How free is 'free' word order in Sanskrit?

AMBA KULKARNI, PREETI SHUKLA, PAVANKUMAR SATULURI *and* DEVANAND SHUKL

Abstract: Sanskrit being inflectionally rich, the conventional wisdom about Sanskrit word order is that it is free. The concept of sannidhi (proximity), one of the necessary factors in the process of verbal cognition, provides a constraint on the word order of Sanskrit. We study the free word order of Sanskrit in the light of the dependency framework. The weak non-projectivity condition on dependency graphs captures the sannidhi constraint. Gillon worked within the framework of phrase-structure syntax and noted that the freeness is constrained by clause boundaries. In an examination of the cases of dislocation observed by Gillon and all verses of the *Bhagavadgītā*, we notice that two relations, viz. adjectival and genitive, are more frequently involved in sannidhi violation. We conclude that the relations involved in sannidhi violation correspond to utthāpya-ākāṅkṣā (expectancy which is to be raised) barring a few exceptional cases.

Keywords: word order, ākāṅkṣā, utthita ākāṅkṣā, utthāpya ākāṅkṣā, sannidhi, projectivity, dislocation, phrase structure, dependency structure

1 Introduction

Sanskrit being inflectionally rich, the conventional wisdom about Sanskrit word order is that it is free. It is also a common understanding among linguists that the free-ness in Sanskrit is confined

by clause boundaries. That is, it is not possible to freely interleave the constituents from subordinate clauses with elements from the main clause or other subordinate clauses. Thus the mixing of elements from different clauses is not allowed. In order to build a parser, this conventional wisdom needs to be formalised. It is necessary to know whether there are any exceptions to it, and if there are what their nature is.

The first systematic work on Sanskrit word order is by Staal (1967). He discussed the distinction between *sambandha* and *abhisambandha* in the beginning of his monograph. The former indicates grammatical relations while the latter refers to word order or arrangement. Pāṇini was very much aware of the fixed order between certain elements such as the order between a stem and a suffix,[1] the order between a prefix and a verb, and the order between the components of a compound. For example he designates a certain component an *upasarjana* and then specifies the position of the component so labeled in a compound. From the classical Sanskrit literature, we are aware of the cases where the position governs the meaning of the particles such as *api*. Speijer (1886) lists several of them.

Staal concluded that almost all Indian theorists regarded word order as free either implicitly or explicitly while the western Sanskritists rejected the free word order theory and classified the types of word order into preferential, traditional, habitual, etc. based on probabilities of occurrence and frequencies. For building a parser, though specifying preferential word order is useful in prioritizing multiple analyses produced, more precise constraints in terms of what is not allowed in grammar would help in pruning out impossible parses.

[1] *pratyayaḥ A.* 3.1.1, and *paraśca A.* 3.1.2, a suffix is placed after the root, nominal base, or item ending in a feminine affix in relation to which it is introduced.

In the same monograph, Staal presented his model of free word order in Sanskrit. He noticed the analogy between the movement of sister nodes in a Calder mobile and the free movement of elements within a phrase in a sentence. Staal called a tree that allowed the free movement of its constituents a 'wild tree'. In this model, the sisters under each node can be freely transposed provided that such transposition induces no tangling. Although Staal provided a model characterising free word order, still he did not give a precise definition that can be used to test the model empirically. Gillon's (1996: 7) empirical study led him to formalise Staal's notion of wild trees. Based on this empirical study, he noticed a pattern in the discontinuities and hence modified Staal's conjecture considering the discontinuity as a movement. He sums up his observations as follows:

> A moved constituent is either
>
> 1. a complement of the verb;
> 2. a complement or modifier of the subject noun phrase; or
> 3. a complement or modifier of a complement of the verb.

These observations provide useful relaxations for building a constituency parser.

While phrase structure grammar is suitable for understanding the constituency structure of a sentence, dependency grammar helps us in understanding the relations between various components and the semantics associated with the sentence. In the field of Natural Language Processing (NLP), recent years have also seen a growing trend towards producing dependency output in addition to constituency trees. The dependency format is preferred over the constituency not only from the point of view of evaluation (Lin 2003) but also because of its suitability for a wide range of NLP

tasks such as Machine Translation (MT), information extraction, question answering etc. (Marneffe, MacCartney, and Christopher D. Manning 2006). The importance of dependency parsing has been well recognised by computational linguists in the recent past (Culotta and Sorensen 2004; Haghighi, Ng, and C. D. Manning 2005; Quirk, Menezes, and Cherry 2005).

