Itaretara Dvandva: A challenge for Dependency Tree semantics

Amba Kulkarni and Vasudha Neelamana
Department of Sanskrit Studies
University of Hyderabad
Hyderabad
ambakulkarni@uohyd.ac.in

Abstract

The *itaretara dvandva* compounds in Sanskrit exhibit two different senses - a conjunctive and a disjunctive. This calls for a splitting a sentence into multiple sentences by distributing the components of such a compound. The use of nested compounds with *itaretara dvandva* as components of bigger compounds further complicate the matter, since the constituency structure of such compounds do not capture the distributive sense of the *dvandva* compounds. In this paper we illustrate the difficulties a reader faces while understanding such compounds, with specific examples from the preliminary text of Āyurveda - Aṣṭāṅgahṛdayam.

1 Introduction

Aṣṭāṅgahṛdayam (AH) is one of the important Indian treatises that deals with Āyurveda. This text authored by Vāgbhaṭa addresses eight branches of Āyurveda viz *kāya* (general medicine), *bāla* (child and woman care), *graha* (idiopathic diseases), *ūrdhvāṅga* (ENT and dental), *śalya* (surgery), *damṣṭra* (toxicology and forensic sciences), *jarā* (geriatrics) and *vṛṣa* (aphrodisiacs). It consists of 120 chapters divided into six sections. This being one of the foundation level text, around first 15 chapters of this text are included in the first year of Bachelor's course on Āyurveda for study. Other chapters, being specialised branches, are included in the syllabus in later years.

With the current emphasis of the National Council for the Indian System of Medicines (NCISM) encouraging students to read original Sanskrit texts in Āyurveda, we planned to develop an e-reader semi-automatically for the first few chapters of AH, using Saṁsādhanī.[1] The e-reader provides us with the following information to the user: The original *śloka*, its sandhi split version, the morphological analysis of each word, also the compound analysis showing the constituency structure and the type of a compound, the dependency graph providing the sentential structure exhibiting the *kāraka* relations among the words, the prose order of the verse and the dictionary meanings of each word.

Students who join a bachelor's course on Āyurveda typically have just a preliminary knowledge of Sanskrit and in some cases students have all their school education through English medium and thus hardly have any exposure to Sanskrit language let alone Sanskrit grammar! In such a scenario, the e-readers come as an aid to the teachers who can concentrate more on teaching students how to 'understand' a text in origin, with the help of the grammatical information provided by the Saṁsādhanī platform.

In this paper, we show how some constructions with *itaretara dvandva* pose problems in understanding the semantics. In the next section, the convention for drawing a dependency tree of a sentence/verse is explained, followed by the description of the associated semantics.

[1] https://sanskrit.uohyd.ac.in/scl

This is followed by a short discussion on various semantics associated with the *itaretara dvandva* compound. The third section discusses six constructions involving *itaretara dvandva* and the challenges they throw while representing the semantics through the dependency trees and the solutions thereof. This is followed by a conclusion summerising the observations.

2 Dependency Tree and Associated Semantics

Saṁsādhanī platform provides e-readers for various texts. These e-readers are built semi-automatically using the existing tools on the platform. The steps followed for generation of the e-readers are as follows.

1. *Padaccheda*: First the given verse is sandhi-split using the Heritage segmenter[2] augmented by the statistical ranking module (Krishnan et al., 2024) which provides a sandhi split version of the text. Machine marks the sandhi split between the compound components with a '-', while the split between the words (*padas*) with a space. This is further manually verified by a human and corrected if necessary.

2. *Vākyaccheda:* In this step, a human being reads the input sentence/verse, and decides to split it into multiple units termed sentences, following the Kātyāyana's definition - *eka-tiṅ-vākyam*, if necessary. Since the dependency parser[3] is developed following the grammarian's theory of verbal cognition, every sentence should have a verb in it in order to get a parse. If the verse does not have a finite verb, then a finite verb such as *asti / bhavati / vartate* etc. is supplied manually. After splitting, if needed, words are borrowed from the previous part.

3. Each of such sentences is then fed separately to the Anusāraka engine of Saṁsādhanī. This module produces complete analysis of the input text, by providing all the possible morphological analysis of each word, and chooses the appropriate morphological analysis in the given context, by carrying out the sentence level analysis. The relations between the words are proposed, and again, based on the heuristics and ranking algorithm, appropriate kāraka/non-kāraka relation between the words is chosen, which is represented as a dependency tree. Kulkarni (2019) and Kulkarni (2021) describe the complete algorithm for this step.

4. The parsed output produced in the previous step might not be perfect. At this stage, human being chooses the correct relation from among the possible relations, and makes sure that the parsed output is correct, and is faithful to the meaning as described in the commentaries.

5. The solutions thus approved by the human beings are then presented to the readers in a format as shown in Figures 1, 2 and 3.

Figures 1, 2 and 3 give a snapshot of a page of the e-reader for the first *gadya-sūtra* of AH. Figure 1 shows the original *gadya-sūtra* at the top. Figure 2 shows its analysis in a table. Figure 3 shows the dependency tree that shows the relations between the words in the input text. The words are linked to the four dictionaries Sanskrit-Hindi (Apte 1890), Sanskrit-English (Monnier-Wlilliam 1899), Sanskrit-French (from Sanskrit Heritage platform of Gérard Huet, 2002-2025) and Cappeller's Sanskrit German (1887) dictionary.

This dependency tree displays semantic information that can be extracted from a sentence using the constraints of *Ākāṅkṣā* (expectancy), *Yogyatā* (mutual congruity) and *Sannidhiḥ* (proximity). The term 'semantics' is understood differently in different contexts. For the purpose of this paper, we define a semantic representation as one that reflects the meaning

[2]https://sanskrit.inria.fr; also avaialble at https://sanskrit.uohyd.ac.in/scl → sandhi splitter
[3]available at https://sanskrit.uohyd.ac.in/scl

अष्टाङ्गहृदयम् 01.001_1

अथात आयुष्कामीयमध्यायं व्याख्यास्यामः।

Figure 1: input gadya-sūtra

Index	Word	Morph In Context	Kaaraka Relation
1.1	अथ	अथ{अव्य}	कालाधिकरणम्,5.1
2.1	अतः	अतः{अव्य}	सम्बन्धः,5.1
3.1	आयुष्कामीयम्	आयुष्कामीय{पुं;2;एक}	विशेषणम्,4.1
4.1	अध्यायम्	अध्याय{पुं;2;एक(अधि_इ4;घञ्;अदादिः)}	कर्म,5.1
5.1	व्याख्यास्यामः	वि_आङ्_ख्या1{कर्तरि;लृट्;उ;बहु;परस्मैपदी;अदादिः}	-

