SANSKRIT WORD FORMATION AND CONTEXT FREE RULES

Brendan S. Gillon
University of Toronto

0 Introduction

In recent years, a number of linguists, among them Selkirk (1982) and Di Sciullo and Williams (1987), have explored the hypothesis that word formation, including compounding and derivation, can be given a syntactic characterization of its own in terms of context free rules which are independent of those used for phrases and sentences. One way to explore the empirical adequacy of this hypothesis is to try to apply it to languages rich in derivational and inflectional morphology as well as in complex compounds. One such language is Classical Sanskrit. Below, I shall investigate how such a hypothesis can be applied to compound formation in Classical Sanskrit, touching only in passing upon its applicability to derivation.

1.0 Context Free Rules for Complex Sanskrit Words

The most striking case for the utility of context free rules in the analysis of word formation is provided by derivationally complex words. Derivational affixes have two salient characteristics: first, that they are bound morphemes; and second, that they frequently change the lexical category of the stems to which they are applied. The second is captured by assigning them lexical categories (Williams 1981), as a result of which derivationally complex words yield, in principle, to an analysis in terms of context free rules; the first is captured by assigning subcategorization frames to derivational affixes' lexical entries. (For further details, see Selkirk 1982 Ch. 3.1.)

The applicability of such rules to the analysis of Sanskrit derivation is shown by the examples below.

(1) A \rightarrow AA (comparative and superlative adjectives):
    (A (A śuci ) (A -tara ))
    pure -er
    purer

(2.1) N \rightarrow AN (abstract nouns):
    (N (A guru ) (N -tva ))
    heavy -ness
    heaviness

(2.2) N \rightarrow AN (adjectival nouns):
    (N (A krśa ) (N Ø ))
    weak one
    a weak one
(3) $V \rightarrow AV$: (denominative verbs):
   $(v\ (\_\ lohita\ ))\ (v\ -aya\ )$
   red
   -en
to redden

(4.1) $A \rightarrow NA$: (adjectives with possessive suffix):
   $(A\ (n\ dh\_i\ ))\ (A\ -man\ )$
   thought
   -possessing
   possessing
   thought

(4.2) $A \rightarrow NA$: (adjectives of appurtenance):
   $(A\ (n\ \_\iva\ ))\ (A\ -a\ )$
   \_iva
   belonging to \_iva

(5) $N \rightarrow NN$: (abstract nouns):
   $(n\ (n\ \_\k\_a\ ))\ (n\ -\_\_va\ )$
   tree
   -ness
   treeness

(6) $V \rightarrow NV$: (denominative verbs):
   $(v\ (n\ putra\ ))\ (v\ -ya\ )$
   son
   to treat like a son

(7) $A \rightarrow VA$: (participles):
   $(A\ (v\ kr\_ )\ (A\ -ta\ ) )$
   make
   -en
   made

(8) $N \rightarrow VN$: (deverbal nouns):
   $(n\ (v\ \_\r\_ )\ (n\ -\_\_\_a\ ) )$
   see
   -ing
   sight

(9) $V \rightarrow VV$: (causative):
   $(v\ (v\ \_\_\r\_ )\ (v\ -aya\ ) )$
   see
   show

What is remarkable is that these are exactly the rules needed for the analysis of English derivation. (See Selkirk 1982 p. 82.)

<table>
<thead>
<tr>
<th>$A$</th>
<th>$AA$</th>
<th>$N$</th>
<th>$AN$</th>
<th>$\star P$</th>
<th>$AP$</th>
<th>$V$</th>
<th>$AV$</th>
</tr>
</thead>
<tbody>
<tr>
<td>$A$</td>
<td>$NA$</td>
<td>$N$</td>
<td>$NN$</td>
<td>$\star P$</td>
<td>$NP$</td>
<td>$V$</td>
<td>$NV$</td>
</tr>
<tr>
<td>$\star A$</td>
<td>$PA$</td>
<td>$\star N$</td>
<td>$PN$</td>
<td>$\star P$</td>
<td>$PP$</td>
<td>$V$</td>
<td>$PV$</td>
</tr>
<tr>
<td>$A$</td>
<td>$VA$</td>
<td>$N$</td>
<td>$VN$</td>
<td>$\star P$</td>
<td>$VP$</td>
<td>$V$</td>
<td>$VV$</td>
</tr>
</tbody>
</table>

Table 1

Context free rules also provide an elegant analysis of Sanskrit compounds. Classical Indian grammarians recognize four principal types of compounds: tatpurusha, or determinative,
compounds; \textit{dvandva}, or copulative, compounds; \textit{bahuvrīhi}, or exocentric, compounds; and \textit{avyayībhāva} compounds. Each of the first two types of compounds, including each of the various sub-types of the first type, is easily analyzed by context free rules, as illustrated below.

(10) \( A \rightarrow AA (\text{viśeśa}-\text{pada}-\text{karmadhāraya}):\)
\[ (\text{\( A \)} (\text{\( A \)} \text{ tulya } ) (\text{\( A \)} \text{ śveta } )) \]
\[ \text{same} \quad \text{white} \]
\[ \text{equally} \quad \text{white} \]
\[ (\text{\( A \)} (\text{\( A \)} \text{ snāta } ) (\text{\( A \)} \text{ anulipta } )) \]
\[ \text{bathed} \quad \text{oiled} \]
\[ \text{bathed and oiled} \]

(11) \( N \rightarrow AN (\text{viśeśa}-\text{pūrva}-\text{pada}-\text{karmadhāraya}):\)
\[ (\text{\( N \)} (\text{\( N \)} \text{ nīla } ) (\text{\( N \)} \text{ utpala } )) \]
\[ \text{blue} \quad \text{lotus} \]
\[ \text{blue} \quad \text{lotus} \]

(12) \( V \rightarrow AV (\text{cvi-gati}):\)
\[ (\text{\( V \)} (\text{\( V \)} \text{ laghu } ) (\text{\( V \)} \text{ kṛ } )) \]
\[ \text{light} \quad \text{make} \]
\[ \text{make} \quad \text{light} \]

(13.1) \( A \rightarrow NA (\text{upamāna}-\text{pūrva}-\text{pada}-\text{karmadhāraya}):\)
\[ (\text{\( A \)} (\text{\( N \)} \text{ anala } ) (\text{\( A \)} \text{ uṣṇa } )) \]
\[ \text{fire} \quad \text{hot} \]
\[ \text{hot like fire} \]

(13.2) \( A \rightarrow NA (\text{vibhakti}-\text{tatpuruṣa}):\)
\[ (\text{\( A \)} (\text{\( N \)} \text{ go } ) (\text{\( A \)} \text{ hita } )) \]
\[ \text{cow} \quad \text{beneficial} \]
\[ \text{beneficial to cows} \]

(14.1) \( N \rightarrow NN (\text{vibhakti}-\text{tatpuruṣa}):\)
\[ (\text{\( N \)} (\text{\( N \)} \text{ dadhi } ) (\text{\( N \)} \text{ odana } )) \]
\[ \text{curd} \quad \text{rice} \]
\[ \text{rice with curd} \]

(14.2) \( N \rightarrow NN (\text{upapada}-\text{tatpuruṣa}):\)
\[ (\text{\( N \)} (\text{\( N \)} \text{ kumbha } ) (\text{\( N \)} -\text{kāra } )) \]
\[ \text{pot} \quad \text{maker} \]
\[ \text{pot maker} \]

(14.3) \( N \rightarrow NN (\text{avadhārana}-\text{pūrva}-\text{pada}-\text{karmadhāraya}):\)
\[ (\text{\( N \)} (\text{\( N \)} \text{ rāja } ) (\text{\( N \)} \text{ rśi } )) \]
\[ \text{king} \quad \text{sage} \]
\[ \text{sage-king} \]

(14.4) \( N \rightarrow NN (\text{upamāna}-\text{uttara}-\text{pada}-\text{karmadhāraya}):\)
\[ (\text{\( N \)} (\text{\( N \)} \text{ puruṣa } ) (\text{\( N \)} \text{ vyāghra } )) \]
\[ \text{man} \quad \text{tiger} \]
\[ \text{tiger of a man} \]

- 17 -
(14.5) \( N \rightarrow NN \) (dvandva):
\( (n \ (n \ sukha \ ) \ (n \ duhka \ ) \)
happiness pain
happiness and pain