Gillon's observation that discontinuities are clause-bounded, implies that all the dependents of the head of the clause are still within the clause boundary, though they are free to move within a clause. Dependency parsing has an adjacency principle, also known as the projectivity principle, which is similar to this ban on discontinuous structures in phrase structure grammar. We would like to see how various cases of dislocation in phrase structure grammar translate into the dependency framework. While phrase structure analysis reveals extraposition from various positions, the facts observed so far raise some questions which lead us to take up this study further. Does extraposition lead to violation of the adjacency principle? Does it lead to any tangling in the dependency structure? Do poetic constructions bring in any tangling? Is verse more free than prose?

In what follows, we explain the concept of ākāṅkṣā (expectancy) and sannidhi (proximity) as discussed in the Indian grammatical tradition. Sannidhi imposes certain restrictions on word order. We compare the concept of sannidhi with the projectivity principle and the weak non-projectivity conditions of dependency trees. We attempt to formalise this notion of sannidhi and the violation of sannidhi with regard to two types of ākāṅkṣā (expectancy), viz., utthita (risen) and utthāpya (that which is to be raised) in order to test the constraint empirically. All the examples of dislocation discussed by Gillon are studied from the point of view of sannidhi. Further, we test this constraint on all the verses of the *Bhagavadgītā* (*BhG.*).

2 Word order in Sanskrit

Let us look at a few Sanskrit examples of possible and impossible word orders. For the sentence

(1) *rāmaḥ grāmam gacchati.*
Rama{nom.} village{acc.} go{3p. sg.}.
Rama goes to a village.

with 3 words, we get 3! (= 6) possible word orders as shown below:

(1.1) *rāmaḥ grāmam gacchati.*
(1.2) *rāmaḥ gacchati grāmam.*
(1.3) *gacchati rāmaḥ grāmam.*
(1.4) *grāmam rāmaḥ gacchati.*
(1.5) *grāmam gacchati rāmaḥ.*
(1.6) *gacchati grāmam rāmaḥ.*

These sentences convey the same overall meaning. But they differ from each other with respect to some additional meaning such as topicalisation, focus, etc. In a given context, one of them may be more suitable than the others. Some orders may be less frequent than the others. For example, consider the sentence

(2) *śvetaḥ aśvaḥ dhāvati.*
White{nom.} horse{nom.} run{3p. sg.}.
White horse runs.

The six possible word orders with the same overall meaning are

(2.1) *śvetaḥ aśvaḥ dhāvati.*
(2.2) *śvetaḥ dhāvati aśvaḥ.*
(2.3) *aśvaḥ śvetaḥ dhāvati.*
(2.4) *dhāvati śvetaḥ aśvaḥ.*
(2.5) *aśvaḥ dhāvati śvetaḥ.*

(2.6) *dhāvati aśvaḥ śvetaḥ.*

Though all these are acceptable, and convey the same overall meaning that 'a white horse runs', (2.1) and (2.4) are more probable than others. In (2.2) and (2.5) the modifier or the modified is added as an after-thought. In (2.3) and (2.6) the modifier appears after the modified, as if it is an after-thought.

Let us look at another example (3).

(3) *rāmaḥ dugdham pītvā śālām gacchati.*
Rama{nom.} milk{acc.} drink{absolutive} school{acc.} go{3p. sg.}.
Rama goes to school after drinking milk.

This sentence has 5 words. Are all the 5! (= 120) combinations meaningful?

(4) **rāmaḥ śālām dugdham gacchati pītvā.*
*Rama{nom.} school{acc.} milk{acc.} go{3p. sg.} drink{absolutive}.
*Rama to school milk goes drinking.

Sentence (4) which is obtained by permuting the words in sentence (3) does not lead to any verbal cognition, and thus it shows that all permutations need not be meaningful. So the question is, which permutations are meaningful and which are not? We look at the Indian grammatical theories that deal with these problems.

3 Indian theories of expectancy and proximity

The process of verbal cognition involves analysis of a sentence, and this analysis typically involves non-determinism. The problem of non-determinism was well recognised by the mīmāṁsakas[2]

[2]Kumārila Bhaṭṭa in his *Tantravārttika* (Sastri 1903: 505–6) mentions 3 factors necessary for the understanding of the correlation between words:

who proposed three factors, viz., ākāṅkṣā (expectancy), yogyatā (mutual compatibility), and sannidhi as necessary conditions for proper verbal cognition. We discuss here ākāṅkṣā and sannidhi, which are relevant for studying the problem of dislocation and word order.

