Annotation Scheme for Kāraka Level Tagging and Guidelines

Prepared by Prof. K. V. Ramkrishnamacharyulu, Sheetal Pokar, Devanand Shukla, and Amba Kulkarni on behalf of Sanskrit Consortium (2008-12) Modified further by Amba Kulkarni and Sanjeev Panchal

July 17th, 2019

Contents

0	Bacl	kground	4
	0.1	What is the intended level of semantic tagging?	4
	0.2	Which relation to tag and which to not?	4
	0.3	द्योतकपद् versus वाचकपद्	5
	0.4	Convention for marking the relations	5
	0.5	Granularity	6
	0.6	Unit for Tagging	6
	0.7	Proposed কাংক-tag-set for Sanskrit	7
1 कारकसम्बन्धाः(Kaaraka Relations)		7	
	1.1	कर्ता	7
		1.1.1 प्रयोजककर्ता	8
		1.1.2 प्रयोज्यकर्ता	8
	1.2	कर्म	9
		1.2.1 गौणकर्म and मुख्यकर्म	9
		1.2.2 गतिकर्म	10
		1.2.3 वाक्यकर्म(Sentence connector इति)	10
	1.3	करणम्	11
	1.4	सम्प्रदानम्	11

	1.5	• * * * * * * * * * * * * * * * * * * *			
	1.6	अधिकरणम्	11		
		1.6.1 कालाधिकरणम्	12		
		1.6.2 देशाधिकरणम्	12		
		1.6.3 विषयाधिकरणम्	12		
2	कारके	तरसम्बन्धाः	12		
	2.1	भावलक्षणसप्तमी_अनन्तरकालः	12		
	2.2				
	2.3				
	2.4	पूर्वकालः			
	2.5	•			
	2.6	भविष्यत्समानकालः			
	2.7	सहायकिकया			
	2.8	सम्बोध्यः			
	2.9	हेतुः			
	2.10	प्रयोजनम्			
	2.11	कर्तृसमानाधिकरणम्			
	2.12	र् कर्मसमानाधिकरणम्			
	2.13	क्रियाविशेषणम्			
	2.14	प्रतिषेधः			
	2.15	षष्ठीसम्बन्धः			
	2.16	•			
	2.17	वीप्सा			
		विशेषणम्			
		सम्बोधनसूचकम्			
		विभक्तम्			
		अभेदः			
	2.22	ल्यप्कर्माधिकरणम्			
		निर्धारणम्			
	2.24	अत्यन्तसंयोगः	21		
	2.25	अपवर्ग_सम्बन्धः	21		
	2.26	सम्बन्धः			
		2.26.1 सम	22		
		2.26.2 इव	22		
		2.26.3 एव	22		
		2.26.4 अपि	22		
	2.27	वाक्यकर्मचोतकः	23		

3	उपपदसम्बन्धाः (Upapada-Relations)				
	3.1	विभक्त्याधारित-उपपद्विभागः			
	3.2	अर्थाधारि	रेत-उपपद्विभागः	23	
		3.2.1	सन्दर्भविन्दुः (Reference point for direction)	23	
		3.2.2	तुलनाबिन्दुः (Reference point for comparison)	24	
		3.2.3	विषयाधिकरणम् (Locus showing the domain)	24	
		3.2.4	निर्धारणम् (Determination)	24	
		3.2.5	प्रयोजनम् (Purpose)	24	
		3.2.6	उद्गारवाचकः (Exclamatory)	24	
		3.2.7	सहार्थः (Association)	24	
		3.2.8	विनार्थः (Dis-association)	25	
		3.2.9	स्वामी (Possessor-possessee)	25	
		3.2.10	स्रोतः (Source)	25	
4	वाक्ये	तरसम्बन्ध	I: (Inter Sentential Relations)	25	
	4.1	Conne	ctors such as किन्तु, परन्तु, etc	25	
			ctors which occur in pairs	26	
	4.3	समुचित	ः/सुप्_समुचितः (Conjunction)	28	
5	History 2				
A	Tagset of Dependency Relations 3				

0 Background

Manually annotated corpus at various levels has become now an essential resource for analysis of language texts using computers. Such a resource is not only useful for machine learning but also comes handy as a test data for rule-based systems. To extract various kinds of relations between the words in a sentence, it is necessary to have a corpus tagged at the level of word relations. Three natural questions arise while tagging the word relations in a text.

- 1. What is the intended level of semantic tagging?
- 2. Which relation to mark and which to not?
- 3. How to treat function (द्योतक) words?

Here are a few examples of such relations:

0.1 What is the intended level of semantic tagging?

In the sentence, स्थाली पचित what is the relation between स्थाली and पचित? Is स्थाली an अधिकरण (locus) of the action पचित or is it a कर्ता? Taking into account the reality, one would like to mark the relation as अधिकरण. The relation of अधिकरण is a better representation of the अर्थजगत् whereas the relation of कर्ता is faithful to what has been coded by the morphemes, thereby representing the शब्दजगत. Thus there are two distinct levels of tagging. The relation of कर्ता can be marked just by looking at the suffix involved, whereas to mark the relation of अधिकरण, one needs to know the 'पदार्थ'.

We, at this point in time, decide to mark only the information coded by morphemes, and thus confine ourselves to the হাল্বসান, and not to the अर्थजगत.

0.2 Which relation to tag and which to not?

Some relations are marked explicitly (by means of a suffix in Sanskrit), while some are not. For example, in the sentence, रामः दुग्धम् पीत्वा शालाम् गच्छिति the relation between रामः and पीत्वा is not marked explicitly by any suffix. On the other hand, गच्छिति, being in कर्तिरप्रयोग marks the कर्ता. Here राम, which is the कर्ता of the activity of going, thus takes the प्रथमाविभक्ति. In other words, the relation of कर्ता between रामः and गच्छिति is marked in the अभिहितत्व (agreement). But the relation between रामः and पीत्वा is not marked explicitly by any means. पाणिनि provides a special rule 'समानकर्तृकयोः पूर्वकाले', which states that in case one activity precedes the other one and they share the कर्ता, then the preceding activity takes 'त्तवा' suffix. Therefore the knowledge that रामः is the कर्ता of पीत्वा as well, is not marked by any morpheme but is the result of the inference: since पीत्वा has त्तवा suffix and this activity that of the going, the कर्ता of गच्छित and पीत्वा should have been the same. In other words, the knowledge that रामः is कर्ता of पीत्वा is a पार्थिकबोध and hence we do not code this relation directly.

From the two relations viz. the पूर्वकालसम्बन्ध between पीत्वा and गच्छित and the कर्तृसम्बन्ध between रामः गच्छित, appealing to the rule समानकर्तृकयोः पूर्वकाले, one may infer automat-

ically that रामः is also the कर्ता of पीत्वा. Thus, to decide the कर्ता of the verb in त्त्वा, one should know which activity it precedes.