Figure 2: grammatical analysis

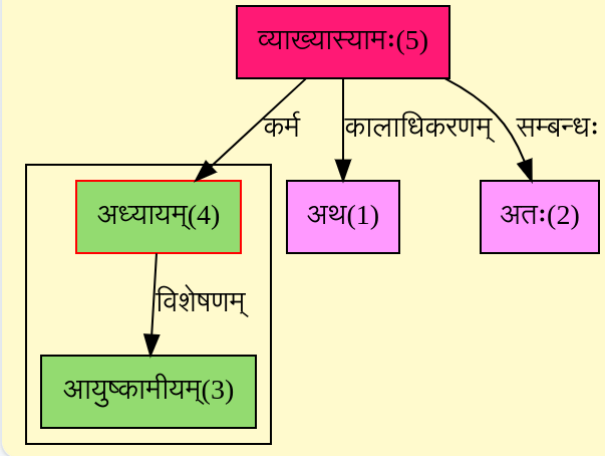

Figure 3: Dependency tree

of the text as it is understood by a language speaker(Abend and Rappoport, 2017). The fundamental component of semantic representation of a text is the argument structure - who did what to whom, where, when, why, how, etc.. Pāṇini's *kāraka* theory provides the basic semantics of a sentence. The dependency parse tree produced by Saṁsādhanī marks these *kāraka* relations and also several non-*kāraka* relations. The tagging guidelines[4] provide a complete list of all tags that are currently being used.

The semantics associated with various syntactico-semantic relations (*kāraka* as well as non-kāraka) is provided by Pāṇini in Aṣṭādhyāyī through their definitions. These definitions provide the semantics associated with the labels on the edges of the tree. However, the semantics associated with the nodes which represent the concepts is not marked in a dependency tree. In this sense, a dependency tree does not completely represent the semantics associated with a sentence. Nevertheless, assuming that the reader deciphers the meaning associated with the words and the fact that a dependency tree provides an argument structure associated with the sentence, a reader can understand the meaning of the sentence provided s/he can decipher the concepts associated with the words.

A few lines about the conventions followed in drawing the dependency tree are in order. The relations are between the meanings associated with the stems, and not the *padas*. The suffixes associated with the stems are the indicators of various relations. In the dependency trees, we mark the relations between the *padas* and not between the stems - nominal (*prātipadikas*) or verbal (*dhātus*) roots, because these diagrams are meant for the readers to understand the original text. If the node labels correspond to the stems, a reader without good grammar knowledge may face difficulties in linking the stems with the word forms in the given text. The edges are directed, with the head of an arrow pointing towards the node having the role denoted by the label on the edge. Thus an edge labeled *karma* of an activity points to a node which is the *karma* of that activity.

Thus from the dependency tree such as Figure 3, one can get semantics of the input text as *āyuṣkāmīyam* is an adjective of *adhyāyam*, which is the *karma* (goal) of the activity of elaboration, the indeclinable *atha* attached to the verb by *kālādhikaraṇam* (location of time) states that the activity will begin now. Finally, the other indeclinable *ataḥ* connected to the verb is a discourse element that marks a relation of this sentence to the previous one, indicating that the activity of elaboration is the result (of some curiosity). Since a dependency tree shows the relations of words within a single sentence, this relation is marked simply by a generic word *sambandhaḥ* (relation). The table in Figure 2 shows the morphological analysis of each word, followed by the dependency analysis. The colors indicate the Part of Speech(POS) category viz. an indeclinable, a noun and a verb. Within nouns, different colors are used to indicate different case terminations (*vibhakti*s). Thus a student who has 'understood' the grammar but not memorised the word forms etc. still with the help of the analysis table shown in Figure 2 can understand the meaning of the input text. But the word meanings are not marked in the dependency tree, for which s/he can rely on the linked dictionaries. This is very close to the *śābdabodha*, the understanding, one gets after hearing a sentence.

While to a large extent these dependency trees help in 'understanding' the original text, during the development of an e-reader for the AH, we came across some constructions a) that use *dvandva* (copulative) compounds, especially the *iteratara dvandva*, and b) the words *kramāt* or *krameṇa*, similar to 'respectively readings' (Chaves, 2012) in English, that pose problems in faithful syntactic representation providing the desired semantics.

[4] `https://sanskrit.uohyd.ac.in/scl/GOLD_DATA/Tagging_Guidelines/`

2.1 Semantics of *Dvandva* compound

Dvandva, according to Pāṇini, is employed in the sense of '*ca*' (and).[5] This *ca*, as Joshi and Roodbergen (1997, p70) note,

> ... may mean either the two (or more) components are jointly (or simultaneously) involved in an action or that they are involved each independently (or separately) of each other. In the former case '*ca*' takes a conjunctive sense and in the latter case it takes a disjunctive sense.

As illustrated by them further, in the sentence, *rāmalakṣmaṇau gacchataḥ*, both *rāma* and *lakṣmaṇa* go together and not independently. For the latter usage, they provide an example from the Aṣṭādhyāyī. In the *sūtra saptamīviśeśaṇe bahuvrīhau* (A 2.2.35), the *dvandva* compound *saptamīviśeśaṇe* provides the disjunctive sense viz. either a component ending in the seventh case termination or a component functioning as a qualifier, in the case of *bahuvrīhi* compound is placed at the beginning of a compound. We notice one more usage of *dvandva* in the Aṣṭādhyāyī, where the *sūtra yathā saṅkhyam anudeśaḥ samānam* (A 1.3.10) governs. In the *sūtra eco'yavāyāvaḥ* (A 6.1.78), the four letters denoted by the *pratyāhāra e**C*** viz. *e,o,ai,and au* change to *ay*, *av*, *āy* and *āv* respectively, if followed by a vowel, in close proximity. Here the *dvandva* compound *ayavāyāvaḥ* is not only disjunctive, but there is also a sense of respectively. We notice the same disjunctive usage and the sense of 'respectively' in the instances of *dvandva* found in AH.

3 Challenging Syntactic Structure in AH

AH deals with Āyurveda where all the discussions revolve around the three *doṣas* (humors) viz. *vāta*, *pitta* and *kapha*. This naturally results in the use of a *dvandva*. We also see a very prominent use of *krameṇa* or *kramāt* either explicitly or implicitly (through *adhyāhāra/anuvṛtti*) throughout this text. We look at a few constructions we came across in AH with these two features and discuss the problems in representing their semantics compactly without deviating from the syntax.

A) Consider the following hemistich from AH.