(15) \( V \rightarrow NV \) (cvi-gati):
\( (v \ (n \ shreni \ ) \ (v \ bhuh \ ) \)
line be
to line up

(16) \( A \rightarrow PA \) (pradi-tatpuru\(\text{s}\)a):
\( (a \ (r \ ati \ ) \ (a \ bala \ ) \)
beyond strong
very strong

(17) \( N \rightarrow PN \) (pradi-tatpuru\(\text{s}\)a):
\( (n \ (p \ adhi \ ) \ (n \ raja \ ) \)
over king
over-lord

(18) \( V \rightarrow PV \):
\( (v \ (p \ adhi \ ) \ (v \ shi \ ) \)
upon lie
to lie upon

One type of tatpuru\(\text{s}\)a compound which yields especially nicely to the approach under consideration here are upapada-
tatpuru\(\text{s}\)a compounds. This compound behaves like other tatpuru\(\text{s}\)a compounds, except that one of its immediate constituents is a bound form, such as the final occurring elements of -jha (knowing) and -kara (doing) or the initial occurring elements of su- (good/well) and ku- (bad/badly). On the approach here, nothing special needs to be said concerning such compounds beyond what must be stipulated for the lexical entry of each such bound form, namely, that it has a certain subcategorization frame.

The rules for compound formation set out above are virtually identical with those required for English compound formation, given below (Selkirk 1982 p. 16):

\[
\begin{align*}
A & \rightarrow AA & N & \rightarrow AN & \ast P & \rightarrow AP & \ast V & \rightarrow AV \\
A & \rightarrow NA & N & \rightarrow NN & \ast P & \rightarrow NP & \ast V & \rightarrow NV \\
A & \rightarrow PA & N & \rightarrow PN & \ast P & \rightarrow PP & V & \rightarrow PV \\
\ast A & \rightarrow VA & N & \rightarrow VN & \ast P & \rightarrow VP & \ast V & \rightarrow VV
\end{align*}
\]

Table 2

Interestingly, while there are discrepancies, they are not great. The first discrepancy accrues to compounds of the forms \( V \rightarrow AV \) and \( V \rightarrow NV \). Both types of compounds are well attested and productive in Sanskrit; however, only three verbs can participate in such formations: the verb \( kf \) (to do), the verb \( bhuh \) (to be),
and the verb as (to be). Yet such formations are not strangers to English either; thus, dry-clean appears to be of the form V → AV, while hand-carry appears to be of the form V → NV. However, in spite of their frequency, it is the view of Selkirk (1982 p. 17), following Marchand (1969: pp. 58-65), that such compounds are backformations from compounds of the form A → AA (dry-cleaned) and of the form A → NA (hand-carried) respectively.

Be that as it may, there is at least one indisputable discrepancy: English compounds of the form N → VN, which are completely absent from Sanskrit. English compounds of this form are not all that rare. Those which are paraphrasable as "an N which or who V's" appear to be productive (Bauer 1983: Ch. 7.2.1.2). They include not only the four mentioned by Selkirk (1982: p. 14), whet-stone, rattle-snake, swear-word, scrub-woman, but also many others such as draw-bridge, draw-knife, tow-truck, glow-worm, drive-shaft, stop-watch, and turn-table. Other compounds of the same form, which are paraphrasable as "that which or one who V's an N" -- such as tug-boat, turn-key, pick-pocket, break-water, and spoil-sport -- appear to be, at best, marginally productive (Bauer 1983: Ch. 7.2.1.2). And finally there are such oddities as think tank, rap session, and washroom.

Two of the four principal types of compounds have not been included above: they are: bahuvr̥hi, or exocentric, compounds, and avyāyabhāva compounds. Avyāyabhāva compounds have the form (⟨r x, n y⟩) and function as adverbs. I shall defer discussion of them until later and concentrate now, instead, on bahuvr̥hi, or exocentric, compounds.

1.1 Bahuvr̥hi, or Exocentric, Compounds

A bahuvr̥hi compound is said to be exocentric because it typically modifies an element, an inflected noun, external to it. Indeed, it modifies a noun exactly as an adjective does, agreeing with the noun it modifies in case, number, and gender. This fact suggests the obvious hypothesis that bahuvr̥hi compounds are compounds which have undergone adjectival conversion, wherein a phonetically null, possessive, adjectival suffix -- virtually synonymous with, but having a different distribution from, the phonetically overt, possessive, adjectival suffix -ka -- has been added to a constituent compound noun. This hypothesis provides a straightforward account not only for how a bahuvr̥hi compound acquires its inflectional morphology but also for the fact that a bahuvr̥hi compound may undergo any compounding or derivation which any adjective may undergo. For example, just as an adjective such as kr̥ṣa (thin) may be converted into an abstract noun, kr̥ṣa-tva (thin-ness), through the addition of the suffix -tva (-ness), so may a bahuvr̥hi compound such as dīrgha-kaṇṭha (long-neck-ed; cf., level-head-ed) be turned into dīrgha-kaṇṭha-tva (long-neck-ed-ness; cf., level-headed-ness).
The inverse of bahuvrīhi compound formation is the conversion given above which turns adjectives into common nouns. Thus, an adjective such as kṛśaḥ (thin) can function, as its English translation can, as a common noun, meaning the same thing as its English nominal counterpart, the thin. (See Dash and Gillon 1987 for reasons supporting the view that adjectives and nouns are distinct lexical categories in Sanskrit.) These two conversions, taken together, account for the common noun-like usage of bahuvrīhi compounds, such as:

(19.1) NBT 48.4
vyutpanna<saṁketasya>-B-N,
glossed by the commentator as,

(19.2) NBTP 49.1-2
vyutpannaḥ jñātaḥ saṁketaḥ yena saḥ vyutpanna-saṅketaḥ
arisen known convention by whom he arisen-conventioned
One for whom the conventions of language have arisen
(jñāta (known) glosses vyutpanna (arisen).)

This hypothesis implies that every bahuvrīhi compound of the form x₁ corresponds with a compound of the form x-ka₁. And, in fact, this correspondence is often exploited in the Sanskrit commentaries of Medieval Classical Indian philosophy (B&T 5.3.15). It should be noted, however, that not everything of the form x-ka₁ corresponds with a bahuvrīhi compound of the form x₁; for, unlike the suffix -ka, the phonetically null suffix for bahuvrīhi compounds is sensitive to the structure of its left sister. In particular, its left sister must be complex, of the form x-y, where x is either the prefix (sa-), the privative prefix (a-), or a lexical item which can denote a predicatable attribute of what y denotes, either literally (so-called saṁānādikaraṇa-bahuvrīhi compounds), as in the case of dīrgha-kaṇṭha (long-neck-ed, or figuratively so-called śaṣṭi-vyadhikaraṇa-bahuvrīhi compounds), as in the case of candra-mukha (moon-faced, i.e., having a face like the moon). There are a few cases where what is denoted by x is not a predicatable attribute of what is denoted by y. These bahuvrīhi compounds, so-called saptami-vyadhikaraṇa-bahuvrīhi compounds, are exceptional and infrequent.

The role of complexity in determining the availability of a compound is not confined to bahuvrīhi compounds. Derivational complexity also plays a role in the acceptability of numerical, appositional compounds, so-called dvigu compounds: a numeral cannot be a left sister in a compound, unless the compound is a proper name, has a collective sense, or is itself the left sister of some other element, affixal or lexical. (A 2.1.51; B&T 4.24).

In addition, it is worth remarking that English too has bahuvrīhi compounds. They are formed with the adjectival, possessive suffix, -ed. However, unlike the phonetically null
Sanskrit bahuvricihi suffix, the English suffix attaches not only to compounds, thereby yielding such formations as long-leg-ed, literal-mind-ed, and two-foot-ed, but also to stems of simple words, thereby yielding such formation as principled, wing-ed, and verandah-ed. Yet, at the same time, the English suffix, like its Sanskrit counterpart, is subject to certain restrictions. First, like its Sanskrit counterpart, when it is appended to complex word, the left sister of the complex word denotes an attribute which is predicatable of what is denoted by the complex word's right sister. Again, like its Sanskrit counterpart, the predication may be either literal (single-member-ed or figurative (duck-bill-ed, pot-belly-ed). In addition, though, the English suffix -ed, unlike its Sanskrit counterpart, seems to be confined to cases expressing some kind of relation of inalienable possession between what is denoted by that to which the suffix is attached and what is denoted by the word modified by the word formed with the suffix. (See Bauer 1983 pp. 30, 57-60, and 93 and references cited there.)

