Generation of Sanskrit Compounds

Pavankumar and Amba Kulkarni

Department of Sanskrit Studies,
University of Hyderabad,
Hyderabad

Sanskrit Compounds

Samāsa(Compound) is a technical name assigned by Pāṇini.
Samāsa:

1. Unification (In the sense of activity / process) and
2. Unified (In the sense of result of action)

Classification of Sanskrit Compounds

Sanskrit Compounds are broadly classified into four types:

- Avyayībhavaḥ (Adverbial)
- Tatpuruṣaḥ (Endocentric)
- Bahuvrīhiḥ (Exocentric)
- Dvandvaḥ (Conjunctive)

Features of Sanskrit Compounds

- It is a single word (ekapadam).
- It has a single case suffix (ekavibhaktikam) with an exception of *aluk* compounds such as yudhiṣṭiraḥ, where there is no deletion of case suffix of the first component.
- It has a single accent(ekasvaraḥ).
- The order of components in a compound is fixed.
- No words can be inserted in between the compounds.
- The compound formation is binary with an exception of dvandva and bahupada bahuvrīhi.
- Euphonic change (sandhi) is a must in a compound formation.
- Constituents of a compound may require special gender or number different from their default gender and number. e.g. pāṇipādam, pācikābhāryaḥ, etc.

Generation of Compounds

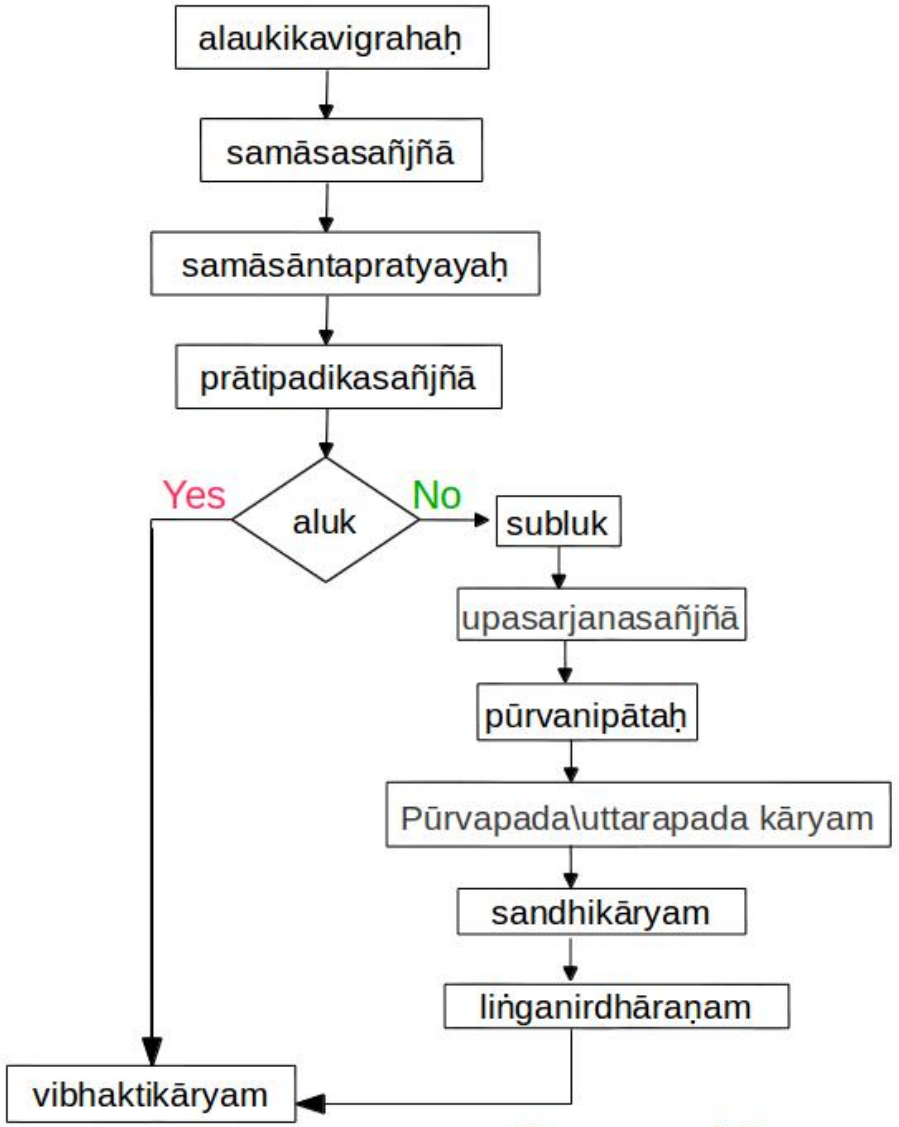
Input: a) Concept
Blue Lotus
nīlam utpalaṃ

