The possessive relation in Sanskrit bahuvrīhi compounds:
Ellipsis or movement?

DAVIDE MOCCI, *SCUOLA UNIVERSITARIA SUPERIORE IUSS PAVIA*

ABSTRACT

Many Sanskrit bahuvrīhis involve a possessive relation whereby one of the bahuvrīhi-members is the possessum and an expression not mentioned within the bahuvrīhi is the corresponding possessor: e.g., *ugra-putra-* (ṚV 8.67.11), not 'mighty son(s)' but 'Aditi having mighty sons' or 'Aditi whose sons are mighty'. This study addresses the following research question: how is this possessive relation established in Sanskrit bahuvrīhis? We consider two possible strategies. According to the first strategy, a linguistic unit which conveys the meaning 'having' and undergoes ellipsis combines with the bahuvrīhi stem: e.g., the combination of this elided unit with *ugra-putra-*, which per se would convey the meaning 'mighty son(s)', yields the meaning 'having mighty sons'. According to the second strategy, the possessor starts out within the phrase projected by one of the bahuvrīhi-members: e.g., *áditi-* (i.e., the Sanskrit term for 'Aditi') starts out as the specifier of the phrase projected by *putrá-* in the above example; in this configuration *áditi-* is read as the possessor of *putrá-*; only subsequently will *áditi-* exit the bahuvrīhi. We argue that the second strategy is superior because only it captures certain restrictions on the internal order of bahuvrīhis.

1 Introduction

Compounding is an extremely productive phenomenon in Sanskrit: in fact, "almost any meaning that can be expressed using two or more separate words can also be expressed using a single compound" (Lowe 2015:72) in this language. For this reason, special attention has been devoted to Sanskrit compounding in the Indian grammatical tradition (starting from Pāṇini's section A 2.1-2) as well as in modern linguistic treatises such as Wackernagel (1905). The development of formal theories of language in the second half of the Twentieth Century has offered a fresh take on this subject matter. Thus Gillon (2008) capitalises on theories of word-formation such as Selkirk (1982) and Di Sciullo and Williams (1987) to formulate explicit constraints on the interpretation of Sanskrit bahuvrīhi compounds. On the other hand, Lowe (2015) analyses Sanskrit major compound types within the framework of Lexical-Functional Grammar, showing that they are generated by syntactic rules (i.e., the rules that combine words into sentences), rather than by morphological rules (i.e., the rules that combine morphemes into words).

Pursuing this line of research, in this work we provide a formal analysis of Sanskrit bahuvrīhi compounds (or simply bahuvrīhis), an exemple of which is *ugra-putra-* in (1) (for an explicit definition of bahuvrīhis, see §2 below).[1] The bahuvrīhi *ugra-putra-* is formed by combining *ugrá-* 'mighty' with *putrá-* 'son(s)'. However, *ugra-putra-* does not denote a mighty son, but rather a possessor of mighty sons. The information as to who this possessor is (i.e., the identity of this possessor) is not contained within the bahuvrīhi *ugra-putra-*: this information is supplied by a

---

[1] We use the hyphen at the end of a Sanskrit form (e.g., *ugrá-*) to cite that form as a stem (for an explicit definition of stem, see §2.1 below). *ugra-putra-* is accented on its first syllable (i.e., *úgraputre*) in (1) because it is a pāda-initial vocative (Macdonell 1910:81, 97). Since the accent displayed by a form inflected in the vocative and occupying the pāda-initial position need not coincide with the normal accent of that form, and since (1) constitutes the only occurrence of *ugra-putra-* as a bahuvrīhi in accented sources, we cannot be certain as to which syllable *ugra-putra-* is normally accented on. We therefore omit the accent of *ugra-putra-* when we quote this form outside (1). More on the accent of bahuvrīhis in §2.2 below.

bahuvrīhi-external expression. Thus, *adite* in the passage reported in (2) reveals that the possessor of mighty sons is indeed the goddess Aditi.

(1)   úgra-putre   (adite)   (ṚV 8.67.11)
       mighty-son.VOC.SG.F   Aditi.VOC.SG.F
       'O Aditi, having mighty sons (or 'O Aditi, whose sons are mighty').

(2) *utá tvắm adite mahi/ aháṃ devi úpa bruve/ sumṛḷīkắm abhíṣṭaye// párṣi dīné gabhīrá ắ̃/ úgra-putre jíghāṃsataḥ/ mắkis tokásya no riṣat//*
'And, great goddess Aditi, I entreat you, the very merciful, (for us) to prevail. Deliver us, whether we're in the shallows or the deep, from someone who wishes to smite us, o you who have powerful sons. Let none of our offspring be injured.' (ṚV 8.67.10-11; tr. Jamison and Brereton 2014).

The bahuvrīhi-external expression that reveals the identity of the possessor denoted by the bahuvrīhi is known as the external referent of the bahuvrīhi. Thus, *adite* is the external referent of *ugra-putre* in (1). The brackets enclosing *adite* in (1) signal that *adite* is not actually adjacent to *ugra-putre* in the line quoted. We shall follow the same practice when exemplifying bahuvrīhis throughout this work: we shall systematically pair bahuvrīhis with their external referent, and bracket this external referent in case it is not adjacent, in the textual passage quoted, to the bahuvrīhi it is associated with.

A possessive relation is implied in *ugra-putra-*: *putrá-* 'son(s)' is the possessum and the external referent *adite* is the corresponding possessor in (1).[2] However, no overt linguistic unit signals the presence of the possessive relation in this example: there is no overt verb or suffix which conveys the meaning 'having' in (1), and no relative pronoun corresponding to *whose* of the English translation shows up either. We then wonder: how is the possessive relation that is perceived in bahuvrīhis like (1) established? At least two solutions come to mind, namely (3a) and (3b) (a third solution is considered in §2.4 below).

(3) a. <u>The Elided Possessive Unit Strategy</u>
     A linguistic unit that conveys the meaning 'having' undergoes ellipsis: to wit, the meaning 'having' conveyed by this unit is perceived, although this unit is not phonologically realised. Let us call this elided unit B. The combination of the bahuvrīhi stem, which conveys the meaning 'X', with B causes the bahuvrīhi to acquire the meaning 'having X'.
     b. <u>The Movement Strategy</u>
     An expression E_1 may be read as the possessor of another expression E_2 (E_2 being the possessum of E_1) when E_1 occupies the specifier of the phrase projected by E_2. The possessor (E_1) subsequently moves out of the phrase projected by E_2, so that no possessor of E_2 results as being visible within the phrase projected by E_2 itself.

According to (3a), the bahuvrīhi stem *ugra-putra-* ((1)), which would convey the meaning 'mighty son(s)', combines with B, thereby yielding *ugra-putra-*B, where B is silent qua elided: in this way, *ugra-putra-*B comes to denote the property of having mighty sons. The external referent of the bahuvrīhi will then supply the bearer of the property denoted by the bahuvrīhi itself: in (1), the goddess Aditi denoted by *adite* bears the property, denoted by *ugra-putra-*B, of having mighty sons. Hence, *adite* is interpreted as the possessor and *putrá-* as the corresponding Possessum in (1). To be noted that we are using 'elided' in (3a) in the sense of 'which has undergone ellipsis'. Ellipsis as we use it here is a cover term for all the linguistic phenomena in which "there is more meaning comprehended than what is presumably directly conveyed by the components of an expression" (Deshpande

---

[2] The possessive relation is merely one of the several types of relations that may hold between the external referent of the bahuvrīhi and one of the bahuvrīhi-members in Sanskrit. The traditional examples illustrating this point are reported in Cardona (1997:220). See also Pontillo (2021).

1989:103; see also Candotti and Pontillo 2022:2-5): in the context of (3a), the meaning 'having' is perceived but is not directly conveyed by any of the overt components of (1).[3]

On the other hand, according to (3b), the external referent *adite* starts out within the bahuvrīhi *ugra-putra-*, as the specifier of the phrase projected by *putrá-*: in this configuration *adite* is read as the possessor and *putrá-* as the corresponding possessum. *adite* subsequently moves to a bahuvrīhi-external position, so that the possessor of *putrá-* is not visible within *ugra-putra-*.

In this work we argue that the possessive relation perceived in Sanskrit bahuvrīhis like (1) is established by means of The Movement Strategy (3b), and that The Elided Possessive Unit Strategy (3a) is in fact unavailable. The argument that we build for achieving this conclusion capitalises on the internal order of bahuvrīhis (the internal order of a compound is the order in which that compound's members are arranged): only the Movement Strategy predicts the internal orders in (4), (5), (6), and (7) (the syntactic functions specified in the glosses are commented on in §2.3).

(4) úgra-putre   (adite).   (ṚV 8.67.11)
     [mighty]-[son].VOC.SG.F   [Aditi].VOC.SG.F
     [Predicate]-[**Subject**]   [Possessor of **Subject**]
     'O Aditi, having mighty sons' (or 'O Aditi, whose sons are mighty').

(5) vṛṣaṇ-aśvéna   (ráthena).   (ṚV 8.20.10)
     [bull]-[horse].INS.SG.M   [chariot].INS.SG.M
     [Predicate]-[**Subject**]   [Possessor of **Subject**]
     'With the chariot having bull-like horses' (or 'with the chariot whose horses are like bulls').

(6) vṛtrá-putrā   (dā́nuḥ).   (ṚV 1.32.9)
     [Vṛtra]-[son].NOM.SG.F   [Dānu].NOM.SG.F
     [Subject]-[**Predicate**]   [Possessor of **Predicate**]
     'Dānu, having Vṛtra as her son (or 'Dānu, of whom Vṛtra is the son').

(7) índro   vájra-hastaḥ   (ṚV 1.173.10)
     [Indra].NOM.SG.M   [mace]-[hand].NOM.SG.M
     [Possessor of **Predicate**] [Subject]-[**Predicate**]
     'Indra, having a mace in his hand (or 'Indra, in whose hand is a mace').

Some preliminary remarks are in order at this point. First, the bahuvrīhis in (4)-(7) belong to Vedic Sanskrit, namely the earliest stage of Sanskrit (1500-600 BCE): we only consider bahuvrīhis from Vedic Sanskrit in this work. Second, by saying that we predict a bahuvrīhi's internal order, we mean that we make a claim to the effect that a certain bahuvrīhi, with a certain internal order, will be found in the Vedic texts hosted by the DCS (The Digital Corpus of Sanskrit—Hellwig 2010-2021) and by TITUS (Thesaurus Indogermanischer Text- und Sprachmaterialien—Gippert et al. 2016). Third, the uploading of Vedic texts to the DCS and TITUS has not yet been completed, so any of our statements of (non-)attestation can only be partial, and valid until proven otherwise. Last, the derivation which allows us to predict the internal orders in (4)-(7) will be cast in the framework of generative grammar.

The remainder of this paper is organised as follows. §2 provides an explicit definition of the relevant technical terms concerning Sanskrit compounding, and contains an argument for choosing between the Elided Possessive Unit Strategy and the Movement Strategy. §3 sets the stage for the Movement Strategy (3b), by reviewing some arguments in favour of a syntactic derivation of Sanskrit compounds. §4 advances a formal derivation of bahuvrīhis that is based on the Movement Strategy, and shows that the restrictions on bahuvrīhis' internal order intertwine with locality conditions on movement. §5 concludes the paper.

---

[3] Ellipsis in the sense used here is not to be confused with the notion 'copy' developed within the copy-theory of movement (Chomsky 1995:183-194; 2019:26). Simplifying somewhat, while copies result from an application of the movement operation, the ellipsis of suffix B described in (3a) does not involve movement.
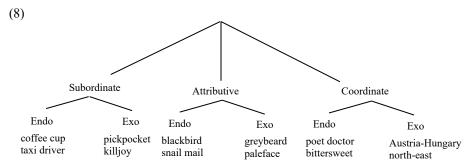
2 A definition of bahuvrīhi

In this section we define bahuvrīhis. Then, we spell out the argument for choosing between The Elided Possessive Unit Strategy (3a) and The Movement Strategy (3b).

2.1 What is a bahuvrīhi?

A compound is a lexical category Z made up of two other lexical categories X and Y (Guevara and Scalise 2009:107, Scalise and Vogel 2010:6). More formally: [X Y]$_Z$, where X and Y are said to be the internal members of Z. Lexical categories include nouns, adjectives, verbs, and prepositions.[4] The morphological realisation of X and Y changes from language to language: thus, X and Y may be roots, stems, inflected words (Bauer 2017:4).

An important characteristic of Sanskrit (Latin, Ancient Greek, etc.) compounds is that X of "[X Y]$_Z$" is realised as a stem, while Y is a fully inflected word (Lowe 2015:72, 93): e.g., *aśvaśapháḥ* 'horse's hoof' (cf. *aśvaśaphéna* in ŚB 13.3.4)[5] is analysed as [$_Z$ [$_X$ aśva]-[$_Y$ śaphá-ḥ]], where *aśva-* 'horse' is a stem and *śapháḥ* 'hoof' a word inflected in the NOM.SG.[6] Given an independent word W, the stem of W is what is left when any inflectional ending of W is dropped (Lieber and Štekauer 2009:5; Kastovsky 2009:324): e.g., the stem of the independent word *aśvasya* is *aśva-*, obtained by dropping the GEN.SG nominal ending *-sya* of *aśvasya*.

Scalise and Bisetto (2009:§4) proposed an influential classification of compounds, which may be summarised by means of the following tree (from Scalise and Vogel 2010:7):[7]

(8)

However, there is no single node, in (8), into which the whole category of Sanskrit bahuvrīhis could fit: indeed, Sanskrit bahuvrīhis may be subordinate exocentric (e.g., *vájra-hasta-* of (7)), attributive exocentric (e.g., *ugra-putra-* of (4)), as well as coordinate exocentric (e.g., *somendrá-* TS 2.3.2.6 '[the oblation] belonging to Soma and Indra'). Defining Sanskrit bahuvrīhis as exocentric compounds would not work either; this is because while all bahuvrīhis are exocentric, not all exocentric compounds are bahuvrīhis (Bauer 2017:64-71): e.g., so-called *upapadasamāsa*s like *mayo-bhū́-* (AVŚ 1.5.1) '(the waters) by means of which delight obtains' would count as exocentric in accordance with the hyponymy test, but are not treated as bahuvrīhis by the Indian grammatical tradition. 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 many bahuvrīhis in Sanskrit do

---

[4] See Corver (2013) on the distinction between lexical and functional categories.

[5] As the attentive reader may have noted, the accent on the initial *a-* of *áśva-* has disappeared in *aśva-śaphá-*. This illustrates a general tendency of compounds, namely that they have exactly one accent. For a general overview of compound accentuation, see Whitney (1899:484-485), Wackernagel (1905:40-43), and Macdonell (1910:91-97).

[6] Note that, in accordance with Pāṇini's rule A 2.4.71, *aśvaśaphénaḥ* is to be analysed as [[aśva]-[śapha]]-ḥ, where both *aśva-* and *śapha-* are fully inflected words (*pada*s) whose nominal endings have been zero-replaced: *-ḥ* qualifies as the ending of the whole compound-stem *aśvaśaphá-* rather than as the ending of *śapha-* (see Cardona 1997:21-23, 186, 207; Kiparsky 2009:67, 81-82; Candotti and Pontillo 2019:31).

