appears to the right of the *vo* head noun ((1a,b)) and one in which the *jo* clause precedes the *vo* noun ((1c)). Srivastav, whose analysis is primarily semantic, identifies the former as straightforward relative clauses, the latter as true correlatives.

[1]Urdu and Hindi are structurally essentially identical. For the sake of brevity, we only refer to Urdu, but all observations in this paper apply to Hindi as well.

(1) a. [**vo** laṛki [**jo** kʰaṛ-i hɛ]
that girl.F.Sg.Nom which stand-Perf.F.Sg be.Pres.3.Sg

hɑs rɑh-i hɛ]
laugh stay-Perf.F.Sg be.Pres.3.Sg
'Who is standing, that girl is laughing.' (Srivastav 1991:642)

b. [**vo** laṛki hɑs rɑh-i hɛ
that girl.F.Sg.Nom laugh stay-Perf.F.Sg be.Pres.3.Sg

[**jo** kʰaṛ-i hɛ]]
which stand-Perf.F.Sg be.Pres.3.Sg
'That girl is laughing, who is standing.' (Srivastav 1991:642)

c. [**jo** kʰaṛ-i hɛ]
which stand-Perf.F.Sg be.Pres.3.Sg

[**vo** laṛki hɑs rɑh-i hɛ]
that girl.F.Sg.Nom laugh stay-Perf.F.Sg be.Pres.3.Sg
'Who is standing, that girl is laughing.' (Srivastav 1991:642)

With respect to true correlatives as in (1c), no standard LFG analysis exists to date. Here, we depart from Srivastav's analysis and instead follow Bhatt (1997), who argues that Hindi correlatives must be understood as the equivalent of free relative clauses in European languages. Unlike Bhatt, however, we do not treat the correlative as an adjunct, but as a specifier of DP. The relevant evidence comes from the interaction with quantifiers and demonstratives, topicalization and the behavior of multi-head correlatives. Our analysis therefore builds on existing argumentation from a primarily semantic perspective (Srivastav 1991) and from within Minimalism (Bhatt 1997), but ultimately differs in the syntactic treatment of correlatives.

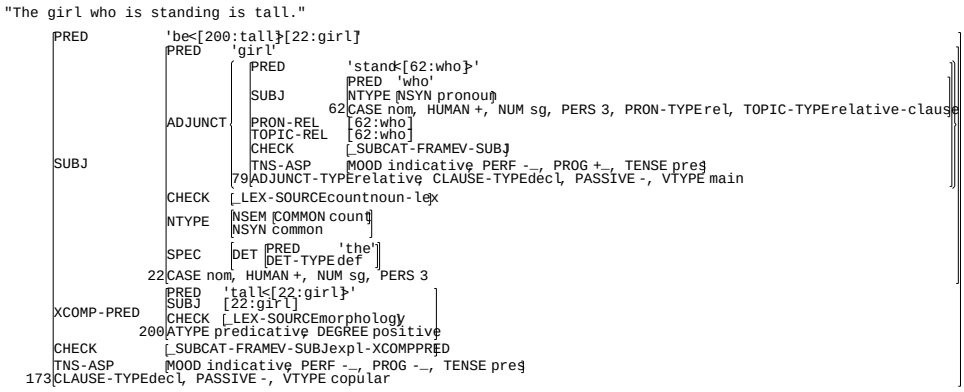
2 Standard Analyses of Relative Clauses

Linguistic typology (e.g., Lehmann 1984) generally distinguishes three classes of relative clauses: free and bound relative clauses, with the latter divided into restrictive and non-restrictive relative clauses. Bound relative clauses appear either adjacent to the phrase that they modify or extraposed at the end of the sentence. Within the ParGram project, bound relative clauses are analyzed as NP-modifying adjuncts of the c-structure category *CPrel* (Butt et al. 1999): The lexical requirements of the embedded finite verb must be fulfilled, meaning that arguments corresponding to the verb's subcategorization frame must be provided. A sample ParGram f-structure analysis of an English simple (non-extraposed) bound relative is shown in (2).

The relative pronoun, which must be an argument of the relative clause's predicate or the argument of a prepositional adjunct modifying the relative clause's

predicate, is encoded as the TOPIC-REL of the relative clause if it appears in preposed (topicalized) position, such as in English, German or French relatives. The functional structure projected by the relative clause is encoded as an adjunct (with 'ADJUNCT-TYPE relative') of the relative head.[2] Extraposed bound relative clauses are adjoined at f-structure to a single NP via functional uncertainty.

(2)



Non-restrictive relative clauses, such as (3) receive essentially the same structural analysis as restrictives, but are marked with an additional feature 'RESTR −' in order to flag them for a different semantic interpretation.

(3) Mary, who is standing there, is tall.

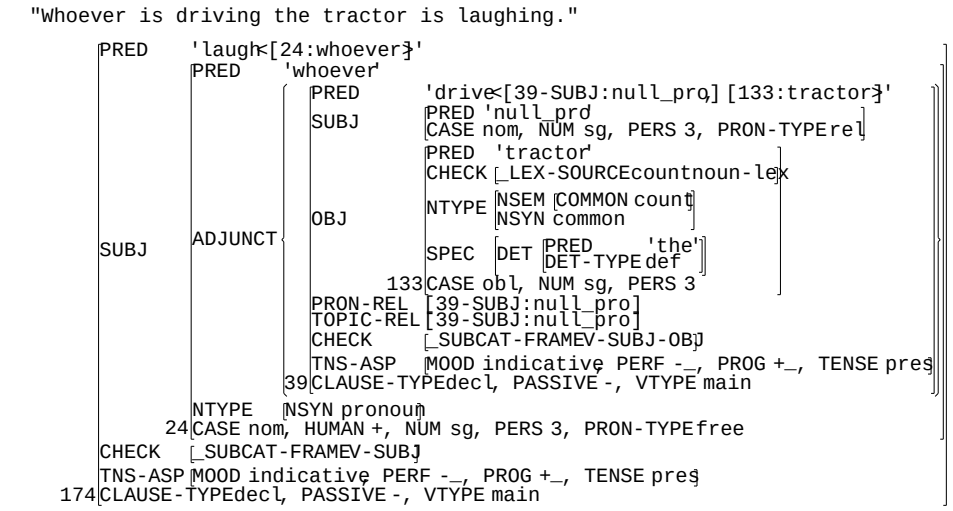
Since non-restrictive relative clauses, like appositives, perform a different illocutionary act than the proposition signified by the matrix clause, they do not have any truth-conditional value regarding the interpretation of their relative head, whereas restrictive relative clauses and their heads form a semantic unit via set intersection. Note that it is not always trivial to classify a relative clause as restrictive or not. In English, non-restrictive relative clauses are often marked by distinct punctuation, changes in intonation or special lexical items (e.g., *incidentally* or *by the way*).

