

The Syntax of Sanskrit bahuvrīhis

John J. Lowe

University of Oxford

Davide Mocci

Scuola Universitaria Superiore IUSS Pavia

Proceedings of the LFG'22 Conference

Miriam Butt, Jamie Y. Findlay and Ida Toivonen (Editors)

2022

CSLI Publications

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

Abstract

Lowe (2015b) provided an LFG-based analysis of Classical Sanskrit compounds, including bahuvrīhis of the adjective-noun type. In this paper we show that Lowe’s (2015b) analysis cannot account for the full range of bahuvrīhi types attested in Sanskrit; we improve and extend Lowe’s account to cover the major types of bahuvrīhi. We also formalize a broader generalization on the ordering constraints of the members of bahuvrīhis.

1 Introduction

The old Indo-Aryan language Sanskrit shows a remarkable productivity and frequency in its use of compounding. While compounding is usually treated as a morphological phenomenon, a lexical process of forming word-sized units from two or more (usually) lexical stems, Lowe (2015b) argues that certain properties of Sanskrit compounds suggest rather a syntactic status, though a syntactic status somewhat different from the ‘ordinary’ rules of phrasal syntax. Lowe provides an LFG analysis of Classical Sanskrit compounding which assumes that compounds are formed in the syntax, and utilizes non-projecting categories to model the more constrained syntactic properties of compound sequences.

Lowe (2015b) addressed only the five main compound types in Sanskrit – dvandva, tatpuruṣa, karmadhāraya, bahuvrīhi and avyayībhāva – and of these only the most common subtypes. In particular, Lowe presented an analysis of only one type of bahuvrīhi. In this paper we explore the major types of bahuvrīhi compounds in Sanskrit, and reveal some generalizations which could not be accommodated by Lowe’s original analysis.¹ We revise the LFG-based analysis of Lowe (2015b) to account for the wider range of data and capture the relevant generalizations.

2 Compounding in Sanskrit

In Sanskrit, compounds may be productively formed from many different phrase types, including noun phrases consisting of a head noun modified by an adjective or another noun, verb phrases, coordinated noun phrases, etc.² For illustration, consider (1) and (2):

[†]This paper is the result of a joint research work entirely discussed and shared by both authors. Merely for the sake of (Italian) academic requirements, §1 and §§6–8 are attributed to John Lowe, and §§2–5 to Davide Mocci. We thank the audiences at SE-LFG32 (21 May 2022) and at LFG22 (13 July 2022) and two anonymous reviewers for their insightful comments.

¹Given constraints of space, we cannot cover the full range of bahuvrīhis here; the most prominent types of bahuvrīhi omitted here are those whose left-hand members are particles, including negative bahuvrīhis and those with a prepositional left-hand member. We plan to treat these types elsewhere.

²For standard accounts of compounding in Sanskrit, see Whitney (1896, 480–515) and Wackernagel (1905).

- (1) *śaphād áśvasya*
 hoof.ABL horse.GEN
 ‘from the horse’s hoof’ (*Rgveda* 1.117.6).
- (2) *aśva-śaphéna*
 horse-hoof.INS
 ‘by means of the horse’s hoof’ (*Śatapathabrāhmaṇa* 13.3.4)

A (partitive) possessive relation holds between *áśva-* and *śaphá-* in both (1) and (2). However, only in (1) is such a relation overtly expressed, by the GEN.SG ending on *áśva-*; in (2) *áśva-* appears in bare stem form, and nothing else overtly marks the possessive relation. Following Guevara and Scalise (2009, 107) we can define a compound as a lexical category Z such that: i. Z is the output of combining a lexical category X with another lexical category Y; ii. some morphosyntactically covert semantic relation *r* holds between X and Y in Z. Formally: [X *r* Y]_Z.

The category Z in the compound scheme [X *r* Y]_Z is typically understood to be a lexical category; that is, Z is a word, and its formation is therefore a morphological process, not a syntactic process. As a lexicalist theory of syntax, LFG assumes the STRONG LEXICALIST HYPOTHESIS, according to which “no syntactic rule can refer to elements of morphological structure” (Lapointe 1980, 8). Syntactic rules relate words and phrases (i.e., units constituted of words) to other words and phrases within a sentence, but do not relate words or phrases to sub-word units – see Williams (2007, 354), Molina-Muñoz (2013, 187), and Lowe (2015b, 76). Since a compound (i.e., Z) is a word, the internal elements (i.e., X and Y) of a compound should be unable to participate in syntactic relations, i.e., relations with compound-external words and phrases. This corresponds with the standard treatment of compounding as a morphological process; for treatments within LFG, see e.g. Ørsnes (1996), Butt et al. (1999, 94–97), Baker and Nordlinger (2008), and Lee and Ackerman (2011).³

As in most other languages, compounds in Sanskrit are traditionally analysed as words, and do show features consistent with this analysis. Like single words, Sanskrit compounds bear only one accent and are endowed with only one case ending. Thus, both *áśva-* and *śaphá-* preserve their accent and are case-marked in (1), where they are not compounded, while only *śaphá-* (or rather, the compound as a whole) is accented and case-marked in (2). Also, derivational processes may apply to compounds just as to single words: e.g. the possessive suffix *-ín-* may attach to a single word like *rátha-*, producing *rathín-* ‘possessing a chariot, charioteer’ (*Rgveda* 6.47.31), but also to compounds like *śiti-kákṣa-* ‘white belly’ (*Atharvaveda-Śaunakīya* 5.23.5), producing *śiti-kakṣín-* ‘possessing a white belly, white-bellied’ (*Taittirīya-Saṃhitā* 5.5.20.1).

However, Lowe (2015b, 75–86) provides several arguments to show that non-final members of Sanskrit compounds can entertain relations with compound-exter-

³Note that none of these authors analyse bahuvrīhis. For an overview of recent theoretical approaches to compounding beyond LFG, see Lowe (2015b, 87–88) and Eik (2019, 83–146).

nal words or phrases (including sentences), in apparent violation of the strong lexicalist hypothesis; in the next section we report some of these arguments.

3 The permeability of Sanskrit compounds

Here we discuss three of the arguments which Lowe (2015b, 75–86) provides to show that compounding is a syntactic, not morphological process, in Sanskrit. First, in Sanskrit it is possible for words or phrases external to a compound to modify the left-hand member of that compound, and not the compound as a whole. This is called ‘asamartha’ compounding (see Gillon 1994, 2007; Molina-Muñoz 2013; Lowe 2015b, 76–77). Second, while there is a strong crosslinguistic tendency for words to be anaphoric islands (that is, outbound/inbound anaphora to/from a unit smaller than a word is crosslinguistically extremely rare – see Postal 1969), productively formed compounds in Sanskrit plainly violate anaphoric islandhood (Kiparsky 2009, 83; Molina-Muñoz 2013, 191–192; Lowe 2015b, 78). Both these features are illustrated in the following Classical Sanskrit example:⁴

- (3) *apratibaddhasya [tat- abhāve] sarvatra [abhāva-*
unconnected.GEN_i that_i- absence.LOC everywhere absence-
asiddeḥ]
nonestablishment.ABL
‘due to the nonestablishment of the absence of a thing_i that is unconnected
in every absence of it_i’ (*Pramānavārttikasvavṛtti* 12.23; Lowe 2015b, 78).

Here the left-hand member of the compound *abhāva-asiddeḥ* is modified by the compound-external NP *apratibaddhasya*; furthermore, the demonstrative pronoun *tat*, which occurs as the left-hand member of the compound *tat-abhāve*, is coreferential with *apratibaddhasya*, which is not contained in the compound.