[7] Endo and Exo in (8) stand for 'endocentric' and 'exocentric', respectively. We shall only consider semantic endo-/ exocentricity here: given the compound schema "[X Y]$_Z$", a compound Z is endocentric if and only if Z is a hyponym of either X or Y; conversely, Z is exocentric if and only if Z is not a hyponym of either X or Y. See Scalise and Fábregas (2010:111) and Bauer (2017:37, 64-65).

not involve any part-whole relation between the external referent and a bahuvrīhi-member: e.g., *kṛtá-brahman-* (ṚV 2.25.1) '(he) by whom the sacred formulation was created' (see Pontillo 2021:507-509 for discussion). Accordingly, we shall propose a working definition of bahuvrīhis modelled on the relative clause analysis of the Indian grammatical tradition (Cardona 1997:220; Lowe 2015:74). Taking again "[X Y]$_Z$" as our starting point, we define bahuvrīhis as those compounds Z whose paraphrasis involves a relative clause such that:[8]

(9) i. the relative pronoun is canonically inflected in a Case other than nominative;
    ii. and 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.

For instance, *ugra-putra-* of (1) is paraphrased as in (10). Here the relative pronoun is inflected in the genitive, in compliance with (9i). Moreover, the internal members of *ugra-putra-*, namely *putrá-* and *ugrá-*, serve as the subject and the predicate of the relative clause in (10): this satisfies condition (9ii). Since both conditions in (9) are satisfied, *ugra-putra-* qualifies as a bahuvrīhi.

(10) ugrāḥ   putrā   yasyā   aditiḥ.
      mighty.NOM.PL.M   son.NOM.PL.M   REL.GEN.SG.F   Aditi.NOM.SG.F
      'Aditi whose sons are mighty'.

The same remarks apply to *vṛṣaṇ-aśvá-* ((5)), *vṛtrá-putra-* ((6)), and *vájra-hasta-* ((7)). Let us briefly examine these compounds.
    vṛṣaṇ-aśvá- denotes a possessor of bull-like horses; this possessor is identified as the Maruts' chariot (*rátha-*) in ṚV 8.20.10.[9] Thus, *vṛṣaṇ-aśvá-* may be paraphrased as in (11). Since the relative pronoun is inflected in the genitive in (11), (9i) is satisfied. Since *áśva-* 'horse' and *vṛ́ṣan-* 'bull' are the internal members of *vṛṣaṇ-aśvá-* but also the subject and the predicate of the relative clause in (11), (9ii) is also satisfied. Accordingly, *vṛṣaṇ-aśvá-* qualifies as a bahuvrīhi.

(11) vṛṣaṇo   'śvā   yasya   rathaḥ.
      bull.NOM.PL.M   horse.NOM.PL.M   REL.GEN.SG.M   chariot.NOM.SG.M
      'The chariot whose horses are bulls'.

Turning now to *vṛtrá-putra-*, this compound denotes the possessor—qua mother—of a son whose name is Vṛtra; this possessor is identified as Dānu in ṚV 1.32.9.[10] This allows us to paraphrase *vṛtrá-putra-* as in (12). Since the relative pronoun is inflected in the genitive in (12), (9i) is met. Furthermore, (9ii) is also met because the internal members of *vṛtrá-putra-*, namely *vṛtrá-* 'Vṛtra' and *putrá-* 'son', serve as the subject and the predicate of the relative clause in (12). Consequently, *vṛtrá-putra-* qualifies as a bahuvrīhi.

(12) putro   vṛtro   yasyā   dānuḥ.
      son.NOM.SG.M   Vṛtra.NOM.SG.M   REL.GEN.SG.F   Dānu.NOM.SG.F
      'Dānu of whom Vṛtra is the son'.

---

[8] The 'paraphrasis' of a compound corresponds to the vigraha of the Indian grammatical tradition, i.e., the combination of inflected words into which the compound is solved or analysed. The first formulation of (9i) is due to Kātyāyana's *Vārttika* 19 *ad* A 2.2.24 (see also Benfey 1852:273). Pāṇini's definition of bahuvrīhis, contained in A.2.2.23-24, differs from the definition adopted in the text. See Candotti and Pontillo (2019:33-36); (2022) for an illustration of Pāṇini's model of bahuvrīhis, and Pontillo (2021) for a critical assessment of how that model was received by modern scholarship.

[9] *vṛsanaśvéna maruto vṛ́ṣapsunā/ ráthena vṛ́ṣanābhinā/ ā́ śyenā́so ná pakṣíṇo vṛ́thā naro/ havyā́ no vītáye gata*// 'O Maruts, with your chariot with its bullish horses, bullish breath, bullish wheel-naves, come here at will, like winged falcons, to pursue our oblations, o men.' (tr. Jamison and Brereton 2014).

[10] *nīcā́vayā abhavad vṛtráputra/ índro asyā áva vádhar jabhāra/ úttarā sū́r ádharaḥ putrá āsīd/ dā́nuḥ śaye sahávatsā ná dhenúḥ*// 'The strength of Vṛtra's mother ebbed; Indra bore his weapon down upon her. The mother was above; the son below: Dānu lies like a milk-cow with her calf.' (tr. Jamison and Brereton 2014).

Likewise, *vájra-hasta-* denotes the (inalienable) possessor of a mace-holding hand; such a possessor is identified as Indra in ṚV 1.173.10.[11] Thus, the paraphrasis in (13) properly captures the meaning of *vájra-hasta-*. The relative pronoun is inflected in the genitive in (13), in compliance with (9i). In addition, *vájra-* 'mace' and *hásta-* 'hand', namely the internal members of *vájra-hasta-*, serve as the subject and the predicate of *vájra-hasta-* in (13), so that (9ii) is also complied with. Since both conditions in (9) are satisfied, *vájra-hasta-* qualifies as a bahuvrīhi.

(13) haste   vajro   yasyendraḥ.
      hand.LOC.SG.M   mace.NOM.SG.M   REL.GEN.SG.M_Indra.NOM.SG.M

   'Indra in whose hand is a mace'.

Besides having the defining characteristics in (9), Sanskrit bahuvrīhis also display adjectival agreement (i.e., agreement in Case, gender, and number) with their external referent; moreover, Sanskrit bahuvrīhis show a peculiar accentuation which distinguishes them from other compound types. We examine the accentuation of bahuvrīhis in the next subsection, while we address the agreement facts in §3.5 below.

2.2   The accent of bahuvrīhis

The fundamental rule on bahuvrīhi accentuation was formulated by Pāṇini in rule A 6.2.1: *bahu-vrīhau prakṛtyā pūrvapadam* 'the left-hand member retains its original accent in bahuvrīhis' (see, e.g., Wackernagel 1905:291; Katre 1987; Cardona 1997:385-386). This rule is meant to capture the fact that, when Z of "[X Y]$_Z$" is a bahuvrīhi, then Z is typically accented on the left-hand member X, more precisely on the same syllable on which X is accented when X occurs outside Z (Macdonell 1910:92). For illustration, consider *rāja-putra-*, formed from *rā́jan-* 'king' and *putrá-* 'son(s)'. When this compound is endocentric, such as in ṚV 10.40.3, it shows up as *rāja-putrá-* 'the king's two sons' (referring to the divine Aśvin twins); conversely, when this compound is a bahuvrīhi, such as in ṚV 2.27.7, it shows up as *rā́ja-putra-* '[Aditi] whose sons are kings' (see Candotti and Pontillo 2022:10-11, 19). The reader may easily verify that the bahuvrīhis *vṛtrá-putra-* ((6)) and *vájra-hasta-* ((7)) comply with A 6.2.1. As for the accent of *vṛṣaṇ-aśvá-* ((5)) and *ugra-putra-* ((4)), see below in this subsection and fn. 1 above.

Indeed, A 6.2.1 allows for several exceptions: on the one hand, many bahuvrīhis are accented on the right-hand member and, on the other hand, many endocentric (hence, non-bahuvrīhi) compounds are accented on the left-hand member. The list of rules formulated by Pāṇini to capture these exceptions is reported in Cardona (1997:385-391). A full discussion of these exceptions will not be attempted here; the interested reader is referred to the standard treatments of accent in Sanskrit compounds, notably Whitney (1899:480-515), Wackernagel (1905), and Macdonell (1910:91-97). Here we limit ourselves to mentioning a couple of exceptions that are relevant for the examples quoted in §4.3 below as well as for *vṛṣaṇ-aśvá-* ((5)).

When the left-hand member of a bahuvrīhi is *víśva-* 'all', that bahuvrīhi is accented on the second syllable of *víśva-* (Wackernagel 1905:292; Macdonell 1910:92): e.g., *viśvá-bhānu-* (ṚV 4.1.3) 'whose brightness is on everything', 'all-radiant'. When the left-hand member of a bahuvrīhi is instead *dús-* or *sú-* 'well', that bahuvrīhi is accented on the right-hand member, typically on the same syllable on which the right-hand member is accented outside compounds (Whitney 1899:508-509; Wackernagel 1905:293-295; Macdonell 1910:93): e.g., *su-bhága-* (ṚV 1.86.7) 'whose portion is good', 'well-portioned' (cf. *bhága-* 'portion'). When a compositional suffix (also known as *samāsānta* suffix) is attached to a bahuvrīhi, the bahuvrīhi's accent may be on the suffix, in which case the suffix is said to be dominant (Whitney 1899:505; Macdonell 1910:86-87; Kiparsky 2010:32): e.g., *śiti-kakṣ-ín-* (TS 5.5.20.1; from *śiti-* + *kákṣa-* + *-ín-*) 'whose belly is white' (cf. *śiti-*

---

[11] *víṣpardhaso narā́ṃ ná śáṃsair/ asmā́kāsad* <u>*indaro*</u> <u>*vájrahastaḥ*</u>/ *mitrāyúvo ná pū́rpatiṃ súśiṣṭau/ madh-yāyúva úpa śikṣanti yajñaíḥ*// 'The contenders—(crying) "Indra with mace in hand will be ours!"— seeking (the one in) the middle [=Indra], try to win him over with their sacrifices, as those seeking an ally, (approaching) a lord of strongholds in regard to his good command, (try to win him over) with lauds of men.' (tr. Jamison and Brereton 2014).

kákṣa- 'id.' in AVŚ 5.23.5);[12] *vārdhrā-ṇas-á-* (TS 5.5.20.1; from *vārdhra-* + *nás-* + *-á-*) 'whose nose is leathern' (cf. *ŕ̥jū-nas-* lit. 'whose nose is straight' in ṚV 8.52.2). Finally, the accent of *vṛṣan-aśvá-* does not coincide with either the accent of *vŕ̥ṣan-* or that of *áśva-*: see Whitney (1899:506); Wackernagel (1905:298).[13] *vṛṣan-aśvá-* may perhaps be treated along the lines of *tri-bandhú-* (ṚV 7.37.7) 'whose relations are triple' in Kiparsky's (2010) model of compound accentuation (see especially ibid., p. 32): first, the bahuvrīhi stem (**trí-bándhu-*; **vŕ̥ṣaṇ-áśva-*) combines with a dominant zero compositional suffix, which deletes the original accent of the bahuvrīhi-members (**tri-bandhu-Ø*; **vṛṣaṇ-aśva-Ø*); next the Oxytone Rule assigns default final accent to the bahuvrīhi stem (*tri-bandhú-Ø = tri-bandhú-*; *vṛṣaṇ-aśvá-Ø = vṛṣaṇ-aśvá-*).[14]

Earlier we considered the question as to whether the possessive relation perceived in some Sanskrit bahuvrīhis is established via The Elided Possessive Unit Strategy (3a) or The Movement Strategy (3b). We now construct an argument which makes it possible to settle this question on empirical grounds.

2.3 The Elided Possessive Unit Strategy vs. The Movement Strategy

In analyses of bahuvrīhis that resort to The Elided Possessive Unit Strategy (3a), the elided possessive unit B expressing the meaning 'having' combines with a pre-formed compound-stem CS: e.g., B would combine with *ugra-putra-* as a whole in (1). Consider the possibility in which CS is the stem of an endocentric compound. This possibility is advocated by Gillon (2008:3) for at least one class of bahuvrīhis: so-called samāna-adhikaraṇa-bahuvrīhis, in which an attributive relation holds between the bahuvrīhi-members. He assumes that a (samāna-adhikaraṇa-)bahuvrīhi is the combination of an endocentric compound with B (cf. also Bopp 1827:357; Whitney 1889:501-502; Marchand 1967:335; Kiparsky 1982:139). E.g., by combining *śiti-* 'white' with *kákṣa-* 'lurking place', 'belly', we obtain the endocentric compound *śiti-kakṣa-* which expresses the meaning 'white belly'.[15] Then, by attaching B to the stem of the endocentric compound *śiti-kakṣa-*, the omophonous bahuvrīhi *śiti-kákṣa-* (= *śiti-kakṣa-*B) is formed which expresses the meaning 'having a white belly' (AVŚ 5.23.5).[16]

The Elided Possessive Unit Strategy looks particularly plausible in the light of the existence of *śiti-kakṣín-* (TS 5.5.20.1). *śiti-kakṣín-* is synonymous with *śiti-kákṣa-* insofar as *śiti-kakṣín-* and *śiti-kákṣa-* both denote an entity having a white belly in the texts quoted here: this entity is identified as the worms in the case of *śiti-kákṣa-* but as a bird species in the case of *śiti-kakṣín-*.

Now that we have established the synonymity between *śiti-kákṣa-* and *śiti-kakṣín-*, consider that *- ín -* is a possessive suffix which attaches to a nominal X to form a derivative nominal with the meaning 'endowed with X': cf. *rátha-* 'chariot' + *-ín-* 'possessing' > *rathín-* 'possessing a chariot', 'charioteer'. One may then take the B occurring in such bahuvrīhis as *śiti-kákṣa-* as the elided counterpart to the *-ín-* of *śiti-kakṣín-*: this would straightforwardly account for the synonymy between *śiti-kákṣa-* and *śiti-kakṣín-*, and for the establishment of the possessive relation between the external referent of the bahuvrīhi and one of the bahuvrīhi-members. Thus, the postulation of the elided possessive unit B by the Elided Possessive Unit Strategy (3a) appears to be well-motivated.

This is not the end of the story, however. We wonder: can the Elided Possessive Unit Strategy (3a) predict the internal orders in (4)-(7)? To repeat, the Elided Possessive Unit Strategy takes it that

---

[12] See §2.3 below on the synonymity of *śiti-kakṣ-in-* and *śiti-kákṣa-*. The accent of *śiti-kákṣa-* also deviates from the general rule A 6.2.1 for bahuvrīhi accentuation. Indeed, *śiti-kákṣa-* belongs to the set of bahuvrīhis which have a disyllabic adjective ending in *-i* or *-u* as their left-hand member: these bahuvrīhis are commonly accented on the same syllable on which the right-hand member is accented outside compounds (Wackernagel 1905:296-297; Macdonell 1910:92-93; Kiparsky 2010:31-32).

[13] The accent required by A 6.2.1, i.e., *vŕ̥ṣaṇ-aśva-*, is attested in MS 2.5.5 and ŚB 3.3.4.18.

[14] The accentuation of compounds may be used to diagnose their internal constituency (Kiparsky 2010:34; 2016:16-17). We shall not pursue this interesting line of analysis here.

[15] Although *śiti-kakṣa-* is not actually attested as endocentric compound, we assume, for the sake of argument, that this form was nonetheless possible in principle.