Consider another sentence, बालकः हस्तेन कदलीफलम् छित्वा खादन्तम् वानरम् पश्यित. Here the activity of छिद् precedes both – that of खाद् as well as that of पश्य. So unless one finalises the precedence, one can not decide the कर्ता of छिद्. Therefore, we decide NOT to mark this relation at the syntactico-semantic level of annotation.

0.3 द्योतकपद versus वाचकपद

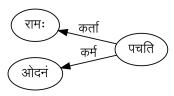
We mark the relations between पदs, where पद is a 'सुप्तिङन्तम् पदम्'. For example, the relation of गच्छित with राम: is that of कर्ता. A पद may be either a वाचक or द्योतक. Consider a sentence रामेण सह सीता वनम् गच्छित. In this sentence, 'सह' marks the relation between रामेण and सीता.

0.4 Convention for marking the relations

We mark the relations by using a directed label arrow. The direction of an arrow decides the name of a relation. For example, the relation between रामः and पचित is called कर्ता and is marked with an arrow from पचित ending onto रामः.

The relation between ओदनं and पचित is called कर्म and is marked with an arrow from पचित ending onto ओदनं. We name the relations using प्रथमान्त words.

From these diagrams, one can get various **মান্দ্ৰ**নীম্ব following different schools, by focussing on appropriate nodes.



For example, starting with रामः covering all nodes, and terminating at the main verb, one gets the वैयाकरण's शाब्दबोध as रामकर्तृक-ओद्नकर्मक-पाकानुकूलव्यापारः. If we traverse the diagram ending in the प्रथमान्त word रामः, we get नैयायिकs शाब्दबोध as ओद्नकर्मक-पाकानुकूल-कृतिमान् रामः. Though one can generate the शाब्दबोध following different schools, the diagram will have one pre-dominant node from where arrows emerge. This node is called a 'root' node and denotes the मुख्यविशेष्य. Since we will be following the वैयाकरण's शाब्दबोध, typically this will be the main verb in the sentence which will be the मुख्यविशेष्य according to वैयाकरणs.

Since we follow वैयाकरणs शाब्दबोध, we expect a तिङन्त to be present. In case तिङन्त is not present in a sentence, we insert an appropriate verb अस्ति/ भवति. This is in tune with पतञ्जलि's statement अस्तिभवन्तीपरः प्रथमपुरुषोऽप्रयुज्यमानोऽप्यस्ति.¹

Since for computational purpose we require a text file and not the graphics images, we propose the following scheme of annotation for text.

¹ महाभाष्यम् (2.3.1)

The words in a sentence are written one word per line and are numbered. The relation of a word with respect to the other is marked by its name followed by a number indicating the word with which it is related. We decided to choose between a word and its number, the number. Because in a sentence, the same word can occur more than once, and hence may lead to an ambiguity. The numbers, on the other hand, are unique. This also means that the head of the sentence (also known as मुख्यविशेष्य), will not have any relation marked against it. Thus, the relations in the sentence रामः ओदनं पचित are represented as:

```
1 रामः कर्ता 3
2 ओदनं कर्म 3
3 पचति
```

0.5 Granularity

The relations proposed by Prof. K V Ramkrishnamacharyulu (2009) are given in the AppendixA. As one can see from the relations, they are very fine-grained, each of the कारक is being sub-divided into many. Though the fine-grained कारक-analysis is necessary for deeper analysis, as well to handle cases of divergences between languages, it also needs a good understanding of व्याकरण on the part of an annotator. We suggest 3-tier tagging as follows:

- 1. Level 1: Coarse grain annotation, as suggested in this draft,
- 2. Level 2: Sharing of relations, and fine-grained annotation (This may be done mechanically),
- 3. Level 3: Semantic level annotation.

Level 1: In the sentence रामः दुग्धम् पीत्वा शालाम् गच्छति, रामः will be marked as कर्ता of गच्छति at the first level of tagging. In the second level of tagging, the machine will mark the relation between रामः and पीत्वा as कर्ता automatically.

Level 2: In the sentence घटः नश्यति, घटः will be marked as कर्ता at the first level. At the second level, घटः by looking at the verb can be marked as अनुभवी-कर्ता mechanically.

Level 3: In the sentence स्थाली पचित। स्थाली will be marked as a कर्ता in the first level. In the third level, one can then further mark स्थाली as an अधिकरण.

In what follows we discuss only the first level of tagging.

0.6 Unit for Tagging

Before we start the discussion on tagging, let us also decide what is the unit for tagging. Since we are now talking about the relations between words, it is natural to think of a **sentence** as a unit. Then, the natural question is how do we define a sentence?

From a computational point of view, we may define a sentence as one that is terminated by either a full stop or a question mark. So here are some examples of sentences:

```
• रामः वनम् गच्छति।
```

• रामः प्रतिदिनं शालां गच्छति।

- किन्तु सः पाठं न पठति।
- रामः खादति पिबति च।
- यदि त्वं इच्छिसि तर्हि अहं भवतःगृहम् आगमिष्यामि।
- यत्र नार्यः तु पूज्यन्ते रमन्ते तत्र देवताः।

Among these, the first three have only one तिङन्तपदम, while the rest have more than one तिङन्तपदs.

0.7 Proposed कारक-tag-set for Sanskrit

The tags may be broadly classified into two types:

- 1. वाक्याभ्यन्तरः (Intra-sentential): Each of the तिङन्तपद्s will have its own आकाङ्का. The relations within the domain of a तिङन्त are termed as intra-sentential relations.
- 2. वाक्येतरः (Inter-sentential): Relations which join the two तिङन्तs or the arguments in the domain of one तिङन्त with those of the other are called inter-sentential relations.

Each of these can be further subclassified looking at the lexical category of the words involved, or the semantics of the relations involved.

वाक्याभ्यन्तरः (Intra-sentential)

- 1. कारक-सम्बन्धः
- 2. कारकेतर-सम्बन्धः
- 3. उपपदसम्बन्धः
- 4. कर्मप्रवचनीयसम्बन्धः ²

वाक्येतरः (Inter-sentential)

- 5. Relations marked by sentence-connecting words.
- 6. Relations marked by relative pronouns.