> [1] Skt: *taiḥ bhavet viṣamaḥ tīkṣṇaḥ mandaḥ ca agniḥ samaiḥ samaḥ* (AH.Su.1.8.2)
> Eng Tr: These three *doṣas* result in three types of digestive fires, viz. *viṣama* (unsteady or erratic), *tīkṣṇa* (increased) and *manda* (decreased). When the three *doṣas* are balanced in the body, the digestive fire is also balanced.

For proper interpretation, we need to borrow the word *kramāt* from the previous part of the *śloka*. Further this part of the verse consists of two sentences[6] viz.

> [2] Skt: *taiḥ bhavet viṣamaḥ tīkṣṇaḥ mandaḥ ca agniḥ (kramāt).*
> [3] Skt: *samaiḥ samaḥ (agniḥ bhavet).*

The first part borrows the word *kramāt* from the previous *śloka* and the second part borrows *agniḥ bhavet* from the first part. The dependency tree corresponding to the first part is shown in Figure 4. In this sentence, the pronoun *taiḥ* refers to the three *doṣas* viz *vāta*, *pitta* and *kapha*. The borrowed word *kramāt* is enclosed in parentheses to indicate that this is not part of the original verse. *Mandaḥ*, conjoined with *tīkṣṇaḥ* and *viṣamaḥ* is the predicative adjective (*vidheya viśeṣaṇa*) of *agniḥ*.[7] Therefore, the above sentence is semantically equivalent to a set of three sentences, viz.

[5] *cārthe dvandvaḥ (A 2.2.29)*

[6] We follow the definition of sentence as *ekatiṅvākyam* as given by Kātyāyana.

[7] Here all the three viz. *mandaḥ*, *tīkṣṇaḥ* and *viṣamaḥ* together are the predicative adjectives, and hence they are stored in a box. The arrow labeled *vidheyaviśeṣaṇa* is pointing towards *mandaḥ*. For more details regarding the representation of *ca* in the dependency parser, refer to Panchal and Kulkarni (2019)

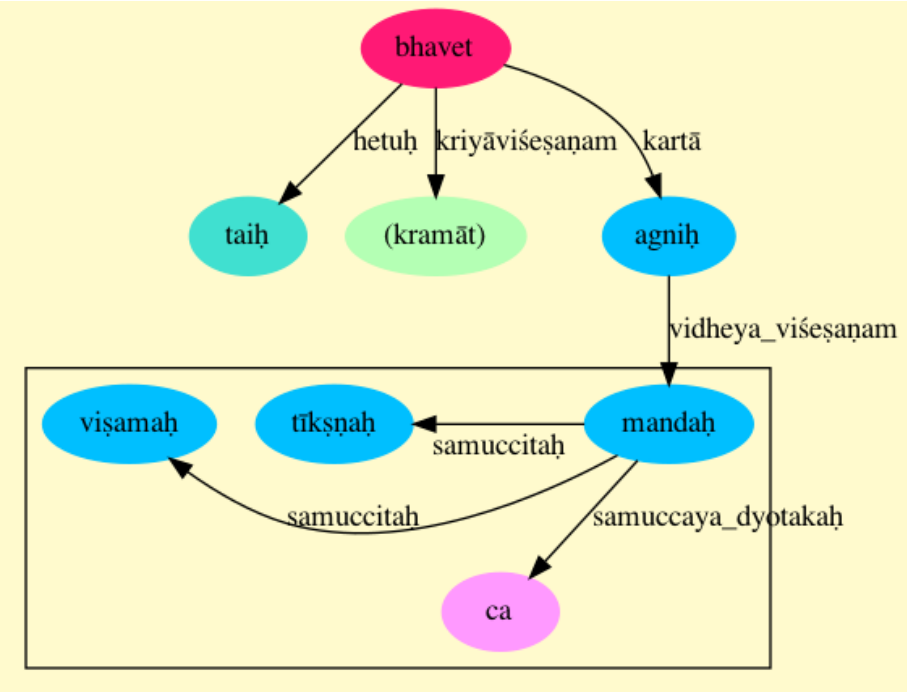

Figure 4: Dependency tree for Sentence [2]

[4] Skt: *vātena agniḥ viṣamaḥ bhavet.*
[5] Skt: *pittena agniḥ tīkṣṇaḥ bhavet.*
[6] Skt: *kaphena agniḥ mandaḥ bhavet.*

The dependency trees corresponding to these three sentences are shown in Figure 5 which together are semantically equivalent to the dependency trees in Figure 4. Here, the presence

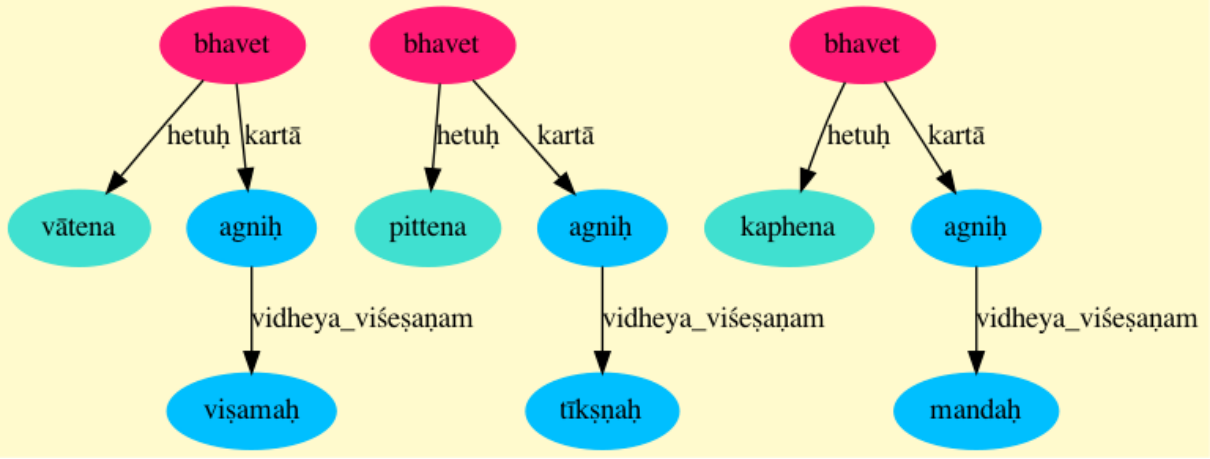

Figure 5: Dependency tree for sentence [2] after redistribution

of three different words, viz *viṣamaḥ*, *tīkṣṇaḥ* and *mandaḥ* facilitated the division of the verse into three separate sentences with repetition of the verb *bhavet* along with *agni*. Note that this information cannot be compactly represented in a single dependency tree, since there is no direct relation between the components of a compound and the three predicative adjectives. Secondly, for the development of any reasoning or question answering system following the knowledge base approach, we need this information explicitly marked. From

the Machine Translation point of view, however, this does not pose a problem, since such constructions typically go across the languages.