Third, the left-hand member of a Sanskrit compound can take clausal scope when it is realized as an interrogative pronoun, as illustrated in the following Classical example (see Kiparsky 2009, 83; Lowe 2015b, 80):

- (4) *[kim- lakṣaṇam] punas tad brahma*
what- definition.NOM but that.NOM sacred.formulation.NOM
‘But what is the definition of that sacred formulation?’ (lit. ‘Having-what-
as-its-definition is that sacred formulation?’) (*Śarīrakamīmāṃsābhāṣya*
1.1.2; Lowe 2015b, 80).

Examples (3)–(4) show that non-final members of Sanskrit compounds can entertain syntactic relations with elements outside the compound they appear in (an NP in (3), the sentence as a whole in (4)). This is evidence either that Sanskrit

⁴Classical Sanskrit is a standardized form of Sanskrit codified on the basis of Pāṇini’s *Aṣṭādhyāyī* (4th c. BCE) and other works of the indigenous grammatical tradition.

compounds violate strong lexicalism, or that constituent members of Sanskrit compounds must themselves be words, and compounds must be syntactically formed phrases, rather than words. The analysis developed by Lowe (2015b) assumes the latter, treating compounds as syntactically formed units, and thereby conforming to the strong lexicalist assumptions of LFG. Yet it also recognizes a (rather less lexicalist) degree of gradience in the distinction between word and phrase, and models Sanskrit compounds as inhabiting a kind of middle-ground between words and full syntactic phrases. In doing this, however, it makes use only of existing LFG mechanisms independently posited to account for other syntactic phenomena.⁵

As discussed in §2 above, there is some evidence for the lexical status of compounds, and there are also major differences between ordinary phrasal syntax and the constitution of compounds which suggest that if the latter does involve syntax, it is a more restricted syntax than ordinary phrasal syntax. In particular, word order is almost entirely free in the ordinary phrasal syntax of Sanskrit, while the order of compound members is (almost) absolutely fixed; in addition, any full word can appear in ordinary phrasal syntax, while only stem forms of lexical classes can appear in non-final positions of compounds.

To capture the mixed phrasal and lexical features of Sanskrit compounds, Lowe (2015b) has recourse to the concept of *non-projecting categories* (Toivonen 2003). A non-projecting category, represented as e.g. \hat{X} , is a category type which cannot project a phrase, and therefore cannot take another phrase as specifier, complement or adjunct. \hat{X} categories can adjoin to projecting $X(X^0)$ heads, and following Duncan (2007) and Arnold and Sadler (2013), also to other non-projecting categories:

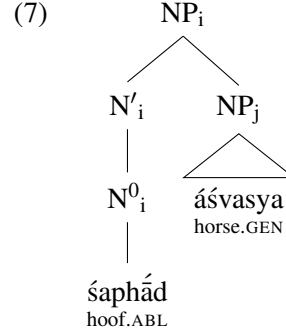
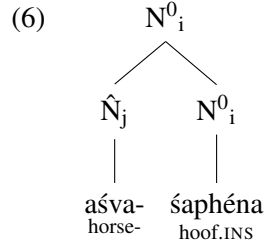
- (5) a. $X^0 \rightarrow \hat{Y} X^0$
 b. $\hat{X} \rightarrow \hat{Y} \hat{X}$

Granted rules of these forms, sequences of non-projecting words may have the crucial feature of syntactic phrases, i.e., hierarchical phrase-structural relations between members, but also a crucial feature of words, i.e., being dominated by an X^0 node. Applied to Sanskrit compounds, then, the top node of any compound ‘phrase’ is of category X^0 , that is the same as the category of lexical words, while non-final members of compounds are adjoined words of category \hat{Y} , which accounts for the phrase-like features of Sanskrit compounds (notably, the visibility of non-final members of Sanskrit compounds to compound-external material); since adjunction of \hat{Y} to X^0 is recursive, the potentially unrestricted length of Sanskrit compounds is easily captured (see Lowe 2015b, 80–83 for relevant examples).

Thus, analysing the compound *aśva-śaphéna* ‘by means of the horse’s hoof’ (2) along the lines of (5a), we get (6), which contrasts with (7), the tree for *śaphād*

⁵We take this to be in principle preferable to significantly augmenting or altering the LFG theory to account for the data discussed above.

ásvasya ‘from the horse’s hoof’ (1):⁶



Lowe (2015b, 107–108) confines his claim about the syntactic status of Sanskrit compounds to Classical Sanskrit, noting that data like (3)–(4) are much rarer in the earlier Vedic Sanskrit language. However, examples comparable to (3)–(4) are already found in Vedic Sanskrit, so we consider it reasonable to assume a syntactic treatment of Vedic Sanskrit compounds, too. For example, asamartha compounding is found even in the earliest Vedic Sanskrit, the *Ṛgveda*: e.g. *devānām vaśa-nīh* ‘one who leads (*nīh*) at the will (*vaśa*) of the gods (*devānām*)’ (*Ṛgveda* 10.16.2d); we also find pronouns used in compounds in the *Ṛgveda* in exactly parallel manner to those illustrated in (3) and (4) above, e.g. *tād-anna-* ‘whose food is this’ and *tād-apas-* ‘whose work is this’ (*Ṛgveda* 8.47.16a), where both instances of *tād* ‘this’ refer to a word outside the compounds; and *kād-artha-* ‘having what purpose’ (*Ṛgveda* 10.22.6b) where, as in (4), the question word has clausal scope. In light of these data, we assume the same analysis for compounding in both Vedic and Classical Sanskrit, and draw and analyse examples from both in the rest of this paper.

4 Bahuvrīhis: introduction

4.1 Defining a bahuvrīhi

The seminal classification of compounds by Scalise and Bisetto (2009, §4) lacks a category that matches precisely the scope of bahuvrīhis in Sanskrit. Bauer’s (2017, 65) definition of bahuvrīhis as “compounds which canonically label a part of the whole which the compound denotes” is problematic, too, because bahuvrīhis that do not involve any part-whole relation are not infrequent in Sanskrit (see Cardona 1997, 220 and Pontillo 2021, 507–509). For Sanskrit, at least, we adopt the following definition of bahuvrīhis, which reflects the way this compound type was analysed within the ancient Indian grammatical tradition:⁷ a bahuvrīhi is a com-

⁶We are assuming that modifiers, including possessors like *ásvasya* of (7), are projected in phrasal syntax as specifiers of the head noun they modify. Nothing rides on this.

⁷See especially Kātyāyana’s *Vārttika* 19 ad A 2.2.24, the *Mahābhāṣya*’s discussion around this *Vārttika*, as well as the remarks in the *Kāśikā-vṛtti* ad A 2.2.24 (see also Benfey 1852, 273). Note

pound Z that may be paraphrased with a relative clause such that i. the relative pronoun is canonically inflected in a case other than nominative, and ii. the functions of subject and predicate are canonically fulfilled, in the relative clause, by the same lexical categories that serve as the internal members of Z. Lowe (2015b, 74) treats the following example:

- (8) [*dīrgha- karṇo*] *devadattaḥ*
 long- ear.NOM.SG.M Devadatta.NOM.SG.M
 ‘Long-eared Devadatta’

The compound *dīrgha-karṇaḥ*⁸ qualifies as a bahuvrīhi since it satisfies both conditions (i)–(ii): it can be (and traditionally is) paraphrased as in (9), where the relative pronoun is inflected in the genitive, and the functions of subject and predicate are respectively fulfilled by *karṇau* and *dīrghau*, i.e., the same lexical categories which *dīrgha-karṇaḥ* is made up of. Note that all bahuvrīhis, including *dīrgha-karṇaḥ*, are semantically exocentric:⁹ e.g., the denotatum of *dīrgha-karṇaḥ* (i.e., an entity whose ears are long) is not a hyponym of either *dīrgha-* ‘long’ or *karṇa-* ‘ear’. *devadattaḥ*, which agrees in case, gender, and number with *dīrgha-karṇaḥ* in (8), specifies the identity of the compound’s denotatum (i.e., tells us who the entity whose ears are long is), and is known as the external referent of the compound (Lowe 2015b, 103).