Within ParGram, free relatives are analyzed quite differently from bound relative clauses. In English, free relatives have the distribution of an NP and are thus treated as such, whereas in German they cannot, like other finite clauses, appear clause internally and are thus treated as a special category *CPfreerel*. The relative pronoun in both languages takes the double function of relative clause head (i.e. the relative clause predicate is attached at f-structure as an ADJUNCT) as well as that of an argument of the matrix clause predicate. The existence of an empty argument of the matrix verb is deduced at f-structure from information provided either by the c-structure construction, as in German, or by the lexical entry of the

[2]The term 'relative head' refers to the noun phrase that is modified by the relative clause. We separate externally-headed clauses from internally-headed clauses. Both external and internal heads are denoted as 'relative heads' in our paper.

free relative pronoun, such as English *whoever* (Butt et al. 1999:96).[3] This allows the relative pronoun to take the grammatical function of the missing matrix argument and project the feature 'PRON-TYPE free' to its f-structure. (4) provides an example, again for English.

(4)



Note that free relatives in English as well as German are semantically ambiguous between singular definite and generic readings. The f-structure encoding of the free relative does not handle this semantic ambiguity although the f-structure provides all necessary information for further semantic processing: for example, tense/aspect information that disambiguates the reference. A generic reading is unavailable if the verb does not license it, which would be the case with progressive aspect, as for example in (5), in which the *-ever* does not lead to a free choice generic reading, but simply implies the uncertainty or irrelevance of the subject's identity. It is assumed that such constraints are handled within a separate semantic projection. A semantic consideration, however, that does have syntactic import is that free relatives do not allow stacking of further restricting or appositive relative clauses, as shown in (6).

(5) Whoever is driving the tractor is laughing.

(6) *Whoever drives the tractor, who is happy, is laughing.

Furthermore, no non-restrictive interpretation of the free relative clause itself is possible. These restrictions are significant within the context of our paper as they apply to Urdu correlative clauses as well (section 4.1).

[3]Note that not only *-ever* pronouns can function as free relative pronouns in English: (i) is also an example for a typical free relative.

i. I eat what you eat.

3 Relativization in Urdu

Urdu, the national language of Pakistan and an official language of India, is an Indo-Aryan language spoken by around 60-80 million native speakers today. It is an SOV-language with relatively free word order, a split-ergative case system and correlative clauses. Since its grammar is largely identical to that of Hindi and large portions of vocabulary are shared, Hindi-Urdu is commonly regarded by linguists as a single language, in contrast to their constitutional status.

Urdu, like Sanskrit, preserves the old Indo-European distinction between relative (Urdu j-class), interrogative (k-class), proximal demonstrative (y-class) and distal demonstrative (v-class) pronouns. It furthermore retains a remnant of the correlative clausal structure that, in Sanskrit, was used to express all kinds of clausal relations, such as relatives, conditionals or sentential complementation in a paratactic manner. Although it has seen some modification and appears in a more constrained distribution than the Sanskrit correlative (for a comparison, see Davison 2006), the Urdu correlative nevertheless retains some of the properties that separate it from the English-type postposed relative clauses, which also exist in Urdu.

Modern Urdu left-adjoined relatives as in (1c), repeated in (7), are generally called correlative clauses after their Sanskrit ancestors and are found in a number of Indo-Aryan languages, such as Bengali, Sindhi, Punjabi, Marathi, Gujarati and Urdu, but also in Hittite, Latin, Ancient Greek, Medieval Russian, and Old English as well as modern Hungarian, Bulgarian and Serbo-Croatian (Bhatt 2003).

(7) [**jo** kʰɑɽ-i hɛ]
which stand-Perf.F.Sg be.Pres.3.Sg

[**vo** lɑɽki hɑs rɑh-i hɛ]
that girl.F.Sg.Nom laugh stay-Perf.F.Sg be.Pres.3.Sg
'Who is standing, that girl is laughing.' (Srivastav 1991:642)

As already mentioned in the introduction, whereas the Urdu embedded and right-adjoined relatives manifest typical properties of restrictive relative clauses and are covered by the same analyses proposed for these structures in the other ParGram languages, previous analyses provide evidence that left-adjoined relatives form a distinct class, rather than being an instance of left-extraposition of an NP modifier. In the generative literature, there has been much discussion of whether so-called preposed, embedded and postposed relative clauses derive from the same underlying structure (e.g., Subbarao 1984) or are base-generated in their respective positions (McCawley 2004). We follow Srivastav (1991) in considering embedded and postposed relative clauses of the *vo-jo* pattern as being structurally and functionally identical, whereas the *jo-vo* correlative pattern in (7) is analyzed as a different construction, based on the following evidence: 1) felicity of internal heads; 2) requirement of an overt demonstrative/quantifier; 3) compatibility with the inclusive focus particle *bʰi*; 4) strictly non-restrictive interpretation; 5) impossibility of relative clause stacking; 6) multiple relativization.

We go through some of the relevant data in the next sections and then, in section 4 proceed to analyze correlatives as in (7) on a par with free relatives (following Bhatt 2003) that appear to be situated in SPECDP (contra Bhatt 2003).[4]

3.1 Structural Differences — Headedness

Embedded and extraposed restrictive relative clauses modify an external head, which means that the head NP is not allowed to appear in the relative clause itself. This is demonstrated in (8) and (9), where the relevant NP head(s) are underlined.[5] In contrast, correlative clauses may realize the full head NP in either clause, neither clause, or both clauses. This is shown in (10).