[16] Lowe (2015:103) similarly assumes that (samāna-adhikaraṇa-)bahuvrīhis are the combination of an endocentric compound with a special unit, but such a unit does not bear any possessive meaning in his account, and only serves the purpose of endowing bahuvrīhis with adjectival agreement properties (see §3.5 below). For an overview of recent analyses of bahuvrīhis, see Candotti and Pontillo (2022:5-9).

the elided possessive unit B attaches to the stem of an endocentric compound to form a bahuvrīhi. But B does not affect the internal order of the compound-stem to which B attaches. This allows us to make a clear empirical prediction on the internal order of bahuvrīhis:

(14) The internal order of a bahuvrīhi will be always the same as the internal order of the endocentric compound corresponding to that bahuvrīhi (the endocentric compound corresponding to a bahuvrīhi is the endocentric compound to whose stem B attaches to form that bahuvrīhi).

The prediction in (14) is highly problematic, however. For one thing, none of the compounds in (4), (6), and (7) (i.e., *ugra-putra-*, *vṛtra-putra-*, *vajra-hasta-*) is attested as endocentric; in point of fact, it is very rare that, if a form is attested as an endocentric compound, then that form is also attested as a bahuvrīhi (and vice versa)—see Renou (1961:114). Therefore, the accuracy of (14) cannot be tested against the internal orders of (4), (6), and (7). Furthermore, *vṛṣaṇ-aśvá-* ((5)) does have an endocentric counterpart, namely *aśva-vṛṣá-*. Indeed, direct evidence for the endocentricity of *aśva-vṛṣá-* comes from the context in which this form occurs, i.e., ŚB 14.4.2.7-8 (= BĀU 1.4.4):[17] *aśva-vṛṣá-* here denotes the masculine analogue of *váḍavā-* 'mare', hence conveys the meaning 'male horse'; the meaning 'male horse' is in turn a more specific version of the meaning 'bull-like horse', on the grounds that a bull-like horse is a horse bearing the properties typical of bulls, and that such properties include masculinity in the Vedic-speaking world. Thus, *aśva-vṛṣá-* properly conveys the meaning 'bull-like horse' in the passage at stake. Since *aśva-vṛṣá-* conveys the meaning 'bull-like horse', *aśva-vṛṣá-* is hyponym of *aśva-* 'horse', which means that *aśva-vṛṣá-* is endocentric. Recall now that the bahuvrīhi *vṛṣaṇ-aśvá-* conveys the meaning 'having bull-like horses': it is then reasonable to take *aśva-vṛṣá-* as the endocentric compound stem to which the elided possessive unit B (conveying the meaning 'having') attaches to form the bahuvrīhi *vṛṣaṇ-aśvá-*. But in accordance with prediction (14), the internal order of *vṛṣaṇ-aśvá-* should match the internal order of *aśva-vṛṣá-*, contrary to fact.

Therefore, (14) is either wrong or not testable when the internal orders in (4)-(7) are at stake. But (14) is all that the Elided Possessive Unit Strategy has to say about the internal order of bahuvrīhis. We conclude that the Elided Possessive Unit Strategy fails to predict the internal orders in (4)-(7). The question is then whether The Movement Strategy (3b) succeeds in predicting these internal orders. To tackle this question, let us go back to (4)-(7), now focusing on the specification of the syntactic functions fulfilled by the bahuvrīhi-members.[18]

Interestingly, the position of the bahuvrīhi-internal predicate in (4)-(7) depends on whether the bahuvrīhi's external referent is construed with that predicate: the predicate precedes the subject when the bahuvrīhi's external referent is construed with (i.e., is interpreted as the Possessor of) the subject (as in (4)-(5)) but not when this external referent is construed with the predicate (as in (5)-(6)).[19] We may thus express the restrictions on the internal orders in (4)-(7) by means of the following generalisation:

(15) i. In a bahuvrīhi in which the predicate is nominal (i.e., an adjective or a noun) and the external referent is interpreted as the Possessor of the <u>subject</u>, the predicate must **precede** the subject;

ii. In a bahuvrīhi in which the predicate is nominal (i.e., an adjective or a noun) and the external referent is interpreted as the Possessor of the <u>predicate</u>, the predicate must **follow** the subject.

---

[17] *sā gaurábhavat vṛṣabha ítarastāṃ sámevàbhavattáto gắvo 'jāyanta// váḍavétarábhavat <u>aśvavṛṣa</u> ítaro gardabhī̀tarā gardabha ítarastāṃ sámevàbhavattáta ékaśaphamajāyata.* 'She became a cow, the other became a bull and was united with her; from that cows were born. The one became a mare, the other a stallion; the one became a she-ass, the other be-came a he-ass and was united with her; from that one-hoofed animals were born.' (tr. Mādhavānanda 1950).

[18] By saying that an expression α is 'Possessor of Subject' or 'Possessor of Predicate' (see (4)-(7) above), we mean that α is interpreted as the possessor of the entity denoted by another expression β, where β fulfills the function of Subject or Predicate.

[19] Note that the restriction that compound-external material can only be construed with (or 'view') the right-hand member of a compound (see Di Sciullo and Williams's model 1987, especially their pp. 24, 30-31) does not hold in Sanskrit: see Gillon (2008); Molina-Muñoz (2013); Lowe (2015:76-77); Candotti and Pontillo (2022:18). Thus, one cannot appeal to this restriction to account for the different positioning of the bahuvrīhi-internal predicate in (4)-(5) vs. (6)-(7).

This way, the question as to whether the Movement Strategy predicts the internal orders in (4)-(7) reduces to the question: can the Movement Strategy account for generalisation (15)? §4 is devoted to providing an affirmative answer to the latter question.

Before concluding this section, we should like to explore a third means of establishing the possessive relation in bahuvrīhis. We shall see that, albeit theoretically plausible, this third possibility faces the same empirical problems as the Elided Possessive Unit Strategy.

2.4 Bahuvrīhis and figures of speech

The Elided Possessive Unit Strategy (3a) and the Movement Strategy (3b) are not in fact the only ways in which the possessive relation between the external referent of the bahuvrīhi and one of the bahuvrīhi-members may be established. Indeed, Bauer (2008:59) (see also Bauer 2016:462-463; 2017:65; cf. Pennanen 1982:245-246; Barcelona 2008:210; Lundquist and Yates 2018:2120 for similar views) proposes that English bahuvrīhis such as *red-cap* (which names people who are abitually associated with a red cap) are regular endocentric compounds interpreted by synecdoche (the 'pars pro toto' figure of speech). For instance, the simplex *crown* properly denotes a headdress possessed by, or being part of ($\subseteq$), a monarch; by resorting to the synecdoche, one may use *crown* (the part) as a name for *monarch* (the whole). Likewise, the compound *red-cap* properly denotes a kind of head covering possessed by, or being part of ($\subseteq$), specific groups of people, typically the military policemen in the UK or the railway porters in the US; as such, *red-cap* is an endocentric compound because it is hyponym of the compound-member *cap*. By synecdoche, however, *red-cap* (the part) names those very groups of people that possess the red head covering (the whole). The synecdoche thus turns the groups of people possessing the red head covering into the external referent of *red-cap*: this way, a possessive relation is established between the external referent of *red-cap* and the head covering denoted by *cap*.

Although this may well be the correct analysis of bahuvrīhis in English (and possibly other languages considered by Bauer 2008:57-60), it cannot be extended to Sanskrit. Indeed, Bauer's synecdoche-based analysis and the Elided Possessive Unit Strategy (3a) alike commit themselves to the prediction in (14). Consider why. In the synecdoche-based analysis, bahuvrīhis are obtained by applying the synecdoche to their corresponding endocentric counterparts; but the synecdoche (much like the elided possessive unit B) is unable to affect the internal order of the endocentric compound to which it applies; accordingly, the internal order of a bahuvrīhi must be the same as the internal order of the endocentric compound corresponding to that bahuvrīhi ((14)).

As already seen, this prediction is either wrong or not testable when it comes to the internal orders in (4)-(7) (see §2.3 above).[20] Given that Bauer's synecdoche-based analysis makes the same predictions on the internal orders in (4)-(7) as the Elided Possessive Unit Strategy (3a), we shall disregard Bauer's analysis in what follows. However, whenever we claim that the Elided Possessive Unit Strategy (3a) fails to predict the internal order of Sanskrit compounds, this criticism can be extended to Bauer's analysis.

3 For a syntactic derivation of bahuvrīhis

In this section we report some evidence that justifies a syntactic derivation for Sanskrit bahuvrīhis. Specifically, we argue that the internal members of Sanskrit bahuvrīhis are phrases, and as such can undergo ordinary syntactic operations such as movement.

3.1 Sanskrit compound-members as words

Lowe (2015) provides several arguments to the effect that both members of Sanskrit compounds are words. These arguments can be divided into two families.

---

[20] An additional reason against extending the synecdoche-based analysis to Sanskrit bahuvrīhis is that the accent of bahuvrīhis typically differs from the accent of endocentric compounds in Sanskrit (see §2.2 above). It is not clear to us how the application of the synecdoche to an endocentric compound may effect this difference in compound accentuation (see Candotti and Pontillo 2022:7).

The first family is targeted at showing the word status of the left-hand member: more precisely, the left-hand member is claimed to be a 'non-projecting word' (i.e., a word of a special sort which does not project either complement or specifier—ibid. p. 91), but nonetheless a word. This first family consists of several arguments, but we only mention two of them here, drawing our data from Vedic Sanskrit (see ibid., pp. 75-86, for the complete list of these arguments).[21]

3.2 Sanskrit compounds are not anaphoric islands

The first argument of the first family capitalises on the fact that the left-hand member of Sanskrit compounds may be a pronoun that is coreferential with a compound-external expression. (16) illustrates this fact.

(16)   tád$_i$-annāya   tád$_i$-apase/   tám
       this$_i$-food.DAT.SG.M   this$_i$-work.DAT.SG.M   this.ACC.SG.M
       bhāgám   upasedúṣe/   tritắya
       portion.ACC.SG.M   approach.PRF.PTCP.DAT.SG.M   Trita.DAT.SG.M
       ca   dvitắya   *ca*/
       and   Dvita.DAT.SG.M   and
       *úṣ*o   duṣvápn*iyam*$_i$ vah*a*/
       dawn.VOC.F.SG   bad_dream.ACC.SG.N$_I$ carry.IMP.2SG

   'To the one who has it as his food and as his work, to the one reverently approaching it as his portion, to Trita and to Dvita, o Dawn, carry the bad dream.' (ṚV 8.47.16a-d; tr. Jamison and Brereton 2014).

The demonstrative pronoun *tád-* 'this' figures as the left-hand member of the two bahuvrīhis *tád-anna-* 'he of whom this is the food' and *tád-apas-* 'he of whom this is the work': *tád-* here is coreferential with *duṣvápnya-* 'bad dream', which shows up in this same verse, as well as in ṚV 8.47.14 (see Jamison and Brereton 2014:1126-1127).

Thus, Sanskrit compounds can contain pronouns, and these compound-internal pronouns are coreferential with an expression placed outside the compound itself. But this is tantamount to saying that Sanskrit compounds are not anaphoric islands (Lowe 2015:78; see also Kiparsky 2009:83; Molina-Muñoz 2013:191-192). English compounds appear to obey different constraints:

(17) a. Whenever he meets an [old woman]$_i$ Pete says he hates young ones$_i$.
     b. *Whenever he meets an [old woman]$_i$ Pete says he is a one$_i$-hater.

(Postal 1969:216)

In (17b) the pronominal element *one*, which is the left-member of the compound *one-hater*, cannot be coreferential with *old woman*, which occurs outside the compound itself (cf. the well-formedness of (17a), in which *ones* is not part of a compound). This shows that a pronominal element contained within a certain compound cannot be coreferential with an antecedent placed outside that compound in English: to wit, English compounds, unlike Sanskrit compounds, are anaphoric islands. Now, anaphoric islandhood is considered a defining property of 'word'. Indeed, it has been assumed since Postal (1969:227) that words are anaphoric islands crosslinguistically (see Lowe 2015:77 and the references cited therein).[22] This is illustrated by the following contrast:

(18) a. I got a divorce$_i$ from Sally and I'll get one$_i$ from Louise too.
     b. *I [divorce]$_i$-d] Sally and I'll get one$_i$ from Louise too.

(Postal 1969:210)

---

[21] Indeed, Lowe (2015:107-108) confines the claim that the members of Sanskrit compounds are words to Classical Sanskrit, a standardised form of Sanskrit codified on the basis of Pāṇini's Aṣṭādhyāyī (4th c. BCE). However, to the extent that his arguments can still apply to Vedic Sanskrit compounds, we consider it reasonable to extend that claim to Vedic Sanskrit compounds, too.

[22] The claim that words and compounds are anaphoric islands in English is indeed an oversimplification and needs to be remodulated in the light of the counterexamples provided by Bauer (2017:20-21).

In (18b) *one* cannot be coreferential with the antecedent noun *divorce* that is contained within (i.e., is part of) the verb *divorced* (a denominal derivative verb formed from the noun *divorce*), *divorced* being a word. However, coreference between *one* and *divorce* becomes possible when the noun *divorce* is itself a word not contained within any other word, as in (18a). Thus, the ill-formedness of (17b) follows on the assumption that English compounds are words, and hence that compound-members are subparts of words in English (see Postal 1969:225-227). But then, Lowe (2015:80) argues, the well-formedness of (16) can be traced back to the fact that the left-hand member of Sanskrit compounds is a word, rather than a subpart of word. Note that this argument is also consistent with the conclusion that the left-hand member of Sanskrit compounds (e.g., *tád-* in the compound *tád-apas-* of (16)) is a phrase that does not contain complements or specifiers.

3.3 Sanskrit compounds and clausal scope

The second argument of the first family is based on the following fact: the left-hand member of Sanskrit compounds can be an interrogative pronoun. Thus, *kád-* is the stem form of the interrogative pronoun (*káḥ*, *kā́*, *kád/kím*), and shows up as the left-hand member of the bahuvrīhi *kád-artha-* (lit. 'whose purpose is what') in the following example:

(19) ádha   gmántā́   uśánā   pr̥chate
      PCL   come.AOR.PTCP.ACC.DU.M   Uśanā.NOM   ask.IND.PRS.3SG
      vām/   kád-arthā   na   ā́
      you.ACC.DU   what-purpose.NOM.DU.M   we.GEN   to
      gr̥hám/   ā́   jagmathuḥ   parākā́d/
      house.ACC   to   come.IND.PRF.2DU   afar.ABL
      diváś   ca   gmáś   ca
      heaven.ABL   and   earth.ABL   and
      mártiyam//
      mortal.ACC

'Then Uśanā asks you both [=Indra and Kutsa] on your coming, "With what purpose (have you come) to our house? You have come from afar, from heaven and earth, to a mortal."' (R̥V 10.22.6; tr. Jamison and Brereton 2014).

kád- turns the sentence whose predicate is *jagmathuḥ* into an interogative sentence ('With what purpose have you come to our house?') in (19); technically, *kád-* takes clausal scope (see Kiparsky 2009:83; Lowe 2015:80). Since words and phrases—but not subparts of words—are known to have the ability to take clausal scope in Sanskrit (much like in English), *kád-*, which displays this ability in (19), should also qualify as a word. Accordingly, the left-hand member (*kád-*) of the bahuvrīhi *kád-artha-* is a word, rather than a subpart of word. In this case, too, the argument would stand even if we were to take *kád-* as a phrase (which does not contain complements or specifiers).