- (9) *dīrghau* *karṇau* *yasya* *devadattaḥ*
 long.NOM.DU.M ear.NOM.DU.M who.GEN.SG.M Devadatta.NOM.SG.M
 ‘Devadatta, whose ears are long’

Endocentric compounding typically involves a variable or underspecified relation *r* between the compound members, the precise value of which is determined contextually.¹⁰ In endocentric compounding, moreover, one member of the compound is the (semantic and morphosyntactic) head. This is the basis of the formalization of endocentric compounds by, e.g., Butt et al. (1999, 94–97), involving an f-structure feature COMPOUND (or MOD, in current ParGram) which embeds the f-structure of the non-head compound member within the f-structure of the head.

Bahuvrīhis are different. In bahuvrīhis neither compound member is the head, and in addition the relationship between the compound members is constant: *r* always represents a predicative relationship between the compound members. The variable relationship in the case of bahuvrīhis is the relationship of the external referent with the bahuvrīhi-internal subject or predicate (or, according to Lowe 2015b,

that this traditional analysis of bahuvrīhis is different from the definition found in Pāṇini’s *Aṣṭādhyāyī* itself, for which see Candotti and Pontillo (2019, 33–36; 2022), Pontillo (2021).

⁸The final *-aḥ* of *dīrgha-karṇaḥ* appears as *-o* in the example due to sandhi in context before the following word. We leave such alternations unmarked below.

⁹The reverse does not hold, however: not all exocentric compounds are bahuvrīhis; see Bauer (2017, 64–71).

¹⁰Cf. Allen’s (1978) “Variable R” condition (see ten Hacken 2009, §3.2). A compound Z is semantically endocentric if Z is interpreted as a hyponym of a member of Z.

with the whole predicative structure made up of the bahuvrīhi-internal subject and predicate – see §4.3 below), which in the relative clause type parse of the compound is modelled in the variable case marking of the relative pronoun. In terms of the f-structure of bahuvrīhis, then, we require an analysis rather different from previous analyses of endocentric compounds, which captures the predicative nature of the relation between the compound members. Before presenting this analysis, we compare alternative analyses of bahuvrīhis.

4.2 Previous analyses of bahuvrīhis

Previous analyses of bahuvrīhis can be subdivided into two subclasses: i. silent-suffix-based analyses, according to which bahuvrīhis derive from endocentric compounds via the application of a silent suffix (see, e.g., Whitney 1896, 501–502; Marchand 1967, 335; Kiparsky 1982, 139; Gillon 2007, 3); and ii. synecdoche-based analyses, according to which bahuvrīhis derive from endocentric compounds via a figure of speech, typically the synecdoche (e.g., Pennanen 1982, Bauer 2008).

A convenient way of describing silent-suffix-based analyses is by considering examples like *dīrgha-karṇa-ka-* (Lowe 2015b, 84): here the possessive suffix *-ka-* attaches to the stem *dīrgha-karṇa-* to produce a bahuvrīhi conveying the same meaning as the bahuvrīhi *dīrgha-karṇa-* ‘long-eared’ (see Gillon 2007, 3–4). This is parallel to the synonymy between English *red-head* and *red-headed*. To account for this synonymy, one may assume that *red-head* features a silent counterpart to the overt suffix *-ed* (which here has a possessive meaning), and that likewise *dīrgha-karṇa-* features a silent counterpart to the overt possessive suffix *-ka-*. On this analysis, therefore, the bahuvrīhi stem *dīrgha-karṇa-* would be understood as [[*dīrgha-karṇa-*]-Ø]. Now at least in principle, it is possible to form an endocentric compound *dīrgha-karṇa-* ‘(a) long ear’ (though this is in fact unattested). It is then possible to assume that the formation of a bahuvrīhi involves (overt or null) suffixation of a possessive suffix to an endocentric compound: the exocentricity of bahuvrīhis would then be a mere reflex of this suffixation.

A particularly clear example of figure-of-speech-based analyses of bahuvrīhis is Bauer (2008). He notes that the semantic exocentricity of bahuvrīhis is not specific to this special class of compounds, but also to single words. For example, the single word *crown* may denote not its ‘literal’ referent, but a monarch, i.e., a person habitually associated with wearing the ‘literal’ referent of the word. The shift in the denotation of *crown* is standardly explained as synecdoche; Bauer (2008, esp. 59) argues that the apparent shift in the denotation of bahuvrīhis can be explained in the same way. On this analysis, therefore, the bahuvrīhi-hood of a compound like *red-head* or *dīrgha-karṇa-* is contextually, rather than structurally, determined; from a structural point of view, bahuvrīhis would be identical to semantically endocentric compounds.

A critical assessment of silent-suffix-based analyses and figure-of-speech-based analyses cannot be attempted here (but see Candotti and Pontillo 2022, 5–9 on this point). Here we limit ourselves to noting that both these types of analysis posit

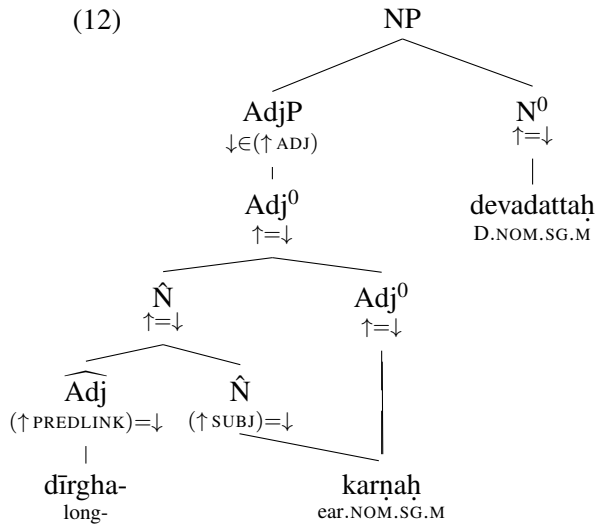
the existence of an endocentric compound serving as a sort of ‘substratum’ for the bahuvrīhi: it is to that substratum that a figure of speech or a silent possessive suffix applies to yield a bahuvrīhi. At least in Sanskrit, the postulation of such an endocentric substratum is problematic, as was already pointed out by Renou (1961, 114): it is very rare that, if a form is attested as bahuvrīhi, then that form is also attested as an endocentric compound (and vice versa). It is not therefore plausible to assume that the latter necessarily and productively serve as the base for the former.

4.3 A syntactic analysis of Sanskrit bahuvrīhis: Lowe (2015b)

Lowe’s (2015b) analysis of bahuvrīhis is based on Classical Sanskrit adjective-noun bahuvrīhis such as *dīrgha-karṇaḥ*, for which he proposes the following c- and f-structure (Lowe 2015b, 100–106):¹¹

$$(10) \quad \text{Adj}_{[_pr]} \rightarrow \begin{array}{ccc} \hat{N} & & \text{Adj}_{[_pr]} \\ \uparrow=\downarrow & & \uparrow=\downarrow \\ \widehat{\text{Adj}} \in \text{CAT}(\uparrow \text{ PREDLINK}) & & \end{array}$$

$$(11) \quad \hat{N} \rightarrow \begin{array}{ccc} \widehat{\text{Adj}} & & \hat{N} \\ (\uparrow \text{ PREDLINK}) = \downarrow & & (\uparrow \text{ PRED}) = \text{‘null-be} \langle \text{SUBJ, PREDLINK, OBL}_\theta \rangle \text{’} \\ & & (\uparrow \text{ SUBJ}) = \downarrow \\ & & (\uparrow \text{ DIS}) = (\uparrow \text{ OBL}_\theta) \\ & & (\uparrow \text{ DIS PRED}) = \text{‘pro’} \end{array}$$