In this connection, it should be noted that there is a small class of exocentric compounds in English which fail to have the -ed suffix. They are all nouns; and, for the most part, there is a corresponding compound with the -ed suffix, whose sense is closely allied to the sense of the same word without the -ed suffix: red-head (red-headed); dim-wit (dim-witted); hard-back (hard-backed), etc. Such compounds, discussed by Selkirk (1982 p. 26), are not productive in English; and I think that they are best viewed as lexicalized common nouns, obtained from the elision of the -ed suffix.

1.2 Adverbs and Compounds

One question which can reasonably be asked is whether or not the context free rules require, beyond the four basic lexical categories, the category of adverb. There are two facts which militate in favor of this addition. First, there is a number of compounds whose left sisters seem to require an adverbial interpretation; second, there are compounds which function as adverbs, so-called avyayiḥbhāva compounds.

In assessing the cogency of these facts, let me begin by discussing some compounds, which, though susceptible of an adverbial interpretation, do not require that this category be introduced into the context free rules for their structure to be properly characterized. Consider the compounds:

(20.1) PSED: sv
    nisarga<nipuṇa
    nature-clever
    naturally clever
(20.2) SGSC 21.14 <= Mb 2.7.3>
kāma<gama
will-moving
moving at will

(20.3) PSED: sv
māsa<deya
month-payable
payable in a month

(20.4) Co p. 189
prayatna<prekṣanīya
effort-discernible
discernible with effort.

In spite of what the English translations of these compounds might suggest, there are no adverbs here. This is evident from two facts: first, the subordinate member is a noun; and second, the compounds themselves have acknowledged paraphrases in which the subordinate member appears in an oblique case. These paraphrases are: [NP3 nisargena] nipuṇa (clever by nature), [NP3 kāmena] gama (moving at will), [NP3 māsena] deya (payable in a month), and [NP3 prayatnena] prekṣanīya (to be discerned with effort). Similar considerations bear on such compounds as bahya-antar (outside and inside) (W art. 1259) as well as compounds such as antar-dhā (to place inside); in each case, the putatively adverbial elements can be taken as nouns.

More problematic are compounds whose left sister falls into the category of adjective and yet has an adverbial interpretation. Thus, for example, tulya (same) is an adjective and serves as an adjectival modifier in the compound tulya<dharma (same dharma). Yet, it also serves as an adverbial modifier in such compounds as tulya<śveta (equally white) and tulya<bala (equally strong). The same holds for the compounds udagra<ramanīya (intensely lovely) (Co pp. 92-3) and madhura<ukta (sweet spoken) (Co pp. 188-9).

Compounds which have the same constituency and whose leftmost immediate constituent is an adjective having an adverbial interpretation occur in English as well: for example, fresh-cut-potatoes and soft-spoken-guy. The difference between Sanskrit and English is that Sanskrit tolerates any adjective in the right-hand position whereas English tolerates only present and past participles. Thus, while the adjective equal, like its Sanskrit translation tulya, may occur in a compound, for example, equal-opportunity-employer; it may not do so felicitously, if the sister to its right is not a participle, which is borne out by the unacceptability of English compounds such as equal-strong.

The fact that a subordinate adjectival element in a compound is liable to an adverbial interpretation does not require that this element be an adverb: after all, one might simply retain the syntactic description of the compounds above as (A A A) and consign the treatment of the adverbial interpretation of the
first element to that part of compound analysis which provides for their interpretation.

This solution might be baulked at, since there are words which are adverbs, when taken in isolation, and which occur as the subordinate element of a compound, retaining their adverbal interpretation. Such words are those which end in -\textit{thā}, such as \textit{tathā} (thus), \textit{anyathā} (otherwise), \textit{mithyā} (in vain), as well as \textit{punar} (again), \textit{evam} (so), etc. Compounds of this ilk are these:

\begin{itemize}
\item \textbf{(21.1)} Co p. 93
\begin{itemize}
\item \textit{anyathā}<sambhāvana
\item otherwise-supposing
\item supposing otherwise
\end{itemize}
\item \textbf{(21.2)} NBTP 49.16; PSED: sv
\begin{itemize}
\item \textit{anyathā}<anupapatti
\item otherwise-impossible
\item being impossible otherwise
\end{itemize}
\item \textbf{(21.3)} PSED: sv
\begin{itemize}
\item \textit{punar}<ukta
\item again-spoken
\item spoken again
\end{itemize}
\item \textbf{(21.4)} PSED: sv; SGSC 50.19
\begin{itemize}
\item \textit{tathā}<gata
\item thus-happened
\item happened thus
\end{itemize}
\end{itemize}

But even these facts are not decisive, since these very same words, while having only an adverbal interpretation while outside of a compound, are also susceptible to an adjectival interpretation within some compounds:

\begin{itemize}
\item \textbf{(22.1)} NBTP 29.11
\begin{itemize}
\item \textit{tathā}<pravṛtti
\item such-adversion
\item such adversion
\end{itemize}
\item \textbf{(22.2)} NBTP 29.13
\begin{itemize}
\item \textit{tathā}<vṛtti
\item such-activity
\item such activity
\end{itemize}
\item \textbf{(22.3)} NBTP 32.11
\begin{itemize}
\item \textit{tathā}<vṛtti
\item such-meaning
\item such meaning
\end{itemize}

\end{itemize}
(22.4) SGSC 50.1: K108
mithyāmāhātya<br>garva
false-greatness-pride
pride of false greatness

(22.5) SGSC 204.12: K108
evan<br>prāya
such-kind
such kind

(22.6) NBTP 67.5
a-punar<br>āvṛtti
non-again-return
no further return.

There would be no need to expand the categories used in the context free rules, at least not on account of these compounds, should one hypothesize that these adverbs have in fact two forms, one free and one bound and that the bound forms are adjectives while the free forms are adverbs. As a result, the bound forms, like all adjectives, are liable to an adverbial interpretation, when they occur in compounds.

Finally, it should be remarked that prepositions in compounds, so-called prādi-tatpuruṣa compounds, behave in a similar way. When the right sister of a preposition in a compound is a noun, as in ati-bhāra (over-load: i.e., excessive load), the preposition has an adjectival interpretation; and when it is an adjective, as in ati-balā (overly or excessively strong), it has an adverbial one. The same is attested in English. Compare over-dose (i.e., excessive dose) with over-abundant (i.e., excessively abundant).

2.0 Regularities of Sanskrit Compound Formation

Having seen that there is a good prima facie case to be made for the analysis in terms of context free rules of Sanskrit words in general and Sanskrit compounds in particular, I shall turn to a more detailed examination of the empirical adequacy of this analysis of Sanskrit compounds. In particular, I shall show this kind of analysis not only captures in a simple way the seven basic regularities known to characterize Sanskrit compounds but also extends their scope.

What are these regularities? They are:

(23.1) The accentuation of a compound is that of a simple word, not that of a phrase (A 6.1.155; A 6.1.223ff and Ca p. 164; A 6.2.2; A 6.2.43).

(23.2) Compounds are subject to the inflectional and derivational morphological forms of simple words (A 2.4.71; A 6.3.1). In particular, inflection occurs at the end of compounds,
not within them; derivational suffixes can be added as easily to compounds as they can be to words.

(23.3) Compounds are of unbounded complexity. (4) Elements of a compound, unlike elements of phrases, have a fixed linear order (A 1.2.43; A 2.2.30ff; A 1.2.36). In general, whereas no two immediate constituents of a compound can be permuted and the sense of the compound retained for its members; any two immediate constituents of a phrase can be permuted and the sense of the phrase retained.

(23.5) A compound is usually analyzable into two immediate constituents (A 2.2.4; Ca pp. 259-260); and if there is a head, it is the second immediate constituent (A 1.2.43; A 2.2.30; Ca pp. 261-263). (6) A compound has a typical, and for Pāṇini a canonical, phrasal paraphrase such that if a compound has the form \((x\cdot y)_i\) then its phrasal paraphrase has the form \(x_i\cdot y_i\) (where "i" and "j" denote one of the seven Sanskrit cases). Moreover, the head of a canonical phrasal paraphrase is the head of the compound being paraphrased.