3.1 Ākāṅkṣā (expectancy)

We first come across the term *ākāṅkṣā* in the definition of a sentence in Jaimini's *Mīmāṁsāsūtra* 2.1.46.

> *Arthaikatvāt ekam vākyam sākāṁkṣam cet vibhāge syāt.*

A group of words forms a sentence i) if when separated the words have mutual expectancy, and ii) the words serve a single purpose. Thus ākāṅkṣā or syntactic expectancy among words is a necessary condition for a group of words to form a sentence. Literally ākāṅkṣā is the desire on the part of a listener to know (*jñātum icchā*) other words in a sentence for complete understanding. Now if ākāṅkṣā is the curiosity (*jijñāsā*) on the part of a listener then after listening to a verbal form such as *ānayati* 'brings', a listener will have a curiosity to know who brings, what s/he brings, how s/he brings, and so on. Further if the object of bring is, say, a cow, the listener may have further curiosities to know what the color of the cow is, what the purpose of bringing her is, and so on. There is no end to such curiosities. These curiosities are more of a psychological nature than a syntactic one.

The Naiyāyikas made a clear distinction between psychological and syntactic expectancy. Ākāṅkṣā according to Naiyāikas is the syntactic expectancy a word has in order to correlate with

---

Ākāṅkṣā sannidhānaṁ ca yogyatā ceti ca trayam;
sambandhakāraṇatvena kḷptaṁ nānantaraśrutiḥ.

another. For example, in a word *dvāram* 'to the door', the stem *dvāra* 'door' denotes an object in the real world, and the *am* suffix (an accusative marker) marks an expectancy of a verb whose karman (object) can be *dvāra* 'door'. This expectancy which arises from the knowledge of suffix is a syntactic one, and it allows one to connect the word *dvāram* with *pidhehi* 'close'. This expectancy is not one way, but *mutual*. It is also not psychological. It is based on the usages of the verbs in a sentence and is thus syntactic in nature. As an another illustration, consider two verbs *gaṃl̥* and *vā*. The verb *gaṃl̥* is used in the sense of motion (*gaṃl̥ gatau* (*SK.* 982). The verb *vā* is used in two meanings, viz., gati 'motion' and gandhana 'pointing out' (*SK.* 1050). Though both *gaṃl̥* and *vā* are used in the sense of motion, *gaṃl̥* is sakarmaka (transitive) while *vā* is akarmaka (intransitive). This requirement of a karman (object) for *gaṃl̥* is not psychological but is based on the usage of the verb. The expectancies which are mutual, direct and natural are termed *niyata* or *utthita ākāṅkṣā* (restricted or risen expectancy) (Raja 1963). The expectancy between a verb and the words denoting kārakas or between relational words falls under this category.[3]

In contrast to mutual expectancy, the expectancy that is unilateral is called *aniyata* or *utthāpya ākāṅkṣā* (unrestricted or to be raised). This is aroused only if necessary. So it is potential. For example, in a phrase such as *white cow*, the ākāṅkṣā of *white* for a substantive is natural, but the ākāṅkṣā of *cow* to have an adjective is potential. It gets aroused in the presence of an adjective such as *white*. Even a noun in apposition may arouse an expectancy. The example discussed in the *Rāmarudrī* commentary on the *Dinakarī* commentary on the *Nyāyasiddhāntamuktāvalī* is *udayati candraḥ kumudabāndhavaḥ* (Jere 2002) 'Rises the moon, the friend of a lotus (that opens its petals during night and closes

[3] *Niyatākāṅkṣā: yathā kriyākārakapadānāṁ parasparākāṅkṣā.* (Jere 2002: 2–3).

them in the morning)'. After hearing *udayati candraḥ* 'the moon rises' all the expectancies are fulfilled. And thus understanding is complete. Now when one hears *kumudabāndhavaḥ* 'friend of a lotus' then this word needs to be related to one of the words uttered earlier because there can not be freely hanging words in a meaningful sentence. This word has an expectancy of a substantive, and thus it gets related to *candraḥ*. Both these cases are examples of unilateral expectancy.