¹¹We correct a typo in (11) = Lowe (2015b, 103, ex. 63), and use DIS in place of REL-TOP for consistency with Dalrymple et al. (2019). For Lowe’s analysis of other compound types, including dvandvas, tatpuruṣas, and avyayībhāvas, see Lowe (2015b, 94–100, 106).

$$(13) \left[\begin{array}{l} \text{PRED } 'Devadatta' \\ \text{ADJ } \left\{ \begin{array}{l} \text{DIS } \boxed{1} [\text{PRED } 'pro'] \\ \text{PRED } 'NULL-BE \langle \text{SUBJ, PREDLINK, OBL}_\theta \rangle' \\ \text{SUBJ } [\text{PRED } 'ear'] \\ \text{PREDLINK } [\text{PRED } 'long'] \\ \text{OBL}_\theta \boxed{1} \end{array} \right\} \end{array} \right]$$

The stem form *dīrgha-* instantiates the $\widehat{\text{Adj}}$ of (11), while the nominative case *karnaḥ* instantiates two nodes by the mechanism of Lexical Sharing (Wescoat 2002): the \hat{N} daughter in (11), and the $\text{Adj}_{[\text{pr}]}$ daughter in (10).¹² This $\text{Adj}_{[\text{pr}]}$ corresponds to Gillon’s (2007) silent possessive suffix *but*, unlike that suffix, does not bear any possessive meaning: $\text{Adj}_{[\text{pr}]}$ only serves the purpose of endowing bahuvrīhis with adjectival agreement properties. The use of the complex category for the top node of the bahuvrīhi is to parametrize projection/non-projection: in (13) the bahuvrīhi is not embedded in a larger compound, but if it were non-final in a larger compound, the $\text{Adj}_{[\text{pr}]}$ would be realised as $\widehat{\text{Adj}}$. Other technical details of this analysis, including the CAT predicate, are examined in §6 below.

5 Bahuvrīhis: types and patterns

As discussed above, Lowe (2015b) provides an analysis of only one type of bahuvrīhi, the Adj-N type such as *dīrgha-karna-* ‘whose ears are long’ (equivalently, ‘long-eared’), or *ugra-putra-* ‘whose sons are mighty’. There are two other major types of bahuvrīhi which we analyse in this paper: V-N bahuvrīhis, that is where the left-hand member of the compound is a verb, and N-N bahuvrīhis. The former also shows a set of forms with the order N-V, while the N-N type can be divided into at least two subgroups.

V-N bahuvrīhis are highly productive in Sanskrit; they typically involve a non-finite verb form, usually a perfect passive participle (PPP), as the left-hand member alongside a noun as right-hand member, which is construed as the subject of the action expressed by the (passive) verb. The external referent is usually interpreted as the agent of the passive expression, but may be interpreted as having other roles (see Candotti and Pontillo’s 2022, 18–19 discussion of *kṛtā-brahman-* in the *Rgveda*). For example: *rātā-havis-* ‘by whom an oblation (*havis-*) has been offered (*rātā-* to *rā* ‘give, offer’); *āhita-agni-* ‘by whom the fire (*agni-*) has been placed (*āhita-* to *ā-dhā* ‘place’) (on the altar).¹³

¹²Words that show lexical sharing are specified in the lexicon as associating with two terminal nodes in the c-structure.

¹³Since the external referent in such compounds has an argument role (agent, location, recipient, goal, etc.) with respect to the participle, and since the ordering constraint on such compounds is different from that of Adj-N bahuvrīhis, it is better to treat the participle in such compounds as V than as Adj (as suggested by an anonymous reviewer).

Building on Moro (1997, 2017) and Dikken (2006), we subdivide N-N bahuvrīhis into two subtypes, one in which the left-hand member is the subject of the predication and the right-hand member the predicate (call it the 'Subj-Pred' subtype), and one in which this subject-predicate relation is reversed (call it the 'Pred-Subj' subtype), with the external referent being interpreted as the possessor of either the subject or the predicate in both these subtypes. An example of the Pred-Subj subtype is *vṛṣaṇ-aśvā-* 'whose horses are bulls': it is the right-hand member 'horses' which is the subject, to which is ascribed the property of being '(like) bulls'. In contrast, *tvát-pitr-* 'whose father you are' exemplifies the Subj-Pred subtype of N-N bahuvrīhi: it is the left-hand member *tvád-* 'you' which is the subject, to which is ascribed the property of being a father (specifically, the father of the external referent); analogously, in *vṛtrá-putra-* 'of whom Vṛtra is the son' the left-hand member *vṛtrá-* is the subject to which the property of being a son (the son of the external referent) is ascribed (see Wackernagel 1905, 274–275). *vájra-hasta-* 'in whose hand is a mace' patterns with *tvát-pitr-* and *vṛtrá-putra-* in this respect: following Whitney (1896, 507–508), we take *vájra-hasta-* to be a 'locational' bahuvrīhi, where the right-hand member is interpreted as a locative predicate expressing the location of the subject left-hand member (cf. Bopp 1827, 320 n. 3 for an alternative analysis).

Notably, in both the Subj-Pred and the Pred-Subj subtype of N-N bahuvrīhi, it is the right-hand member of the compound which is interpreted as the possessum of the external referent ('whose horses', 'whose father', 'whose son', 'in whose hand'), regardless of whether that right-hand member is the subject or predicate of the bahuvrīhi itself. Adj-N bahuvrīhis show the same patterning as N-N bahuvrīhis: in *úgra-putra-* 'whose sons are mighty', it is the left-hand member which serves as the predicate while the right-hand member serves as the subject. In all cases it is the right-hand member of the compound with which the external referent enters a dependency: it is the right-hand member which is interpreted as the possessum of the external referent.

However, the situation is different when we consider V-N bahuvrīhis. In the standard type, e.g., *rātá-havis-* 'by whom an oblation has been offered', *áhita-agni-* 'by whom the fire has been placed (on the altar)', it is the left-hand member, the verb, which serves as the predicate of the bahuvrīhi and on which the external referent is directly dependent (usually, as mentioned above, as agent). With V-N bahuvrīhis it is also possible, though less common, to invert the order, e.g. *agny-áhita-* (*Mahābhārata* 12.281.21) which is synonymous with *áhita-agni-* (see Wackernagel 1905, 302–303 for other relevant examples). In the inverted order, it is the right-hand member which serves as the bahuvrīhi-internal predicate, and on which the external referent is dependent.¹⁴

¹⁴The external referent of such bahuvrīhis as *jāta-danta-* 'whose teeth have been born' (i.e. 'who has teeth'; *Vaikhānasasmārtasūtra* 7.2) and *hata-putra-* 'whose sons have been killed' (*Kauṣṭhikībrāhmaṇa* 4.6.7) is apparently dependent (as a possessor) on the bahuvrīhi-internal N (i.e., *danta-* and *putra-*) and not to the participle (i.e., to *jāta-* and *hata-*). Interestingly, these bahuvrīhis also have synonymous inverted counterparts: *danta-jāta-* (*Vaikhānasasmārtasūtra* 5.9); *putra-hata-*

These facts show that the internal order of Sanskrit bahuvrīhis obeys nontrivial restrictions, which can be descriptively phrased as follows (see Mocci 2022):

(14) ORDERING GENERALIZATION:

(i) Bahuvrīhis whose predicate is nominal (Adj or Noun): If a bahuvrīhi's external referent is interpreted as a dependent of the bahuvrīhi member X, X is allocated to the right-hand slot of the bahuvrīhi.