b) Intermediate Linguistic Expression (alaukikavigraha)
nīla + su utpala + su type : karmadhāraya

Output: nīlotpalam

Our choice is b.
This process involves several steps.

Process of Compound Generation



(B1+S1 B2+S2)

(B1+S1 B2+S2) samāsa

(B1+S1 B2+S2 S3)

(B1+S1 B2+S2 S3) prātipadikam

(B1+[S1] B2+[S2] S3)

Decide the order

B1 ⟶ B3 /_ B2

B2 ⟶ B4 / B1_

B1' + B2' = B'

Decide the Gender

Process of Compound Generation

B = Base
S = Suffix

Contd ...

The rules pertaining to the compound formation are primarily of two types.

- sañjñāsūtrās and
- vidhisūtrās

Context Free Rules

The sañjñāsūtrās are implemented as a Context Free Rules.

Regular Grammar

The vidhisūtrās are implemented as a Regular Grammar.

Grammar for Compound Generation

padam:		subantam
	|	tiṅantam
	;	
subantam:		stem sup
	;	
stem:		$stem_{NT}$
	|	$stem_{T}$
	;	
$stem_{NT}$:		kṛt
	|	taddhitaḥ
	|	samāsaḥ
	;	
samāsa:		alaukikavigrahaḥ
	;	
alaukikavigrahaḥ:		stem ‘+’ sup stem ‘+’ sup
	|	avyaya stem ‘+’ sup
	|	stem ‘+’ sup avyaya
	;	

Lexer for Compound Generator

```
sup:    su|au |jas
      | am |auṭ |śas
      | tā |bhyām |bhis
      | ṅe |bhyām |bhyas
      | ṅasi |bhyām |bhyas
      | ṅas |os |ām
      | ṅi |os |sup
      ;
avyaya  su
      | upa
      | prati
      ;
stem_T  [a-zA-Z]+
      ;
```

Contd ...

Vidhisūtrās are further classified into three types:

1. Right context
2. Left context
3. Extended right context

Right Context

$W_1\ W_2 \rightarrow W_3\ W_2$.

W_1 changes to W_3 in the context of W_2.

pādasya padājyātigopahateṣu P6.3.52

pāda $\rightarrow$ *pada* / *āji* |*āti* |*ga* |*upahata*

Left Context

$W_1\ W_2 \rightarrow W_1\ W_3$.

W_2 changes to W_3 in the context of W_1.

jyotirāyuṣaḥ stomaḥ P8.3.83

jyotis |*āyus* / *stoma* → *ṣṭoma*

Extended Right Context

$W_1\ W_2 \rightarrow W_1\ W_2\ W_3$.

In the context of W_1 and W_2, W_3 is appended to the input string.

antarbahirbhyām ca lomnaḥ P5.4.117

antar |*bahir loman* → *antar* |*bahir loman ap*

Contd ...

$$B_1 \rightarrow B_3 \ / \ \_ \ B_2$$

$$B_2 \rightarrow B_4 \ / \ B_1 \ \_$$

Most of the sūtrās require sequence of phonemes but there are some sūtrās which require a rich structure involving following features:

- The morphological analysis of a word,
- The semantic category of a word,
- The meaning of a word, and
- The relation between two components.