3.4 Word or phrase?

We have reviewed two of Lowe's (2015) arguments to the effect that the left-hand member of Sanskrit compounds is a word (for other arguments, see Lowe 2015:80-86). We have also noted that these arguments are consistent with an alternative conclusion: that Sanskrit compounds' left-hand member is a phrase rather than a word. On this alternative conclusion, the stem *áśva-* occurring in the compound *aśva-śaphá-* 'horse's hoof' would be represented as a bare phrase (i.e., a phrase which does not contain complements or specifiers) in the phrase marker.

Let us then assume that a compound's left-hand member that is morphologically a stem qualifies as a bare phrase from the syntactic point of view. We identify the head of such a phrase with the stem itself: thus, the structure of the *áśva-* contained within *aśva-śaphá-* would be [NP [N° aśva-]], which amounts to a bare phrase. However, we still need to explain why the phrase corresponding to the compound's left-hand member does not contain complements or specifiers and lacks inflection. According to Lowe (2015:91-93), the lack of inflection of a certain compound-member and its inability to contain complements or specifiers both follow from this compound-member being a non-projecting word. An alternative explanation that does not appeal to the notion of non-projecting

word may be the following. It is standardly assumed that fully inflected words are endowed with some functional projection above the lexical level (e.g., above N[oun] P[hrase]/ A[djectival] P[hrase]/ V[erb] P[hrase]).[23] Thus, a natural way to capture the absence of inflection in Sanskrit compounds' left-hand members is to say that these members only consist of the projection headed by the lexical head (N°/ A°/ V°/ P°), lacking instead the functional layers stacked above such a head. We return to this point in §4.2, where the inability of a compound-member to take complements and specifiers is also traced back to the fact that it lacks functional layers.

3.5 The right-hand member of Sanskrit compounds

Let us now turn to the second family of arguments, i.e., the arguments targeted at showing the word-status of Sanskrit compounds' right-hand member.

Lowe (2015:72, 93, 97-99) analyses Sanskrit endocentric compounds such as *aśvaśapháḥ* 'horse's hoof' as [aśva-[śapha-ḥ]], where the nominal ending -*ḥ* belongs to the right-hand member *śaphá*-. But this means that the right-hand member of endocentric compounds (e.g., *śapháḥ* of *aśvaśapháḥ*) is formally identical to a full-fledged inflected word which occurs outside compounding (e.g., *śapháḥ* is also attested as an independent word in MS 2.5.6.6). Then, the formal identity of Sanskrit endocentric compounds' right-hand member with full-fledged inflected words constitutes crucial evidence, in Lowe's analysis, for considering this right-hand member as a word rather than as a subpart of word. Things get more complicated when it comes to bahuvrīhis. Indeed, the right-hand member of a bahuvrīhi is not formally identical to an independent word:

> nouns have inherent grammatical gender in Sanskrit, but at the end of a bahuvrīhi a noun can be inflected in any gender, since it must agree with the compound's external referent. So an inherently masculine noun, for example, which cannot otherwise appear in neuter or feminine forms, can appear in such forms at the end of a bahuvrīhi. Therefore a noun form used at the end of a bahuvrīhi can be considered a rather different type of word; it is not, in fact, a noun of the standard type. It is a noun with adjectival agreement properties or, to put it another way, a noun that is partly adjectival. (Lowe 2015:103)

In fact, Lowe (2015:103-105) analyses a bahuvrīhi like *ugraputraḥ* '(he/ she) whose sons are mighty' ((4)) as having the following structure:

(20) i. Adj° → N̂ Adj°
    ii. N̂ → Âdj N̂

Adj° attaches to the whole compound structure, which is dominated by N̂. The stem form *ugrá*- instantiates Âdj of (20ii), while *putraḥ* instantiates two nodes by the mechanism of Lexical Sharing: the N̂ daughter of (20ii), as well as the Adj° daughter of (20i) (see ibid., pp. 104-105, for the technical details of this analysis). Therefore, according to this analysis the right-hand member of bahuvrīhis is a complex unit—constituted of a non-projecting word (N̂) and a word (Adj°)—which bears a special kind of inflection (i.e., adjectival inflection). Since the presence of inflection is a defining characteristic of words, the right-hand member of bahuvrīhis qualifies as a word.

However, we cannot immediately adopt Lowe's conclusion in this case. Indeed, we analyse bahuvrīhis such as *ugraputraḥ* as [[ugra-putra-]-ḥ], where both *ugrá*- and *putrá*- are in stem form, and the nominal ending -*ḥ* is affixed to the whole compound stem [ugra-putra-], rather than to the right-hand member only. In this way, -*ḥ* does not qualify as the inflection of the right-hand member, which thus results as being morphologically identical to the left-hand member.[24] This means that we can no longer appeal to the presence of inflection on the right-hand member of bahuvrīhis to justify the word status of this right-hand member. Nonetheless, there is still reason to maintain that

---

[23] For instance, nominal and adjectival projections in noncompound syntax may be endowed with the Agr(eement) P(hrase) of Cinque (2005:317-318, 321 n. 24), with the KP ('Case Phrase') of Bittner and Hale (1996), or, more generally, with a D(eterminer) P(hrase) (decomposable as in Zamparelli 2000). Verbal projections are endowed with T(ense) P(hrase), and C(omplementiser) P(hrase) at least, TP and CP being further decomposable along the lines of Cartography (see Rizzi 2013).

[24] This analysis of *ugraputraḥ* as [[ugra-putra-]-ḥ] is reminiscent of Pāṇini's analysis of compounds. See the references of n. 6 above.

a bahuvrīhi's right-hand member is a word or a phrase, and not a subpart of word.

Perfect passive participles (PPP, also known as *-ta-* adjectives) typically occur as left-hand members in Sanskrit bahuvrīhis. Thus, we find: *jāta-danta-* (VaikhGS 7.2) '(the boy) whose teeth have grown', where *jātá-* 'grown' is a PPP; *ā́hitāgni-* (TS 2.2.2.2) '(he) by whom the sacred fire has been established', where *ā́hita-* 'established' is a PPP; *hata-putra-* (KB 4.6.7) '(Vasiṣṭha) whose sons were killed', where *hatá-* 'killed' is a PPP. These bahuvrīhis fall outside the scope of general-isation (15) because they involve a verbal predicate (i.e., the PPP itself); nonetheless, they are very interesting in that their internal order may be inverted without any apparent change in meaning. Thus, the following bahuvrīhis are also attested, with the PPP occurring as the right-hand member of the compound: *danta-jāta-* (VaikhGS 5.9) '(the boy) whose teeth have grown'; *agny-āhita-* (MBh. 12.281.21) 'he by whom the sacred fire has been established';[25] *putra-hata-* (PB 4.7.3) '(Vasiṣṭha) whose sons were killed'—more examples of bahuvrīhis having a PPP as their right-hand member are given in Wackernagel (1905:302-303).

Now, if the left-hand member of a Sanskrit compound is to be considered a word or a phrase for the reasons spelled out above (see §§3.1-4), then the PPP occurring as the left-hand member of the bahuvrīhis *jāta-danta-*, *ā́hitāgni-*, and *hata-putra-* also counts as a word or a phrase. We have seen that the PPP can optionally occupy the right-hand slot in these bahuvrīhis (thereby yielding *danta-jāta-*, *agny-āhita-*, and *putra-hata-*): this should be impossible if these bahuvrīhis' right-hand slot were smaller than a word or a phrase. Therefore, these bahuvrīhis' right-hand slot must be at least as big as a word or a phrase. In keeping with the preceding discussion, we conclude that both the left-hand member and the right-hand member of the bahuvrīhis *jāta-danta-*, *ā́hitāgni-*, and *hata-putra-* (as well as of *danta-jāta-*, *agny-āhita-*, and *putra-hata-*) is a phrase.

This argument cannot be easily reproduced for proving the phrasal status of the bahuvrīhis in (4)-(7): the linear order of *ugra-putra-*, *vṛṣaṇ-aśvá-*, *vṛtrá-putra-*, and *vájra-hasta-* cannot be in-verted (**putrogra-*, **aśva-vṛṣan-*, **putra-vṛtra-*, and **hasta-vajra-* are all unattested as bahuvrīhis). Nevertheless, the non-invertibility of the linear order of these bahuvrīhis will be shown to follow from independent reasons in §4. We may then assume that the right-hand member of the bahuvrīhis in (4)-(7) also has phrasal status.

In §3.4 we proposed to explain the stem status of Sanskrit compounds' left-hand member as a result of the fact that this left-hand member is a bare phrase headed by a lexical category and lacking functional structure (e.g., an NP rather than a DP). We have now claimed that, in all Sanskrit bahu-vrīhis, both the left-hand member and the right-hand member are stems from the morphological point of view, and phrases from the syntactic point of view. Accordingly, we suggest that the phrases corresponding to the left- and right-hand members of Sanskrit bahuvrīhis are always headed by lexical categories and always lack functional structure.

A brief remark is in order at this point. We have made two important assumptions in this sub-section: i. that the right-hand member of the bahuvrīhis in (4)-(7) is a phrase; ii. that the phrases corresponding to the left- and right-hand members of Sanskrit bahuvrīhis are always bare phrases projected by lexical heads and lacking functional structure. As a matter of fact, these assumptions do not follow from any logical necessity: they are tenable only to the extent that they allow decom-posing complex empirical facts into simpler abstract entities.[26] In this specific case, these assump-tions can be maintained only to the extent that they make it possible to reduce the well-formedness or ill-formedness of Sanskrit bahuvrīhis' internal order to the interaction of independently motivated syntactic constraints. Other assumptions are advanced in §4: they should all be considered in this perspective.

3.6 Summary

Section §3 contained several arguments to the effect that both members of Sanskrit compounds are words or phrases. These arguments capitalise on the fact that a compound-member can enter

---

[25] Although we could not find any earlier attestation of *agny-āhita-*, it is still safe to claim that this form was common at the time when Pāṇini lived (4th c. BCE), i.e., during the late Vedic era, insofar as Pāṇini's rule A 2.2.37 (*vāhitāgnyādiṣu*) alludes to it. See Cardona (1997:223).

[26] As remarked by Perrin (1913) (quoted in Moro 2017:41), the task of the scientist is to "explain what is visible and complicated by means of what is simple and invisible".

typically syntactic relations in Sanskrit (e.g., coreferentiality, clausal scope): only words and phrases are known to be able to enter such syntactic relations in Sanskrit; the left-hand member of Sanskrit compounds enters these relations; hence, Sanskrit compounds' left-hand member is a word or phrase. When independent restrictions are factored out, the left-hand member can invert with the right-hand member in Sanskrit bahuvrīhis, as in *jāta-danta-* vs. *danta-jāta-*: to make sense of this kind of inversion, we have suggested that the right-hand member of Sanskrit bahuvrīhis also has word- or phrasal status. Concerning instead the right-hand member of Sanskrit endocentric (hence, non-bahuvrīhi) compounds, the fact that it is formally identical to a word makes it possible to assign word- or phrasal status to it. We are then left with the following picture: the left- and right-hand members of Sanskrit compounds are words or phrases; consequently, the syntactic operations applying to these compound-members should be comparable to the syntactic operations applying to non-compound words and phrases. In §4 we shall argue that movement is among the syntactic operations which apply to compound-members in Sanskrit. Furthermore, we have assumed that the left-hand member and the right-hand member of Sanskrit bahuvrīhis are not words, but rather bare phrases projected by lexical categories and lacking functional projections: this assumption will allow us to implement the Movement Strategy (3b) in terms of phrasal movement, thus dispensing with word- (i.e., head-)movement.[27]

We are now ready to explore a syntactic derivation of bahuvrīhis.

4 Movement and bahuvrīhis: The Movement Strategy

In this section we lay out a formal derivation of Sanskrit bahuvrīhis couched within the framework of generative grammar. This derivation incorporates the Movement Strategy (3b), which involves the movement of the possessor from the phrase projected by a compound-member to a position outside the compound. The restrictions on the internal order of bahuvrīhis will be seen to follow from independently motivated constraints on this movement of the possessor.

4.1 The predicative link within Sanskrit bahuvrīhis

Dikken (2006:10-12) proposed that every ascription of a property by a phrase XP to another phrase YP is mediated in syntax by a predicative link. E.g., *the cause of the riot* ascribes the property of having caused the riot to *a picture of the wall* in (21) (Moro 1997:24; 2017:164); therefore, a predicative link must occur between *the cause of the riot* and *a picture of the wall*.

(21) John renders [$_{SC}$ [$_{DP}$ a picture of the wall] [$_{DP}$ the cause of the riot]].

The predicative link is the link between the predicate and the subject-of-predication (Moro 2019:1). The predicative link is codified, at the beginning of the derivation, by a symmetrical structure referred to as SC 'small clause' (Moro 2000:71; 2019:4; Chomsky 2013:42; cf. Bittner and Hale 1996:8-10). The small clause is made up of two components α and β (namely, the subject-of-predication and the predicate) such that: i. both α and β are maximal projections; ii. α c-commands β and vice versa (Moro 2000:27; 2004:§5).[28] But how can we understand which of the two components of the small clause is the predicate and which is the subject-of-predication?

Indeed, the phrase XP that ascribes properties to another phrase YP within the small clause is the predicate, YP being the subject-of-predication (see Moro 2017:49-52, 164, building on Geach 1980:60, 80; 1987:9-10). Therefore, *the cause of the riot* is the predicate and *a picture of the wall* the subject-of-predication in (21). Now, an alternative way exists of telling subject-of-predication

---

[27] This is reminiscent of Cinque's (2005), (2010) work on adjectives. He shows that a series of generalisations on the ordering and interpretation of adjectives (including Greenberg's Universal 20) are captured by assuming that only phrasal movement is available crosslinguistically, even when the unit that moves up the tree is formally identical to a simple word. See especially Cinque (2005:317 n. 6); (2010:37-41). Cf. also Ntelitheos's (2022:11) proposal that the formation of synthetic compounds in Modern Greek involves phrasal movement, as opposed to head-movement, of the compound's left-hand member.

[28] A node α c-commands a node β iff: i. α does not dominate β; ii. the first node that dominates α also dominates β. Dominance is instead defined as follows: let T be a syntactic tree, let α and β be nodes of T; then α dominates β in T if it is possible to draw a continuous line towards the bottom from α to β.

and predicate apart: the predicate contains (or coincides with) a phrase which 'selects' the subject-of-predication. For an expression E₁ to select another expression E₂ essentially means that E₁ sets up an abstract scenario and specifies the semantic role played by the denotatum of E₂ in that scenario (see Chomsky 1986:86-87; Cinque 1990:41). Thus, *cause of the riot* (i.e., a phrase contained within the DP *the cause of the riot*) selects *a picture of the wall* in (21): *cause of the riot* sets up a scenario, describable as 'some entity bears the property of causing the riot', and specifies that the picture denoted by *a picture of the wall* serves as the property bearer in that scenario. We may then define the predicate as the component of a small clause that contains (or coincides with) the selector of that small clause's remaining component, known as subject-of-predication (see Moro 1997:115 for the original observation that predicates select subjects, not vice versa). We shall see that the notion 'selection' plays a crucial role in the following subsections.

Let us summarise the results of these paragraphs:

(22) If a phrase XP ascribes some property to another phrase YP, then: i. XP and YP form a small clause at the beginning of the derivation; ii. XP selects (or contains a selector of) YP; iii. XP is the predicate of YP.