(ii) Bahuvrīhis whose predicate is verbal: If a bahuvrīhi's external referent is interpreted as a dependent of the bahuvrīhi member X, X is typically allocated to the left-hand slot of the bahuvrīhi, but may sometimes be allocated to the right-hand slot of the bahuvrīhi.

In brief, the constraint on the ordering of Adj-N and N-N (but not V-N) bahuvrīhis is determined by the dependency between the external referent of the bahuvrīhi and one of the compound members. This fact cannot be captured by the f-structure proposed by Lowe (2015b) for bahuvrīhis, shown in (13), since it does not codify any dependency between the bahuvrīhi's external referent and one particular bahuvrīhi member. In what follows we propose modified f-structures for Adj-N and N-N bahuvrīhis to resolve this issue.

6 Analysis

Consider again (10)–(13) from Lowe (2015b, 103–104). Lowe's (2015b) analysis assumes two PS-rules, (10) and (11), which are constrained to co-occur by means of the CAT predicate in (10).¹⁵ This works, at least as an ad hoc means for ensuring co-occurrence of the PS-rules, in the case of Adj-N bahuvrīhis, but cannot be immediately extended to all bahuvrīhi types. And if the principle of constraining co-occurrence in this way is extended to other types, we will end up with at least three pairs of rules, one for each of the three major bahuvrīhi types, with no single, general rule for the bahuvrīhi category as a whole.

An additional concern with Lowe's (2015b) analysis is the label \hat{N} as the left-most daughter in (10) and mother node in (11). This implies an underlying endocentric base to bahuvrīhis (Lowe 2015b, 103); as discussed in §4.2, this is a problematic assumption. Furthermore, (11) in fact produces an exocentric f-structure, in the sense that neither daughter node in (11) projects to the same f-structure as the mother. The label \hat{N} therefore appears to be unwarranted. What we want, rather, is an exocentric c-structure node. Such an exocentric node will be necessarily non-projecting, as it occurs only within the non-projecting syntax of compound phrases. We propose the label B for this exocentric node. This means: a) we can have a single phrase structure rule in place of (10) which introduces all types of bahuvrīhi; b)

(*Pañcaviṃśabrāhmaṇa* 4.7.3). These bahuvrīhis deserve further study.

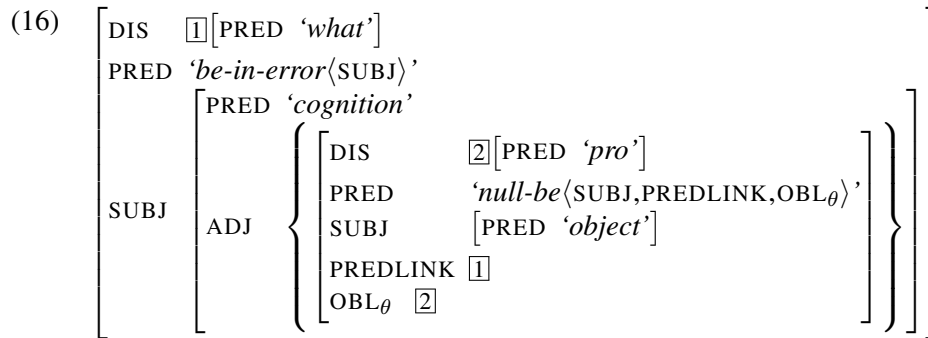
¹⁵The CAT predicate requires that the f-structure (\uparrow PREDLINK) be related to a c-structure node \widehat{Adj} by the relation ϕ^{-1} (for the definition of CAT see Dalrymple et al. 2019, 247–250). Granted that this can only be satisfied by the rule in (11), (11) must necessarily expand the \hat{N} in (10).

there is no need for the CAT predicate to constrain the possible expansions of the bahuvrīhi node, and c) the exocentricity of the bahuvrīhi is appropriately captured in the c-structure.¹⁶

The second major difference in our approach from that of Lowe (2015b) concerns the f-structure relation between the bahuvrīhi and its external referent. As in his analysis of participles (Lowe 2015a), Lowe (2015b) assumes a null-relative analysis of bahuvrīhis; that is, as an adjectival modifier the bahuvrīhi constitutes a null-headed relative clause in the f-structure. But as with participles (Haug and Nikitina 2012), an alternative option is to assume a cyclical f-structure. In the case of Sanskrit bahuvrīhis, there is some evidence in favour of a cyclical f-structure. The following example is given by Lowe (2015b, 79):

- (15) *[yad_i viṣayā]_{bv} buddhir na vyabharati tat_i sat*
RELPRO- object.NOM cognition.NOM not be_in_error.3SG that real
‘A thing_i which, when a cognition_j that has it_i as its_j object is not in error, that thing_i is real.’
(Lit. ‘a cognition having-what_i-as-its-object is not in error, that_i is real.’)

Lowe (2015b) does not analyse this example, but if it were analysed according to his proposal, we would end up with the following for the relative clause:



Here, the in-clause ‘gap’ corresponding to the relative pronoun *yad* ‘what’ is inside an f-structure which itself contains the head of another long-distance dependency, that which (on this view) models the function of the bahuvrīhi. This violates the Complex NP Constraint (CNPC; Ross 1967, 127), which rules out any kind of

¹⁶Our understanding of B is as the equivalent, within the non-projecting syntax of compounds, of the exocentric clausal node S in projecting (noncompound) syntax. We denote this as B, and not as e.g. \hat{S} or \hat{B} , following the implications of the proposals of Lovstrand and Lowe (2017) and Lowe and Lovstrand (2020), according to which non-projection and exocentricity are incompatible; that is, we cannot have an explicitly non-projecting exocentric label like \hat{B} or \hat{S} . An alternative to assuming a new c-structure category type, suggested to us by Mary Dalrymple, would be to treat our B as a metacategory; bahuvrīhis would then involve a ternary branching structure, but as a metacategory B would enable us to state generalizations and constraints over the first two daughters.

extraction from a relative clause.¹⁷

Since the outer relative clause in (15) is explicit, it appears preferable to analyse the bahuvrīhi in such a way that the whole structure does not violate the CNPC, which means abandoning the null-relative analysis of Lowe (2015b). We therefore assume a cyclical f-structure in the analyses given below.¹⁸

6.1 Adj-N

We begin by presenting our rules for the analysis of Adj-N bahuvrīhis, the only type treated by Lowe (2015b). As discussed above, we propose a single main rule for all bahuvrīhis, introducing an exocentric category B, in place of (10):

$$(17) \quad \text{Adj}_{[\text{pr}]} \rightarrow \begin{array}{cc} \text{B} & \text{Adj}_{[\text{pr}]} \\ \uparrow=\downarrow & \uparrow=\downarrow \end{array}$$

For Adj-N bahuvrīhis, we assume the PS-rule in (18), which for the phrase *úgra-putre adite* ‘Aditi, whose sons are mighty’ produces the f-structure in (19).

$$(18) \quad \text{B} \rightarrow \begin{array}{cc} \widehat{\text{Adj}} & \hat{\text{N}} \\ (\uparrow \text{PREDLINK}) = \downarrow & (\uparrow \text{PRED}) = \text{‘null-be(SUBJ,PREDLINK)’} \\ & (\uparrow \text{SUBJ}) = \downarrow \\ & (\text{ADJ} \in \uparrow) = (\downarrow \text{POSS}) \end{array}$$

$$(19) \quad \left[\begin{array}{c} \text{PRED} \text{ ‘Aditi’} \\ \text{ADJ} \left\{ \begin{array}{l} \text{PRED} \text{ ‘null-be(SUBJ,PREDLINK)’} \\ \text{SUBJ} \left[\begin{array}{c} \text{PRED} \text{ ‘sons’} \\ \text{POSS} \text{ [1]} \end{array} \right] \\ \text{PREDLINK} \left[\text{PRED} \text{ ‘mighty’} \right] \end{array} \right\} \end{array} \right]$$

Lowe (2015b) treats the external referent of Adj-N bahuvrīhis as corresponding to an OBL_θ argument within the bahuvrīhi-internal predicative nucleus. However, given that we need the OBL_θ for the external referent in the case of V-N bahuvrīhis

¹⁷The CNPC was originally formulated as a condition on phrase structure, because the gap corresponding to the relative pronoun and constrained by the CNPC was represented in phrase structure in the framework of Ross (1967). However, since that gap is represented in f-structure in the LFG framework, we take the CNPC to hold at f-structure here. We are thankful to Miriam Butt for drawing our attention to this point.