3.2 Sannidhi (proximity)

Sannidhi is defined in the *Tarkasaṅgraha* as 'an utterance of words without any gap' (*padānām avilambena uccāraṇam.*), or as 'the presentation of word meanings without any intervention' (*avyavadhānena padajanya padārthopasthitiḥ*). From the text processing point of view, the important point is the presentation of word meanings without any intervention. In other words, if the related words are intervened by some unrelated words, then such an utterance does not produce any verbal cognition. To make the point clear, Viśvanātha Pañcānana in his *Nyāyasiddhāntamuktāvalī* (Joshi 1985: 194), gives the following example.

(5) *giriḥ bhuktam agnimān devadattena.*
Hill is eaten fiery by Devadatta.

The words *giriḥ* 'hill' and *agnimān* 'fiery' have mutual expectancy so do *bhuktam* 'eaten' and *devadattena* 'by Devadatta'. *Bhuktam*, being a past participle in the neuter, expects a kartṛ in the instrumental case and a *karman* in the neuter nominative. But this group of four words does not have a neuter nominative. So we interpret either the verb *bhuj* 'to eat' in this context to be an intransitive, or consider the example as a case of an ellipsis of the *karman*. In either case, this group of four words corresponds to two

independent sentences, whose arguments are intertwined. *Bhuktam* intervenes between the words *giriḥ* and *agnimān* which have mutual expectancy. Similarly *agnimān* intervenes between the related words *bhuktam* and *devadattena*. This intervention forms an obstacle to verbal cognition (*śābdabodha*).

The condition of not having intervention is only a necessary condition in the process of śābdabodha. For, even the non-intervention may give rise to more than one śābdabodha – one of them as a true cognition (*pramātmaka jñāna*) and the other one as a false cognition (*bhramātmaka jñāna*) as explained by Viśvanātha Pañcānana (Joshi 1985: 194) with the example in (6).

Consider the group of words

(6) *nīlo ghaṭaḥ dravyam paṭaḥ.*
blue pot thing cloth.

This may lead to two cognitions, viz.:

1. The pot is blue, and the cloth is a thing.
2. The cloth is blue, and the pot is a thing.

Viśvanātha Pañcānana argues that the first one, in the given situation, leads to a true cognition and the second one to a false cognition. The notion of true cognition and false cognition is thus context dependent. The notion of sannidhi on the other hand depends only on the expectancies and the compatibility of meanings of the words involved.

In the next section we compare the notions of sannidhi and ākānkṣā with the projectivity principle and the weak non-projectivity conditions.

4 Dependency parsing and word order

Dependency analysis dates back to Pāṇini. The computational implementation of a dependency parser for Indian Languages based

on Pāṇinian grammatical formalism is described by Bharati, Chaitanya, and Sangal (1994). In the recent years, the seminal work of Tesnière (1959) became the basis for work on dependency grammar. Meaning-text theory (Melĉuk 1988), Word Grammar (Hudson 1984), and Functional generative description (Segall, Hajičová, and Panevová 1986) are some of the flavours of dependency grammar. Bharati, Bhatia, et al. (1998) extended their Pāṇinian grammatical formalism to English. However, the first full-fledged computational implementation of a dependency grammar for English is the Link parser (Sleator and Temperley 1993).

The dependency parse of a sentence establishes relations between the morphemes in the form of nominal and verbal bases (*prātipadika*s and *dhātu*s), through the morphemes in the form of nominal and verbal suffixes, through positional information, or through concord. In the case of Sanskrit, it is predominantly nominal and verbal suffixes which mark relations. In sentence (1) above, the verbal suffix *ti* in *gacchati* establishes the relation between a person whose name is *Rāma* and the activity of going represented by the verbal root *gaṃḷ*. The nominal suffix *am* represents the relation between the entity denoted by the nominal stem *grāma* and the activity of going denoted by *gaṃḷ*. Dependency structure is also termed *governance*, since it tells us which words govern which others.

4.1 Tree traversal and possible word orders

The dependency structure of a sentence is typically represented as a graph whose nodes correspond to the words and whose edges correspond to the relation between nodes. This structure is characterised by two properties: a) the graph is acyclic, and b) every node, except the root node, has exactly one incoming arrow. Thus this structure is best represented in a tree. Tree traversal is the process of visiting each node in a tree structure exactly once in a

systematic manner. Thus with every tree traversal is associated a word order. The traversal is defined in the context of binary trees by the order in which the nodes are visited. There are three distinct traversals:

1. Post-order traversal,
2. In-order traversal, and
3. Pre-order traversal.

If the root node (or the top node), left node and right node are denoted by T, L, and R respectively, then the post-order traversal is L-R-T, the in-order traversal is L-T-R and the pre-order traversal is T-L-R. At each visit of the node, all the sub-nodes under it are descended recursively till all the leaf nodes under that node are visited.