B) Next we see the tenth verse from AH where the predicative adjective is a *dvandva* compound. Since *dvandva* compound is a *nityasamāsa*,[8] splitting such a sentence becomes impossible without rewriting it.

[7] Skt: *taiḥ ca tisraḥ prakṛtayaḥ hīna-madhya-uttamāḥ pṛthak.* (AH.Su.1.10.1)
Eng Tr: Due to the dominance of a single *doṣa* body's constitution is classified into three types *vāta prakṛti*, *pitta prakṛti*, and *kapha prakṛti*, which are *hīna* (poor), *madhya* (moderate), and *uttama* (good), respectively.

Since there is no verb in this part, we supply a verb *bhavanti*. The dependency tree for sentence [7] is as shown in Figure 6.

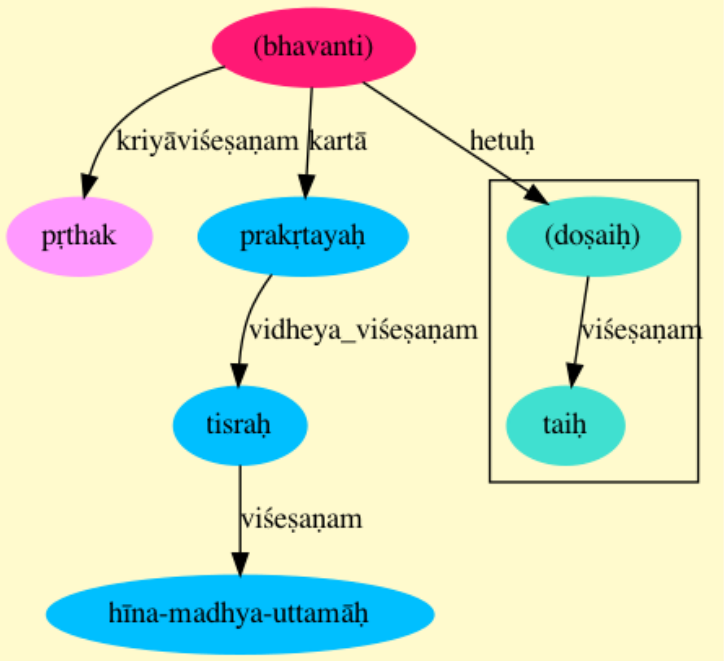

Figure 6: Use of a compound

The pronoun '*taiḥ*' as in the previous case refers to *doṣaiḥ*. The adverb *pṛthak* is responsible for expressing the three different *prakṛtis* resulting from the three *doṣas*. Underlying semantics is collective semantics of the following three sentences viz.

[8] Skt: *vātena prakṛtiḥ hīnā bhavati.*
[9] Skt: *pittena prakṛtiḥ madhyā bhavati.*
[10] Skt: *kaphena prakṛtiḥ uttamā bhavati.*

Since *hīna-madhya-uttamāḥ* is one word, we need to distribute its three components viz. *hīna*, *madhya* and *uttama* over three *doṣas*, the pronoun *tat* refers to. Corresponding three dependency trees would be as shown in Figure 7. As in the previous case, here also we do not see any direct relation between the components of the compounds with three *doṣas*, and thus it is impossible to represent this information compactly in a single dependency graph that can represent the desired semantics. As in the previous case, here also such

[8] *Nitya-samāsa* does not have a *sva-pada-vigraha-vākya*, a meaning-paraphrase.

constructions may not pose problems from Machine Translation point of view, since most of the frquently used languages have such constructions.

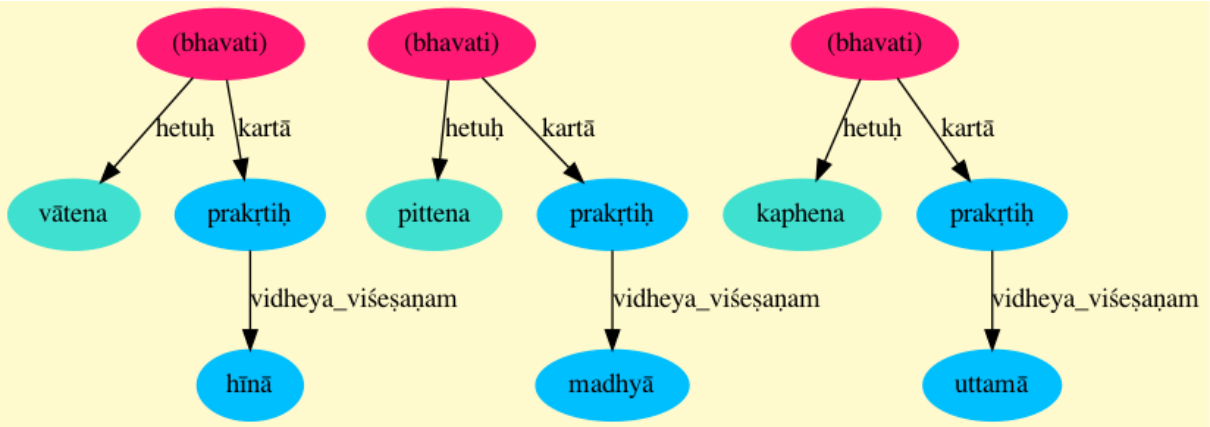

Figure 7: distribution of compound components

C) These cases were easier. Let us now look at some more problematic cases. Consider the verse

[11] Skt: *vayohorātribhuktānāṃ te'ntamadhyādigāḥ kramāt* (AH.Su.1.8.1)

[12] Segmented: *vayaḥ-ahaḥ-rātri-bhuktānāṃ te anta-madhya-ādi-gāḥ kramāt bhavanti.*
Eng Tr: *Vāta*, *pitta* and *kapha* are predominantly present in the last, middle and first stages of age, day, night and digestion respectively. In other words, *vāta* is predominantly present in the last stage of the life (old age), the last stage of the day (evening hours), the last stage of the night (ending hours of the night) and the last stage of the digestion (end of digestion). *Pitta* is predominantly present in the middle stage of the life (middle age), the middle stage of the day (midday), the middle stage of the night (midnight) and the middle stage of the digestion (during the process of digestion). Similarly *kapha* is predominant in the first stage of the life (childhood), the first stage of the day (morning hours), the first stage of the night (starting of night hours), and the first stage of the digestion (beginning of the digestion).