¹⁸One may be tempted to claim that Sanskrit does not obey strong islands like the CNPC on the grounds that a subject island (which is also a strong island) appears to be violated in the following example from Davison (2009, 235) (see Cinque 1990, 1–2, 161 n.2 on the distinction between strong and weak islands): *yát kím ákaram̐ tásmād idám ápat* ‘The fact_i that I did what_i is the reason_i this happened?’ (*Śatapathabrāhmaṇa* 4.1.5.4). However, this example involves extraction from a relative-correlative structure, which is subject to different (and apparently weaker) constraints from ‘ordinary’ relative clauses, as pointed out to us by Mary Dalrymple. Thus, an analysis of Sanskrit bahuvrīhis which avoids a strong island violation is preferable. Note that although (15) also involves a relative-correlative structure, the CNPC-violating extraction is from the (supposed) null-relative clause of the bahuvrīhi, and not from the relative clause involved in the correlative construction.

(below), it makes more sense to treat the external referent here rather as a POSS (i.e., possessor – understood in the widest possible sense) of the bahuvrīhi-internal subject. This will ultimately permit a formalization of the fact, made explicit in generalization (14i) (see §5 above), that the internal order of Adj-N (and N-N) bahuvrīhis is constrained by the dependency holding between the external referent and one of the bahuvrīhi members.

6.2 N-N

In §5 we distinguished two types of N-N bahuvrīhi: Subj-Pred and Pred-Subj. In addition to (17), which applies for all bahuvrīhis, we assume the following PS-rules, corresponding respectively to *tvāt-pitr-* ‘whose father you are’ (Subj-Pred), *vṛṣan-aśvā-* ‘whose horses are (like) bulls’ (Pred-Subj), and *vájra-hasta-* ‘in whose hand is a mace’ (Subj-Pred, locative). The corresponding f-structures follow.¹⁹

$$(20) \quad B \rightarrow \begin{array}{cc} \hat{N} & \hat{N} \\ (\uparrow \text{SUBJ}) = \downarrow & (\uparrow \text{PRED}) = \text{'null-be'} \langle \text{SUBJ, PREDLINK} \rangle' \\ & (\uparrow \text{PREDLINK}) = \downarrow \\ & (\text{ADJ} \in \uparrow) = (\downarrow \text{POSS}) \end{array}$$

$$(21) \quad B \rightarrow \begin{array}{cc} \hat{N} & \hat{N} \\ (\uparrow \text{PREDLINK}) = \downarrow & (\uparrow \text{PRED}) = \text{'null-be'} \langle \text{SUBJ, PREDLINK} \rangle' \\ & (\uparrow \text{SUBJ}) = \downarrow \\ & (\text{ADJ} \in \uparrow) = (\downarrow \text{POSS}) \end{array}$$

$$(22) \quad B \rightarrow \begin{array}{cc} \hat{N} & \hat{N} \\ (\uparrow \text{SUBJ}) = \downarrow & (\uparrow \text{PRED}) = \text{'null-be'} \langle \text{SUBJ, PREDLINK} \rangle' \\ & (\uparrow \text{PREDLINK}) = \downarrow \\ & (\downarrow \text{CASE}) = \text{LOC} \\ & (\text{ADJ} \in \uparrow) = (\downarrow \text{POSS}) \end{array}$$

$$(23) \quad \boxed{1} \left[\begin{array}{l} \text{PRED 'gods'} \\ \text{ADJ} \left\{ \begin{array}{l} \text{PRED 'null-be'} \langle \text{SUBJ, PREDLINK} \rangle' \\ \text{SUBJ} \left[\text{PRED 'you'} \right] \\ \text{PREDLINK} \left[\begin{array}{l} \text{PRED 'father'} \\ \text{POSS } \boxed{1} \end{array} \right] \end{array} \right\} \end{array} \right]$$

¹⁹These rules license a degree of ambiguity, of course, since any of the three rules might equally well apply to any N-N compound. To some extent this is desirable, since ambiguity is, in principle, possible: given an appropriate context, for example, *vájra-hasta-* could conceivably mean ‘whose hand is a mace’. But in most cases, the ambiguity will be constrained at the level of interpretation, since only one f-structure will give a coherent meaning.

$$(24) \quad \left[\begin{array}{l} \text{PRED 'chariot'} \\ \text{ADJ} \left\{ \begin{array}{l} \text{PRED 'null-be' (SUBJ, PREDLINK)'} \\ \text{SUBJ} \left[\begin{array}{l} \text{PRED 'horses'} \\ \text{POSS } \boxed{1} \end{array} \right] \\ \text{PREDLINK} \left[\text{PRED 'bulls'} \right] \end{array} \right\} \end{array} \right]$$

$$(25) \quad \left[\begin{array}{l} \text{PRED 'Indra'} \\ \text{ADJ} \left\{ \begin{array}{l} \text{PRED 'null-be' (SUBJ, PREDLINK)'} \\ \text{SUBJ} \left[\text{PRED 'mace'} \right] \\ \text{PREDLINK} \left[\begin{array}{l} \text{PRED 'hand'} \\ \text{CASE LOC} \\ \text{POSS } \boxed{1} \end{array} \right] \end{array} \right\} \end{array} \right]$$

6.3 V-N

As anticipated in §5 above, V-N bahuvrīhis – where V is typically realized as a (perfect passive) participle – involve different possible semantic roles of the external referent in relation to the event denoted by the participle. For instance, the external referent of *rātá-havis-* is interpreted as an agent of the event of offering denoted by *rātá-*. We represent the variable semantic role of the external referent as OBL_θ in the f-structure.

$$(26) \quad \text{B} \rightarrow \begin{array}{cc} \hat{\text{V}} & \hat{\text{N}} \\ \uparrow=\downarrow & (\uparrow \text{SUBJ}) = \downarrow \\ (\text{ADJ} \in \uparrow) = (\uparrow \text{OBL}_\theta) & \end{array}$$

$$(27) \quad \left[\begin{array}{l} \text{PRED 'person'} \\ \text{ADJ} \left\{ \begin{array}{l} \text{PRED 'offer' (SUBJ, OBL}_\theta\text{)'} \\ \text{SUBJ} \left[\text{PRED 'oblation'} \right] \\ \text{OBL}_\theta \boxed{1} \end{array} \right\} \end{array} \right]$$

In this way, the f-structure for V-N bahuvrīhis like *rātá-havis-* differs from the one for Adj-N and N-N bahuvrīhis: the element which the external referent is matched with is an immediate member of the f-structure set that serves as the value of ADJ in V-N bahuvrīhis; on the other hand, such an element is more deeply embedded in the f-structure set serving as the value of ADJ in Adj-N and N-N bahuvrīhis.

7 Modelling the ordering generalization

Within the analysis we have proposed, the ordering generalization on the internal members of Adj-N and N-N bahuvrīhis can be neatly modelled. That is, we

can formalize the generalization itself as a generalization over all bahuvrīhi types, which provides a deeper level of analysis than simply varying the details of the rules for the different types.