(23.7) Inflected words, which are external to a compound, are not construed with uninflected elements subordinate within it (A 2.1.2; Ca p. 165.)

There are exceptions to every one of the generalizations presented above. To begin with, some compounds, such as dvandva compounds (co-ordinate compounds) formed from the names of gods, do not have the accent of simple words. In addition, inflectional morphology sometimes occurs within a compound (A 6.3.7-8), so-called aluk compounds, or within derivational morphology (A 6.3.17). Such cases are not productive and are best treated, as Pāṇini does, with special rules whose applications are triggered by particular lexical items. (See Ca 2.4.2.3.6; Whitney 1881 art. 1250.) Linear order for compounds is fixed. The few cases where a compound's elements are permutable are cases where two distinct derivations happen to yield two compounds with the same meaning.

All Sanskrit compounds, except dvandva and viṣeṣaṇa-ubhaya-pada-karmadhāraya compounds, copulative compounds, are analyzable into a pair of lexical elements. These compounds appear to have no upper bound on the number of lexical elements which may appear in them; and, moreover, there appears to be no syntactically principled way in which might be grouped together. The dvandva compound deva+gandharva+mānuṣa+uraga+rākṣasāh (gods, fairies, humans, serpents, and ogres) (B&T 4.30), comprises five lexical elements; and there appears to be no non-arbitrary constituent between the five lexical elements and the compound as a whole. Nonetheless, such compounds can be generated by the categorial rules given above, namely, N \(\rightarrow\) NN and A \(\rightarrow\) AA. It is worth noting that the same situation arises, by the way, with conjoined noun phrases, in both English and Sanskrit: they too
appear to have no upper bound on the number of conjuncts which may appear in them; and, moreover, there appears to be no syntactically principled way in which they too might be grouped together.

2.1 Headedness of Sanskrit Compounds

Another regularity stated above is that Sanskrit compounds are, for the most part, head final. In order to assess the extent to which this regularity holds, one needs to pin down the notion of headedness of a compound. Intuitively, the head of a compound, or a phrase for that matter, is a distinguished, immediate constituent of the compound, or phrase. The question arises: in what does the constituent's being distinguished consist? Several criteria have been suggested in answer to this question; and, it will be useful to consider Sanskrit compounds in light of each.

The most commonly encountered criterion is a semantic one: the head of a compound is the stem of the compound which is a hyponym of the compound itself. For example, in the compound rāja-puruṣa (king-man: i.e., servant), puruṣa (man), not rāja (king), is the head, since puruṣa (man), not rāja (king), is a hyponym of rāja-puruṣa (king-man: i.e., servant).

One bad reason for thinking that this criterion is inadequate is the frequently alleged fact that gun, for example, seems to be the head of the compound toy gun but that a toy gun is not a gun. This allegation overlooks an extremely pervasive, and otherwise well recognized, ambiguity: a word may be used to denote not only a canonical set of objects, but also a set of objects which look like the typical kind of object in the canonical set. Thus, one calls an airplane "an airplane", but one also calls a picture of an airplane or a model of an airplane "an airplane". Indeed, this rule seems to be recursive: a picture of a picture of an airplane may be called "an airplane". (See Cohen 1985 and Gillon 1990a for discussion.)

The real problem with using hyponymy as a criterion for determining a compound's headedness is that the clear cases of hyponymy are much more limited in their compass than the clear cases of headedness. Suppose that one stem is a hyponym of another just in case the latter names a kind of what the former names. Then the criterion fails to apply to those compounds whose intuitively distinguished element is a verb or adjective. So, whereas everyone is prepared to say that what is named by rāja-puruṣa (*king-man: servant) is a kind of what is named by puruṣa (man); no one is prepared to say that what is named by go-hita (*cow-beneficial: beneficial to a cow) is a kind of what is named by hita (beneficial); or that what is named by laghu-kr (*light-make: make light) is a kind of what is named by kr (make); yet clearly hita (beneficial) is the distinguished stem in go-hita, and kr (make) in laghu-kr. One might think that the ambit of the criterion could be expanded, should the relation of one thing
being a kind of something else be replaced by the relation of the
extension of one term’s denotation being a subset of the
extension of another’s. But this explication of hyponymy won’t do
either. (See Gillon 1990b for discussion.)

The semantic criterion is antedated by a syntactic one,
which could reasonably be attributed to Pāṇini himself: For
Pāṇini, almost all Sanskrit compounds have a canonical phrasal
paraphrase; and, the compound and its canonical phrasal
paraphrase have a common derivational ancestor. The commonality
of derivational ancestry not only provides an account for the
semantic equivalence between a compound and its canonical phrasal
paraphrase; but it also provides a way of defining the head of a
compound: the head of a compound is the stem in the compound
which is the stem of the head word in the canonical phrasal
paraphrase (A 1.2.43 and A 2.2.30; Ca pp. 261–263). This
definition has two noteworthy consequences: first, for the most
part, the syntactic features of a compound are those of the head
of its canonical (phrasal) paraphrase; and second, Sanskrit
compounds are, by and large, head final.

There are exceptions, however; and Pāṇini treats them as
such. For example, certain tatpuruṣa compounds such as ardha-
pippali (half-pepper) and pūrva-kāya (*fore-body: front of the
body) (A 2.2.1-2) have the canonical phrasal paraphrases of
pūrva kāyasya (front of the body) and ardham pippalyāḥ (half of
a pepper), where the heads of the paraphrases, pūrva and ardham
respectively, correspond to the first stem in each compound.
Other exceptions include: the tatpuruṣa compound rāja-danta
(*king-tooth: king of the teeth, i.e., the front teeth) (A
2.2.3-5 and 2.2.31); certain karmadhāraya compounds such as
prāpta-yauvanā, paraphrased by Pāṇini as prāpta yauvanam (having
arrived at adolescence), where the first stem, prāpta,
corresponds to the head of the canonical paraphrase, prāpta (A
2.2.4); and certain other karmadhāraya compounds such as mayūra-
vyaṃsaka (peacock-cunning: cunning peacock), again where the
first stem mayūra, corresponds to the head of the canonical
paraphrase, mayūraḥ vyaṃsakah (A 2.1.72).

Interestingly, as observed by Roose and Tubb (1981 art.
4.29), the karmadhāraya compounds treated under the same rubric
as prāpta-yauvanā (A 2.2.4) could also have been analyzed as
easily as bahuvrīhi compounds, whose internal structure is that
of a head-final karmadhāraya, namely as, (prāpta<yauvanā)B (one
whose adolescence has been attained).

This syntactic criterion for lexical headedness is certainly
an improvement over the semantic one: it covers almost all cases
of compounds which intuitively can be said to have heads,
including all the cases covered by the semantic criterion.
However, the criterion is less than ideal, for it does not
provide a purely formal criterion for lexical headedness, but
only establishes a formal link between heads of compounds and
heads of their canonical paraphrases, which themselves are not
formally determined. This means that, in the final analysis, the
definition of a compound's head is not a formal one.

Only in the last few years has a fully formal criterion of
lexical headedness emerged. Its inspiration, though, is phrasal.
The criterion takes advantage of the fact that both phrases and
compounds have constituent structure, which can be characterized
in terms of context-free rules.

Now a formally precise and an empirically fairly adequate
definition of a phrasal head can be provided for within X-Bar
theory: Y is the head of Z iff Y is a daughter of Z and there is
an X such that Y = X$^i$ and Z = X$^{i+1}$. This definition, however,
when modified to accommodate the fact that lexical syntactic
categories do not have bar increments, yields the result that all
compounds of the form (x X X X) are double headed. But this is an
implausible analysis of the preponderance of compounds of this
form, and in particular, of tatpurusha compounds, be they English
or Sanskrit; for it is as far-fetched to claim for Sanskrit that
both raja (king) and purusha (man) are both heads of rajapuruṣa
(king-man: i.e., servant) as it is to claim for English that dish
and rack are both heads of dish-rack.