Normal Relative Clause

(8) a. **vo** laṛki [**jo** kʰaṛ-i hɛ]
that girl.F.Sg.Nom which stand-Perf.F.Sg be.Pres.3.Sg
lɑmbi hɛ
tall.F.Sg be.Pres.3.Sg

b. ***vo** [**jo** laṛki kʰaṛ-i hɛ]
that which girl.F.Sg.Nom stand-Perf.F.Sg be.Pres.3.Sg
lɑmbi hɛ
tall.F.Sg be.Pres.3.Sg

c. ***vo** laṛki [**jo** laṛki kʰaṛ-i hɛ]
that girl.F.Sg.Nom which girl.F.Sg.Nom stand-Perf.F.Sg be.Pres.3.Sg
lɑmbi hɛ
tall.F.Sg be.Pres.3.Sg
'The girl who is standing is tall.'

Extraposed Relative Clause

(9) a. **vo** laṛki lɑmbi hɛ [**jo** kʰaṛ-i hɛ]
that girl.F.Sg.Nom tall.F.Sg be.Pres.3.Sg which stand-Perf.F.Sg be.Pres.3.Sg

[4] Most of the data discussed in this paper is based on the previous discussions of Hindi correlative structures found in Srivastav (1991), Dayal (1996) and Bhatt (2003). Additionally, we have checked the data with three Pakistani doctoral students at Konstanz, all native speakers of Urdu.

[5] Mahajan (2000:9) does not agree with Srivastav's judgment with respect to the relative clause in (8c) and considers it grammatical, a view shared by our informants. However, this does not make Srivastav's generalization, nor our argumentation here, invalid. Srivastav's analysis deals with restrictive relative clauses, but the sentence-initial *vo* in (8c) has a clear deictic interpretation: Its reference is fixed and further intersective import (e.g., by restrictive relative clauses) is not admissible. Thus the relative clause in this case must have non-restrictive meaning if the sentence is to be grammatical. The fact that non-restrictive relative clauses can be internally headed is confirmed by McCawley (2004).

b. *__vo__ lʌmbi hɛ [__jo__ laṛki kʰaṛ-i hɛ]
that tall.F.Sg be.Pres.3.Sg which girl.F.Sg.Nom stand-Perf.F.Sg be.Pres.3.Sg

c. *__vo__ laṛki lʌmbi hɛ
that girl.F.Sg.Nom tall.F.Sg be.Pres.3.Sg

[__jo__ laṛki kʰaṛ-i hɛ]
which girl.F.Sg.Nom stand-Perf.F.Sg be.Pres.3.Sg
'The girl is tall, who is standing.'

Correlative Clause

(10) a. [__jo__ kʰaṛ-i hɛ] __vo__ laṛki lʌmbi hɛ
which stand-Perf.F.Sg be.Pres.3.Sg that girl.F.Sg.Nom tall.F.Sg be.Pres.3.Sg

b. [__jo__ laṛki kʰaṛ-i hɛ] __vo__ lʌmbi hɛ
which girl.F.Sg.Nom stand-Perf.F.Sg be.Pres.3.Sg that tall.F.Sg be.Pres.3.Sg

c. [__jo__ laṛki kʰaṛ-i hɛ]
which girl.F.Sg.Nom stand-Perf.F.Sg be.Pres.3.Sg

__vo__ laṛki lʌmbi hɛ
that girl.F.Sg.Nom tall.F.Sg be.Pres.3.Sg
'Which girl is standing, that girl is tall.'

With respect to headedness, the *jo-vo* (correlative) and *vo-jo* (relative) patterns thus differ quite markedly.

3.2 Correlative as an Operator — The Demonstrative Requirement

Coming from a primarily semantic perspective, Dayal (1996:181) analyzes the correlative *jo*-clause as an operator that locally binds a variable in the main clause. The variable must be contained in the interpretation of the determiner of the external head NP. Her reasons for this analysis build on Subbarao's (1984:13) initial observation that the relative clause cannot be adjoined to the left if the main clause NP is indefinite, as shown in (11a). In fact, Dayal shows that correlatives have to observe a more stringent requirement. Given that in Urdu bare NPs can in principle always also function as definites (Dayal 2003), Dayal formulates a 'demonstrative requirement' (Srivastav 1991:649): the matrix clause must contain a demonstrative. This demonstrative can either be overt as in (11b), or can be analyzed as being there implicitly in the presence of quantifiers such as *sab* 'all' ((11c)), *dono* 'both' ((11d)) or *tino* 'all three'.

(11) a. *[__jo__ laṛkiyã kʰaṛ-i hɛ̃] laṛkiyã
which girl.F.Pl.Nom stand-Perf.F be.Pres.3.Pl girl.F.Pl.Nom

lʌmbi hɛ̃
tall.F.Sg be.Pres.3.Pl
'Girls that are standing are tall.'

b. [**jo** laṛkiyã kʰaṛ-i hɛ̃] **vo** laṛkiyã
which girl.F.Pl.Nom stand-Perf.F be.Pres.3.Pl those girl.F.Pl.Nom

lɑmbi hɛ̃
tall.F.Sg be.Pres.3.Pl
'The girls that are standing are tall.'

c. [**jo** laṛkiyã kʰaṛ-i hɛ̃] **sɑb** laṛkiyã
which girl.F.Pl.Nom stand-Perf.F be.Pres.3.Pl all girl.F.Pl.Nom

lɑmbi hɛ̃
tall.F.Sg be.Pres.3.Pl
'All (the) girls that are standing are tall.'

d. [**jo** laṛkiyã kʰaṛ-i hɛ̃] **dono** laṛkiyã
which girl.F.Pl.Nom stand-Perf.F be.Pres.3.Pl both girl.F.Pl.Nom

lɑmbi hɛ̃
tall.F.Sg be.Pres.3.Pl
'Both (the) girls that are standing are tall.'