Morphological analysis of a word

svayam kṛta + su → svayaṁkṛtam
svayam ktena P2.1.25

Semantic category of a word

pūrvāhṇa + ṅi kṛta + su → pūrvāhṇakṛtam
ktenāhorātrāvayavāḥ P2.1.45

Meaning of a word

kim + ṅas rājan + su → kiṃrājā

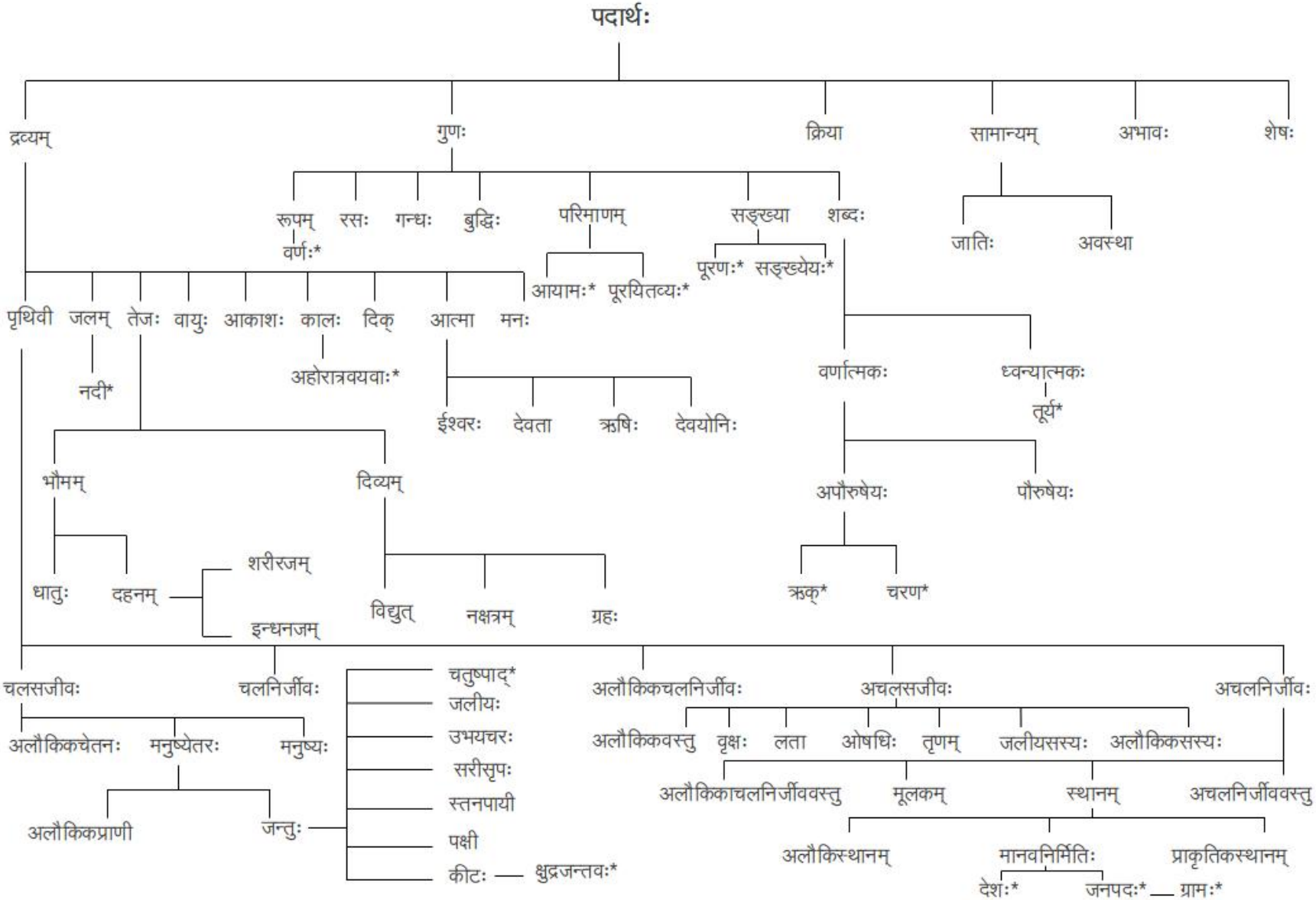
kim kṣepe P2.1.64

The relation between two components

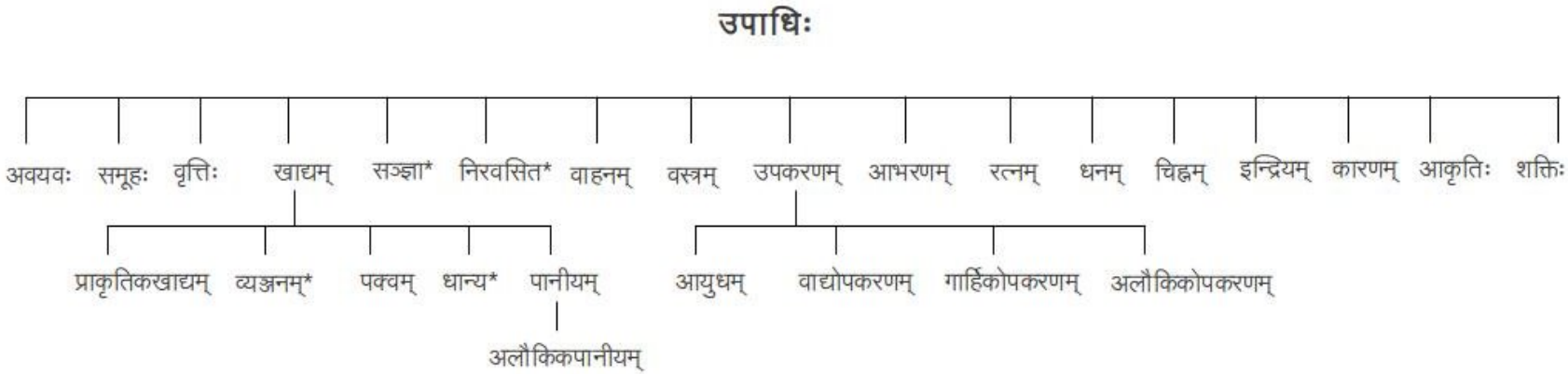
nīla + su utpala + su → nīlotpalam

viśeṣaṇaṃ viśeṣyeṇa bahulam P2.1.57

Ontological Classification



Upādhi


उपाधिः
अवयवः
समूहः
वृत्तिः
खाद्यम्
सञ्ज्ञा*
निरवसित*
वाहनम्
वस्त्रम्
उपकरणम्
आभरणम्
रत्नम्
धनम्
चिह्नम्
इन्द्रियम्
कारणम्
आकृतिः
शक्तिः
प्राकृतिकखाद्यम्
व्यञ्जनम्*
पक्वम्
धान्य*
पानीयम्
अलौकिकपानीयम्
आयुधम्
वाद्योपकरणम्
गार्हिकोपकरणम्
अलौकिकोपकरणम्

Other Relations

Apart from these relations there are other relations like:

- resemblance for comparison
- contradictory/opposition etc.

Resemblance for Comparison

ghana and śyāma undergo compound formation by the rule upamānāni sāmānyavacanaiḥ (P2.1.55) they form a compound word ghanaśyāmaḥ .

Contradictory/Opposition

The words śīta and uṣṇa denote opposite properties, and hence by vipratiṣiddham cānadhikaraṇavāci (P2.4.13) rule, they form a compound word śītoṣṇaṃ.

Extra Linguistic Information

There are some sūtrās which need extra linguistic information for example:
The condition that the components form a compound will refer to a name of a person (sañjña)
yudh + ṅi sthira + su → yudhiṣṭhiraḥ
haladantātsaptamyāḥ sañjñāyām P6.3.9