There is a criterion of lexical head, adopted by Selkirk
(1982 p. 20) and also, in effect, by Di Sciullo and Williams
(1987 p. 24), which avoids the implausible results just noted. It
is that Y is the head of Z iff Y is the right most daughter of Z
and is the same lexical category of Z.

This criterion enjoys not only a theoretical advantage over
the paraphrasal criterion, but in some cases, it also enjoys an
empirical advantage. Thus, on the paraphrasal criterion, the
tatpurusha compounds, ardhapippalī (half-pepper) and pūrva-kāya
(*fore-body: front of the body) (A 2.2.1-2), are head-initial,
since the first words in these compounds are the heads of their
canonical phrasal paraphrases. However, on the lexical criterion,
they are head-final.

This criterion pays a further dividend in terms of
strengthening another correlation. In general, the gender of a
tatpurusha compound is the gender of its final lexical element.
This is also true in the cases of the tatpurusha compounds grouped
with pūrva-kāya (*fore-body: front of the body) (A 2.2.1-2). In
particular, where the gender of compounded stems differ, the
gender of the compound is the gender of the final stem, not the
initial stem. So, for example, the gender of ardhapippalī
(half-pepper) is the gender of the final stem, pippalī, which is
feminine, and not the gender of the initial stem, artha, which is
neuter. Under the paraphrasal criterion, such compounds are
exceptions to the generalization that the features of a compound
are the features of its head; whereas, under the lexical
criterion, they are not. This is not to deny that some of the
items listed as exceptional by Pāṇini are exceptional. Certainly
they are. Rather, this is to affirm that there is a principled way whereby the number of these exceptions can be reduced.

Let us now see how the two syntactic criteria fare when applied to the remaining three great divisions of Sanskrit compounds: bahuvrīhi, or excentric, compounds; avyayībhāva compounds; and dvandva, or copulative, compounds.

Under the paraphrasal criterion, bahuvrīhi compounds are anomalous. On the one hand, the head of the paraphrase is, for the most part, the last lexical element in the compound; on the other hand, the features of the compound are those of the lexical item modified by the compound, not the final lexical item in it. Under the lexical criterion, supplemented by the analysis of these compounds as ones which have undergone adjectivalization through zero conversion, bahuvrīhi compounds proper are headless, though the compounds from which they are derived are typically head final. Thus, a bahuvrīhi compound no more has a lexical head than their English counterparts do, which require, for the most part, the overt suffix -ed for their formation. Thus, for example, mukha (face) is no more the head of ugra-mukha-B (grim-faced) in Sanskrit than face is the head of grim-face-ed in English.

The lexical criterion proves illuminating for such English bahuvrīhi compounds as red-head, dim-wit, hard-back, straight-edge, long-legs, and heavy-weight, mentioned as problematic by Selkirk (1982 p. 26). All of these can be regarded as nominalizations of the bahuvrīhi compounds red-head-ed, dim-wit-ed, hard-back-ed, straight-edge-ed, long-leg-ed, and ?heavy-weight-ed, where the suffix -ed has been elided. What is odd here, then, is not that these compounds are excentric, but that they have managed to be excentric without the overt presence of the usual -ed suffix.

Avyayībhāva compounds bring the paraphrasal criterion into conflict with the generalization of the head-finality of compounds. As stated earlier, this compound results in indeclinables (roughly, adverbs) of the form (ADV (p x )-(N y)) and is paraphrasable into a phrase consisting of the same two words, x and y -- its paraphrase being a prepositional phrase of the form [PR [p x ] [N yi ] ], where i is the case required by x. One such compound is bāhiḥ-grāmam (out-village: i.e., outside of the village), which has the canonical form of the second case (i.e., accusative), singular, neuter. Its paraphrase is bāhiḥ grāmāt, where the word corresponding to the first element of the compound retains its form, being a preposition, and the second word acquires the form of the fifth case (i.e., ablative), singular, masculine -- the form canonically assigned by its accompanying preposition. Under the paraphrasal criterion, these compounds are head-initial, since clearly the head of the paraphrase is the proposition.
Whether or not avyayībhāva compounds are headed in Sanskrit under the lexical criterion depends on what lexical category these compounds are assigned to. The morphology of these compounds is nominal: they terminate with an oblique nominal ending, typically a neuter, singular, second case (i.e., accusative) ending. However, their function is adverbial, much as prepositional phrases in English may function adverbially; and adverbs derived from adjectives terminate with a neuter, singular, second case ending. Two options, then, seem possible: avyayībhāva compounds are adverbs, in which case they are headless, just as bahuvrīhi compounds are headless; or, they are nouns, in which case they are head-final.

The latter option is unattractive, for it would mean that avyayībhāva compounds would be indistinguishable from prādi-tatpurūṣa compounds. Prādi-tatpurūṣa compounds are compounds with immediate constituents of the same lexical categories occurring in the same order as avyayībhāva compounds; and they are clearly of the same lexical category as their final elements, which may be either nouns or adjectives. Thus, for example, pra-ācārya (super-teacher: i.e., eminent teacher) is a noun and the features of the compound are the features of ācārya (teacher); while ut-caṇḍa (*high-violent: highly violent) is an adjective like its final immediate constituent caṇḍa (violent) (Co p. 135). If avyayībhāva compounds are nouns or adjectives, then there would be nothing to account for the fact that they always function as adverbs and never as nouns whereas prādi-tatpurūṣa compounds always function as nouns or adjectives and never as adverbs.

As it turns out, the same analysis carries over to such compounds in English. English too has avyayībhāva compounds: over-head, over-land, over-seas, up-town, up-stairs, up-stream, down-town, down-stairs, down-stream, in-doors, in-land, etc.; for each of these expressions may function as adverbs. Moreover, there is no question that these expression are compounds: common nouns in English require an overt deteriner to form a noun phrase, unless they are either plural or mass; but in several instances, the common noun is singular and not mass, and yet no determiner intervenes between it and the preceding preposition. English avyayībhāva compounds are less productive than Sanskrit avyayībhāva compounds. In particular, English avyayībhāva compounds are more restricted in the kind of preposition and noun they can tolerate: over-head is acceptable, but over-shoulder is not; up-stairs is acceptable, but up-ladder is not. English avyayībhāva compounds clearly function adverbially. And they must be distinguished from English prādi-tatpurūṣa compounds, such as over-dose, which functions as a noun and never as an adverb. At the same time, the very English expressions just identified as avyayībhāva compounds can also occur as left sisters in compounds: over-head-compartment, up-stairs-apartment, in-door-pool, etc. It is not clear whether this is also true of Sanskrit avyayībhāva compounds.
Thus, according to the lexical criterion, avyayībhāva compounds are headless, while prādi-tatpurusā are head-final. Thus, the generalization is retained that all headed compounds are head-final.

Consider now dvandva, or copulative, compounds. The lexical criterion implies that the final element in such a compound is its head. But this conflicts with the universality of the generalization that the features of the compound are those of the head. The grammatical number of these compounds is determined in two ways. On the one hand, when a collective reading of a dvandva compound is required, the number of the compound is singular and the gender neuter, regardless of how many nouns form the compound. This is obligatory for certain canonical pairs or groups: for example, pāñjī+pādam (hands and feet). On the other hand, when no collective reading is required, the grammatical number is that number of nouns conjoined by the compound. If two nouns are conjoined, then the number is dual: for example, Rāma+Krṣṇa (Rama and Krishna); and if more than two, the number is plural: for example, deva+gandharva+mānuṣa+uraga+rākṣasāḥ (gods, fairies, humans, serpents, and ogres). As a result, the grammatical features of a dvandva compound are not usually the grammatical features of its head.

It should be noted, however, that dvandva compound's gender, aside from the neuter gender it is assigned when the compound requires a collective reading, is that of the final noun. In particular, when a dvandva compounds lexical elements have different genders and are permuted, the gender of the compound varies with the gender of the last element. For example, the gender of mayūrī (pea hen) is feminine and that of kukkuṭa (cock) is masculine; while the gender of mayūrī+kukkuṭa (pea hen and cock) is masculine, that of kukkuṭa+mayūrī (cock and pea hen) is feminine (Ca p. 267). This fact may incline one to retain the lexical criterion of compound headedness and modify the generalization pertaining to match of grammatical features of the compound and its head. However, there is some evidence to suggest that this is not the right choice.