Interestingly, Lowe (2015:101) analyses Sanskrit bahuvrīhis as expressing an embedded predication at f(unctional)-structure:[29] thus, the f-structure of a bahuvrīhi such as *ugra-putra-* ((4)) specifies that 'mighty', i.e., the basic meaning of *ugrá-*, is the value of the attribute PREDLINK (corresponding to our notion of predicate) and 'son' (i.e., the basic meaning of *putrá-*) the value of the attribute SUBJECT. This can be recast in the terms of the theory of predication adopted here, along the following lines.

The AP *ugrá-* ascribes the property of being mighty to the NP *putrá-* in *ugra-putra-* ((4)). Equivalently, *ugrá-* selects *putrá-* in this compound: *ugrá-* sets up a scenario describable as 'some entity bears the property of being mighty', and specifies that the sons denoted by *putrá-* serve as the property bearer in that scenario. This means that *ugrá-* is the predicate and *putrá-* the subject-of-predication. Then, in accordance with (22), the structure assigned to *ugra-putra-* at the beginning of the derivation is a small clause:[30]

(23) [sc [NP putra-] [AP ugra-]].

4.2 The place of the possessor within Sanskrit bahuvrīhis

In accordance with the Movement Strategy (3b), the possessor *áditi-* of such examples as (4) starts out as the specifier of the phrase projected by the possessum, hence within the phrase projected by the possessum: more explicitly, the position in which *áditi-* starts out in (4) and the position in which a genitival modifier such as *aditeḥ* starts out in *putrā́so áditeḥ* (ṚV 8.18.5) 'Aditi's sons' are one and the same position. In (4) the possessum is *putrá-*, which is generated within the small clause (23). Now, in keeping with the remarks of §3.4, *putrá-* is a bare NP in (4), i.e., an NP that does not contain complements or specifiers. How can we reconcile the proposal that *áditi-* starts out in the specifier of the phrase projected by *putrá-*, with the assumption that *putrá-* is a bare NP? We submit that complement and specifier positions are indeed in principle available in bare phrases such as [NP [N° putra-]], but that these positions cannot host <u>overt</u> linguistic material at the end of the derivation: this means that some expression may sit in the complement or specifier of, e.g., [NP [N° putra-]] at the initial stage of the derivation, provided this expression vacates that complement or specifier

---

[29] In Lexical-Functional Grammar, "f-structure represents syntactic predicate-argument structure in terms of grammatical functions such as subject and object" (Börjars, Nordlinger and Sadler 2019:13). F-structures are graphically represented as sets of attribute-value pairs.

[30] This analysis, which takes the initial stage of the derivation of a nominal compound to involve a small clause, resembles the analysis proposed by Kayne (1994:106) for *that idiot of a doctor*: *that idiot of a doctor* is a complex nominal, but nonetheless involves a predicative link between *doctor* (the subject-of-predication) and *idiot* (the predicate), and this predicative link is implemented by a small clause at the initial stage of the derivation (see also Dikken 1998; 2006:161-246; Moro 2000:49-61).

during the derivation.[31] Of course, in order for this proposal to be tenable, some independent motivation should be found that justifies the expression's vacating the complement or specifier position of the bare phrase. Let us thus assume that the possessor *áditi-* of (4) starts out as the specifier of [$_{NP}$ [$_{N^\circ}$ putra-]], as depicted in (24) (a revised version of (23)).

(24) [$_{SC}$ [$_{NP}$ aditi- [$_{N'}$ putra-]] [$_{AP}$ ugra-]].

Here *áditi-*, by virtue of being a nominal expression that sits in the specifier of the phrase projected by *putrá-*, constitutes a nominal modifier of *putrá-*, and is interpreted as the possessor of *putrá-*. Qua possessor and nominal modifier of *putrá-*, *áditi-* bears an unvalued Case feature (typically genitive) to be valued in a specific syntactic configuration: more precisely, this feature is valued when *áditi-* is in the specifier of a functional head, call it F°, such that F° is part of the DP-layer with which the possessum is endowed, as in (25).[32]

(25) [$_{DP}$ … [$_{FP}$ aditi- [$_{F'}$ [$_{NP}$ [$_{N'}$ putra-]] F°]]]

In this case the possessum is *putrá-*. But to account for the fact that bahuvrīhi-members are morphologically stems, we concluded above (see §3.4-5) that all the functional structure gravitating around these members (e.g., around *putrá-* in (4)) is missing; accordingly, *áditi-* cannot be hosted in any FP placed above *putrá-*. This in turn means that *áditi-* cannot be Case-marked in (24) (more precisely, its Case feature remains unvalued). We have thus found an independent reason for the obligatory movement of *áditi-* from the specifier of the NP headed by *putrá-* in (24): *áditi-* must move to a position outside such an NP because otherwise (i.e., if it stays in situ) it cannot be Case-marked, hence it cannot be licensed.

To sum up: Sanskrit bahuvrīhis' members are phrases projected by a lexical head, contain complements and specifiers, but lack the functional structure—including FP—gravitating around the lexical head. FP is given over, among other things, to valuing the Case feature of the lexical head's nominal modifiers. Therefore, if a nominal modifier of a bahuvrīhi-member is generated in the complement or specifier position projected by this member, such a modifier has to move to the FP gravitating around this member in order to get its Case feature valued. If FP is missing, the nominal modifier in question has to vacate the bahuvrīhi-internal specifier/ complement position in which it was generated, and reach a higher position in which the modifier's Case feature gets valued. Hence, the fact that Sanskrit bahuvrīhis' members lack the functional structure that ordinarily gravitates around non-compound phrases implies the following: that the specifier and complement positions of the phrases corresponding to Sanskrit bahuvrīhis' members cannot contain any overt nominal modifier at the end of the derivation. A welcome result, as we saw.

The take-home point from this subsection is thus the following. From a single hypothesis, i.e., that the functional structure gravitating around both bahuvrīhi-members is missing, we derive two facts: first, the stem status of the bahuvrīhi-members (i.e., the fact that they do not display inflection); second, the inability of the bahuvrīhi-members to contain overt complements or specifiers.

4.3 The linking element in Sanskrit bahuvrīhis

We have posited the structure in (24) in order to account for the interpretation of (4) (notably, the fact that *áditi-* is interpreted as the possessor of *putrá-* and that *ugrá-* ascribes the property of being mighty to *putrá-*). However, we need more structure than is contained in (24) if the Movement Strategy is to be implemented. For one thing, we should provide a landing site for *áditi-*, which we know has to end up outside the compound at the end of the derivation (as shown by (4)). But we also need some additional empty slot for theory-internal reasons.

---

[31] The idea that an illicit structure can be generated provided that some operation intervenes to rescue it before the end of the derivation is not new in the generative literature: indeed, this idea is at the core of Moro's (2000) principle of Dynamic Antisymmetry and Chomsky's (2013; 2015) Labelling Algorithm.

[32] Longobardi (2001:567) proposes that a nominal N_1 marked with an inflectional genitive ending may occupy either a high or a low position (called GenS and GenO, respectively) in the DP-periphery of the nominal N_2 modified by N_1. Longobardi and Silvestri (2013:100-101) offer a minimalist implementation of this proposal in terms of feature-valuation.

Indeed, small clauses are intrinsically unstable structures: the overt linguistic material contained within the small clause cannot be linearised (Moro 2000), and the small clause itself cannot be labelled (Moro 2009; Chomsky 2013; 2015). But linear order and labels are properties that every linguistic output must possess to qualify as well-formed.[33] Consequently, a small clause like (24) cannot reach, as it stands, the end of the derivation: it must somehow be rescued (i.e., linearised and labelled) during the derivation. In brief, the small clause is rescued if either phrase that makes up the small clause (i.e., the subject-of-predication or the predicate) is moved outside the small clause itself.[34] This means that (24) should be supplemented with a second landing site (over and above the landing site for the movement of *áditi-*), devoted to hosting either the subject-of-predication or the predicate of the small clause.

We submit that this second empty slot or landing site is the specifier of a silent linking element L°. Linking elements are meaningless extensions that occur either between the two internal members of a compound (Lieber and Štekauer 2009:§1.3) or, seldom, at the end of a compound (Okubo 2014:230-231). We propose that, in Sanskrit bahuvrīhis, L° occurs at the end of a compound stem, as in (26). More precisely, L° takes the small clause as complement, and projects an LP: indeed, LP is the syntactic structure of the whole compound stem *ugra-putra-* of (4).[35]

(26) [$_P$ __] … [$_{LP}$ __ [$_{L'}$ [$_{SC}$ [$_{NP}$ aditi- [$_{N'}$ putra-]] [$_{AP}$ ugra-]] L°]]

The P which shows up in (26) is a cover term for any position outside LP (i.e., outside the structure of the compound stem *ugra-putra-*): we identify P as the empty slot in which the movement of *áditi-* (i.e., the external referent of the bahuvrīhi *ugra-putra-*) terminates. The exact location of P within the sentence changes depending on the Case ending taken on by *áditi-*: if *áditi-* surfaces in the nominative, P coincides with some Spec-TP position (i.e., the position where nominative Case is assigned); if *áditi-* surfaces in the accusative, P coincides with the complement of a transitive verb (i.e., the position in which accusative Case is assigned), and so on.

In our account, the main reason for positing the presence of L° in (26) is theory-internal: either the subject-of-predication (i.e., *[$_{NP}$ aditi- [$_{N'}$ putra-]]*) or the predicate (i.e., *[$_{AP}$ ugra-]*) has to move outside the small clause of (26) in order to rescue the small clause itself; the linking element L° provides the landing site for this movement by projecting an empty specifier (i.e., Spec-LP). However, some evidence exists for the presence of a functional head at the end of Sanskrit bahuvrīhis. Thus, various changes take place at the end of bahuvrīhis (see Macdonell 1916:278-279, from which most of the following examples are taken).

First, if the stem of the right-hand member terminates in *-an-*, the latter may be replaced by *-a-*, as in (27). Second, the semantically empty suffixes *-a-*, *-ya-*, and *-ka-* may be added to the right-hand edge of the bahuvrīhi stem, as in (28). Third, the possessive suffix *-in-* (for which see §2.3 above) may also be added to the right-hand edge of the bahuvrīhi stem, as in (29). Last, the final *-i-* of some stems is replaced by *-a-* when such stems occur in the right-hand slot (i.e., at the end) of the bahuvrīhi; conversely, the final *-a-* of some other stems is replaced by *-i-* when these stems occur in the right-hand slot of the bahuvrīhi: see (30).[36]

---

[33] See more recently Moro and Roberts (2020). The reason why a small clause cannot be linearised or labelled is that it is a symmetrical structure (technically, a 'point of symmetry'), while labelling and linearisation are defined in such a way as to obtain in asymmetrical structures only: see Kayne (1994:3-6) (also Moro 2000:15-28) on the algorithm that determines the linear order of terminal nodes, and Chomsky (2013:43-44) on the algorithm that determines the label of nonterminal nodes.

[34] In essence, the small clause component that undergoes movement (e.g., the subject-of-predication) is replaced by a trace which, being phonologically null, need not be linearised with respect to the other small clause component (e.g., the predicate), thus solving the problem of linearisation. In addition, the trace thereby created, being unable to label the small clause, allows the member that remained in situ within the small clause to provide a label to the small clause itself, thus solving the labelling problem as well. See Moro (2000; 2004), Chomsky (2013; 2015), Rizzi (2016), and Moro and Roberts (2020) for in depth-discussion.

[35] The idea that compounds are headed by a linking element is advocated for by Di Sciullo (2009) and Delfitto and Melloni (2009). According to Eik (2019:182), L° contributes to the interpretation of the compound in which it is contained by providing "an instruction that specifies how elements should compose" (see Eik 2019:183 for an implementation of this idea in formal semantic terms).

[36] See §2.2 above for the special accentual properties of the bahuvrīhis in (27)-(30).

(27) *víśva-* 'all' + *kárman-* 'work' > *viśvá-karma-* (ṚV 10.166.4) 'whose action is on everything', 'performing all work'.

(28) a. *sú-* 'well' + *gó-/gáv-* 'cow' + *-a-* > *sugáva-* (ṚV 1.116.25) 'whose cows are good'.
     b. *anyá-* 'other' + *udára-* 'womb' + *-ya-* > *anyódarya-* (ṚV 7.4.8) 'whose womb is another', 'born from another womb'.
     c. *trí-* 'three' + *ambā́-* 'mother' + *-ka-* > *tryàmbaka-* (ṚV 7.59.12) 'whose mothers are three'.

(29) *śiti-* 'white' + *kákṣa-* 'lurking place', 'belly' + *-ín-* > *śiti-kakṣ-ín-* (TS 5.5.20.1) 'whose belly is white'.

(30) a. *daśan-* 'ten' + *aṅgúli-* 'finger' > *daśāṅgulá-* (ṚV 10.90.1) '(extension) whose fingers are ten', '(extension that is) ten fingers long'.
     b. *dhūmá-* 'smoke' + *gandhá-* 'smell' > *dhūmágandhi-* (ṚV 1.162.15) 'in which there is smell of smoke', 'smelling of smoke'.

The hypothesis that Sanskrit bahuvrīhis involve a functional head L° that is located to the immediate right of the compound stem can explain the data in (27)-(30). Consider how.

L° takes the small clause as complement to its left, and the small clause contains the compound-members, as depicted in (31) above. L° can be realised as *-a-*, *-ya-*, *-ka-*, *-ín-*, *-i-*, or Ø. After one of the components of the small clause has raised to Spec-LP, and *áditi-* has moved to P, L° is affixed to the remaining component of the small clause: this component happens to be the right-hand member of the compound, as shown in (31).[37]

(31) [$_P$ aditi$_j$-] … [$_{LP}$ [$_{AP}$ ugra-]$_i$ [$_{L'}$ [$_{SC}$ [$_{NP}$ t$_j$ [$_{N'}$ putra-]] t$_i$] L°]].
    = (aditi-)   ugra-putra-L°

Although we are still far from a full understanding of these facts, it seems reasonable to claim that the affixation of L° to the right-hand member of the bahuvrīhi brings about phonological changes in the final segment of the right-hand member itself: deletion of a stem-final sound, as in (27); addition of a suffix, as in (28)-(29); replacement of a stem-final sound with another stem-final sound, as in (30). If no functional head were present at the end of a bahuvrīhi, the phonological changes in (27)-(30) would be hard to justify.

Let us now return to the well-known fact that Sanskrit bahuvrīhis agree in Case, gender, and number with the external referent of the compound, e.g., with *áditi-* in (4) (see §3.5 above). We should like to show that a relatively simple explanation emerges for this fact if the presence of the linking element L° is posited in the structure of Sanskrit bahuvrīhis.
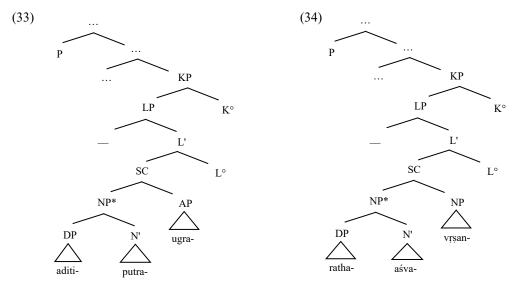
To begin with, if LP constitutes the structure of the bahuvrīhi stem (see (26) and (31)), where should the Case ending of a bahuvrīhi be housed? Following Bittner and Hale (1996), we consider Case endings as being the spell-out of a functional head (the Case head K°) which takes a nominal stem as complement. Simplifying somewhat, a Case-marked nominal such as *sákhy-uḥ* (< *sakhi-uḥ*) 'companion-GEN.SG' is analysed as [$_{KP}$ [$_{NP}$ sakhi] [$_{K°}$ uḥ]]: that is, the stem *sákhi-* is an NP, while the Case ending *-uḥ* is a K°, K° being a full-fledged syntactic head; *-uḥ* takes *sakhi-* as complement, and projects a KP; accordingly, the whole Case-marked nominal *sákhy-uḥ* qualifies as a KP headed by *-uḥ* (see Hale and Bittner 1996:3-6). In the same vein, we consider the Case ending of a bahuvrīhi as the spell-out of a Case head K° which takes the whole bahuvrīhi stem LP as complement. For instance, the structure of *ugraputrā* would be the following (based on a slight modification of (31)), with the NOM.SG.F ending *-ā* of *ugraputrā* sitting under K°:

(32) [$_P$ aditi$_j$-] … [$_{KP}$ [$_{LP}$ [$_{AP}$ ugra-]$_i$ [$_{L'}$ [$_{SC}$ [$_{NP}$ t$_j$ [$_{N'}$ putra-]] t$_i$] L°]] K°].
    = (Aditi-)   ugra-putra-L°-ā

---

[37] *t$_x$*, which stands for 'trace bearing index x', indicates the position occupied by a coindexed phrase (e.g., XP$_X$) before that phrase moved to another position. See §4.5 below on the mapping between hierarchical structure and linear order.