Recall that Adj-N and N-N bahuvrīhis are subject to a strict constraint which V-N bahuvrīhis are not: for Adj-N and N-N bahuvrīhis, the internal member on which the external referent is dependent (usually as possessor) must be the right-hand member of the compound. In the case of the V-N bahuvrīhis analysed here, the external referent is always dependent on the V (as a complement), and both V-N and N-V orders are possible.

We note first that, given the analysis we have proposed, there is a difference in the f-structures between Adj-N/N-N and V-N bahuvrīhis. In the former case (28), the f-structure of the external referent inside the bahuvrīhi is embedded within another f-structure (f) which itself has some grammatical function (GF) within the f-structure of the bahuvrīhi, and this f is different from the f-structure g which represents the left-hand member of the bahuvrīhi. In the latter case (29), the equivalent f-structure f is the outer f-structure for the bahuvrīhi (that is, the external referent is less deeply embedded), and this is also the same as the f-structure for the (usual) left-hand member of the compound (i.e. here the verb).

$$(28) \quad \left[\begin{array}{c} \text{PRED } 'Devadatta' \\ \vdots \\ \text{ADJ } \left\{ \begin{array}{l} \text{GF } g \left[\begin{array}{c} \vdots \\ \vdots \end{array} \right] \\ \text{GF } f \left[\begin{array}{c} \vdots \\ \vdots \\ \text{GF } \boxed{1} \end{array} \right] \end{array} \right\} \end{array} \right] \quad (29) \quad \left[\begin{array}{c} \text{PRED } 'Devadatta' \\ \text{ADJ } \left\{ f, g \left[\begin{array}{c} \vdots \\ \vdots \\ \text{GF } \boxed{1} \end{array} \right] \right\} \end{array} \right]$$

We can then capture the difference between the Adj-N/N-N and V-N bahuvrīhis by reference to f-precedence:

- (30) F-precedence (Bresnan et al. 2016, 213):
 f f-precedes g if and only if the rightmost node in $\phi^{-1}(f)$ precedes the rightmost node in $\phi^{-1}(g)$.

Rather than arbitrarily constrain the order of daughters only in the Adj-N and N-N bahuvrīhis, we can treat the ordering of daughters in bahuvrīhis of all types as unspecified, but we add the following constraint to what we have given above as the leftmost daughters of each type:

$$(31) \quad (\text{GF} \quad (\text{ADJ} \in \uparrow)) \not\prec_f \downarrow \\ ((\text{GF} \leftarrow) \text{GF}) = \downarrow$$

The left-hand side of this constraint finds the f-structure which contains the f-structure of the external referent inside the bahuvrīhi, that is f in (28) and (29). This is complicated because by the cyclic f-structure we have assumed, the f-structure

of the external referent appears in two places, within the bahuvrīhi and outside, containing the bahuvrīhi, and also because the relation between the f-structure for the left-hand member and that of the external referent inside the bahuvrīhi is variable. Essentially, (GF (ADJ ∈ ↑)) finds an f-structure containing the f-structure for the external referent (ADJ ∈ ↑). The off-path constraint ensures that the f-structure found is the one that occurs inside the bahuvrīhi, and not any f-structure that contains the external referent further out. It does this by ensuring that there is a local relation between the f-structure in question and the f-structure representing the left-hand member of the bahuvrīhi. Constraint (31) then requires that this f-structure (*f* above) does not f-precede the f-structure representing the left-hand member of the bahuvrīhi (*g* above). Take first the case of an N-N bahuvrīhi:

$$(32) \quad B \rightarrow \begin{array}{l} \hat{N} \\ (\uparrow \text{ PREDLINK}) = \downarrow \\ (\text{ GF } (\text{ ADJ } \in \uparrow)) \not\prec_f \downarrow \\ ((\text{ GF } \leftarrow) \text{ GF}) = \downarrow \end{array} \quad \begin{array}{l} \hat{N} \\ (\uparrow \text{ PRED}) = \text{ 'null-be(SUBJ, PREDLINK)'} \\ (\uparrow \text{ SUBJ}) = \downarrow \\ (\text{ ADJ } \in \uparrow) = (\downarrow \text{ POSS}) \end{array}$$

$$(33) \quad \boxed{\text{I}} \left[\begin{array}{l} \text{ PRED 'chariot'} \\ \text{ ADJ } \left\{ \begin{array}{l} \text{ PRED 'null-be(SUBJ, PREDLINK)'} \\ \text{ SUBJ } f \left[\begin{array}{l} \text{ PRED 'horse'} \\ \text{ POSS } \boxed{\text{I}} \end{array} \right] \\ \text{ PREDLINK } g \left[\text{ PRED 'bull'} \right] \end{array} \right\} \end{array} \right]$$

Since *f* must not f-precede *g*, we cannot invert. The same will apply to the other N-N bahuvrīhis, and to the Adj-N bahuvrīhis. In contrast, in the case of V-N bahuvrīhis, *g* = *f*, so necessarily *f* does not f-precede *g*:

$$(34) \quad B \rightarrow \begin{array}{l} \hat{V} \\ \uparrow = \downarrow \\ (\text{ GF } (\text{ ADJ } \in \uparrow)) \not\prec_f \downarrow \\ ((\text{ GF } \leftarrow) \text{ GF}) = \downarrow \end{array} \quad \begin{array}{l} \hat{N} \\ (\uparrow \text{ SUBJ}) = \downarrow \\ (\text{ ADJ } \in \uparrow) = (\uparrow \text{ OBL}_\theta) \end{array}$$

$$(35) \quad \boxed{\text{I}} \left[\begin{array}{l} \text{ PRED 'person'} \\ \text{ ADJ } \left\{ \begin{array}{l} f, g \left[\begin{array}{l} \text{ PRED 'offer(SUBJ, OBL}_\theta) \text{' } \\ \text{ SUBJ } \left[\text{ PRED 'oblation'} \right] \\ \text{ OBL}_\theta \boxed{\text{I}} \end{array} \right] \end{array} \right\} \end{array} \right]$$

Therefore the order of the daughters is not constrained, licensing both V-N and N-V orderings, as attested (e.g., *āhita-agni-* vs. *agny-āhita-*).

8 Conclusion

Following Lowe (2015b), Sanskrit compounding is best treated as a syntactic process, but one that is somewhat less fully syntactic than the standard phrasal syntax

of the language. This is captured by recourse to nonprojecting categories, which allow us to model syntactic structures which do not involve full XP phrases and which are headed by lexical (X^0) categories. Revising Lowe's (2015b) analysis of Sanskrit bahuvrīhi, we have a) modeled all major types of Sanskrit bahuvrīhi, Adj-N, N-N and V-N in a consistent way, without positing an unsubstantiated endocentric core to bahuvrīhi; b) avoided the problems associated with a null-relative analysis of bahuvrīhi; and c) captured the ordering constraint on the members of bahuvrīhi as a deeper generalization rather than simply as an ad hoc difference between rules for different bahuvrīhi types.