Here, the pronoun *te* refers to *doṣas* which are three in number. The other two nouns are compounds, with three and four components each, with one compound having a genitive relation with the other one. Thus, the dependency tree for this *śloka* is as shown in Figure 8.

From this dependency tree, we understand that, *te*, i.e. the *doṣas*, respectively (*kramāt*) are present at the end, in the middle or in the beginning. Here we have taken the distributive meaning, due to the presence of the word *kramāt*. There is an expectancy: end of what, middle of what and beginning of what. This expectancy is fulfilled by the compound with genitive case termination - *vayaḥ-ahaḥ-rātri-bhuktānām* (of the age, of the day, of the night and of the digestion process). The three positions viz. the beginning, middle, and the end are to be distributed over time-slots of the life, day, night, and the duration of the digestion process.

The meaning then is *vāta*, *pitta*, and *kapha* dominate the end, the middle and the beginning part of one's life, of the day, of the night and of the digestion process. There is a disjoint reading of *te* with the three components of the compound *anta-madha-ādi* respectively. The components of the compound *vayaḥ-ahaḥ-rātri-bhuktānām* get attached to each of the component of the compound *ādi-madhya-anta-gāḥ* resulting into 3*4=12

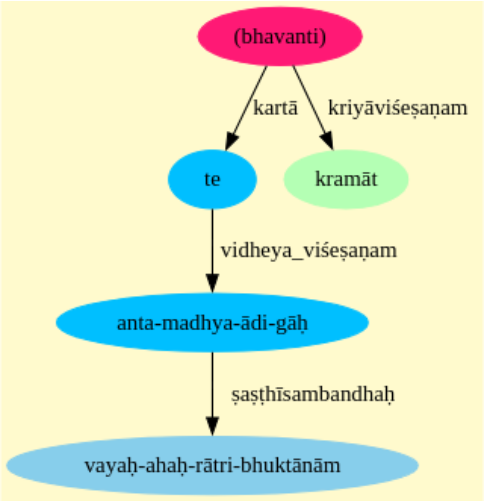

Figure 8: Dependency tree for AH 8.1

combinations. The dependency tree in Figure 8 may be redrawn, by replacing *te* with *vāta-pitta-kapha*, as in Figure 9, explicitly marking the relation between the components of various compounds. However, as we notice, the graph becomes very clumsy from a readers point of view. Also, the role of the component *gāḥ* in this representation is not clear.

The question one may ponder upon now is - Is this explosion into various possibilities by multiplication due to different numbers of components in the two compounds, or due to the genitive marker, or due to the meanings involved, or due to the extra-linguistic context? Discussion pertaining to these questions is out of scope of this paper.

As discussed earlier, here also from the reader's perspective, it may be better to represent these as twelve different trees. This also facilitates the Questions-Answering system. Further from the machine translation point of view, not all languages, such as English, allow the distribution of the components of a compound in genitive with the components of a compound with which the genitive compound is related to. This is obvious from the English translation of the verse.

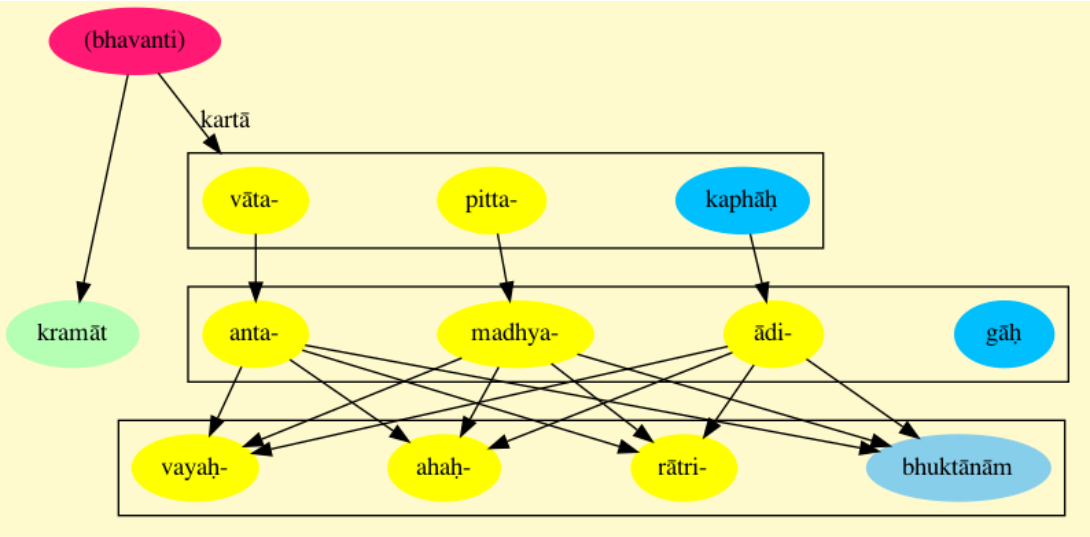

Figure 9: Replacing the pronoun by its referent

D) Consider the following part of the seventh *śloka* from AH.

[13] Skt: *vikṛtāvikṛtā dehaṃ ghnanti te vartayanti ca.* (AH.Su.1.7.1)
Segmented: *vikṛta-a-vikṛtāḥ dehaṃ ghnanti te vartayanti ca.*
Eng Tr: These three humors cause diseases in the vitiated state and keep the body in healthy condition when they are in the equilibrium state.

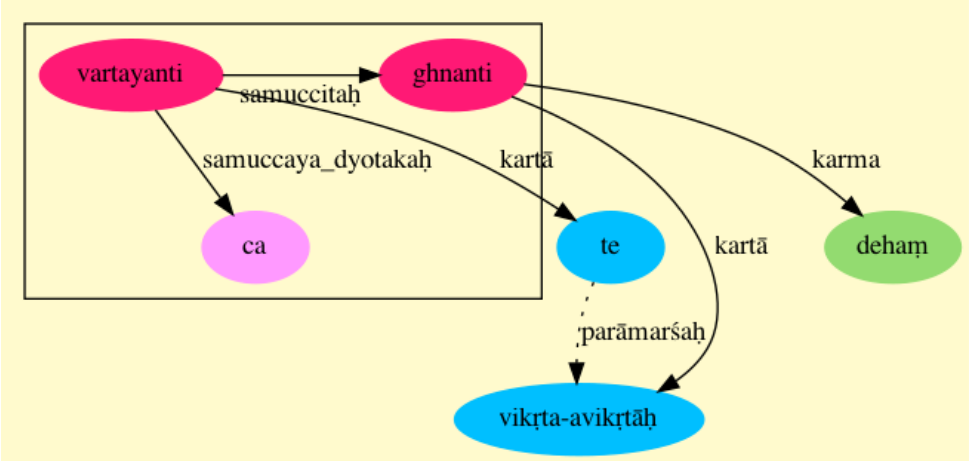

Figure 10: Dependency tree for Sentence (14)

The dependency tree for this sentence is shown in Figure 10. In this tree *vikṛtāvikṛtāḥ* is marked as a *kartā* of *ghnanti* and the pronoun *'te'* which is *kartā* for *vartayanti* refers to *vikṛtāvikṛtāḥ*. Hence from the graph one gets the interpretation that both the *vikṛta* and *avikṛta doṣas* kill the body and also reside in the body.