वाक्याभ्यन्तरः (Intra-sentential)

1 कारकसम्बन्धाः(Kaaraka Relations)

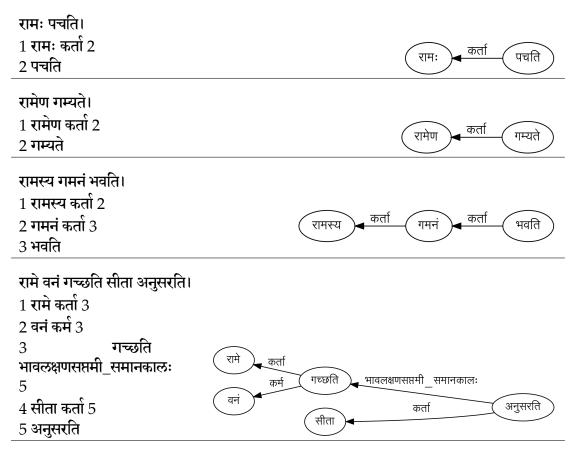
1.1 कर्ता

If the कर्ता is अभिहित, it is in first case otherwise it is in third case.

Not implemented

The कर्ता of the क़दन्त verbs may be in sixth case.

When the verb is in सतिसप्तमी, the कर्ता will have third or seventh case.



1.1.1 प्रयोजककर्ता



1.1.2 प्रयोज्यकर्ता

The प्रयोज्यकर्ता is by default in तृतीयाविभक्ति. In the case of verbs belonging to गतिबुद्धिप्रत्यवसानार्थशब्दकर्मक (1.4.52) the प्रयोज्यकर्ता is in द्वितीयाविभक्ति.

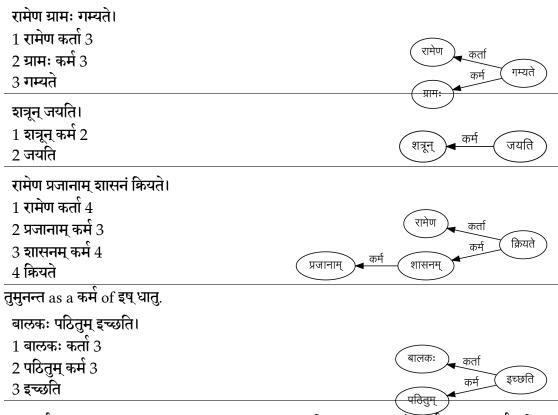




कर्म 1.2

If the कर्म is अभिहित, then it is in प्रथमा, otherwise, it is in द्वितीयाविभक्ति.

The कर्म of a कृदन्त takes षष्टीविभक्ति.



The कर्मसंज्ञा in this sentence is justified by the सन्-विधायकसूत्र "धातोः कर्मणः समानकर्तृकादिच्छायाम् वा" (3.1.7).

गौणकर्म and मुख्यकर्म

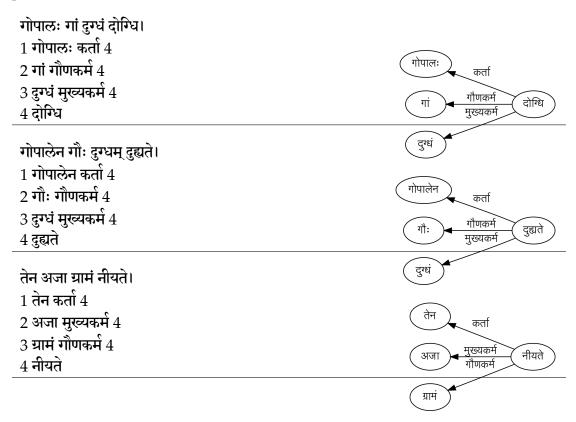
In Sanskrit there are certain verbs mentioned in the following कारिका—

दुह्याच्पच्दण्डूधिप्रच्छिचिब्रूशासुजिमथ्मुषाम्। कर्मयुक् स्यादकथितं तथा स्यान्नीहकुष्वहाम्॥3

which are द्विकर्मकs. Out of these one कर्म is गौण and the other is प्रधान/मुख्य. We decided to mark them as गौणकर्म and मुख्यकर्म, instead of just कर्म. This is to ensure that only

सिद्धान्तकौमुदी, अकथितं च (1.4.51) इति सूत्रव्याख्याने

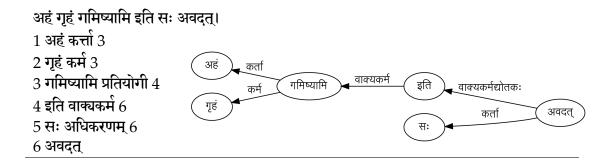
द्विकर्मकs have more than one कर्म, and no other धातुs allow more than one कर्म. Further the information of गौण and मुख्य is also useful in transforming and active sentence to passive.



1.2.2 गतिकर्म



1.2.3 वाक्यकर्म(Sentence connector इति) .

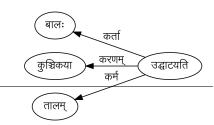


⁴ Not marked in the parser

1.3 करणम्

बालः कुञ्चिकया तालम् उद्घाटयति।

- 1 बालः कर्ता 4
- 2 कुञ्चिकया करणम् 4
- 3 तालम् कर्म 4
- 4 उद्घाटयति



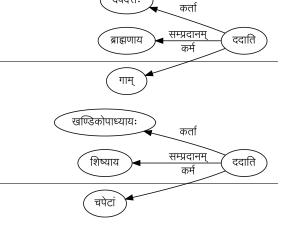
1.4 सम्प्रदानम्

देवदत्तः ब्राह्मणाय गां ददाति।

- 1. देवदत्तः कर्ता 4
- 2. ब्राह्मणाय सम्प्रदानम् 4
- 3. गाम् कर्म 4
- 4. ददाति

खण्डिकोपाध्यायः शिष्याय चपेटां ददाति।

- 1 खण्डिकोपाध्यायः कर्ता 4
- 2 शिष्याय सम्प्रदानम् 4
- 3 चपेटां कर्म 4
- 4 ददाति

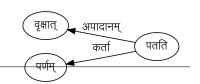


देवदत्तः

1.5 अपादानम्

वृक्षात् पर्णं पतित।

- 1. वृक्षात् अपादानम् 3
- 2. पर्णम् कर्ता 3
- 3. पतति



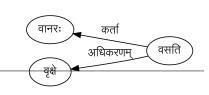
1.6 अधिकरणम्

If the पदार्थ denotes देश , काल or विषय we mark them as देशाधिकरणम्, कालाधिकरणम्, and विषयाधिकरणम् respectively. The default marking is अधिकरणम् .