It seems that all Sanskrit compounds of the form (x X X x) are ambiguous. There are only two types of such compounds: one type where X is A, and the other where X is N. As we saw before, in each case, the final element may or may not be distinguished with respect to the initial one. In the case of adjectives, the final element is distinguished when the initial element adverbially modifies the second; otherwise, neither element is distinguished, rather both are construed as modifiers of whatever it is the compound itself modifies. So, the compound tulya-sveta can mean not only equally white but also equal and white. An analogous ambiguity obtains in the case of nouns as well. Thus, karma-phala (action-result) could be karmaphala (actions and results) or kamaphal (results of actions) (B&T 5.3.16). Surely, it is the headedness of the compound which determines the reading, in each case. But the identity of the head is not, in
these cases, reducible to identity of lexical categories, since all the lexical categories involved are the same.

As anathema as it might be to some reductionist-minded linguists, perhaps headedness cannot be defined away -- perhaps it is just a primitive diacritic, applied freely, subject to the following language specific parameters: (1) No non-head occurs to the right of a head; (2) if x is the head of y, then x and y must be the same category; (3) only non-heads bear thematic roles; (4) the features of a head of a compound are the features of a compound.

2.2 Interpenetration of Lexical and Phrasal Syntax

Recall that Pāṇini's treatment of compounds is to pair them with canonical phrasal paraphrases with which they share a common derivational ancestor; in addition to their semantic relation, they bear the syntactic relations of having the same heads and of having the same constituency. Hence, the constituency of compounds mirrors that of their canonical phrasal paraphrases. A condition on compound formation is that two elements cannot undergo compounding, the deletion of morphology from the subordinate element, unless the two elements form a constituent (A 2.1.2). A consequence of this is that inflected lexical items exterior to a compound are not construable with lexical elements subordinate within it. The applicability of this rule is illustrated both by Patañjali, in his Mahābhāṣya, or Great Commentary, on Pāṇini's Asṭādhyāyī (at A 2.1.1), and by Bhartṛhari, in his work on the semantics of Sanskrit (VP 3.14.46), with the following example:

(24.1) [NP1 ((ṛddha<rāja)<puruṣaḥ) ]
(24.2) *[NP1 [AP6 ṛddhasya ] (rāja<puruṣaḥ) ]
rich king-man
servant of a rich king

An analogous fact in English,

(25.1) ((city<destruction)<story)
(25.2) *(destruction<story) of the city

has prompted DiSciullo and Williams (1987: p. 30) to conclude that "the arguments of the nonhead are no part of the argument structure of the compound."

Unfortunately, while their generalization holds of English compounds, it does not of Sanskrit compounds. Counter-examples are furnished both by Patañjali (Mahābhāṣya on A 2.1.1) and by Bhartṛhari (VP 3.14.47):
Indeed, compounds appearing in configurations such as that in (26.3) are given a special name by Sanskrit grammarians: they call them *asamartha* compounds (i.e., non-constituent compounds). Moreover, these compounds are attested in the classical literature. A study of approximately three-hundred sentences, chosen essentially at random from the Classical Sanskrit corpus, reveals thirteen clear cases of *asamartha* compounds. (See Part I of Appendix I below.) And a study of the first approximately five-hundred sentences of a single text reveals forty-three clear cases. (See Part II of Appendix I below.)

One person to attempt to meet the challenge presented by these compounds to Pāṇini’s grammar of Sanskrit was Bhartrhari, who suggested that *asamartha* compounds are limited to cases where the subordinate element in the compound expresses a relation. It is tempting, following Bhartrhari’s lead, to conjecture that the subordinate lexical element in the compound has a thematic role which is assigned to the inflected lexical item external to the compound. A survey of the cases mentioned above as culled from the literature shows that, for the most part, such is the case: the subordinate lexical element is a deverbal noun or adjective, either subcategorizing for an NP complement or having associated with its verbal root a thematic argument, which the external lexical item satisfies. This suggests that while, in English, "the arguments of the nonhead are no part of the argument structure of the compound"; in Sanskrit, they do.

Moreover, the accessibility of elements subordinate within a compound to elements external to a compound is not confined to cases of the satisfaction of subcategorization frames and argument structure. Contrary to what Di Sciullo and Williams (1987: pp. 40, 50, 58-9, 90) observed about English, an observation anticipated by Paul Postal, subordinate elements within Sanskrit compounds may enter into anaphoric relations with lexical items external to them, either in the form of a compound containing a pronoun and the pronoun requiring an antecedent or in the form of a subordinate element within a compound providing an antecedent for a pronoun external to the compound. Moreover, every personal pronoun, including the reflexive pronoun ("sva") and the third person personal pronoun ("tāt"), may occur as a subordinate element within a compound.

(26.1) [NPs [NPs Devadattasya ] guruh ] kulam
(26.2) (Devadatta<guru><kulam
(26.3) [NPs Devadattasya ] (guru<kulam) Devadatta’s teacher’s family

(27.1) PVS 10.3
[s [ABC [NPs [NPs rāgasya l] (an-upayoge) ] ] [ADV katham ]
  passion non-use how
[NPis (tat₁<śaktiḥ) ] [vṛ upayujyate ]
  it-potentiality use
If there is no causal role for passion <in the production of speech>, in what way does its potentiality have a causal role?
(27.2) PVS 11.1
[s [NP1s [NP3 [NP6 anvayasya ] niścayena ]
concomitance ascertainment
[NP6 (((viruddhah)+((tat1<pakṣyāṇām)) ] nirāsah ] ]
contrary+it-similar rejection
There is the rejection of contraries and ones like them by the ascertainment of concomitance.

(28.1) PVS 22.7
[ABC [NP7 [NP6 kāryasya 1] [PP [NP2 (sva<klāraṇam) ] antareṇa ]
effect self-cause without bhāve ] ]
existence
if an effect were to exist without its cause

(28.2) PVS 19.25
[s [NP1s tau 1] [VP [NP2 (sva<kratibaddham) ] nivartayataḥ ] ]
the two self-connected exclude
The two exclude what is connected with them.

Both of these cases, the one where lexical elements external to a compound can be construed as arguments to lexical elements subordinate within a compound, and the one where lexical elements external to a compound can bear anaphoric relations to elements subordinate with it, can be handled within the Di Sciullo and Williams view, provided that argument structure is permitted to percolate to the compound as a whole from its subordinate elements.

However, even this modification of the Di Sciullo and Williams view runs into trouble. First, in the sentence cited above, the word guru, unlike its English translation teacher, is not a deverbal noun, and so one would not think of it as having a thematic role associated with it. One solution would be to expand argument structure to include not only argument positions associated with a verbal root, but also elements associated with certain kinds of relational nouns which cannot be said to be derived from verbal roots, such as brother, king, etc.

Even so, problematic cases remain. There are cases where the lexical element external to a compound is an adverb and the lexical element subordinate within is a perfect passive participle.

(29) SGSC 9.1.3 <= V 2.134>
[AP1 [ADV ṛṇḍham ] [PR7 khulu ] [NP7 tvayi ] (baddha<bhāvā>-B ] ]
firmly indeed on you fixed-affection-ed
[NP1s Urvāśī ].
Urvāśī
Indeed, Urvāśī is one whose affection is firmly fixed on you. Remark: bādh subcategorizes for an NP7.
(See Appendix I for other examples.) There is also a case where the lexical element external to a compound is in apposition to an element subordinate within:

(30) PVS 15.20
[s yena [n5 [n6 (vacana-ādeḥ) ] (((kim-cit)<mātra)<sādharmyāt)
so that speech-etc. something-only-similarity
] [n1s ((sarva<ākāra)<sāmyam) ] [vp anumīyete ]
all-respect-similarity should be inferred.
... so that, from the commonality of merely something, say of speaking, similarity in every respects should be inferred.

And, in fact, the appositive element may be an entire clause:

(31) PVS 17.23
[n5 [ [s [n5 ((eka<dharma)>1<sadbhāvāt) ] [vp [n3 (tat1<anyena)
one-property-presence it-another
api ] bhavitavyam ] iti ] (niyama<abhāvāt) ]
too must be present PRT necessity-absence
... because of the lack of a necessity <to the effect> that because of the existence of one property, another too must exist.

(See Appendix I for other examples.)