The commentaries, on the other hand, state that the two components in this *dvandva* compound are distributed over the two verbs. The dependency tree in Figure 11 represents the interpretation in the commentary.

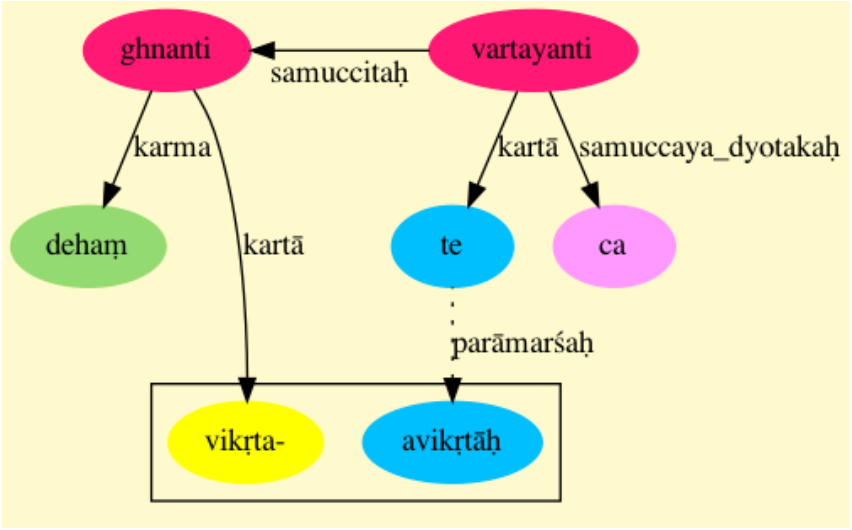

Figure 11: Dependency Tree showing associated semantics

In this example, in addition to having a distributive use of the components of *dvandva*, there is an additional anomaly that the pronoun *te* refers to only a part of the compound (*eka-deśa-parāmarśa*) as shown in Figure 11!

E) Consider the following *śloka*

[14] Skt: *kālārthakarmanāṃ yogo hīnamithyātimātrakaḥ*
samyagyogaśca vijñeyo rogārogyaikakāraṇam (AH.Su.1.19)
Segmented: *kāla-artha-karmanāṃ yogaḥ hīna-mithyā-atimātrakaḥ*
samyak-yogaḥ ca vijñeyaḥ roga-ārogya-eka-kāraṇam
Eng Tr: Less, more, or wrong unison of time, senses, and functions is the reason for the disease, and right unison of these three factors is the reason for the healthy state.

Here the compound *roga-ārogya-eka-kāraṇam* has two components which themselves are compounds viz. *roga-ārogya* and *eka-kāraṇam*. (See Figure 12). The constituency analysis for the compound is ((roga-ārogya)-(eka-kāraṇam)) and in the sentential analysis, the components of sub-ordinate compound, *roga* and *ārogya*, get distributed over the second constituent, *eka-kāraṇam*, of the matrix compound, resulting into *rāga-eka-kāraṇam* and *ārogya-eka-kāraṇam* respectively.

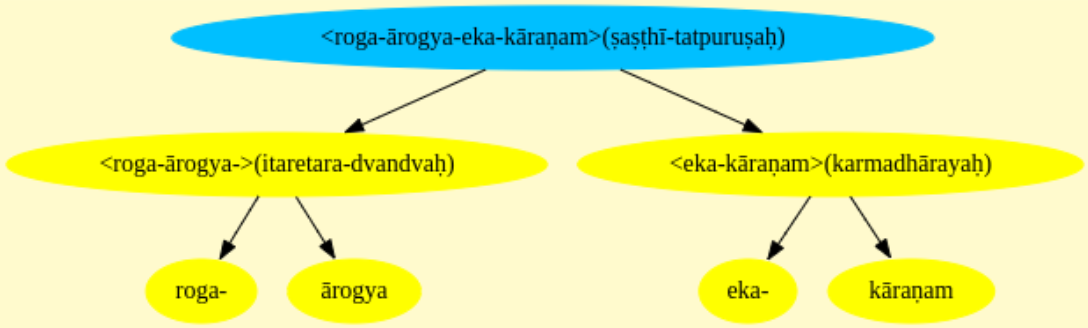

Figure 12: compound structure

If we look at the meaning of this verse, we note that two reasons are being described - one for the diseased state (*roga-eka-kāraṇam*) and the other one for the healthy state (*ārogya-eka-kāraṇam*). In the constituency tree, we do not have nodes corresponding to *roga-eka-kāraṇam* and *ārogya-eka-kāraṇam* as in the previous example. Hence in this case we need to split the sentences into two as below.

[15] Segmented: *kāla-artha-karmanāṃ yogaḥ hīna-mithyā-atimātrakaḥ vijñeyaḥ roga-eka-kāraṇam*

[16] Segmeted: *samyak-yogaḥ ca vijñeyaḥ ārogya-eka-kāraṇam*

The first part has a compound that is modified by another compound with genitive marker, each of them having three components, thus amounting to nine combinations!

The dependency tree for this *śloka* is represented in Figure 13.

This is a case, where the constituency structure does not help in understanding the underlying meaning of the compound! This necessitates the splitting of the input text into separate sentences. It is also obvious that modern Indian languages and English do not allow such constructions, which is again a challenge for Machine Translation.

F) Here is the final example from the AH.