The trees in Figure 1 show these three traversals for sentence (1) in the first row and the traversals after the leaf nodes are transposed in the second row.[4] The dotted lines show relations between words, while numbered thick lines (in red) show the order of tree traversal. As is clear from the figures, the tree traversal orders shown correspond to the six possible word orders listed above in (1.1) through (1.6).

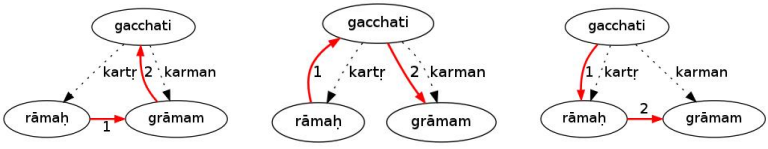
4.1.1 Generalising tree traversal

For a tree with more than 2 leaf nodes, we generalise the traversal as follows:

Let A_1 be the root node, and $A_2, \ldots, A_n$ be the leaf nodes. Let $\{A_2, \ldots, A_n\}$ stand for all possible permutations of $n-1$ nodes. Then the only possible traversals with n nodes are the following:

1. **Pre-Order Traversals** A_1 $\{A_2, \ldots, A_n\}$.

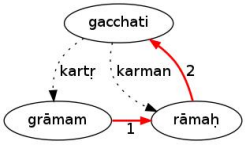
[4]For ease of reading, we label the node by a word including the suffix, and not with a stem.



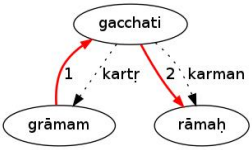
(a) *Post-order (1.1)*

(b) *In-order (1.2)*

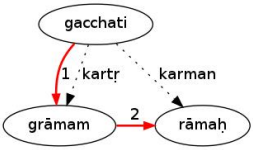
(c) *Pre-order (1.3)*



(d) *Post-order (1.4)*



(e) *In-order (1.5)*



(f) *Pre-order (1.6)*

Figure 1

Traversals for sentence (1) with and without transposition

2. **In-Order Traversals** $\{A_2, \ldots, A_k\}\ A_1\ \{A_{k+1}, \ldots, A_n\}$, where k varies from 2 to $n-1$.
3. **Post-Order Traversals** $\{A_2, \ldots, A_n\}\ A_1$.

When any of the A_2 to A_n is itself a tree, then at each of these nodes, we recursively traverse the sub-trees in any of these three orders to get possible word orders.

Now we look at the example (3) above. The dependency graph for this sentence is shown in Figure 2. We notice that the word order in (4) can not be produced from this dependency tree by any of the above traversal methods. While the different traversals give us a computational device to generate all possible word orders, they do not give an efficient mechanism to decide whether a given word order can be obtained through one of the traversals or not. The governance relation does not capture the word order or the precedence. In dependency structures, the precedence is captured by projecting the nodes of the governance structures onto the linear representation of the words in a sentence form. There have

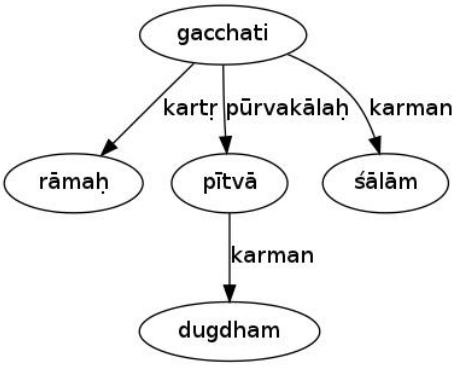


Figure 2
Dependency graph for sentence 3

been several efforts in the domain of dependency framework to study the relation between governance and precedence (Bodirsky, Kuhlmann, and Möhl 2005; Havelka 2005; Nivre 2006). These studies have resulted in various types of constraints on dependency graphs. These constraints ban some word orders. The strongest among these is the projectivity principle. The less stringent ones are weak non-projectivity and well-nestedness.