3.0 Conclusion:

A very intriguing aspect of the foregoing examination is the way in which compounds of one type in Sanskrit embed within compounds of a corresponding type in English (i.e., every Sanskrit compound of one type has a unique well-formed English correspondent of the corresponding type), and vice versa. Thus, the word formation rules for Sanskrit form a proper subset of those for English, coming very close to co-inciding with them. At the same time, those rules which apply in Sanskrit and whose application in English is not as wide, typically differ with respect to identifiable sub-regularities. For example, both English and Sanskrit have compounds of the form N -> NN, yet the application of the rule in English is systematically more constrained than in Sanskrit. Yet, while Sanskrit permits the unconstrained formation of headless compounds of this sort, so-called dvandva, or copulative, compounds; English tolerates such a compound only when its constituents are proper names (e.g., Alsace-Lorraine). At the same time, Sanskrit permits the unconstrained formation of headed compounds of the same sort, including compounds of the form \((n (n x ) (n y ))\), where \(x\) denotes the possessor of what is denoted by \(y\); English does not tolerate such compound. Thus, guru<kulam> is acceptable in Sanskrit; its translation, teacher-family is not in English. Analogous remarks apply to compounds of the form A -> AA. Such compounds are permitted in both English and Sanskrit, but Sanskrit permits such compounds to be construed in a way that the first element adverbially modifies the second, regardless of whether the second adjective is derived from a verb or not;
whereas English permits such compounds only when the second adjective is derived from a verb. Finally, headless PN compounds (i.e., ADV -> PN), so-called avayabhava compounds, are limited to prepositions of a certain kind, such as up and down, in and out, and over and under.

This reciprocal embedding of compounds of various types is evidence, I believe, of the theoretical utility of approach of word formation explored here, for, where the embedding of one type of compound within another is proper, the residual compounds are characterizable in terms of independently recognized syntactic or semantic properties -- a fact highly suggestive of parametric variation. At the same time, however, the Sanskrit facts suggest the need for some revisions in the details of the approach as propounded first by Selkirk (1982) and later by Di Sciullo and Williams (1987). For example, we have seen some evidence that headedness may not be definable in terms of lexical constituent structure alone. Moreover, we have seen that more information reaches the complex formations from the non-head elements than is permitted by either Selkirk or Di Sciullo and Williams.
Appendix: Examples of Non-Constituent Compounds

Part I: Study of a Random Sample of Classical Sanskrit Prose

(1) § 3.9.16 <= SGSC 3.1.6
[NP1 [NP3 [NP7 tasyām ] (snigdha<drṣṭyā) ] (sūcita<abhilāṣah)B ]
on her with fixed-gaze indicated-affection-ed
... whose affection was indicated by his gaze being fixed on her
Remark: snigdha subcategorizes for NP7.
NP3 is the agent of "sūcita".

(2) K119.14 <= SGSC 3.1.12
[NP6 [NP6 ... (nija<balasya) ] (viśrāma<hetoḥ) ]
own-army rest-reason
... in order to rest his own army
Remark: NP6 is the theme (karma) of viśrām

(3) § 6.4.4 <= SGSC 6.1.4) <Prakṛt>
[AP1 (a-(śruta<pūrvah)) [NP3 āvābhyaḥ ] ] [NP1s [AP1 eṣah ]
non-heard-previous by us two this
vṛttāntaḥ]
news
This news had not be previously heard by the two of us.
Remark: NP3 is the agent of śruta.

(4) K 356.13 <= SGSC 8.2.1
[NP1 [NP5 (janma<karmataḥ) ] ((malina-tara)<janam>-B ]
than action+deed dirti-er<people-ed
... where people are dirtier than their actions or deeds
Remark: -tara subcategorizes for an NP5.

(5) K 356.13 <= SGSC 8.2.1
[NP1 [NP5 janataḥ ] ((nistriṁśa-tara)<(loka<hṛdayam)>-B ]
than people pitiless-more-worldly-heart-ed
... where worldly hearts are more pitiless than the people
Remark: -tara subcategorizes for an NP5.

(6) K 356.13 <= SGSC 8.2.1
[NP1 [NP5 (loka<hṛdebyaḥ) ]
worldly-hearts
(((nirghṛna-tara)<<(sarva<samvyavahāra)<<(samasta<puruṣam)>-B ]
cruel-more-entire-behavior-all-people-ed
... where every person's behavior is always more cruel than their
worldly hearts.
Remark: -tara subcategorizes for an NP5.
(7) V 2.134 <= SGSC 9.1.3
[AP1 [ADV dr̥ham ] ... [NP7 tvayi ] (baddha<bhāvā>-B ] ) [NP1s firmly on you fixed-affection-ed
Urvaśī ],
Urvaśī
Urvaśī is one whose affection is firmly fixed on you.
Remark: baddh subcategorizes for an NP7.
ADV modifies baddha.

(8) V 2.129 <= SGSC 14.1.4
na asmi [NP1 [NP7 bhavatyoḥ ] ((īśvara<niyoga><pratyarthī) ] 
not I am of you two master-charge-hostile.
I am not hostile to the charge to you two of <your> master.

(9) 3.15.6 <= SGSC 14.2.1
[AP3 [NP6 ... janasya] (ārti<hareṇa) ] [NP3 rajñā ]
people suffering-remover king
bhavitavyam.
must be.
A king must be one who removes the suffering of ... people ...
Remark: NP6 is the agent of ārti.

(10) 3.14.6 <= SGSC 28.1.4
[NP1 [NP3 (rśi<janena) ] (sarva<damanaḥ) iti ]
by the sages all-queller as (krta<nāmadheyah] ] asi
made-appellation you are
... You are called the queller of all by the sages.
Remark: NP3 is the agent of krta.

(11) K 294.5 <= SGSC 28.2.1
... [NP1s [AP1 [NP6 anyayoḥ api] ((a-darśana)<krtaḥ ] kleśāḥ ]
those two too not-seeing-made pains
The pains brought about by not seeing those two as well ...
Remark: NP6 is the theme (karma) of darśana.

In addition, there are a number of cases where an inflected element external to a compound might be construed with one of its subordinate elements, but need not be. These cases may be divided into two, depending on whether the inflected element external to the compound is an argument or not of the compound’s subordinate element. The former cases are: Ve 2.26.4 <= SGSC 1.1.5>, V 3.46.18 <= SGSC 2.2.1>, K 292.6 <= SGSC 5.1.7>, V 3.20 <= SGSC 18.1.1>; the latter cases are: § 3.5.2 <= SGSC 3.1.4>, M 1.12.7 <= SGSC 6.1.2>, V 2.129 <= SGSC 14.1.4>, § 1.7.2 <= SGSC 22.1.2>, and P 1.88.12 <= SGSC 27.1.6>.

Part II: Study of a Single Text of Classical Sanskrit Prose

A comprehensive list of cases based on a systematic study of the first four hundred and fifty or so sentences of PVS. Again, these are divided up into two groups, depending on whether the
inflected element external to the compound is an argument or not of the compound's subordinate element.

1. Arguments

(1) PVS 1.7
[AP1 [NP7 atra ] (anubaddha<spṛham>-B ]
  this attached-eagerness-ed
one who has eagerness for this
Remark: anubaddha subcategorizes for an NP7.

(2) PVS 2.2
[NP5 ... [NP7 drṣṭānta-dharminī ] (sattva<siddheḥ>) ]
corroborating instance presence-establishment
because of the establishment of <its> presence in the
  corroborating instance

(3) PVS 2.4
[NP7 (sādhya<abhāve) ] ((a-sattva)<vacana>-vat
  establishable-absence non-presence-statement-like
like the statement of <its> non-presence in the absence of what
is to be established

(4) PVS 2.11
[NP5 [NP6 vaktuh ] (abhiprāya<vaśāt) ] ]
speaker intention-dint
by dint of the intention of the speaker.