[17] Skt: *saṃsargaḥ sannipātaśca taddvitrikṣayakopataḥ.* (AH.Su.1.12)

[18] Segmented: *saṃsargaḥ sannipātaḥ ca tat-dvi-tri-kṣaya-kopataḥ.*
Eng Tr: The vitiation of any of the two humors is called *saṃsarga* and the vitiation of all three humors is called *sannipāta.*

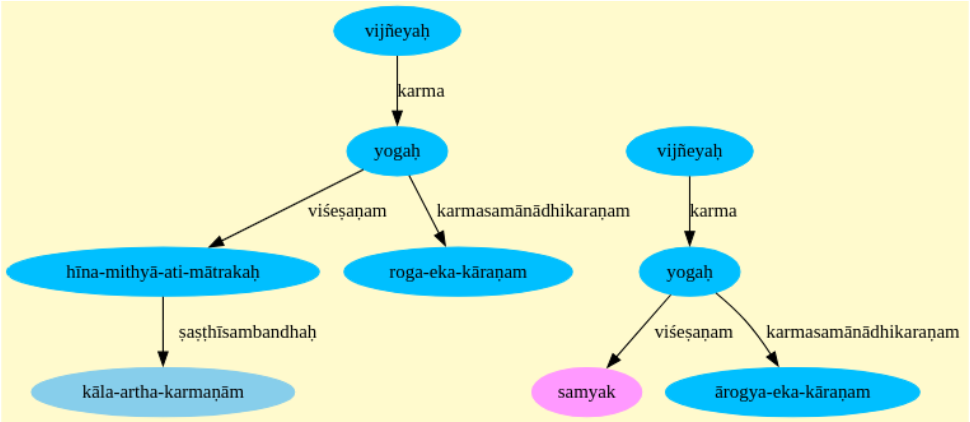

Figure 13: Dependency Tree for AH 19

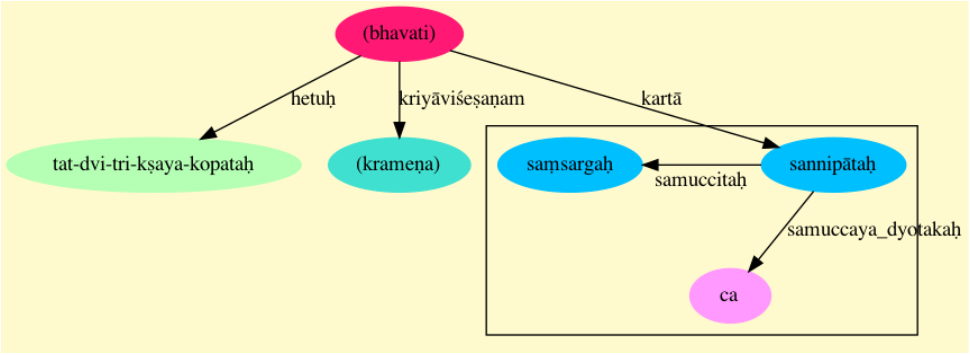

Figure 14: Dependency Tree for example 15

The dependency tree for this is shown in Figure 14.

There are two substantives viz. *saṃsargaḥ* and *sannipātaḥ* that are collectively *kartā* for the supplied verb *bhavati*. What is problematic in this example is the compound. It has 5 components with two *dvandva* and two *ṣaṣṭhī-tatpuruṣas* (See Figure 15). The first *dvandva* is between two numerals, viz. *dvi* and *tri*. It is obvious from the semantics of the words involved that the interpretation of <*dvi-tri*> is disjunctive 2 **or** 3, and not conjunctive 2 **and** 3. The second *dvandva* compound is *kṣaya-kopataḥ*, again these two words have opposite meanings. Hence here also there is a disjunctive reading . Further we have a *ṣaṣṭhaṭhī tatpuruṣa* of these two compounds resulting into ((*dvi-tri*)-(*kṣaya-kopataḥ*)). This compound is interpreted in the current context as *dvi-kṣaya-kopataḥ* **or** *tri-kṣaya-kopataḥ*, and not as *dvi-tri-kṣayataḥ* **or** *dvi-tri-kopataḥ*! Thus, we note that here, as in the previous example, the constituency analysis of a compound is not sufficient to interpret the compounds. We further need the context and the domain knowledge to interpret them. In the context of AH, from the commentaries, we gather that this hemistich means *kṣaya* (loss) or *kopa* (imbalance) of two *doṣas is termed saṃsarga* and the *kṣaya* (loss) or *kopa* (imbalance) of three of them is termed *sannipāta*. Thus, in order for a reader to get the correct reading, one may represent it as in Figure 16. We notice that such constructions are again problematic from the point of view of translation as well as understanding.

4 Conclusion

The *dvandva* compounds may have either a conjunctive or disjunctive meaning. It is the linguistic and sometimes even extralinguistic context that is instrumental in deciding the appropriate sense of the *dvandva* compound. All the examples from AH we discussed exhibit the disjunctive sense.

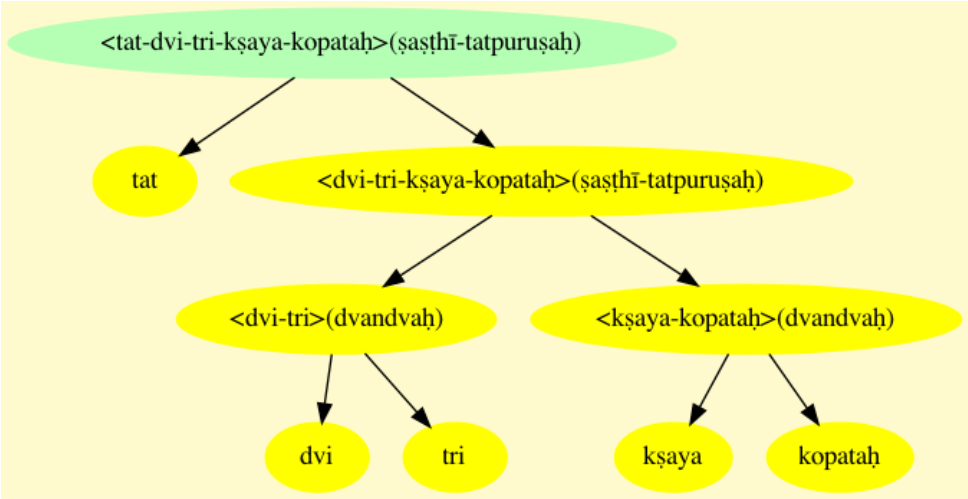

Figure 15: constituency structure

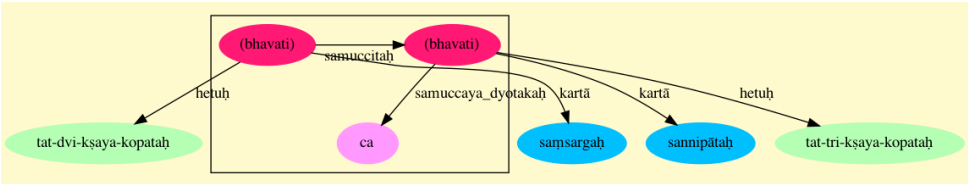

Figure 16: Compound splitting

In some cases we saw *eka-deśa-parāmarśa*, where a pronoun is used to refer to a component of a compound. In some other cases the disjunctive use resulted into distribution of components leading to new compounds that are not part of the constituency structure, making it impossible to represent the syntax using dependency tree without splitting the given verse/sentence into two. We conclude that in all these cases of *iteratara dvandva* with a disjunctive sense, it is appropriate to divide the sentence into multiple sentences to provide transparent semantics.