वानरः वृक्षे वसति। 1 वानरः कर्ता 3

2 वृक्षे अधिकरणम् 3

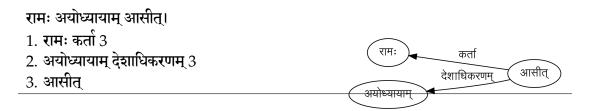
3 वसति



1.6.1 कालाधिकरणम्



1.6.2 देशाधिकरणम्



1.6.3 विषयाधिकरणम्

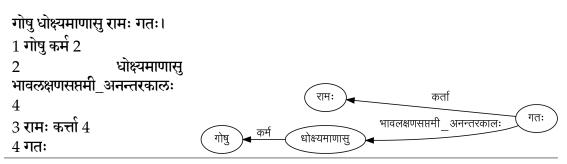


2 कारकेतरसम्बन्धाः

कियाकियासम्बन्धाः

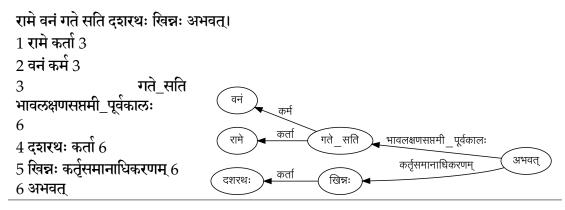
2.1 भावलक्षणसप्तमी अनन्तरकालः

An action denoted by the verb in शत् or शानच् suffix, that comes in the place of छूट्, and is in the seventh case suffix indicates an action which will take place later with respect to another relation. The relation here, therefore, is called भावलक्षणसप्तमी_अनन्तरकालः.



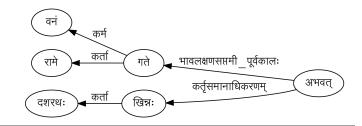
2.2 भावलक्षणसप्तमी_पूर्वकालः

An action denoted by a verb in a क्त with 7th case suffix precedes the action denoted by the main verb. This relation is marked as भावलक्षणसप्तमी_पूर्वकालः.



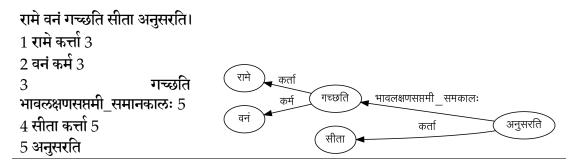
Here the relation of गते_सित with अभवत is marked as भावलक्षणसप्तमी_पूर्वकालः. सित may be absent, as in -

रामे वनं गते दशरथः खिन्नः अभवत्।



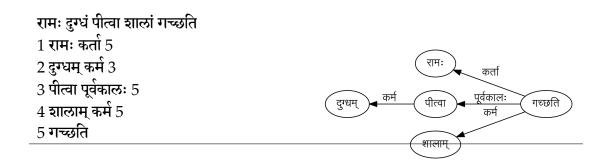
2.3 भावलक्षणसप्तमी समानकालः

An action denoted by a verb in হানু or হানিच with a 7thcase suffix in the place of নৃত্ indicated the simultaneity of the activity with the main activity. This relation is marked by as भावलक्षणसप्तमी_समानकाल:



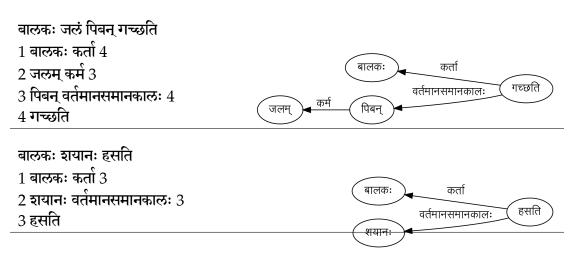
2.4 पूर्वकालः

An action denoted by the verb in कृद्न्त suffix त्त्वा, indicates the precedence relation with respect to another verb. Such relations are marked as पूर्वकालः.



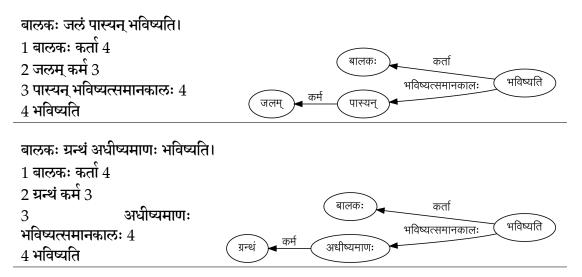
2.5 वर्तमानसमानकालः

An action denoted by the verb in कृद्न्त suffix रातृ/ शानच् in place of लट् and it is related to an action denoted by another verb, the two actions are simultaneous. Hence the relation here is called वर्तमानसमानकालः.



2.6 भविष्यत्समानकालः

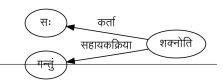
If कृदन्त suffix शतृ/ शानच् is used in place of लृट् then the relation is called भविष्यत्समानकालः.



2.7 सहायकिकया

सः गन्तुं शकोति

- 1. सः कर्ता 2
- 2. गन्तुं सहायकक्रिया 3
- 3. शकोति



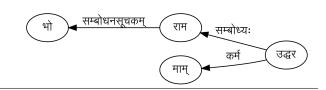
क्रियासम्बन्धाः

2.8 सम्बोध्यः

The relation of a word in सम्बुद्धि is marked as सम्बोध्यः of the corresponding verb. Words such as भो, अपि, हे, अरे etc. are the सम्बोधनसूचकाव्ययs and hence are marked as सम्बोधनसूचकम्.

भो राम माम् उद्धर।

- 1. भो सम्बोधनसूचकम् 2
- 2. राम सम्बोध्यः 4
- 3. माम् कर्म **4**
- 4. उद्धर



2.9 हेतुः

The relation of 'cause' is marked by हेतु:, Sanskrit uses either तृतीया or पश्चमीविभक्ति to mark this relation.

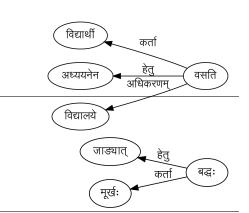
विद्यार्थी अध्ययनेन विद्यालये वसति।

- 1 विद्यार्थी कर्ता 4
- 2 अध्ययनेन हेतुः 4
- 3 विद्यालये अधिकरणम् 4
- **4** वसति.

जाड्यात् मूर्खः बद्धः। 1 जाड्यात् हेतुः 3

2 मुर्खः कर्ता 3

3 बद्धः



In the case of अस्त्यर्थधातुs the relation of हेतुः might be between two nouns as in the following example.