Note now that the Case ending sitting under K° 'views' (i.e., is sensitive to) the gender of the NP stem which K° takes as complement. Thus, when K° takes the masculine NP stem *sákhi-* as complement and K° is GEN.SG, then K° is realised as *-uḥ* (i.e., *sákhy-uḥ*); when instead K° takes the feminine NP stem *sakhī-* as complement and K° is GEN.SG, then K° is realised as *-āḥ* (i.e., *sakhy-āḥ*). However, K° takes LP—and not an NP stem—as complement in (32). Therefore, K° cannot see the gender of any compound-internal NP (e.g., K° cannot see the masculine gender of the NP stem *putrá-* in (32)). Hence the bahuvrīhi's Case ending, which is the spell-out of K°, can acquire the gender of the external referent of the bahuvrīhi: e.g., the Case ending *-ā* acquires the feminine gender of *áditi-* in (32). This explains away the fact that Sanskrit bahuvrīhis have adjectival agreement (i.e., agree in Case, gender, and number with the external referent of the bahuvrīhi). This explanation crucially rests on the intervention of L° between a bahuvrīhi-internal NP stem (e.g., *putrá-* in (32)) and K°: thus, insofar as it is tenable, this explanation lends further support to the postulation of a linking element within Sanskrit bahuvrīhis.

All in all, the proposal that there is a silent functional head L° that takes a small clause as complement in such Sanskrit bahuvrīhis as (4)-(7) is at least compatible with the data found within the Sanskrit language system. Of course, this compatibility does not per se suffice to legitimate the postulation of a silent head: the crucial ingredient needed for legitimating a silent head is that its presence allows us to explain some data (i.e., to decompose complex empirical phenomena into simpler abstract entities). In this case, we shall show that the presence of that head makes it possible to derive the restrictions on the internal orders in (4)-(7) (a complex empirical phenomenon) from the interaction of independently motivated syntactic principles (simpler abstract entities). Thus, we take (26) (= (33)) as the initial stage of the derivation of *ugra-putra-* ((4)).
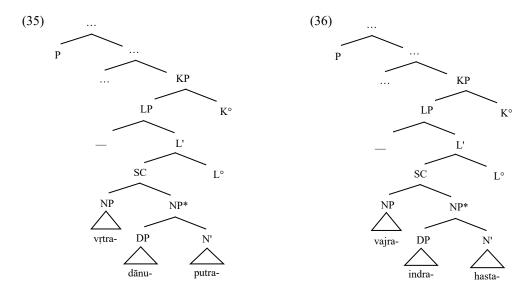
(33)

(34)

Following the same reasoning, we posit an exactly parallel structure for the initial stage of the derivation of *vṛṣan-aśvá-* ((5)), *vṛtrá-putra-* ((6)), and *vájra-hasta-* ((7)).

vŕṣan- 'bull' ascribes the properties typical of bulls (notably, the prominence) to *áśva-* 'horse' in *vṛṣan-aśvá-*. Equivalently, *vŕṣan-* selects *áśva-* in *vṛṣan-aśvá-*: *vŕṣan-* sets up a scenario describable as 'some entity bears the prominence typical of bulls', and specifies that the horses denoted by *áśva-* serve as the prominence bearer in that scenario.

Likewise, *putrá-* 'son' ascribes the property of being a son (specifically, a son of Dānu) to *vṛtrá-* 'Vṛtra' in *vṛtrá-putra-*. Equivalently, *putrá-* selects *vṛtrá-* in *vṛtrá-putra-*: *putrá-* sets up a scenario describable as 'some entity bears the property of being a son of Dānu', and specifies that the individual denoted by *vṛtrá-* serves as the property bearer in that scenario.

By the same token, *hásta-* 'hand' ascribes the property of being in one's hand (specifically, in Indra's hand) to *vájra-* 'mace' in *vájra-hasta-*. Equivalently, *hásta-* selects *vájra-* in *vájra-hasta-*: *hásta-* sets up a scenario describable as 'some entity bears the property of being in Indra's hand', and specifies that the mace denoted by *vájra-* is the property bearer in that scenario.

These bits of semantic information translate, by generalisation (22), into the following structural claims: at the initial stage of the derivation of *vṛṣaṇ-aśvá-*, *vṛ́ṣan-* and *áśva-* form a small clause in which *vṛ́ṣan-* is the predicate and *áśva-* the subject-of-predication; at the initial stage of the derivation of *vṛtrá-putrá-*, *putrá-* and *vṛtrá-* form a small clause in which *putrá-* is the predicate and *vṛtrá-* the subject-of-predication; at the initial stage of the derivation of *vájra-hasta-*, *hásta-* and *vájra-* form a small clause in which *hásta-* is the predicate and *vájra-* the subject-of-predication. After the insertion of the linking element L° and the Case head K°, we get the phrase markers (33)-(36); in each of these, the bahuvrīhi's external referent (*áditi-*, *rátha-*, *dā́nu-*, and *índra-*) is represented as a DP to indicate that it bears an unvalued Case feature that needs to be valued later in the derivation.

The trees in (33)-(36) represent the initial stage of the derivation of the bahuvrīhis in (4)-(7). In the remainder of this paper, we shall show that the restrictions on the internal orders in (4)-(7) follow from the constraints at play in the derivation of such bahuvrīhis.

4.4 The Subjacency Condition

Movement has been discovered to obey nontrivial conditions, technically known as locality conditions. One of these is the so-called Subjacency Condition, which constrains movement from maximal projections. We adopt the following version of the Subjacency Condition, due to Moro (1997:49-57, 115; 2017:108, 163), who builds on Cinque (1990:40-43):

(37) Let H be a head nondistinct from [+V]. Then, nothing can move from a maximal projection XP
    unless one of the following two conditions is met:
    i.   XP is a predicate;
    ii.   H locally c-commands and selects XP.

The crucial notions involved in (37) are local c-command and selection. A node α locally c-commands a node β iff: i. α c-commands β; and ii. there is no node γ such that γ c-commands β and is c-commanded by α (see Rizzi 1990:7; Rizzi and Shlonsky 2007:139). An element α selects an element β if α sets up an abstract scenario and specifies the semantic role played by the denotatum of β in that scenario (see §4.1 above).

Crucially, Moro (1997:115-127) has shown that a head nondistinct from [+V] which does not per se select an XP may indeed derivatively select such XP. This happens whenever condition (38) obtains. This condition essentially states that, if a head is close enough to a certain selector, then that head inherits the selectional capacities of that selector.

(38) Let H be a head nondistinct from [+V] and let XP be a maximal projection. Then, H derivatively selects XP iff: the selector of XP locally c-commands H.

Moro (1997:124) capitalises on (38) to explain the contrast between (39) and (40):

(39) a. [DP The [NP* cause of the riot]]i was [SC [a picture of the wall] ti].
    b. *[Which wall]j do you think [DP the [NP* cause of the riot]]i was [SC [a picture of tj] ti]?

(40) a. [There]i was [SC [a picture of the wall] ti].
    b. [Which wall]j do you think [there]i was [SC [a picture of tj] ti] ?

The NP* *cause of the riot* is the selector of *a picture of the wall* in (39): NP* sets up a scenario describable as 'some entity bears the property of causing the riot', and specifies that the picture denoted by *a picture of the wall* serves as the property bearer in that scenario; note that in this case the selector (NP*) is contained within the predicate (the whole DP *the cause of the riot*), therefore does not coincide with the predicate. Likewise, *there* is the selector of *a picture of the wall* in (40): *there* sets up a scenario describable as 'some entity bears the property of existing', and specifies that the picture denoted by *a picture of the wall* serves as the property bearer in that scenario (see Moro 1997:142-148 for a precise characterisation of existential meaning); in this case, the selector (*there*) coincides with the predicate. Both (39) and (40) are instances of inverse copular sentences, in the sense that the predicate (the DP *the cause of the riot* and *there*, respectively) raises to the specifier of *was* (see Moro 1997: Ch. 1 and 2 for in-depth discussion). So why is the extraction of *which wall* possible in (40b) but not (39b)?

In a nutshell, *was* is a head nondistinct from [+V] in both (39) and (40), by definition. Moreover, *was* does not select the subject-of-predication (*a picture of the wall*) in (39)-(40), as the subject is already selected by the predicate (*there*), or by a phrase contained within the predicate (the NP* *cause of the riot*). Note however that, while *there* locally c-commands *was* in (40), the NP* *cause of the riot* does not locally c-command *was* in (39): the DP dominating *the* intervenes between NP* and *was* in (39), thereby preventing NP* from c-commanding *was*. If we now bear in mind that the NP* *cause of the riot* and *there* are the selectors of *a picture of the wall* in (39) and (40), respectively, the following picture emerges: only in (40) does the selector of *a picture of the wall* locally c-command *was*. This means that *was* can derivatively select the subject-of-predication *a picture of the wall* in (40) but not in (39), in accordance with generalisation (38).

Since *a picture of the wall* is not a predicate in (38)-(39), movement from it is only possible if *a picture of the wall* is locally c-commanded and selected by *was* (which is a head nondistinct from [+V]), in compliance with the Subjacency Condition. Indeed, *a picture of the wall* is locally c-commanded and (derivatively) selected by *was* in (40) but not (39): consequently, movement from *a picture of the wall* is permitted in (40) but not in (39). Thus, the movement of *which wall* in (40b) yields a well-formed sentence because this movement does not violate the Subjacency Condition. On the contrary, the movement of *which wall* in (39b) gives an ill-formed result because it violates the Subjacency Condition (see Moro 1997:115-127 for further details).
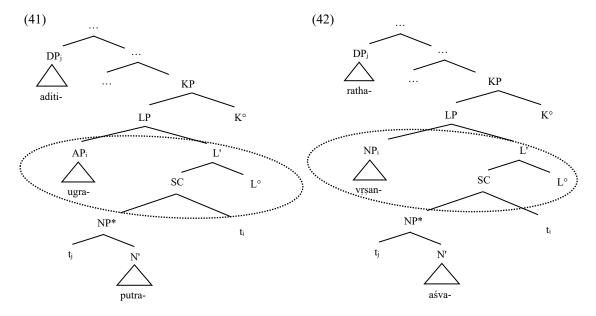
The claim made at the end of the previous subsection, to the effect that the restrictions on the internal orders in (4)-(7) are a consequence of the constraints at play in the derivation of Sanskrit bahuvrīhis like (4)-(7), can now be made more precise: such restrictions follow from the Subjacency Condition. We attempt to show this in the next subsection.

4.5 A unified derivation for Sanskrit bahuvrīhis

The structures in (33)-(36) are unstable in two senses. First, they contain a point of symmetry (i.e., the small clause); second, they contain an element (i.e., the bahuvrīhi's external referent: *áditi-*, *rátha-*, *dā́nu-*, and *índra-*) whose Case feature cannot be valued in situ (i.e., in the specifier of NP*). The problem concerning the point of symmetry is solved if one of the two small clause components (e.g., the NP* *[[aditi-] putra-]* or the AP *ugrá-* in (33)) is displaced to Spec-LP; the problem regarding the unvalued Case feature is instead solved if the bahuvrīhi's external referent (e.g., *áditi-* in (33)) moves to P, which stands for any position in which the Case feature of this external referent can be valued (e.g., Spec-TP).

Now, the bahuvrīhi's external referent cannot directly move to P in (33)-(34). Consider why. The bahuvrīhi's external referent is contained within NP* (a maximal projection) in (33)-(34). Crucially, NP* serves as the subject-of-predication in (33)-(34). Movement from a maximal projection that does not qualify as a predicate complies with the Subjacency Condition (37) only if that maximal projection is selected and locally c-commanded by a head bearing verbal features. We assume that, by virtue of taking a clausal complement (i.e., the small clause), L° bears some verbal feature, on a par with the copula *was* in (39)-(40) above. Now, L° locally c-commands NP*, but does not select it in (33)-(34) (or in (35)-(36)): indeed, NP* is already selected by the small clause predicate, represented as the right-hand component of the small clause (i.e., *ugrá-* and *vŕṣan-*) in (33)-(34). Since L° fails to select NP*, and since NP* is the subject-of-predication (hence, does not qualify as a predicate) in (33)-(34), the direct movement of the external referents *áditi-* and *rátha-* from within NP* to P in (33)-(34) violates the Subjacency Condition. Therefore, this direct movement should be excluded. But if *áditi-* and *rátha-* do not reach P, the bahuvrīhis *ugra-putra-* ((4)) and *vṛṣan-aśvá-* ((5)) cannot be formed: *áditi-* and *rátha-* must end up outside the bahuvrīhis at the end of the derivation, otherwise the Case feature of *áditi-* and *rátha-* cannot be valued. Hence, something must happen in (33)-(34) that enables the movement of *áditi-* and *rátha-* to P. Recall that either component of the small clause must raise to Spec-LP to rescue the small clause. We consider both possibilities in turn.

When it is the right-hand component (i.e., the AP *ugrá-* in (33); the NP *vŕṣan-* in (34)) which raises to Spec-TP, (41)-(42) are obtained.

This raising has dramatic consequences. The AP *ugrá-* and the NP *vŕṣan-* now locally c-command L° in their respective trees; moreover, such phrases are the selectors (as well as the predicates) of the NP*s (subjects-of-predication) *[[aditi-] putra-]* and *[[ratha-] aśva-]*, respectively. But this means that L° derivatively selects NP* in both (41) and (42), in accordance with generalisation (38). Since L° also locally c-commands NP* in (41)-(42), the Subjacency Condition is now satisfied (see specifically condition (37ii)): movement from NP* is possible in (41)-(42) insofar as NP* is locally c-commanded and (derivatively) selected by L°, a head which we have assumed is nondistinct from [+V] (i.e., L° bears some verbal feature). Therefore *áditi-* and *rátha-* can now reach P, yielding (41)-(42), which are the structures corresponding to *ugra-putra-* and *vṛṣan-aśvá-*.

We abstract away from the ordering of *áditi-* with respect to *ugra-putra-*, and from the ordering of *rátha-* with respect to *vṛṣan-aśvá-*: *áditi-* may in principle precede or follow *ugra-putra-* in (4), much as *rátha-* may in principle precede or follow *vṛṣan-aśvá-* in (5). Focusing instead on the internal order of the bahuvrīhis *ugra-putra-* and *vṛṣan-aśvá-*, we must make sure that *ugrá-* linearly precedes *putrá-* in (41), and that *vŕṣan-* linearly precedes *aśva-* in (42), otherwise the correct order for *ugra-putra-* and *vṛṣan-aśvá-* is not obtained. This result is achieved without further stipulation

by appealing to the following considerations on Sanskrit word order.