The use of *iteratata dvandva* alone is not problematic from the Machine Translation point of view. But the presence of two or more such compounds in the same verse/sentence, or the use of embeded *itaretara dvandva* in other larger compounds, is problematic for translation into another language. Sanskrit allows complex compound structures, which are problematic from understanding point of view. Since modern Indian languages and also English do not allow such complex compound constructions, they are problematic from the translation point of view as well.

Acknowledgement

This research is carried out as a part of the ongoing project 'Sanskrit Knowledge Accessor' under the Natural Language Translation Mission BHASHINI, 2024-26. We thank all the Sanskrit teachers from various Ayurveda colleges who are helping us to develop an e-reader for Aṣṭāṅgahṛdayam.

References

Omri Abend and Ari Rappoport. 2017. The state of the art in semantic representation. In Regina Barzilay and Min-Yen Kan, editors, *Proceedings of the 55th Annual Meeting of the Association for*

Computational Linguistics (Volume 1: Long Papers), pages 77–89, Vancouver, Canada, July. Association for Computational Linguistics.

A.G.B. ter Meulen B. H. Partee and R.E. Wall. 1990. *Mathematical Methods in Linguistics*. Kluwer Academic Publishers.

R P Chaves. 2012. Conjunction, cumulation and respectively readings. *Journal of Linguistics*, 48(2):297–344.

R.E. Wall D. R. Dowty and S. Peters. 1981. *Introduction to Montague Semantics*. Kluwer Academic Publishers.

S. D. Joshi and J. A. F. Roodbergen. 1997. *The Aṣṭādhyāyi of Pāṇiṇi*, volume 6. Sahitya Academi.

Sriram Krishnan, Amba Kulkarni, and Gérard Huet. 2024. Normalized dataset for sanskrit word segmentation and morphological parsing. *Language Resources and Evaluation*.

Yusuke Kubota and Robert Levine. 2016. The syntax-semantics interface of 'respective' predication: A unified analysis in hybrid type-logical categorial grammar. *Natural Language and Linguistic Theory*, 34:911–973.

Amba Kulkarni and Anil Kumar. 2011. Statistical constituency parser for sanskrit compounds. In *Proceedings of Indian Conference on NLP*.

Amba Kulkarni and K. V. Ramakrishnamacharyulu. 2013. Parsing sanskrit texts: Some relation specific issues. In *Proceedings of fifth International Conference on Sanskrit Computational Linguistics Symposium*.

Amba Kulkarni, Soma Paul, Malhar A. Kulkarni, Anil Kumar, and Nitesh Surtani. 2012. Semantic processing of compounds in indian languages. In *Proceedngs of International Conference on Computational Linguistics*.

Amba Kulkarni, Pavankumar Satuluri, Sanjeev Panchal, Malay Maity, and Amruta Malwade. 2020. Dependency relations for sanskrit parsing and treebank. In *19th International Workshop on Treebanks and Linguistic Theories*.

Amba Kulkarni. 2019. *Sanskrit Parsing Based on the Theories of Śābdabodha*. D. K. Printworld.

Amba Kulkarni. 2019-20. Appropriate dependency tagset for sanskrit analysis and generation. *Acta Orientalia*, 80:401–425.

Amba Kulkarni. 2021. Sanskrit parsing following indian theories of verbal cognition. *Transactions on Asian and Low-Resource Language Information Processing*, 20:1–38.

Anil Kumar, Vipul Mittal, and Amba Kulkarni. 2010. Sanskrit compound processor. In *Sanskrit Computational Linguistics*.

Anil Kumar, Amba Kulkarni, and Nakka Shailaj. 2024. Start: Sanskrit teaching; annotation; and research tool – bridging tradition and technology in scholarly exploration. In *Proceedings of seventh International symposium of Sanskrit Computational Linguistics*, pages 113–124.

Anil Kumar. 2012. *Sanskrit Compound Processor*. Ph.D. thesis, University of Hyderabad.

Richard Montague. 1970. Universal grammar. *Theoria*, 36(3):373–398.

Ranajit C O, Dr.Apurva Sarma, Umakant Tiwari, and Malini P. 2022. *Ashtangahridayam-1*. Sanskrit Promotion Foundation.

Sanjeev Panchal and Amba Kulkarni. 2019. Co-ordination in sanskrit. *Indian Linguistics*, 80(1-2):59–76.

Pt. Hari Sadashiv Shastri Paradkar. 2016. *Astanga Hrdaya - A Compendium of The Ayurvedic System of Vagbhata with The Commentaries of Sarvangasundara of Arunadatta & Ayurvedarasyana of Hemadri*. Chaukhamba Surbharati Prakashan, Varanasi.

Halvorsen P.K and Ladusaw W.A. 1979. Montague's 'universal grammar': An introduction for the linguist. *Linguistics and Philosophy*, 3:185–223.

N.S. Ramanuja Tatacharya. 2006a. *Śābdabodhamıimāṃsā: An Inquiry into Indian Theories of Verbal Cognition 1*. Institut francis de pondichéry, Rashtriya Sanskrit Sansthan.

N.S. Ramanuja Tatacharya. 2006b. *Śābdabodhamıimāṃsā: An Inquiry into Indian Theories of Verbal Cognition 2*. Institut francis de pondichéry, Rashtriya Sanskrit Sansthan.

N.S. Ramanuja Tatacharya. 2006c. *Śābdabodhamıimāṃsā: An Inquiry into Indian Theories of Verbal Cognition 3*. Institut francis de pondichéry, Rashtriya Sanskrit Sansthan.

N.S. Ramanuja Tatacharya. 2006d. *Śābdabodhamıimāṃsā: An Inquiry into Indian Theories of Verbal Cognition 4*. Institut francis de pondichéry, Rashtriya Sanskrit Sansthan.

Saee Vaze and Amba Kulkarni. 2024. Inter sentential discourse relations. In *Publications of 7th International Sanskrit Computational Linguistics Symposium*, pages 67–83.

Dr. R. Vidyanath. 2013. *Illustrated Aṣṭāṅga Hṛdaya Text with English translation and apppendices*. Chaukhamba surbharati prakashan, Varanasi.

B. H. Partee with Herman Hendriks, 1997. *Handbook of Logic and Language*, chapter Montague Grammar, page 5–92. MIT Press.