References

- Allen, Margaret Reece. 1978. *Morphological investigations*. Ph. D.thesis, University of Connecticut, Storrs.
- Arnold, Doug and Sadler, Louisa. 2013. Displaced dependent constructions. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG13 Conference*, pages 48–68, Stanford, CA: CSLI Publications.
- Baker, Brett and Nordlinger, Rachel. 2008. Noun-adjective compounds in Gunwinyguan languages. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG08 Conference*, pages 109–128, Stanford, CA: CSLI.
- Bauer, Laurie. 2008. Exocentric compounds. *Morphology* 18, 51–74.
- Bauer, Laurie. 2017. *Compounds and compounding*. Cambridge: CUP.
- Benfey, Theodor. 1852. *Vollständige Grammatik der Sanskritsprache. Zum Gebrauch für Vorlesungen und zum Selbststudium*. Leipzig: Brockhaus.
- Bopp, Franz. 1827. *Ausführliches Lehrgebäude der Sanskrita Sprache*. Berlin: Dümmler.
- Bresnan, Joan, Asudeh, Ash, Toivonen, Ida and Wechsler, Stephen. 2016. *Lexical-functional syntax*. Oxford: Wiley-Blackwell.
- Butt, Miriam, King, Tracy Holloway, Niño, María-Eugenia and Segond, Frédérique. 1999. *A Grammar Writer's Cookbook*. Stanford, CA: CSLI.
- Candotti, Maria Piera and Pontillo, Tiziana. 2019. Lexical subordination in derivation and compounding. *Studi e Saggi Linguistici* LVIII(2), 155–175.
- Candotti, Maria Piera and Pontillo, Tiziana. 2022. Dispensing with zero in the analysis of Sanskrit bahuvrīhi: Resurfacing, testing and assessing Pāṇini's model. *Journal of South Asian Linguistics* 12(1), 1–22.
- Cardona, George. 1997. *Pāṇini: His work and its tradition, vol. 1: Background and introduction*. New Delhi: Motilal Banarsidass.
- Cinque, Guglielmo. 1990. *Types of \bar{A} -dependencies*. Cambridge, MA: MIT Press.
- Dalrymple, Mary, Lowe, John J. and Mycock, Louise. 2019. *The Oxford reference guide to Lexical Functional Grammar*. Oxford: Oxford University Press.
- Davison, Alice. 2009. Sanskrit and Hindi/Urdu correlatives. In Anikó Lipták (ed.), *Correlatives cross-linguistically*, pages 223–262, Amsterdam: Benjamins.

- Dikken, Marcel den. 2006. *Relators and linkers: The syntax of predication, predicate inversion, and copulas*. Cambridge, MA: MIT Press.
- Duncan, Lachlan. 2007. Analytic noun incorporation in Chuj and K'ichee' Mayan. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG07 Conference*, pages 163–183, Stanford, CA: CSLI Publications.
- Eik, Ragnhild. 2019. *The morphosyntax of compounding in Norwegian*. Ph.D.thesis, NTNU, Trondheim.
- Gillon, Brendan S. 1994. Bharṭṛhari's solution to the problem of *asamartha* compounds. In Saroja Bhate and Johannes Bronkhorst (eds.), *Bharṭṛhari, philosopher and grammarian*, pages 117–133, Delhi: Motilal Banarsidass.
- Gillon, Brendan S. 2007. Exocentric (bahuvrīhi) compounds in Classical Sanskrit. In Gérard Huet and Amba Kulkarni (eds.), *Proceedings of the First International Sanskrit Computational Linguistics Symposium, Paris, October 29-31, 2007*, pages 1–12, INRIA.
- Guevara, Emiliano and Scalise, Sergio. 2009. Searching for universals in compounding. In *Universals of language today. Studies in Natural Language and Linguistic Theory 76*, pages 101–128.
- Haug, Dag T. T. and Nikitina, Tanya. 2012. The many cases of non-finite subjects: The challenge of “dominant” participles. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG12 Conference*, pages 292–311, Stanford, CA: CSLI Publications.
- Kiparsky, Paul. 1982. From cyclic phonology to lexical phonology. In Harry van der Hulst and Norval Smith (eds.), *The structure of phonological representations, vol. 1*, pages 131–175, Dordrecht: Foris.
- Kiparsky, Paul. 2009. On the architecture of Pāṇini's grammar. In Gérard Huet, Amba Kulkarni and Peter Scharf (eds.), *Sanskrit computational linguistics*, pages 33–94, Berlin: Springer.
- Lapointe, Steven G. 1980. *The theory of grammatical agreement*. Ph.D.thesis, University of Massachusetts Amherst.
- Lee, Leslie and Ackerman, Farrell. 2011. Mandarin resultative compounds: A family of lexical constructions. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG11 Conference*, pages 320–338, Stanford, CA: CSLI.
- Lovestrand, Joseph and Lowe, John. 2017. Minimal c-structure: Rethinking projection in phrase structure. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG17 Conference*, pages 285–305, Stanford, CA: CSLI.
- Lowe, John. 2015a. *Participles in Rigvedic Sanskrit: The syntax and semantics of adjectival verb forms*. Oxford: Oxford University Press.
- Lowe, John. 2015b. The Syntax of Sanskrit compounds. *Language* 91(3), e71–e115.
- Lowe, John and Lovestrand, Joseph. 2020. Minimal phrase structure: a new formalized theory of phrase structure. *Journal of Language Modelling* 8(1), 1–52.
- Marchand, Hans. 1967. Expansion, transposition, and derivation. In D. Kastovsky (ed.), *Studies in syntax and word-formation. Selected articles by H. Marchand*,

- pages 322–337, München: Fink.
- Mocci, Davide. 2022. The possessive relation in Sanskrit bahuvrīhi compounds: Ellipsis or movement? *Journal of South Asian Linguistics* 12(2), 1–30.
- Molina-Muñoz, Adriana. 2013. Sanskrit compounds and the architecture of the grammar. In S.-F. Chen and B. Slade (eds.), *Grammatica and verba. Glamour and verve: Studies in South Asian, historical, and Indo-European linguistics in honor of Hans Henrich Hock on the occasion of his seventy-fifth birthday*, pages 181–201, Ann Arbor: Beech Stave Press.
- Moro, Andrea. 1997. *The raising of predicates: Predicative noun phrases and the theory of clause structure*. Cambridge: Cambridge University Press.
- Moro, Andrea. 2017. *A brief history of the verb to be*. Cambridge, MA: MIT Press, translated by B. McClellan-Broussard.
- Ørnsnes, Bjarne. 1996. Prominence relations and a-structure - evidence from Danish synthetic compounding. In Miriam Butt and Tracy Holloway King (eds.), *Proceedings of the LFG96 Conference*, Stanford, CA: CSLI Publications.
- Pennanen, Esko V. 1982. Remarks in syntagma and word-formation. *Folia Linguistica* 16, 241–261.
- Pontillo, Tiziana. 2021. Did the Sanskrit model bring “true enlightenment to European scholars” when they analysed and classified the bahuvrīhi compounds? *Studi Classici e Orientali* 67, 497–514.
- Postal, Paul. 1969. Anaphoric islands. *Chicago Linguistic Society* 5, 205–239.
- Renou, Louis. 1961. *Grammaire sanscrite*. Paris: Adrien-Maisonneuve, second edition: revised, corrected and augmented.
- Ross, John Robert. 1967. *Constraints on variables in syntax*. Ph.D.thesis, MIT.
- Scalise, Sergio and Bisetto, Antonietta. 2009. The classification of compounds. In R. Lieber and P. Štekauer (eds.), *The Oxford handbook of compounding*, pages 34–53, Oxford: Oxford University Press.
- ten Hacken, Pius. 2009. Early generative approaches. In R. Lieber and P. Štekauer (eds.), *The Oxford handbook of compounding*, pages 54–77, Oxford: Oxford University Press.
- Toivonen, Ida. 2003. *Non-projecting words: A case study of Swedish verbal particles*. Dordrecht: Kluwer.
- Wackernagel, Jacob. 1905. *Altindische Grammatik. Band II:1 – Einleitung zur Wortlehre, Nominalkomposition*. Göttingen: Vandenhoeck & Ruprecht.
- Wescoat, Michael Thomas. 2002. On lexical sharing. Ph.D. Thesis, Stanford University.
- Whitney, William Dwight. 1896. *A Sanskrit Grammar: Including both the Classical language, and the older dialects, of Veda and Brahmana*. Leipzig: Breitkopf & Härtel, third edition.
- Williams, Edwin. 2007. Dumping lexicalism. In G. Ramchand and C. Reiss (eds.), *The Oxford handbook of linguistic interfaces*, pages 353–381, Oxford: Oxford University Press.