दण्डेन घटः अस्ति।

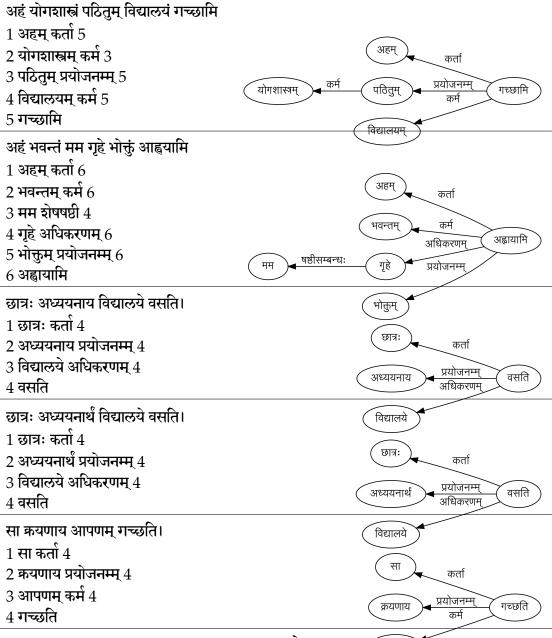
- 1 दण्डेन हेतुः 2
- 2 घटः कर्ता 3
- 3 अस्ति



2.10 प्रयोजनम्

The relation of a तुमुनन्त verb with the main verb is marked as प्रयोजनम्म्,

Sometimes instead of तुमुन्, चतुर्थी विभक्ति or 'अर्थ' is also used with भावार्थकप्रत्यय to indicate the प्रयोजनम्, These are also marked as प्रयोजनम्,



Sometimes the verb with which a verb has प्रयोजनम्सम्बन्ध is क्षांतिस्क (छप्त). In such cases, the कर्म of the elided verb takes a 4th case by the सूत्र "क्रियार्थोपपदस्य च कर्मणि स्थानिनः" (2.3.14). In such cases also we mark the relation as प्रयोजनम् eg.



2.11 कर्तृसमानाधिकरणम्

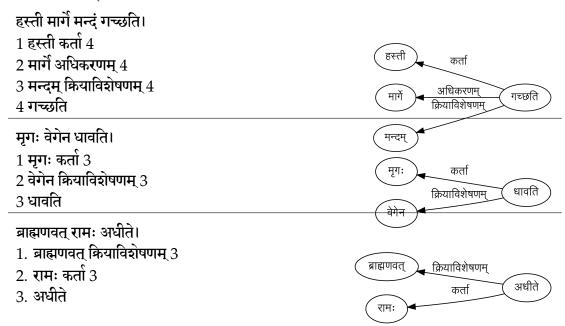
देवदत्तः अध्यापकः अस्ति। 1 देवदत्तः कर्ता 3 देवदत्तः कर्ता 2 अध्यापकः कर्त्तसमानाधिकरणम् 3 अस्ति कर्तृसमानाधिकरणम 3 अस्ति अध्यापकः बालः शयानः अस्ति। 1 बालः कर्त्ता 3 बाल: कर्त्ता 2 शयानः कर्ज्तसमानाधिकरणम् 3 अस्ति कर्तृसमानाधिकरणम् 3 अस्ति

2.12 कर्मसमानाधिकरणम्



2.13 क्रियाविशेषणम्

When a word qualifies either an action or the result of an activity, then it is marked as a क्रियाविशेषणम् .



2.14 प्रतिषेधः

In case of a negation indicating the निषेध, we mark the relation of न with verb as प्रतिषेधः as shown below.



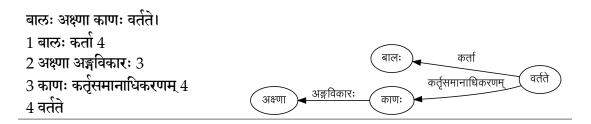
नामपदसम्बन्धाः

2.15 षष्ठीसम्बन्धः

The words with षष्टीविभक्ति which do not indicate कारकविभक्तिs, are marked simply as षष्टीसम्बन्धः.

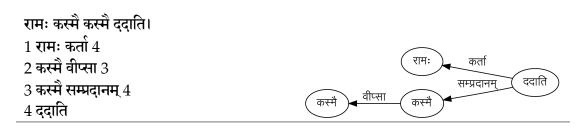


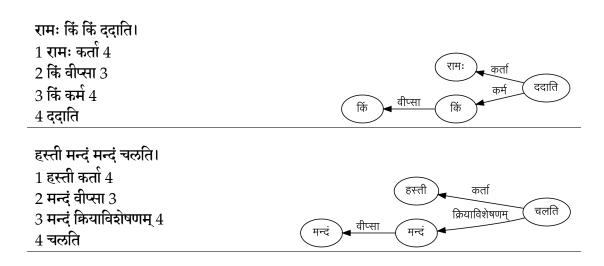
2.16 अङ्गविकारः



2.17 वीप्सा

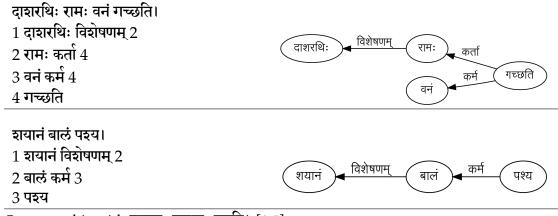
When there is a वीप्सा, the relation of the first word with its repetition as second is marked as वीप्सा and the relation of second word will have a natural कारक/अकारकसम्बन्ध as the case may be. Here are some examples.





2.18 विशेषणम्

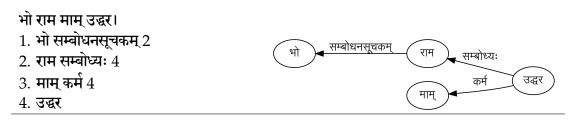
विशेषणs are of two types - those qualifying the उद्देश्य and the other ones which are विधेय. The विशेषणs which qualify the उद्देश्य are called विशेषणs, and the विशेषणs which act as विधेयs will be classified as कर्तृसमानाधिकरणम् or कर्मसमानाधिकरणम्, Here are examples.



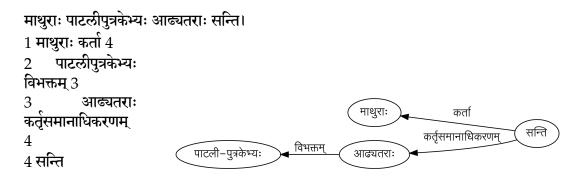
Compare this with बालकः शयानः हसति। [2.5].

2.19 सम्बोधनसूचकम्

The relation of a word in सम्बुद्धि is marked as सम्बोध्यः of the corresponding verb. Words such as भो, अपि, हे, अरे etc. are the सम्बोधनसूचकाव्ययs and hence are marked as सम्बोधनसूचकम्,

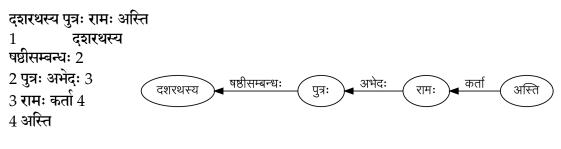


2.20 विभक्तम्

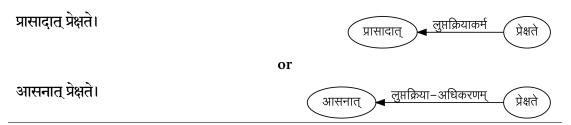


2.21 अभेदः

अभेदः is a special kind of विशेषण. All the उपाधि / relation denoting terms such as आचार्य, पुत्र, पिता, etc., are marked with अभेदसम्बन्ध.