(5) PVS 3.11
[NP5 [NP6 ... bhavataḥ ] (((tat<utpatti)<niyama)<abhāvāt) ] ]
existent it-arising-necessity-absence
because of the absence of necessity of the arising from it of an
  existent

(6) PVS 7.3
[AP7 [NP6 ((sāmagrī<phala)<śaktinām) ] (parināma<anubandhini) ]
totality-result-potentialities development-dependent
which depends on the development of the potentialities for the
result of <its> causal totality

(7) PVS 7.5
[NP5 [NP6 ... śaktinām ] (parināma<apekṣatvāt) ... ]
potentialities development-need
because of the need for the development of the potentialities

(8) PVS 8.1
[NP1 [NP6 (rūpa<(upādāna-hetūnām)> ) ] (pravṛtti<kāraṇam> ]
color-material-causes activation-cause
the cause of the activation of the material causes of color
(9) PVS 8.12
\[NP5 \ [NP6 \ ((a-pratibaddha)\langle svabhāvasya\rangle-B \ ]\]
  non-related-nature-ed
  ((avinābhāva\langle niyama\rangle\langle abhāvāt\rangle \ ] \]
  indispensibility-necessity-absence
because of the absence of the necessity of the indispensability of that whose nature is unrelated.

(10) PVS 12.23
\[ABC \ [NP7 \ ... \ [NP7 \ (kva-cit) \ ] \ (abhāva\langle siddhau \rangle) \ ] \]
somewhere absence-establishment even
even if there is the establishment of an absence in some place

(11) PVS 12.23
\[NP5 \ [NP6 \ (a-pratibaddhasya) \ ] \ ... \ (abhāva\langle a-siddheḥ \rangle) \]
  un-connected absence-non-establishment
because of the non-establishment of the absence of what is unconnected.

(12) PVS 13.4
\[NP5 \ [NP6 \ (ghaṭa-ādīnām) \ ] \ (nairātmya\langle a-siddheḥ \rangle) \]
  pot-etc selflessness-non-establishment
because of the non-establishment of selflessness of such things as pots

(13) PVS 15.4
\[NP3 \ [NP6 \ asya \ api \ ] \ [NP7 \ [AP7 \ (kva-cit) \ ] \ viṣeṣe \ ] \]
  its too any species
  (sambhava\langle āśaṅkayā \rangle)
existence-uncertainty
through uncertainty about the existence of it too in any species

(14) PVS 16.1
\[NP5 \ ... \ [NP6 \ ((a-drṣṭa)\langle viṣeṣānām\rangle-B \ ]\]
  un-observed-special-qualified
  (((vijātīyatva\langle upagama\rangle\langle virodhāt \rangle) \]
  heterogenity-acceptance-contradiction
because of the contradiction of the acceptance of the heterogenity of those things whose special qualities are unobserved

(15) PVS 16.20
\[NP5 \ [NP7 \ (a-(pararṣaṁ\langle artheṣu\rangle)) \ ] \ (saṁghatvatval\langle a-darśanāt \rangle)\]
  non-other-things compositeness-non-observation
because of the non-observation of the compositeness of things not for others

(16) PVS 17.23
\[NP14 \ [NP6 \ sādhanaṣya \ ] \ ((sādhya\langle vyabhicāra\rangle\langle āśaṅkā \rangle) \]
  establisher establishable-deviation-uncertainty
uncertainty about the deviation of the establisher from the establishable.
(17) PVS 19.3
[NP5 [NP6 (arthā<antarasya)]
  thing-other
((tat<anubandha)<niyama><abhāvāt)]
  it-depence-necessity-absence
because of the absence of the necessity of dependence on it of
what is other than it

(18) PVS 19.20
[ADV ... [NP6 anyasya] (prasiddhi<artham)]
  another establishment-sake
for the sake of the establishment of another

(19) PVS 20.16
[NP1s ((avaśyam<bhāva)<niyamaḥ)] ... [NP6 parasya]
  inevitable-presence-necessity of another
the necessity of the inevitable presence of another

2. Non-Arguments

2.1 External Element a Locative Absolute

(20) PVS 2.6
[NP1 [ABC [NP7 [NP5 sāmarthyāt] (arthā<gatau)]]
  meaning-implication
((pratipatti<gaurava)<parihāra><artham>-B)
  understanding-heaviness-exclusion-purpose
whose purpose is to exclude a burden on <one's> understanding,
when one understands the meaning <of the definition> through
implication.

(21) PVS 12.23
[NP5 [NP6 (a-pratibaddhasya)] [ABC [NP7 (tat<abhāve)]
  un-connected
  it-absence
(AP7 sarvatra)] (abhāva<(a-siddheḥ))
  every absence-non-establishment
because of the non-establishment of the absence of what is
unconnected in every absence of it

2.2 External Element an Adverb

(22) PVS 3.17
[NP3 āṃšena] [NP1s (((janya+janaka)-tva)<prasaṅgah)]
  by part produce-producer-nesc-absurdity
the absurdity of partial causation
(23) PVS 21.7
[AP1 [NP3 (sattā<upalambhena)] [ADV sarvadā]
  presence-apprehension always
(((tathā<bhāva)<aṃkā<vipralabdhaḥ)])
  deceived by his apprehension of <a thing’s> presence into the
  suspicion of <its> always being so

2.3 External Element an Apposition

(24) PVS 15.20
[NP5 [NP6 (vacana-ādeḥ)] (((kim-cit)<mātra<sādharmyāt])
  speaking-etc. some-mere-commonality
  because of commonality in merely some respect, such as speaking

(25) PVS 17.23
[NP5 [S [NP5 ((eka<dharma)>sadbhāvāt)] [VP [NP3 (tat<anyena)
  api ] bhavitavyam ] ] iti ] (niyama<abhāvāt)]

(26) PVS 20.12
[NP3 [NP3 . . . a-gatyā] . . . (((iti-ādi)<bhedā)<prayogaiḥ)]
  non-understanding so-etc-differentia-formulation
  through the formulation of such differentia as the
  non-understanding . . .
Notation

Anaphoric Co-indexation: index to the immediate right.
B: appended to bahuvrīhi cpds, so-called exocentric cpds.
Case Marking: occurs at the end of phrasal label or as a
  subscript to a word, following Pāṇinian numbering
    (1 nominative; 2 accusative; 3 instrumental; 4
dative; 5 ablative; 6 genitive; 7 locative)
Parentheses: lexical constituent structure
Square Brackets: phrasal constituent structure
  phrasal labels: [A B C] locative absolute
Subordination: "x<y" means that x is subordinate to y

Abbreviations

A Aṣṭādhyāyī <adhyāya.pāda.sūtra>.
B&T Booose and Tubb 1981.
Co Coulson 1976.
K Kādambarī; Peterson (ed) 1885 <page.line>.
NBT Nyāyabinduṭika; Malvania (ed) 1955.
NBTP Nyāyabinduṭīkapradīpa; Malvania (ed) 1955.
PSED Apte 1890.
PVS Pramāṇavārttika-svavṛttī; Gnoli (ed) 1960 <page.line>.
SGSC Apte 1885 <chapter.exercise.sentence>.
Ś Śakuntalā; Godabole (ed) 1933 <page.verse.line>.
V Vikramorvaśīyam; Pandit (ed) 1889 <act.verse.line>.
VP Vākyapadīya; Iyer (ed) 1973
  <kāṇḍa.samuddeśa.kārikā>.
W Whitney 1881.

Acknowledgements

The initial research for this paper was partially funded by a
grant from the Social Sciences and Research Council of Canada.
The basic research for it was done in 1986 while I was a senior
fellow of the American Institute of Indian Studies, affiliated
with the Centre for the Advanced Study of Sanskrit at Poona
University. While there I had the opportunity to learn from and
work with a number of scholars. I am especially grateful to Dr.
Siniruddha Dash, who devoted so much time to helping me through
the intricacies of the Pāṇinian treatment of Sanskrit compounds.
I was also privileged to receive advice and guidance from Paṇḍit
Śrī T. S. Śrīnivāsa Śāstrī and Śrī Venugopalan of Deccan College.

Earlier versions of this paper were presented 10 February, 1990,
at the "Workshop on Lexical-Syntactic Relations" held at the
University of Toronto, and 9 June, 1990, at the "Twelfth South
Asian Languages Analysis Round Table", held at the University of
California at Berkeley. I wish to thank the audiences in
attendance at these presentations for help comments and
criticism. I also wish to thank Richard Hayes for his extensive
comments on the initial drafts of this paper.

- 43 -
Bibliography


Dash, Siniruddha and Brendan S. Gillon 1986 "The Distinction Between Adjectives and Nouns in Sanskrit". Centre for the Advanced Study of Sanskrit, Poona University, ms.