Srivastav (1991) and Dayal (1996) thus analyze correlative clauses as generalized quantifiers (Cooper 1983) that bind a position inside an IP. Syntactically, she posits the structure in (12) for correlatives as in (11).

(12) $[[\ldots\text{REL-XP}\ldots]_{\text{CP}}\ [\ldots\text{DEM-XP}\ldots]_{\text{IP}}]_{\text{CP}}$

4 Correlatives as Free Relative Specifiers of DP

As shown above, Dayal's (1996) seminal work on relatives and correlatives provides a clear basis for analysis. However, Bhatt (1997, 2003) looks at additional empirical evidence and argues that the *jo-vo* correlatives are better understood as being like free relatives. We discuss his reasons briefly in section 4.1, and adopt his view of correlatives as free relatives, but in a slightly different manner. We present an alternative analysis in section 4.2 by which we analyze the *jo-vo* correlatives as being situated in SPECDP, rather than being adjoined to IP (Srivastav 1991) or to the demonstrative phrase (Bhatt 2003).

4.1 DP Adjunction and Free Relatives

Dayal's analysis of the *jo-vo* correlative as a quantifier within a CP that adjoins to and binds a position inside an IP is challenged by several facts. For example, consider (13), where the correlative clause can appear directly to the left of the external head inside the main clause. This construction is not uncommon, and indicates that an analysis of direct adjunction to DP should be considered (Bhatt 1997, 2003).

(13) hɑsan=ne [**jo** kɪtab tara=ne lɪkʰ-i]
Hassan=Erg which book.F.Sg.Nom Tara=Erg write-Perf.F.Sg

vo pɑsɑnd k-i
that liking do-Perf.F.Sg
'Hassan liked the book which Tara wrote.'

Further evidence for DP-adjunction of the correlative clause comes from question-answer pairs as in (14), which show that the *jo-vo* clause in combination with the required demonstrative makes a perfect short answer to a question, just like a simple DP/NP would.

(14) kɔ̃ a-yi?
who come-Perf.F.Sg?
'Who came?'

[**jo** laṛki vɑhã rɑh-ti hɛ] **vo**
Which girl.F.Sg.Nom there stay-Impf.F.Sg be.Pres.3.Sg that
'Which girl lives there, she'/'The girl who lives there' (Dayal 1996)

Indeed, Wali (1982) already used this fact to argue for DP-adjunction. But Dayal rejects the DP-adjunction analysis in favor of a unified treatment of single correlatives and those with multiple heads (see section 5), by which IP-adjunction is treated as the basic phenomenon and DP-adjunction is analyzed as a case of crosscategorial quantification (Dayal 1996:206).

In contrast, Bhatt (1997) situates the correlative clause primarily within the DP (but see section 4.3 on topicalization facts) and, in particular, as having the functional properties of a free relative clause. As clearly demonstrated by Bhatt (1997), the properties of correlatives and free relatives are strikingly similar. For example, only free relatives (as opposed to restrictives) in English can feature the inclusive focus item *-ever*, a property which we also find in Urdu, where the focus particle *bʰi* 'also' cannot modify a restrictive relative clause ((15a)), but is admissible in correlatives in order to bring out the unspecified identity of the internal head, as shown in (15b).

(15) a. *vo laṛki [jo **bʰi** vɑhã kʰɑṛ-i hɛ]
that girl.F.Sg.Nom which also there stand-Perf.F.Sg be.Pres.3.Pl

nadya=ki sɑheli hɛ
Nadya=Gen.F.Sg friend.F.Sg be.Pres.3.Sg
'That girl, whichever is standing there, is Nadya's friend.'

b. [jo **bʰi** laṛki vɑhã kʰɑṛ-i hɛ]
which also girl.F.Sg.Nom there stand-Perf.F.Sg be.Pres.3.Pl

vo nadya=ki sɑheli hɛ
that Nadya=Gen.F.Sg friend.F.Sg be.Pres.3.Sg
'Whichever girl is standing there is Nadya's friend.'

c. jo **bʰi** laṛki mehnɑt kɑr-ti hɛ
which also girl.F.Sg.Nom effort do-Impf.F.Sg be.Pres.3.Sg

vo safal ho-ti hɛ
that successful be-Impf.F.Sg be.Pres.3.Sg
'Whichever girl makes an effort is successful.'

Note that the presence of *bʰi*, just like *-ever*, forces a generic reading if one is possible, as shown in (15c). The acceptability of the generic reading, however, is dependent on the aspect of the relative clause predicate. If information from tense/aspect does not provide information about genericity, the standard interpretation of the correlative clause is definite, regardless of the presence of the focus item (as is generally the case with free relatives; Jacobson 1995).

As Bhatt (1997, 2003) further shows, the correlative can even take the form of a true free relative without a demonstrative 'correlate' if the case marking of the internal as well as the external head is unmarked (nominative in Urdu). The demonstrative cannot be omitted if either the demonstrative or the correlate is overtly marked by a case clitic (e.g., *ne, ko*), but can be left out if the surface form matches or, in the case of Urdu, has no surface form, as shown in (16).

(16) [**jo** laṛki kʰaṛ-i hɛ] hɑs rɑh-i
which girl.F.Sg.Nom stand-Perf.F.Sg be.Pres.3.Sg laugh stay-Perf.F.Sg

hɛ
be.Pres.3.Sg
'Which girl is standing, is laughing.'

This form of surface matching is known from German free relative constructions, which also require a resumptive demonstrative/determiner if the case marking differs overtly, as is the case in (17a), but not in (17b,c).

(17) a. **Wer** dich nicht mag, ***(den)** mag ich
who.Nom you.Acc not like.Pres.3.Sg. **that.Acc** like.Pres.1.Sg I.Nom

auch nicht.
also not
'Who doesn't like you, I don't like either.'

b. **Wen** du magst, **(den)** will ich auch
who.Acc you.Nom like.Pres.2.Sg **that.Acc** want.Pres.1.Sg. I.Nom also

treffen.
meet
'I also want to meet the one who you like.'

c. **Was** du magst **(das)** gefällt mir auch.
what.Acc you.Nom like.Pres.2.Sg **that.Nom** please.Pres.3.Sg I.Dat also
'Whatever you like also pleases me.'