2.22 ल्यप्कर्माधिकरणम्



Here the words प्रासादात or आसनात do not have any direct relation with प्रेक्षते. So in order to have a proper शाब्दबोध of these sentences, it is necessary to supply the missing verb such as आरुह्म or उपविश्य. After supplying this missing किया, sentences change to प्रासादम् आरुह्म प्रेक्षते or आसने उपविश्य प्रेक्षते and then the relation of प्रासादम् with ल्यबन्त word आरुह्म is that of कर्म. Similarly the relation of आसने with उपविश्य is that of ल्यप्-अधिकरणम्. Whenever ल्यप् word is missing in a sentence we mark the relation of आसनात with प्रेक्षते as ल्यप्कर्माधिकरणम् following the वार्तिक ' ल्यब्लोपे कर्मण्यधिकरणं च'.



2.23 निर्धारणम्

रामः क्षत्रियेषु श्रूरः अस्ति।
1. क्षत्रियेषु निर्धारणम् 2
2. श्रूरः कर्तृसमानाधिकरणम् 3
3. रामः कर्ता 4
4. अस्ति

2.24 अत्यन्तसंयोगः

When a कालवाचि and अध्ववाचि takes an accusative case-marker, we mark the relation of such words with the verb as अत्यन्तासंयोगः.



2.25 अपवर्ग_सम्बन्धः

When a कालवाचि and अध्ववाचि takes an instrumental cases, we mark the relation as such words with the verb as अपवर्गः.

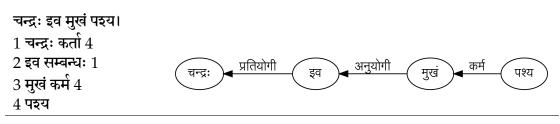


2.26 सम्बन्धः

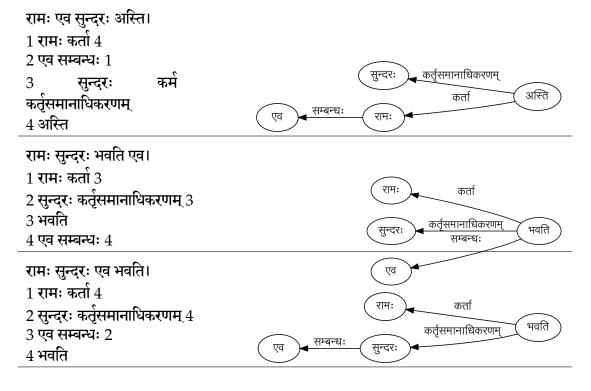
There are certain words such as न, इव, एव, इति etc. whose relations are decided by the meaning of these words. There is no other suffix indicating their relations. For example, the word न marks the negation, the word स्म indicates the past tense, इव indicates the similarity. However, some words such as इति, एव etc. indicate a variety of relations. For example, इति sometimes is used to indicate the शब्दस्वरूप, while sometimes it is used to indicate समाप्ति, sometimes it is used to indicate the कर्म. The word एव sometimes indicate बलाधान (emphasis), sometimes अवधारण. When such words are related to two words, and one of the relations is a कारक /कारकेतर, then the other relation is marked as सम्बन्ध. Otherwise, we mark the relations as प्रतियोगी and अनुयोगी and when they are related to a single word, the relation is marked as सम्बन्धः. The word सम्बन्धः here stands for यत-पदम-तत्-योग्य-सम्बन्धः. That is, if the relation is with एव, then it is एव-योग्य-सम्बन्धः, if it is with न, then it is न-योग्य-सम्बन्धः

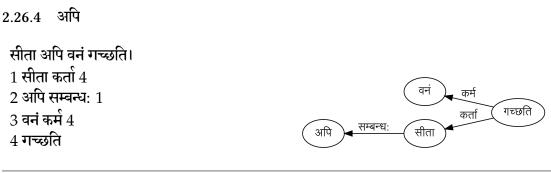
2.26.1 सम





2.26.3 एव





2.27 वाक्यकर्मद्योतकः

अहं गृहं गिमध्यामि इति सः अवदत्।
1 अहं कर्ता 3
2 गृहं कर्म 3
3 गिमध्यामि वाक्यकर्म 4
4 इति वाक्यकर्मधोतकः 6
5 सः अधिकरणम् 6
6 अवदत्

3 उपपदसम्बन्धाः (Upapada-Relations)

उपपदs are related to other nouns or verbs by specific relations such as कारक relation or विशेषण etc.

Upapadas may be classified in two different ways.

- विभक्त्याधारित (Based on their morpho-syntactic properties)
- अर्थाधारित (Based on the semantics they express)

We discuss each of them in detail.

All the उपपद्s demand a specific vibhakti on the preceding noun. But when we look at their relation with other words, we see that they fall under 2 categories.

3.1 विभक्त्याधारित-उपपद्विभागः

There are six types of उपपद्विभक्तिs namely उपपद्दितीया, उपपद्तृतीया, etc.

3.2 अर्थाधारित-उपपद्विभागः

Here are some examples:

3.2.1 सन्दर्भविन्दुः (Reference point for direction)

ग्रामं परितः वृक्षाः सन्ति।

1. ग्रामं सन्दर्भविन्दुः 2

2. परितः अधिकरणम् 4

3. वृक्षाः कर्ता 4

4. सन्ति

3.2.2 तुलनाबिन्दुः (Reference point for comparison)

श्यामेन तुल्यः रामः अस्ति।

- 1. श्यामेन तुलनाबिन्दुः 2
- 2. तुल्यः कर्तृसमानाधिकरणम् 4
- 3. **रामः** कर्ता 4
- 4. अस्ति



3.2.3 विषयाधिकरणम् (Locus showing the domain)

सः पठने लग्नः अस्ति ।

- 1. सः कर्ता 4
- 2. पठने विषयाधिकरणम् 2
- 3. **लग्नः** कर्ता 4
- 4. अस्ति



3.2.4 निर्धारणम् (Determination)

गवां प्रसूता अस्ति ।

- 1. गवां निर्धारणम् 2
- 2. प्रसूता कर्ता 3
- 3. पश्यति



3.2.5 प्रयोजनम् (Purpose)

रामाय कुशलं भूयात्।

- 1. रामाय प्रयोजनम् 4
- 2. कुशलं कर्ता 2
- 3. भूयात्



3.2.6 उद्गारवाचकः (Exclamatory)

दुर्जनं धिक् भवतु ।

- 1. दुर्जनम् उद्गारवाचकः 2
- 2. धिक क्रियाविशेषणम् 3
- 3. भवतु



3.2.7 सहार्थः (Association)