First, in accordance with Hale and Bittner's (1996) analysis, which we have adopted here (see §4.3 above), a Case ending is a syntactic head K°, and the nominal to which the Case ending attaches is the complement of K°. Case endings follow the nominal to which they attach in Sanskrit: e.g., *devá-m* (god-ACC.SG). Therefore, Case heads follow their complements in Sanskrit.

Second, interrogative pronouns occupy the specifier of a functional head in the CP-field in Sanskrit; for simplicity's sake, let us call this position Spec-CP, i.e., specifier of the C° head (see Hale 2018:1933 for a finer representation of the left periphery in Vedic). Moreover, the sentence over which the interrogative pronoun takes scope is projected as the complement of C°. Now, interrogative pronouns ordinarily precede the sentence over which they take scope, as shown below.

(43) kím   ichántī   sarámā   prédám
      what.ACC.SG.N   seek.PRS.PTCP.NOM.SG.F   Saramā.NOM.SG.F   PVB_here
      ānaṭ?
      arrive.IMPF.3SG
      'Seeking what has Saramā arrived here?' (ṚV 10.108.1a; tr. Jamison and Brereton 2014).
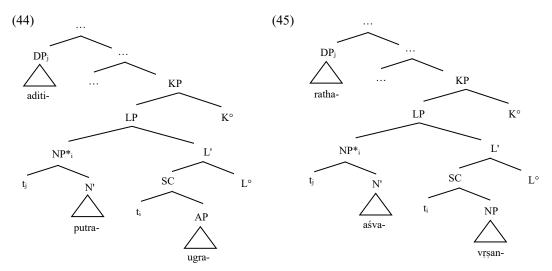
Accordingly, the specifier of C° precedes the complement of C° in Sanskrit. We assume that this state of affairs (i.e., the fact that the specifier of a head precedes the complement of that head) also holds for other phrases in which a specifier can be safely detected. The following picture thus emerges: if X° is a Case-like head, X° follows its complement; the specifier of X° precedes the complement of X° generally. To wit, if we confine our attention to phrases headed by a Case-like head, the internal order of these phrases in Sanskrit is: Specifier-Complement-Head.

Let us now come back to (41)-(42). The lexical items contained in Spec-LP (i.e., *ugrá-* and *vṛ́ṣan-*) linearly precede the lexical items contained in the complement of L° (i.e., *putrá-* and *áśva-*, which are the only overt lexical items dominated by the small clause in (41)-(42)), by virtue of the remarks made above: the specifier of a head precedes the complement of that head in Sanskrit.[38] Consequently, we obtain: *ugra-putra-* and *vṛṣaṇ-aśvá-*. These are the correct internal orders of the bahuvrīhis in (4) and (5).

We now turn to the second possible strategy for rescuing the small clause of (33)-(34), i.e., the strategy in which the left-hand component of the small clause (the NP* *[[aditi-] putra-]* in (33); the NP* *[[ratha-] aśva-]* in (34)) raises to Spec-LP. This is depicted in (44)-(45) below. There is no head that locally c-commands and simultaneously selects NP* in (44)-(45): indeed, L° here does not select NP* derivatively either (in the sense of (38)), because the AP *ugrá-* (i.e., the selector of NP* in (44)) and the NP *vṛ́ṣan-* (i.e., the selector of NP* in (45)) do not locally c-command L° in (44)-(45). As a consequence, nothing can be extracted from NP* in accordance with the Subjacency Condition. We have then reached the following conclusion: the external referents *áditi-* and *rátha-* must move to P to get their Case feature valued; yet, the movement of *áditi-* and *rátha-* to P, shown in (44)-(45), violates the Subjacency Condition. Hence, (44)-(45) are ruled out both if *áditi-* and *rátha-* remain in situ, and if they move to P.
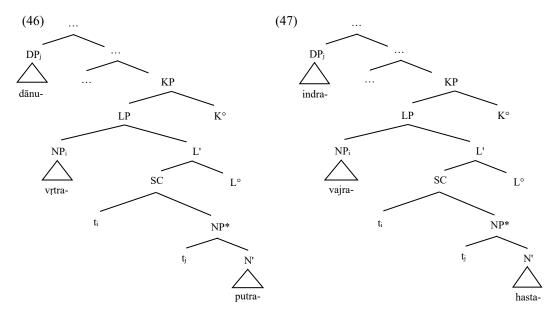
Capitalising on the fact that the specifier of a head precedes the complement of that head, we can map the lexical items populating LP in (44)-(45) to precise linear orders: *putrá-* (which occupies Spec-LP) linearly precedes *ugrá-* (which is contained within the complement of L°) in (44); by the same token, *vṛ́ṣan-* (which occupies Spec-LP) linearly precedes *áśva-* (which is contained within the complement of L°) in (45). To wit, (44)-(45) correspond to the bahuvrīhis **putrogra-* and **aśva-vṛṣan-*, respectively. Since (44)-(45) are illegal structures, **putrogra-* and **aśva-vṛṣan-*, which are derived from (44)-(45), should be ill-formed. This is a welcome result, insofar as **putrogra-* and **aśva-vṛṣan-* are unattested as bahuvrīhis: their non-attestation can now be explained as a result of them violating the Subjacency Condition.

---

[38] When L° is overtly realised (i.e., when L° is spelled out as *-a-*, *-ya-*, *-ka-*, *-ín-*, or *-i-*, as in the examples of §4.3 above), it is akin to a Case ending: L° is a suffix which attaches to a nominal contained in the complement to L°, just as K° is a suffix which attaches to a nominal contained in the complement to K°. Hence, L° linearly follows its complement in (41)-(42), i.e., L° linearly follows the content of the small clause. This may be disregarded here because L° is silent in (41)-(42).

(44)

(45)

Now, (4) and (5) are instances of bahuvrīhis in which the predicate is nominal and the external referent is <u>the Possessor of the subject-of-predication</u>. Clause (i) of generalisation (15) requires that, in such bahuvrīhis, the predicate precede the subject-of-predication; we now have a principled explanation as to why this is so. The internal order in which the predicate precedes the subject-of-predication corresponds to a syntactic configuration that permits the movement of the external referent <u>from within the subject-of-predication</u> not to violate the Subjacency Condition. Consequently, clause (i) of generalisation (15) falls out from the Subjacency Condition.

It remains to show how we get from the phrase markers in (35)-(36) to the bahuvrīhis in (6)-(7). To begin with, the external referents *dā́nu-* and *índra-* must move from within NP* to P to get their Case feature valued in (35)-(36). Moreover, either component of the small clause must raise to Spec-LP to rescue the small clause in (35)-(36). Consider the possibility in which the left-hand component of the small clause (i.e., the NP *vṛtrá-* in (35); the NP *vájra-* in (36)) raises to Spec-LP; after the movement of the external referents *dā́nu-* and *índra-* to P, (35)-(36) become as in (46)-(47). The movement of *dā́nu-* and *índra-* from within NP* to P, depicted in (46)-(47), complies with the Subjacency Condition: indeed, NP* fulfills the function of predicate in (46)-(47), and the Subjacency Condition as formulated in (37) permits movement from a maximal projection that qualifies as a predicate.

(46)

(47)

We now proceed with mapping the lexical items populating LP in (46)-(47) to precise linear orders,

in keeping with our guiding principle that the specifier of a head precedes the complement of that head: (46) maps to *vṛtrá-putra-*, and (47) maps to *vájra-hasta-*. These are the correct internal orders for the bahuvrīhis in (6) and (7).

In contrast, if it is the right-hand component of the small clause (i.e., the NP* *[[dānu-] putra-]* in (35); the NP* *[[indra-] hasta-]* in (36)) that raises to Spec-LP, (35)-(36) become as in (48)-(49) (see below). The phrase markers in (48)-(49) also feature the movement of *dā́nu-* and *índra-* from within NP* to P. Now, NP* is a predicate in (48)-(49) just like in (46)-(47); movement from maximal projections that qualify as predicates is permitted by the Subjacency Condition (37), as we have seen in (46)-(47); accordingly, we would expect the movement of *dā́nu-* and *índra-* from within NP* to P to be Subjacency-compliant in (48)-(49). Note however that NP* is a <u>moved</u> predicate in (48)-(49), unlike in (46)-(47). But movement from a moved predicate is generally difficult (see Moro 1997:50-51; Bianchi and Chesi 2014 and references therein; cf. also Bošković's 2018 study on movement from moved elements); hence, whatever principle bans the movement from a moved predicate will also ban the movement of *dā́nu-* and *índra-* from NP* illustrated in (48)-(49).[39] Since (48)-(49) feature an illegal instance of movement (i.e., the movement of *dā́nu-* and *índra-* from NP*), (48)-(49) are illegal structures.

The LPs in (48)-(49) map to **putra-vṛtra-* and **hasta-vajra-*, respectively, in keeping with the Specifier-Complement-Head scheme. Since **hasta-vajra-* and **putra-vṛtra-* derive from (48)-(49), which are illegal structures, they should be ill-formed: this is a welcome result, insofar as **putra-vṛtra-* and **hasta-vajra-* are unattested as bahuvrīhis.[40]

To sum up, (6) and (7) are instances of bahuvrīhis in which the predicate is nominal and the external referent is interpreted as <u>the Possessor of the predicate</u>. Clause (ii) of generalisation (15) requires that, in such bahuvrīhis, the predicate follow the subject-of-predication; we now have a principled explanation as to why this is so. Indeed, the internal order in which the predicate follows the subject-of-predication corresponds to a syntactic configuration that permits the movement of the external referent <u>from within the predicate</u> not to violate the Subjacency Condition. As a consequence, clause (ii) of generalisation (15) falls out from the Subjacency Condition.

We have thus reduced both clause (i) and clause (ii) of generalisation (15) to an independently motivated principle: the Subjacency Condition.

4.6 Summary and further remarks

---

[39] It is unclear whether the ban on movement from a moved predicate can be completely reduced to the Subjacency Condition. Such a reduction is carried out in Moro (1997:50-57).

[40] This explanation of the ill-formedness of **putra-vṛtra-* and **hasta-vajra-* is only tenable if the movement of *dā́nu-* and *índra-* from NP* to P in (48)-(49) takes place after the movement of NP* to Spec-LP. It remains unclear, however, what prevents the movement of *dā́nu-* and *índra-* from NP* to P from taking place before the movement of NP* to Spec-LP.

In this section we have proposed a unified derivation of Sanskrit bahuvrīhis like (4)-(7). The main ingredients of this derivation are: first, the assumption that the bahuvrīhi's external referent starts out within the phrase projected by one of the bahuvrīhi-members of (4)-(7), but is moved outside the bahuvrīhi (i.e., outside LP) during the derivation, in accordance with the Movement Strategy (3b); second, the identification of the predicate with the bahuvrīhi-member which selects the remaining bahuvrīhi-member; third, the adoption of a mapping principle whereby the lexical items in Spec-LP linearly precede the lexical items contained in the complement of L°.

Next, we have shown that, under the proposed derivation, the restrictions on the internal orders in (4)-(7)—described in generalisation (15)—reduce to an independently motivated syntactic principle, namely Moro's (1997) version of the Subjacency Condition. Specifically, we have argued that the position occupied by the bahuvrīhi-internal predicate with respect to the bahuvrīhi-internal subject in (4)-(7) merely reflects the syntactic configuration needed for licensing the movement of the bahuvrīhi's external referent from the phrase in which this external referent was generated. Thus, the bahuvrīhi's internal order in which the (nominal) predicate precedes the subject-of-predication corresponds to a syntactic configuration that permits the movement of the bahuvrīhi's external referent from within the subject-of-predication not to violate the Subjacency Condition; on the other hand, the bahuvrīhi's internal order in which the (nominal) predicate follows the subject-of-predication corresponds to a syntactic configuration that permits the movement of the bahuvrīhi's external referent from within the predicate not to violate the Subjacency Condition. The derivation proposed in this section is a specific implementation of the Movement Strategy (3b). Insofar the restrictions on the internal orders in (4)-(7) follow from the Subjacency Condition under this derivation, we may say that the Movement Strategy predicts the internal orders in (4)-(7). On the other hand, the Elided Possessive Unit Strategy (3a) fails to predict the internal orders in (4)-(7) (see §2.3 above). We conclude that the Movement Strategy is empirically superior to the Elided Possessive Unit Strategy.

Furthermore, since the proposed derivation of Sanskrit bahuvrīhis makes crucial use of movement, which is an operation characteristic of syntax as opposed to morphology, the success of the proposed derivation in predicting the internal orders (4)-(7) would also speak in favour of Lowe's (2015) conclusion that Sanskrit compounds are formed by syntactic rules (see §3 above).

Finally, it should be pointed out that generalisation (15) only concerns those bahuvrīhis like (4)-(7) in which the predicate is nominal and the external referent is interpreted as the Possessor of one of the bahuvrīhi-members. Thus, the domain of bahuvrīhis whose predicate is verbal (e.g., *kṛtá-brahman-*, mentioned in §2.1 above), and of bahuvrīhis whose external referent is not interpreted as a Possessor, is still to be investigated.

5 Conclusion

Many Sanskrit bahuvrīhis involve a possessive relation whereby one of the bahuvrīhi-members is the possessum (e.g., *putrá-* in (4)) and the external referent of the bahuvrīhi is the corresponding possessor (e.g., *áditi-* in (4)).

In this work we have shown how this possessive relation can be modelled via the Movement Strategy. That is, at the initial stage of the derivation of the bahuvrīhi, the external referent of the bahuvrīhi is contained within the phrase projected by one of the bahuvrīhi-members (e.g., *áditi-* is housed in the specifier of the NP *putrá-* at the initial stage of the derivation of (4)). In this position the external referent is read as the possessor (in a broad sense) of the bahuvrīhi-member in which it is contained (e.g., *áditi-* is read as the possessor of *putrá-* in (4)). Then, in a later stage of the derivation, the external referent moves from the phrase projected by the possessum and lands in a bahuvrīhi-external position, which we have called P (e.g., *áditi-* moves from within the NP *putrá-* to P in (4)). The movement of the bahuvrīhi's external referent from the phrase projected by the possessum to P is constrained by the Subjacency Condition. The internal order of the bahuvrīhi corresponds to a precise syntactic configuration that may or may not violate the Subjacency Condition. This makes it possible to reduce the restrictions on the internal order of Sanskrit bahuvrīhis like (4)-(7), described in generalisation (15), to the Subjacency Condition: certain internal orders are excluded because they correspond to syntactic configutations that violate the Subjacency Condition, while other internal orders are permitted because they correspond to syntactic configurations that comply with the Subjacency Condition. Accordingly, the Movement Strategy predicts the internal

orders in (4)-(7). By contrast, the Elided Possessive Unit Strategy, according to which a linguistic unit that conveys the meaning 'having' and has undergone ellipsis is affixed to an endocentric compound, fails to predict the internl orders in (4)-(7). This means the Movement Strategy is empirically superior to the Elided Possessive Unit Strategy.

We have concluded from this that the possessive relation holding between the external referent of the bahuvrīhi and one of the bahuvrīhi-members in Sanskrit is established via the Movement Strategy, and not via the Elided Possessive Unit Strategy.