In sum, correlative clauses in Urdu have a number of semantic and morphosyntactic properties that are familiar from free relatives in German and English. Within the context of the ParGram project, this points towards the need to find a common underlying analysis for free relatives and correlatives in these languages.

4.2 Specifier of DP

In order to account for similarities between free relatives and correlatives, we treat the correlative clause plus demonstrative constituent as a DP with an f-structure analogous to free relatives in English, since these have comparable semantics and distribution. However, instead of analyzing the relative clause predicate as an adjunct, as Butt et al. (1999) did for free relative clauses, we consider correlative clauses as occupying a specifier position and thus contributing a SPEC attribute to the f-structure. This is done for the following reasons:

- Correlatives cannot be stacked: Whereas normal relative clauses project into an adjunct set, a DP can only be modified by a single correlative.

  (18) *[jo gari tez hɛ] [jo lal hɛ]
  which car.F.Sg.Nom fast be.Pres.3.Sg which red be.Pres.3.Sg
  vo gari sundɑr hɛ
  that car beautiful be.Pres.3.Sg
  'Which car is fast, which car is red, that car is beautiful.'

- Semantically, correlatives function as quantifiers. Thus, they cannot have a non-restrictive interpretation and cannot modify, for example, proper nouns, as this would result in vacuous quantification.

  (19) *[jo vɑhã kʰɑṛ-a hɛ] ram lɑmba hɛ
  who there stand-Perf.M.Sg be.Pres.3.Sg Ram tall.M.Sg be.Pres.3.Sg
  'Who is standing there, Ram is tall.'

- Correlatives appear in complementary distribution with other SPEC material, such as possessors.

  (20) *[jo lal hɛ] [yonas=ki] gari
  which red be.Pres.3.Sg Jonas=Gen.F.Sg car.F.Sg.Nom
  sundɑr hɛ
  beautiful be.Pres.3.Sg
  'Which car is red, Jonas's car is beautiful.'

All of the evidence presented so far points to a DP-internal, non-adjunct analysis of correlatives. In particular, we situate the correlative clause in a specifier position directly left-adjacent to the DP that it quantifies over. However, there is a further set of data that remain to be accounted for.

4.3 Topicalization

In addition to finding correlatives that are directly left-adjacent to the modified constituent, instances of discontinuous correlatives also occur. In both of the examples in (21), the correlative clause is in the regular sentence-initial position, but the demonstrative is embedded further inside the main clause ((20a)) or even embedded within a sentential complement ((20b)).

(21) a. [jo laṛki$_i$ vɑhã hɛ]
which girl.F.Sg.Nom there be.Pres.3.Sg

ram=ne ʊs=ko$_i$ pɑsɑnd ki-ya
Ram=Erg that=Acc liking do-Perf.3.Sg
'Which girl is there, Ram likes her.'

b. [jo laṛki$_i$ ga rɑh-i hɛ]
which girl.F.Sg.Nom sing stay-Perf.F.Sg be.Pres.3.Sg

sita soc^h-ti hɛ [kı vo$_i$ sundɑr hɛ]
Sita think-Impf.F.Sg be.Pres.3.Sg that that beautiful be.Pres.3.Sg
'Which girl is singing, Sita thinks that she is beautiful.'

We propose to analyze this dislocation as an instance of standard topicalization of the correlative clause. This analysis is reasonable, given that Urdu is a discourse-configurational language with basic SOV order that makes heavy use of word order permutations to syntactically encode information structure (cf. Butt and King 1996, Kidwai 2000). The TOPIC function is generally associated with the initial item of the utterance, located in SPECIP. Since Urdu allows not only arguments to be topicalized, but also, unlike English, SPECNP content such as genitive possessors (Mohanan 1994), as shown in (22), it is predicted that correlative clauses should also be able to undergo this dislocation.

(22) [ram=ki] sundɑr hɛ gari
Ram=Gen.F.Sg beautiful be.Pres.3.Sg car.F.Sg.Nom
'Ram's car is beautiful.'

As Bhatt (2003) shows, topicalized correlatives may connect into sentential complements, but are sensitive to island effects, and thus cannot be topicalized from within adjuncts or complex NPs. Furthermore, only one correlative clause can be topicalized. If any other DP in the sentence is modified by a correlative clause, none of these may additionally appear in the front, but must be located in non-initial position within the relevant SPECDP. The only admissible structure that allows two correlative clauses at the beginning of the sentence is one where one correlative is topicalized and the other occupies the specifier position of a sentence-inital DP, as illustrated by (23).

(23) [jo$_i$ kıtab mez=pɑr t^h-i]
which book.F.Sg.Nom table.F.Sg=on be.Past.F.Sg

[[[[jo$_j$ talıb‘ılm hɛ] ʊs$_j$] lɑrke=ne] vo$_i$ lık^h-i]
which student.Nom be.Pres.3.Sg that boy.M.Obl=Erg that write-Perf.F.Sg
‘The boy who is a student wrote the book that was on the table.’

4.4 The LFG Analysis

DP internal correlative clauses are linked to their heads via the functional description expanding the DP node. As topicalized correlatives are discontinuous from their heads, this case is more interesting. As is standard with respect to long-distance dependencies within LFG, topicalized correlatives are linked to their heads via functional uncertainty paths (Kaplan and Zaenen 1989).[6] However, as correlative clauses are not quite standard topics, a little more work needs to be done. The external head of the topicalized correlative is found via the disjunction in (24) (defined in XLE's regular expression notation; Crouch et al. 2006):

(24) CORFUNC =
{SUBJ ∨ {XCOMP ∨ COMP}* {OBJ ∨ OBJ-GOAL } (ADJ-GEN)}

This means that the function CORFUNC is assigned a grammatical function that is either SUBJ or an OBJ or OBJ-GOAL which may be embedded in zero to infinitely many (signified by the Kleene Star) COMPs or XCOMPs, or a genitive possessor of any of these. In the c-structure rule that licenses the topicalized correlatives shown in (25), this function is given a local name (%COR-HEAD) in order to formulate the constraints that must be simultaneously satisfied, such as the demonstrative/quantifier requirement (cf. section 3.2) as well as number and oblique/nominative agreement (coded under NMORPH). The rule in (25) also includes the possibility of a topicalized KP (Urdu Kase Phrase, featuring a DP plus optional case clitic, Butt and King 2004).