रामेण सह सीता वनं गच्छति।

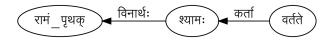
- 1. रामेण सह सहकर्ता 2
- 2. सीता कर्ता 4
- 3. वनं कर्म 4
- गच्छति



3.2.8 विनार्थः (Dis-association)

रामं पृथक् श्यामः वर्तते ।

- 1. रामं_पृथक् विनार्थः 2
- 2. इयामः कर्ता 3
- 3. वर्तते



3.2.9 स्वामी (Possessor-possessee)

सीता गोषु स्वामिनी अस्ति ।

- सीता कर्ता 4
- 2. गोषु स्वामी 2
- 3. स्वामिनी कर्तृसमानाधिकरणम् 4
- 4. अस्ति



3.2.10 स्त्रोतः (Source)

हिमालयात् प्रभृति गङ्गा प्रभवति।

- 1. हिमालयात् स्रोतः 4
- 2. प्रभृति विषयाधिकरणम् 2
- 3. गङ्गा कर्ता 4
- 4. प्रभवति



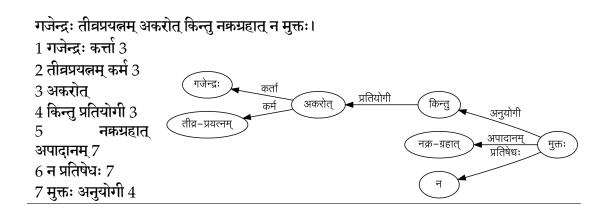
4 वाक्येतरसम्बन्धाः (Inter Sentential Relations)

When a sentence has more than one तिङ्ग्ताs, then the relation between the two বাষ্মs formed by these two तिङ्ग्ताs get established in three ways :

- 1. Connectors such as किन्तु, परन्तु, etc.
- 2. Connectors which occur in pairs.
- 3. Sentence connector इति as वाक्यकर्म.

4.1 Connectors such as किन्तु, परन्तु, etc.

Such connectors join two sentences, which are complete individually. Hence after the first sentence the second sentence start with किन्तु/परन्तुand it connects the two sentences with अनुयोगी and प्रतियोगि relation.

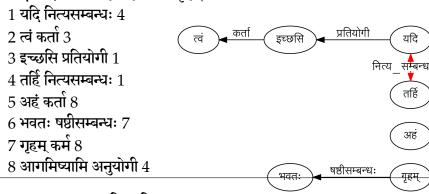


4.2 Connectors which occur in pairs

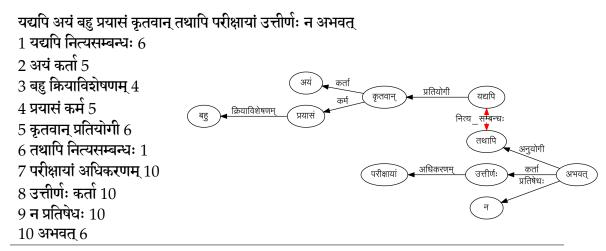
When sentences are connected by pair of connectors such as:

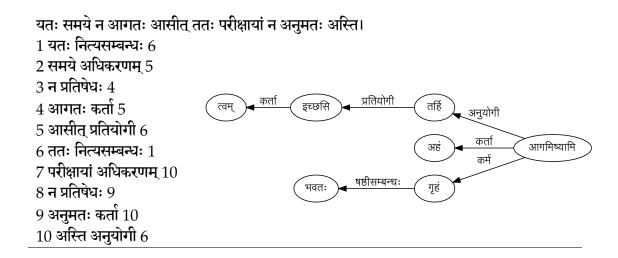
- 1. यदि तर्हि
- 2. यद्यपि तथापि
- 3. यतः ततः/अतः
- 4. यावत् तावत्

here we mark the relations between each of the individual sentences separately, and mark the relations between the main verbs in each of the sentences with यदि and तर्हि, etc. respectively by प्रतियोगी and अनुयोगी and the words यदि-तर्हि etc. are connected with each other by the relation नित्यसम्बन्धः. e.g. यदि त्वं इच्छिस तर्हि अहं भवतः गृहम् आगमिष्यामि।

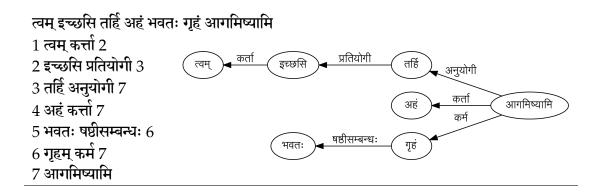


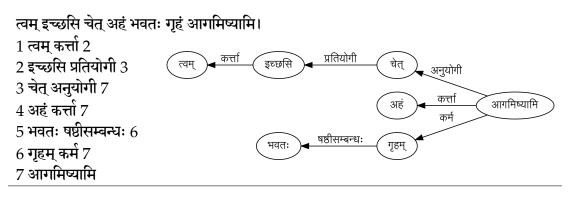
Thus the overall head/ मुख्यविशेष्य of these sentences is आगमिष्यामि.

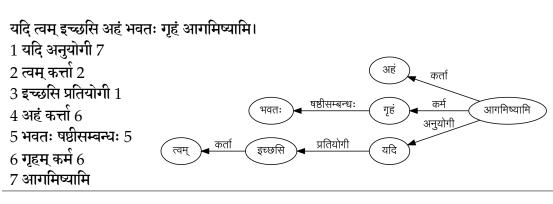


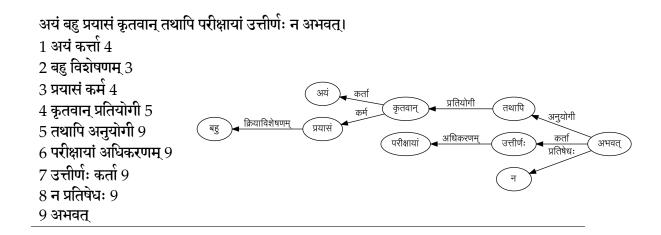


Many a times, one of the two connectives from यदि तर्हि, यद्यपि तथापि , यतः ततः / अतः , यावत् तावत् is absent. In such cases, while annotating the sentence, we provide the missing word in parenthesis as below.