Acknowledgements

References

Primary Literature

A   [Aṣṭādhyāyī]. The Aṣṭādhyāyī of Pāṇini, ed. and tr. by R. N. Sharma, 5 vols. Delhi: Munshiram Manoharlal, 1990-2003.
AVŚ   [Atharvaveda Śaunakīya]. Atharvavedasaṃhitā der Śaunakaśākhā: Eine neue Edition unter besonderer Berücksichtigung der Parallelstellen der Paippalādasaṃhitā, ed. by J-S. Kim. https://www.phil.uni-wuerzburg.de/fileadmin/04080400/2021/Anhang_AVS_Text_06.2021.pdf
BĀU   [Bṛhadāraṇyaka-Upaniṣad]. Bṛhadāraṇyaka-Upaniṣad. In V.P. Limaye and R.K. Vadekar, eds., Eighteen Principal Upaniṣad, Vol. 1. Poona: Vaidika Saṃsodhana Maṇḍala, 1958.
KB   [Kauṣītaki Brāhmaṇa]. Das Kauṣîtaki Brâhmaṇa, ed. by B. Lindner. Jena: H. Costenoble, 1887.
JB   [Jaiminīya Brāhmaṇa]. Jaiminīya Brāhmaṇa of the Sāmaveda, ed. by R. Vira and L. Chandra. Varanasi: Arya Bharati 1954. Reprinted, Delhi: Motilal Banarsidass, 1986.
MBh   [Mahābharata]. The Mahābharata, ed. by V.S. Sukthankar, S.K. Belvalkar, P.L. Vaidya et al., 19 Vols., Poona: Bhandarkar Oriental Research Institute, 1933-1966.
MS   [Maitrāyaṇī Saṃhitā]. Maitrāyaṇī Saṃhitā, ed. by L. von Schroeder, 4 Vols. Leipzig: Brockhaus, 1881-1886. Reprinted, Wiesbaden: Steiner, 1970.
PB   [Pañcaviṃśa Brāhmaṇa = Tāṇḍya-Mahābrāhmaṇa]. The Tāṇḍyamahābrāhmaṇa Belonging to the Sāma Veda with the Commentary of Sāyāṇāchārya, ed. by A. Chinnasvāmi Śāstrī and Paṭṭābhirāma Śāstrī. 2 Vols. Benares: Krishnadas-Gupta, 1935-1936.
ṚV   [Ṛgveda]. Rig Veda: A metrically restored text with an introduction and notes, ed. by B.A. van Nooten and G.B. Holland. Cambridge, MA: Harvard University Press, 1994.
ŚB   [Śatapatha Brāhmaṇa]. The Çatapatha-Brâhmaṇa in the Mâdhyandina-Çâkhâ with Extracts from the Commentaries of Sâyaṇa, Harisvâmin and Dvivedaganga, ed. by A. Weber. Berlin: Dümmler, 1855. Reprinted, Varanasi: Chowkhamba, 1964.
TS   [Taittirīya Saṃhitā]. Die Taittirîya-Saṃhitâ, ed. by A. Weber. Leipzig: Brockhaus 1871-1872.
VaikhGS [Vaikhānasasmārtasūtram]. Vaikhānasasmārtasūtram, the domestic rules of the Vaikhānasa school belonging to the Black Yajurveda, ed. by W. Caland. Calcutta: Baptist Mission Press, 1927.

Secondary Literature

Barcelona, Antonio. 2008. The interaction of metonymy and metaphor in the meaning and form of 'bahuvrihi' compounds. Annual Review of Cognitive Linguistics 6:208-281.
Bauer, Laurie. 2008. Exocentric Compounds. Morphology 18:51-74.

Bauer, Laurie. 2016. Re-evaluating exocentricity in word-formation. In D. Siddiqi and H. Harley, eds., Morphological Metatheory, pages 461-477. Amsterdam: John Benjamins.

Bauer, Laurie. 2017. Compounds and Compounding. Cambridge: Cambridge University Press

Benfey, Theodor. 1852. Vollständige Grammatik der Sanskritsprache. Zum Gebrauch für Vorlesungen und zum Selbststudium. Leipzig: Brockhaus.

Bianchi, Valentina and Cristiano Chesi. 2014. Subject Islands, Reconstruction, and the Flow of the Computation. Linguistic Inquiry 45(4):525-569.

Bittner, Maria and Ken Hale. 1996. The Structural Determination of Case and Agreement. Linguistic Inquiry 27(1):1-68.

Bopp, Franz. 1827. Ausführliches Lehrgebäude der Sanskrita Sprache. Berlin: Dümmler.

Bošković, Želiko. 2018. On Movement out of Moved Elements, Labels, and Phases. Linguistic Inquiry 49(2): 247-282.

Candotti, Maria Piera and Tiziana Pontillo. 2019. Lexical subordination in derivation and compounding: Pāṇini's focusing on the non-head. Studi e Saggi Linguistici. LVIII(2):155-175.

Candotti, Maria Piera and Tiziana Pontillo. 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. Delhi: Motilal Banarsidass.

Chomsky, Noam. 1986. Knowledge of language: Its nature, origin and use. New York: Praeger.

Chomsky, Noam. 1995. The Minimalist Program. Cambridge, MA: MIT Press.

Chomsky, Noam. 2013. Problems of projection. Lingua 130:33-49.

Chomsky, Noam. 2015. Problems of projection: Extensions. In E. Di Domenico, C. Hamann, and S. Matteini, eds., Structures, strategies and beyond: Studies in honour of Adriana Belletti, pages 47-66. Amsterdam: John Benjamins.

Chomsky, Noam. 2019. The UCLA Lectures. Lingbuzz/005485.

Cinque, Guglielmo. 1990. Types of Ā-dependencies. Cambridge, MA: MIT Press.

Cinque, Guglielmo. 2005. Deriving Greenberg's Universal 20 and Its Exceptions. Linguistic Inquiry 36(3):315-332.

Cinque, Guglielmo. 2010. The Syntax of Adjectives: A Comparative study. Cambridge, MA: MIT Press.

Corver, Norbert. 2013. Lexical categories and (extended) projection. In M. den Dikken, ed., The Cambridge Handbook of Generative Syntax, pages 353-424. Cambridge: Cambrisge University Press.

Delfitto, Denis and Chiara Melloni. 2009. Compounds don't come easy. *Lingue e Linguaggio* 8(1):75-104.

Deshpande, Madhav M. 1989. Ellipsis in modern Linguistics and Pāṇini. Annals of Bhandarkar Oriental Research Institute 70:103-124.

Dikken, Marcel den. 1998. Predicate inversion in DP. In A. Alexiadou and C. Wilder, eds., Possessors, Predicates and Movement in the Determiner Phrase, pages 177-214. Amsterdam: John Benjamins.

Dikken, Marcel den. 2006. Relators and Linkers: The Syntax of Predication, Predicate Inversion, and Copulas. Cambridge, MA: MIT Press.

Di Sciullo, Anna-Maria and Edwin Williams. 1987. On the Definition of Word. Cambridge, MA: MIT Press.

Di Sciullo, Anna-Maria. 2009. Why are compounds a part of human language? A view from asymmetry theory. In R. Lieber and P. Štekauer, eds., The Oxford Handbook of Compounding, pages 145-177. Oxford: Oxford University Press.

Eik, Ragnhild. 2019. The Morphosyntax of Compounding in Norwegian. PhD. Thesis, NTNU, Trondheim.

Geach, Peter Thomas. 1980. Reference and Generality: An Examination of Some Medieval and Modern Theories. Ithaca, NY: Cornell University Press.

Geach, Peter Thomas. 1987. Reference and Buridan's Law. Philosophy 62(239):7-15.

Gillon, Brendan. 2008. Exocentric Compounds in Classical Sanskrit. In G. Huet and A. Kulkarni, eds., First International Sanskrit Computational Linguistics Symposium, Oct. 2007, Rocquencourt, France. http://hal.inria.fr/SANSKRIT/fr/. <inria-00202860>

Gippert, Jost, Javier Martínez, Agnes Korn, and Roland Mittmann. 2016. Thesaurus Indogermanischer Text- und Sprachmaterialien (TITUS). https://titus.fkidg1.uni-frankfurt.de/indexe.htm.

Guevara, Emiliano and Sergio Scalise. 2009. Searching for Universals in Compounding. In: S. Scalise, E. Magni, and A. Bisetto, eds., Universals of Language Today. Studies in Natural Language and Linguistic Theory 76:101-128.

Hale, Mark. 2018. The syntax of Indo-Iranian. In J. Klein, B. Joseph, and M. Fritz (in cooperation with Mark Wenthe), eds., Handbook of Comparative and Historical Indo-European Linguistics, Vol. 3, pages 1924-1942. Berlin: de Gruyter.

Hellwig, Oliver. 2010-2021. DCS – The Digital Corpus of Sanskrit. http://www.sanskrit-linguistics.org/dcs/index.php

Jamison, Stephanie W. and Joel P. Brereton, trs., 2014. The Rigveda: The Earliest Religious Poetry of India, 3 Vols. Oxford: Oxford University Press.

Kastovsky, Dieter. 2009. Diachronic perspectives. In R. Lieber and P. Štekauer, eds., The Oxford Handbook of Compounding, pages 323-340. Oxford: Oxford University Press.

Katre, Sumitra M. (translator). 1987. Aṣṭādhyāyī of Pāṇini. Austin: University of Texas Press.

Kayne, Richard S. 1994. The antisymmetry of syntax. Cambridge, MA: MIT Press.

Kiparsky, Paul. 1982. From Cyclic Phonology to Lexical Phonology. In H. van der Hulst and N. Smith, eds., The structure of phonological representations (Part I), pages 131-175. Dordrecht: Foris.

Kiparsky, Paul. 2009. On the architecture of Pāṇini's grammar. In G. Huet, A. Kulkarni, and P. Scharf, eds., Sanskrit Computational Linguistics, ISCLS 2007/2008. Lecture Notes in Computer Science 5402:33-94.

Kiparsky, Paul. 2010. Compositional vs. paradigmatic Approach to Accent and Ablaut. https://web.stanford.edu/~kiparsky/Papers/ucla_IE_09.submitted.new.pdf

Kiparsky, Paul. 2016. The agent suffixes as a window into Vedic grammar. http://web.stanford.edu/~kiparsky/Papers/agentnouns.final.pdf

Longobardi, Giuseppe. 2001. The Structure of DPs: Some Principles, Parameters, and Problems. In M. Baltin and C. Collins, eds., The Handbook of Contemporary Syntactic Theory, pages 562-603. Malden, MA: Blackwell.

Longobardi, Giuseppe and Giuseppina Silvestri. 2013. The Structure of NPs: Some Insights on Case, Empty Categories, and Poverty of Stimulus. In S. Luraghi and C. Parodi, eds., The Bloomsbury Companion to Syntax, pages 88-117. London: Bloomsbury.

Lundquist, Jesse and Anthony D. Yates. 2018. The morphology of Proto-Indo-European. In J. Klein, B. Joseph, and M. Fritz (in cooperation with M. Wenthe), eds., Handbook of Comparative and Historical Indo-European Linguistics, Vol. 3, pages 2079-2195. Berlin: de Gruyter.

Lowe, John J. 2015. The syntax of Sanskrit compounds. Language 91(3):71-114.

Macdonell, Arthur Anthony. 1910. Vedic Grammar. Strassburg: Trübner.

Macdonell, Arthur Anthony. 1916. A Vedic grammar for students. Oxford: Clarendon Press.

Mādhavānanda, Swāmī (translator). 1950. The Bṛhadaraṇyaka Upaniṣad: With the commentary of Śaṅkarācārya. Hollywood, CA: Vedanta Press.

Manzini, Maria Rita. 2017. Passive, smuggling and the *by*-phrase. In N. LaCara, K. Moulton, and A-M. Tessier, eds., A Schrift to Fest Kyle Johnson, pages 233-244. Linguistics Open Access Publications,.

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. Münchener: Fink.

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. 2000. Dynamic Antisymmetry. Cambridge, MA: MIT Press.

Moro, Andrea. 2004. Linear compression as a trigger for movement. In H. van Riemsdijk and A. Breitbarth, eds., Triggers, pages 387-429. Berlin: de Gruyter.

Moro, Andrea. 2009. Rethinking symmetry: A note on labeling and the EPP. Snippets 19:17-18.

Moro, Andrea. 2017. A Brief History of the Verb to Be (tr. by B. McClellan-Broussard). Cambridge, MA: MIT Press.

Moro, Andrea. 2019. The geometry of predication: A configurational derivation of the defining property of clause structure. Phil. Trans. R. Soc. B 375(1791):1-7.

Moro, Andrea and Ian Roberts. 2020. Unstable structures and Generalized Dynamic Antisymmetry. Lingbuzz/005515.

Ntelitheos, Dimitrios. 2022. Compounding in Greek as Phrasal Syntax. Languages 7(151):1-22.

Okubo, Tatsuhiro. 2014. Expletives in words: Linking elements as markers of wordhood. Lingvisticæ Investigationes 37(2):225-239.

Pennanen, Esko V. 1982. Remarks in Syntagma and Word-Formation. Folia Linguistica 16:241-261.

Perrin, Jean-Baptiste. 1913. Les atoms. Paris: Alcan.

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. Proceedings from the Annual Meeting of the Chicago Linguistic Society 5(1):205-239.

Renou, Louis. 1961. Grammaire Sanscrite. Paris: Maisonneuve.

Rizzi, Luigi. 1990. Relativized Minimality. Cambridge, MA: MIT Press.

Rizzi, Luigi. 2013. The functional structure of the sentence, and cartography. In M. den Dikken, ed., The Cambridge Handbook of Generative Syntax, pages 425-457. Cambridge: Cambridge University Press.

Rizzi, Luigi. 2016. Labeling, maximality and the head-phrase distinction. The Linguistic Review 33(1):103-127.

Rizzi, Luigi and Ur Shlonsky. 2007. Strategies of subject extraction. In H-M. Gärtner and U. Sauerland, eds., Interfaces + recursion = language? Chomsky's minimalism and the view from syntax-semantics, pages 115-160. Berlin: de Gruyter

Scalise, Sergio and Antonietta Bisetto. 2009. The Classification of Compounds. In R. Lieber and P. Štekauer, eds., The Oxford Handbook of Compounding, pages 34-53. Oxford: Oxford University Press.

Scalise, Sergio and Antonio Fábregas. 2010. The head in compounding. In S. Scalise and I. Vogel, eds., Cross-Disciplinary Issues in Compounding, pages 109-126. Amsterdam: John Benjamins.

Scalise, Sergio and Irene Vogel. 2010. Why compounding? In S. Scalise and I. Vogel, eds., Cross-Disciplinary Issues in Compounding, pages 1-18. Amsterdam: John Benjamins.

Selkirk, Elisabeth O. 1982. The syntax of words. Cambridge, MA: MIT Press.

Speijer, Jacob Samuel. 1886. Sanskrit Syntax. Leiden: Brill.

Wackernagel, Jakob. 1905. Einleitung zur Wortlehre. Nominalkomposition (Altindische Grammatik, Band II.1). Göttingen: Vandenhoeck & Ruprecht.

Whitney, William Dwight. 1889. A Sanskrit Grammar, including both the Classical language, and the older Dialects, of Veda and Brahmana. Cambridge, MA: Harvard University Press.

Zamparelli, Roberto. 2000. Layers in the Determiner Phrase. New York: Garland.