(25) SPECIP ⟶

CPCORR	∨	KP
(↑CORFUNC)=%COR-HEAD		(↑TOPIC)=↓
(%COR-HEAD SPEC CORR)=↓		@GF
(%COR-HEAD NUM)=(↓TOPIC-REL NUM)		
(%COR-HEAD NMORPH)=(↓TOPIC-REL NMORPH)		
{(%COR-HEAD SPEC DET PRON-TYPE)=c dem		
(%COR-HEAD SPEC DET DEIXIS)=c distal		
∨ (%COR-HEAD SPEC QUANT QUANT-TYPE)=c universal}		

The CPCORR category is defined in analogy to the standard relative clause, CPREL, with the exception that in CPCORR the TOPIC-REL may include a contentful NP. The function SPEC CORR, which encodes the correlative clause itself,

[6]Topicalized possessors as in (22) are, of course, also linked via functional uncertainty equations.

is proposed as an interim solution since the facts presented in Jacobson (1995) and others call for a consequent reanalysis of free relative clauses in the ParGram languages that departs from the ADJUNCT solution proposed in Butt et al. (1999), along with a unified analysis that provides enough information in the f-structure to lead to the correct semantic representation for both correlatives and free relatives.

An unusual but positive aspect to the analysis in (25) is that the subject of the main clause and that of the correlative do not stand in a direct functional relation other than noun agreement. Since the correlative is allowed to contain a full noun phrase in its internal head, it is in principle possible for the internal head and the external head to contain diffent nouns. And precisely this possibility is required by data as in (26).

(26) [**jo** **catr** vahã kʰɑṛ-a hɛ]
which student.M.Sg.Nom there stand-Perf.M.Sg be.Pres.3.Sg

vo **lɑṛka** mera dost hɛ
that boy.M.Sg.Nom I.Gen.M.Sg friend.M.Sg.Nom be.Pres.3.Sg
'Which student is standing there, that boy is my friend.'
(McCawley 2004:300)

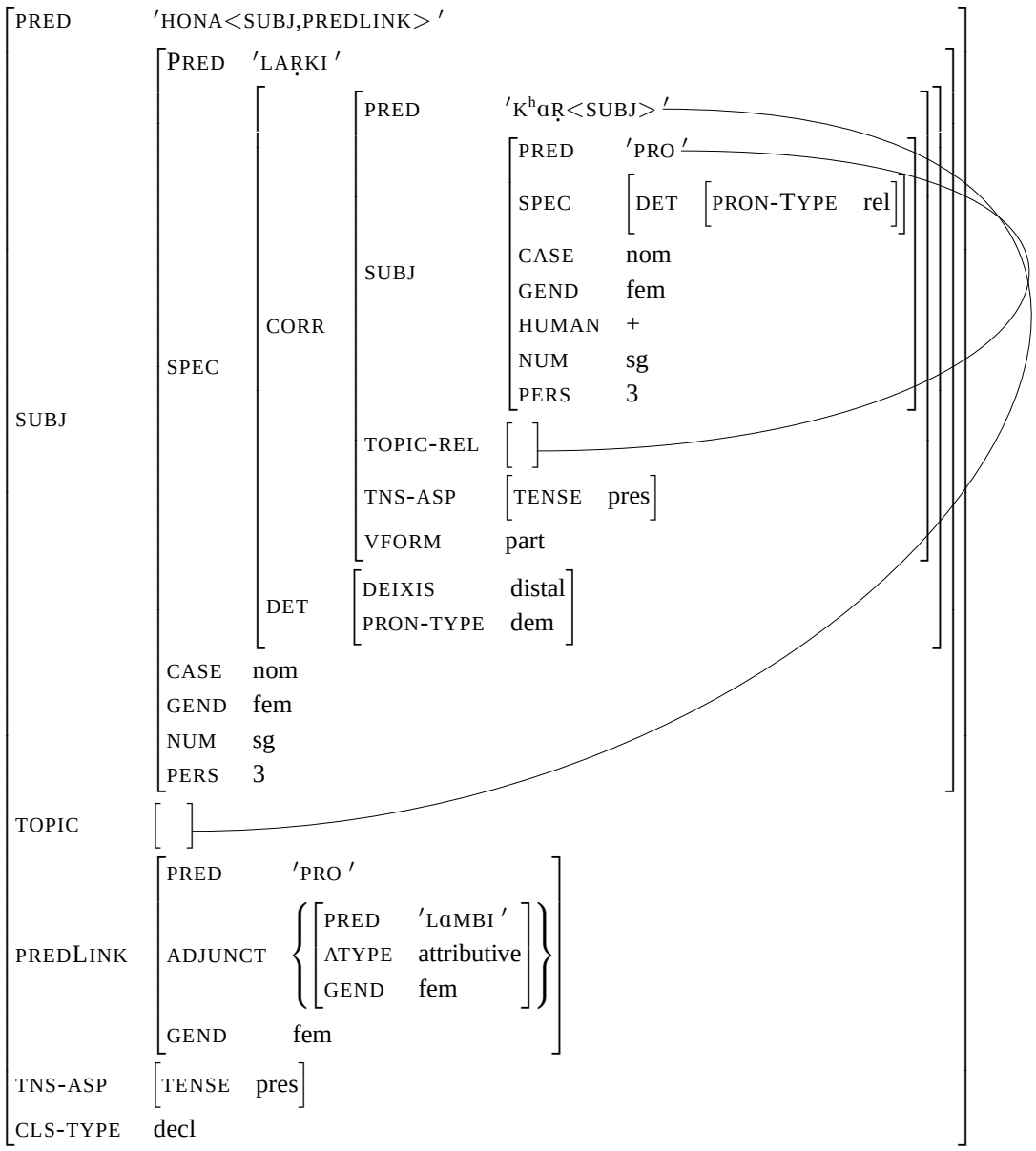
There seem to be semantic constraints on the felicitous choice of the two different nouns (involving synonymy or hyponymy), but previous analyses of these (cf. Dayal 1996:196 and McCawley 2004:300) as well as judgements of our informants leave an inconclusive picture of what relations are acceptable. With respect to our analysis, since the constraints are purely semantic, they are not handled by the syntactic c-structure and f-structure components.