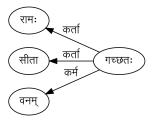






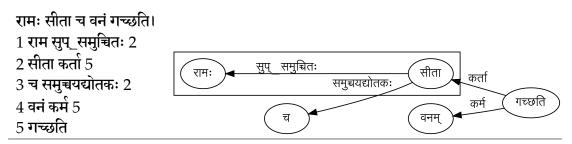
4.3 समुचितः/सुप्_समुचितः (Conjunction)

Consider a sentence रामः सीता च वनम् गच्छतः. In this sentence, both राम and सीता are the कर्ता and वनम् is the कर्म for the धातु गम्. So we may be tempted to mark the relations as



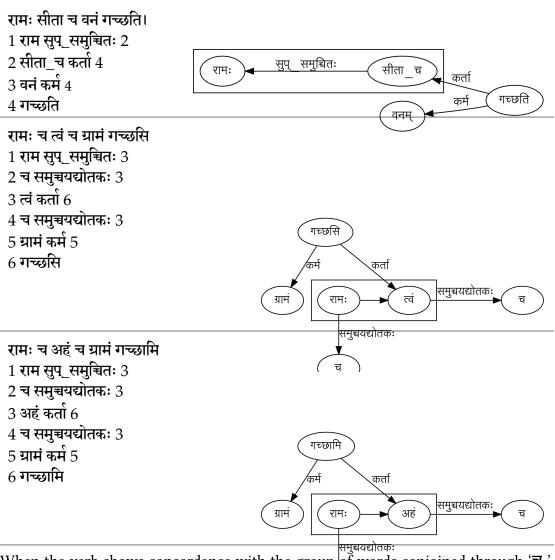
But the कर्तृत्व does not reside in राम and सीता seperately, it resides in both राम and सीता together simultaneously. This is exactly is the meaning of 'च', which indicates समुचय. The कर्तृत्व resides in the समुचय of राम and सीता. We mark the relations as below.

Since 'च' is considered to be a function (द्योतक) word, it just illuminates the meaning and for establishing relation between the words, it is considered to be a meaning to decide the relation between other words. The concordance provides us a clue in the establishing of relations there are two cases. Either the word close to the verb is governed by the verb, or the whole group of words is governed by the verb as one unit. The word which is close the verb is governed by the verb, we mark the relation of the verb with this noun. Consider the following sentence:

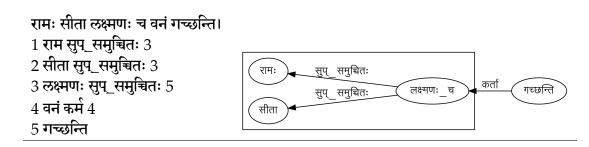


Here the relation between 'च' and सीता is marked as समुचयद्योतकः, and the relation between राम and सीता is marked as सुप्_समृचितः.

When we group the word ₹ with the second conjunct represented as follows.



When the verb shows concordance with the group of words conjoined through 'च ', then the कारकत्व is in the whole group and not in the individual items. Hence in such cases, the relation is marked with the complete group as shown in the annotations of following sentences



5 History

The first tag proposal for kaaraka tagging was prepared by Prof. K V Ramkrishna-macharyulu and was presented in the Third International Sanskrit Computational Linguistics Symposium held at University of Hyderabad, in Jan 2009.

This tagset was compared with the existing tagset of Hindi Tree bank, and a preliminary work of tagging of 100 sentences from सङ्क्षिसरामायण, and the sentences from 15th and 16th sargas of सुन्दरकाण्ड, using this proposed tagset was taken up.

Based on the inputs we received, we had several meetings on kaaraka tagging at Sanskrit Academy and University of Hyderabad. The first meeting was from 24-26th July 2010, the second was from 7-9th Sept 2010, and the third was on 21-22 Oct 2010.

We thank Prof. K V Ramkrishnamacharyulu, who was instrumental in arriving at these guide lines, by providing inputs at various stages of its preparation.

We also thank all the members of the consortium, and especially those who attended the meetings and provided various kinds of inputs by raising questions, providing solutions, participating in the discussions, providing various kinds of feedback on the guidelines, etc.

Following scholars attended one or more meetings on kaaraka tagging-

Prof. K V Ramkrishnamacharyulu	Prof. Veeranarayana Pandurangi		
Prof. S S Murthy	Prof. Tirumala Kulkarni		
Prof. Shrinivas Varkhedi	Prof. Rajadhar Mishra		
Prof. Dipti Mishra Sharma	Prof. Girish Nath Jha		
Prof. Gérard Huet	Prof. Amba Kulkarni		
Dr. Varalakshmi	Dr. Devanand Shukla		
Acharya Ramachandra	Dr. Sheetal Pokar		
Acharya Madhavacharya	Dr. R. Chandrashekhar		
Shri. Pavan Kumar	Shri Anil Gupta		
Ms. Sivaja	Dr. Vibhuti Nath Jha		
Mrs. Preeti Shukla	Ms Monali		
Ms Gayatri	Acharya Deepak		
Shri Madhav Gopal	Shri Nrpendra Pathak		
Shri Jagadish	Shri Lalit		

Our thanks are also due to Prof. Rajeev Sangal, and Prof. Vineet Chaitanya for valuable discussions.

A Tagset of Dependency Relations implemented in the Parser as on July 2019

सम्बन्धाः नामनामसम्बनधाः _{*} षष्ठीसम्बन्धः अङ्गविकारः • कारकसम्बन्धः - कर्ता वीप्सा * विशेषणम् * प्रयोजककर्ता सम्बोधनसूचकम् प्रयोज्यकर्ता * विभक्तम् - कर्म * अभेदः * मुख्यकर्म ल्यप्कर्माधिकरणम् * गौणकर्म निर्धारणम् * वाक्यकर्म अत्यन्तसंयोगः - करणम् अपवर्ग_सम्बन्धः – सम्प्रदानम् वाक्यकर्मद्योतकः – अपादानम् • उपपदसम्बन्धः - अधिकरणम् सन्दर्भबिन्दुः * कालाधिकरणम् – तुलनाबिन्दुः * देशाधिकरणम् - विषयाधिकरणम् * विषयाधिकरणम् - निर्धारणम् • कारकेतरसम्बन्धः – प्रयोजनम् क्रियाक्रियासम्बन्धाः – उद्गारवाचकः * भावलक्षणसप्तमी_अनन्तरकालः – सह-अर्थः * भावलक्षणसप्तमी पूर्वकालः - विना-अर्थः * भावलक्षणसप्तमी_समानकालः – स्वामी * पूर्वकालः वर्तमानसमानकालः - स्रोतः भविष्यत्समानकालः • वाक्येतरसम्बन्धः * सहायकिकया - अनुयोगी - क्रियासम्बन्धाः - प्रतियोगी * सम्बोध्यः - नित्य सम्बन्धः * हेतुः • समुच्चयादिसम्बन्धाः * प्रयोजनम् – समुचितः * कर्तृसमानाधिकरणम् - समुचयद्योतकः * कर्मसमानाधिकरणम्

* क्रियाविशेषणम्

* प्रतिषेधः

– अन्यतरः

अन्यतरद्योतकः