In (28) and (29) we present a sample c-structure and f-structure analysis for the example in (27). The representation of noun phrase structure departs from previous analyses in the Urdu grammar by postulating a DP structure above NP that holds the determiner *vo* (which, in its use as a personal pronoun, accompanies an empty noun head) as well as a SPECDP position potentially containing the correlative.

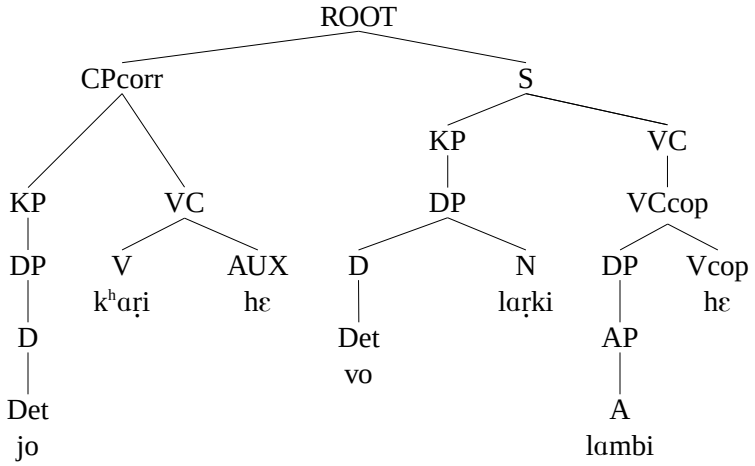
Note that the same sentence can also receive another analysis by which the correlative clause is not topicalized and is contained inside the sentence-inital DP instead of SPECIP. This option, which then lacks the TOPIC function at f-structure, is dispreferred through the use of OT marks (Frank et al. 2001).

(27) [**jo** kʰɑṛ-i hɛ] [**vo** lɑṛki lɑmbi hɛ]
which stand-Perf.F.Sg be.Pres.3.Sg that girl.F.Sg.Nom tall.F.Sg be.Pres.3.Sg
'Who is standing, that girl is tall.'

(28) F-Structure:



(29) C-structure:



5 Further Issues: Multi-Head-Correlatives

So far we have presented an analysis for single-head correlatives (SHC). However, as shown in (30), one of the striking features of Urdu correlatives is that they can appear with more than one relativized element, containing multiple relative pronouns linked to multiple correlate demonstratives in the main clause.

(30) [jɪs$_i$ laɽki=ne jɪs$_j$ laɽke ke satʰ kʰel-a]
which.Obl girl.F.Sg=Erg which.Obl boy.M.Obl with play-Perf.M.Sg

ʊs=ne$_i$ ʊs=ko$_j$ hara-ya
that.Obl=Erg that.Obl=Acc defeat-Perf.M.Sg
'Which girl played with which boy, she defeated him.' (Dayal 1996)

This correlative clause cannot be attached to any single correlate at f-structure, since both internal heads are equally governed by the relative clause predicate, and the functional projection of the predicate cannot be attached to both external heads with the same internal structure. The correlative clause cannot be said to determine either argument of the matrix, but rather determines both by specifying a relation between them. This can be expressed semantically by arguing that multi-head-correlatives (MHC) quantify over ordered tuples rather than individuals (as proposed by Lehmann 1984:344). It can be expressed syntactically by attaching the f-structure of the correlative clause directly to the main clause predicate rather than to one of its arguments. Analogously, Srivastav (1991) and Bhatt (1997) argue for base-generation of the MHC in a position adjoined to IP. Within Dayal's account this means that she presents a unified analysis of SHC and MHC, within Bhatt's account this means that MHC and SHC receive a differing syntactic analysis.

With respect to this issue, we again propose to follow Bhatt's analysis and treat MHC as a separate class for which no analogous construction (such as free relatives

for SHC) exists in languages that do not feature correlatives. Consequently, these sentences cannot be translated straightforwardly into English. Andrews (1975), for example, proposes to translate MHC as conditionals, which gives adequate results as long as the correlative can have a generic interpretation, but this is not always the case. Another suggestion, propably first proposed by Delbrück (1900) for Sanskrit MHC, is to use an indefinite in place of the second relativized phrase, which is anaphorically picked up in the matrix clause (*Whichever girl played with a boy defeated him.*). This translation would also be faithful to the semantics of the construction, but does not do justice to the differing syntactic constraints. As shown in (31), Urdu MHC cannot appear with a matrix predicate of less arity, whereas relatives-cum-indefinite can.

(31) a. [*jɪs lɑṛki jɪs lɑṛke ke sat^{h} k^{h}el-egi]
which.Obl girl.F.Sg.Nom which.Obl boy.M.Obl with play-Fut.F.Sg
vo jit-egi
that win-Fut.F.Sg
'Which girl will play with which boy, she will win.'

b. Whichever girl will play with a boy will win.

MHC are also less constrained in contrast to SHC when it comes to the resumptive pronoun requirement. Even in cases where the demonstrative accompanying a correlative clause could not be dropped, i.e. if there is overt case-marking, they may be dropped with MHC (Bhatt 1997), as shown in (32).

(32) [*jɪs lɑṛke=ne jɪs lɑṛki=ko dek^{h}-a]
which.Obl boy.M.Sg.Obl=Erg which.Obl girl.F.Sg=Acc see-Perf.M.Sg
(ʊs=ne ʊs=ko) pɑsɑnd ki-ya
that.Obl=Erg that.Obl=Acc liking do-Perf.M.Sg
'Which boy saw which girl, he liked her.'

Since the exact nature of the interaction between the constraints of correlative formation and the rampant pro-drop that is generally possible in Urdu (Neeleman and Szendroi 2007) is not yet well understood, our analysis is rather minimal. At c-structure, we assume MHC to be located adjoined above IP. At f-structure, the correlative clause projects an ADJUNCT to the main clause predicate. The anaphoric relation between the relativized elements and possible correlates in the main clause is left to the semantic processing component, which may be tackled once a better understanding of the structure is reached.

6 Conclusion

Building on previous insights by Srivastav/Dayal and Bhatt, we distinguish between relative clauses (*vo-jo*) and correlatives (*jo-vo*), and account for their different internal structure and semantic interpretation. Correlatives are treated as

quantifiers that appear either in the specifier position of the DP they modify or in a topicalized position at the left periphery. At f-structure, they differ from normal relative clauses by projecting to a SPEC structure rather than an adjunct set, which goes along with their quantifier interpretation and their inability to stack. The parallels to free relative clauses suggest that a similar analysis might be argued for in the case of German and English free relatives, which currently receive the same ADJUNCT treatment as standard relatives. The advantages and disadvantages of such a parallel analysis, as well the issue with multi-head-correlatives, can hopefully be understood once a standardized semantic representation has been agreed on within ParGram, and once the existing analysis has been incorporated into the main Urdu grammar in order to investigate interactions with other phenomena, such as pro-drop.

References

Andrews, Avery D. 1975. *Studies in the Syntax of Relative and Comparative Clauses*. Ph. D.thesis, Massachusetts Institute of Technology.

Bhatt, Rajesh. 1997. Matching Effects and the Syntax-Morphology Interface: Evidence from Hindi Correlatives. In *MIT Working Papers in Linguistics*, volume 31, MIT Press.

Bhatt, Rajesh. 2003. Locality in Correlatives. *Natural Language and Linguistic Theory* 21, 485–541.

Butt, Miriam, Dyvik, Helge, King, Tracy Holloway, Masuichi, Hiroshi and Rohrer, Christian. 2002. The Parallel Grammar Project. In *COLING Workshop on Grammar Engineering and Evaluation*.

Butt, Miriam and King, Tracy Holloway. 1996. Structural Topic and Focus without Movement. In *Proceedings of the First LFG Conference*.

Butt, Miriam and King, Tracy Holloway. 2002. Urdu and the Parallel Grammar Project. In *Proceedings of the 19th International Conference on Computational Linguistics (COLING)*.

Butt, Miriam and King, Tracy Holloway. 2004. The Status of Case. In Veneeta Dayal and Anoop Mahajan (eds.), *Clause Structure in South Asian Languages*, pages 153–198, Berlin: Springer Verlag.

Butt, Miriam, King, Tracy Holloway, Niño, María-Eugenia and Segond, Frédérique. 1999. *A Grammar Writer's Cookbook*. CSLI Publications.

Cooper, Robin. 1983. *Quantification and Syntactic Theory*. Dordrecht.

Crouch, Richard, Dalrymple, Mary, Kaplan, Ronald M., King, Tracy Holloway, Maxwell III, John T. and Newman, Paula. 2006. XLE Documentation, Palo Alto Research Center, on-line.

Davison, Alice. 2006. Correlative Clause Features in Sanskrit and Hindi/Urdu, Paper presented at the 9th Diachronic Syntax Conference, Trieste.

Dayal, Veneeta. 1996. *Locality in WH Quantification. Questions and relative clauses in Hindi*. Kluwer Academic Publishers.

Dayal, Veneeta. 2003. Bare Nominals: Non-specific and Contrastive Readings under Scrambling. In Simin Karimi (ed.), *Word Order and Scrambling*, Oxford: Blackwell.

Delbrück, Berthold. 1900. Vergleichende Syntax der indogermanischen Sprachen. Dritter Theil. In Karl Brugmann and Berthold Delbrück (eds.), *Grundriß der vergleichenden Grammatik der Indogermanischen Sprachen*, Trübner.

Frank, Anette, King, Tracy Holloway, Kuhn, Jonas and Maxwell III, John T. 2001. Optimality Theory Style Constraint Ranking in Large-scale LFG Grammars. In Peter Sells (ed.), *Formal and Empirical Issues in Optimality Theoretic Syntax*, pages 367–397, CSLI Publications.

Jacobson, Pauline. 1995. On the Quantificational Force of English Free Relatives. In Emond Bach, Eloise Jelinek, Angelika Kratzer and Barbara Partee (eds.), *Quantification in Natural Language*, Kluwer Academic Publishers.

Kaplan, Ron and Zaenen, Annie. 1989. Long-distance dependencies, constituent structure, and functional uncertainty. In Mark Baltin and Andy Kroch (eds.), *Alternative Conceptions of Phrase Structure*, Chicago University Press.

Kidwai, Ayesha. 2000. *XP-Adjunction in Universal Grammar: Scrambling and Binding in Hindi-Urdu*. Oxford: Oxford University Press.

Lehmann, Christian. 1984. *Der Relativsatz*. Gunter Narr Verlag.

Mahajan, Anoop. 2000. Relative Asymmetries and Hindi Correlatives. In Artemis Alexiadou, Paul Law and Chris Wilder (eds.), *The Syntax of Relative Clauses*, John Benjamins.

McCawley, John D. 2004. Remarks on Adsentential, Adnominal and Extraposed Relative Clauses in Hindi. In Veneeta Dayal and Anoop Mahajan (eds.), *Clause Structure in South Asian Languages*, Berlin: Springer Verlag.

Mohanan, Tara. 1994. *Argument Structure in Hindi*. CSLI Publications.

Neeleman, Ad and Szendroi, Kriszta. 2007. Radical pro drop and the morphology of pronouns. *Linguistic Inquiry* 38(4), 671–714.

Srivastav, Veneeta. 1991. The Syntax and Semantics of Correlatives. *Natural Language and Linguistic Theory* 9(4).

Subbarao, K. V. 1984. *Complementation in Hindi Syntax*. Academic Publications.

Wali, Kashi. 1982. Marathi Correlatives: A Conspectus. *South Asian Review* 6.