4.2 Projectivity principle

The principle of projectivity states a constraint on the dependency tree which bans certain dependency structures. There have been various characterisations of the projectivity principle. Marcus (1967) has shown the equivalence of some of the earlier characterisations e.g. of the ones by Harper and Hays, Lecerf and Ihm, and Fitialov.

> A sentence is projective if and only if we can draw a dependency tree whose every node can be projected

> by a vertical line onto its word form in the surface string without crossing another projection or a dependency edge.

Hudson (1984: 98) calls this projectivity principle the 'adjacency principle'.

Figure 3 shows the projections for sentences (1.1) to (1.6). We notice that the projections shown by dotted lines do not cross the dependency relations shown by thick lines. And hence all the 6 sentences in (1.1) through (1.6) follow the projectivity principle.

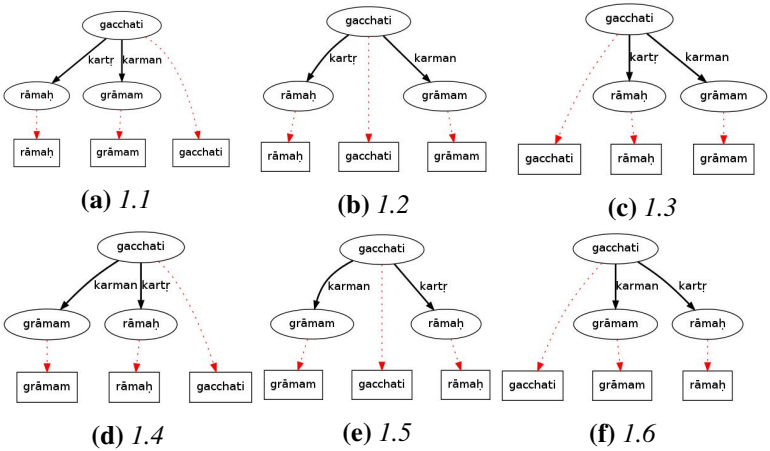


Figure 3

Dependency structure with projections for sentences 1.1 to 1.6

Figure 4 shows the projections of sentences (2.1) to (2.6). We notice that sentences (2.2) and (2.5) have the projection line crossing the governance relation showing the violation of the projectivity principle.

Sentence (4) also violates the projectivity principle. Figure 5 shows the governance relation between *pītvā* and *dugdham* being crossed by the projection of *gacchati*.

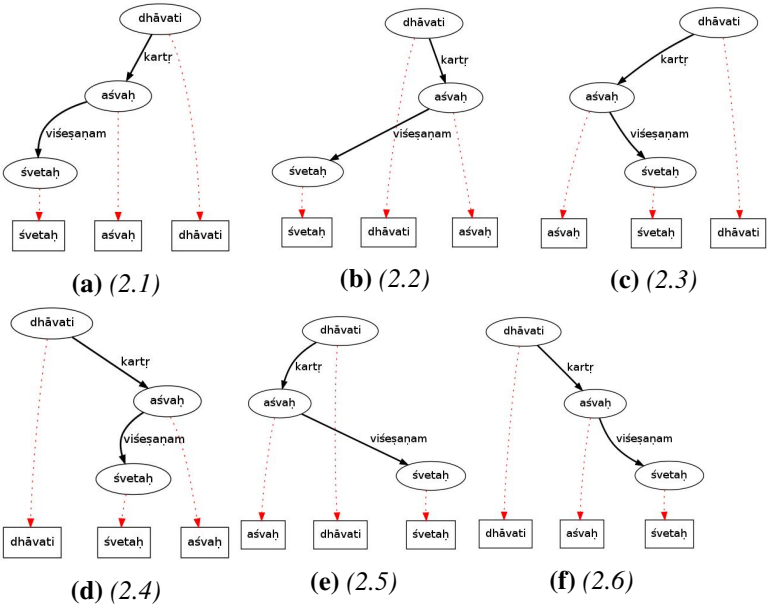


Figure 4
Projections for sentences 2.1 to 2.6

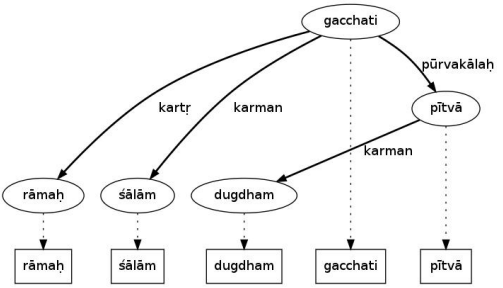
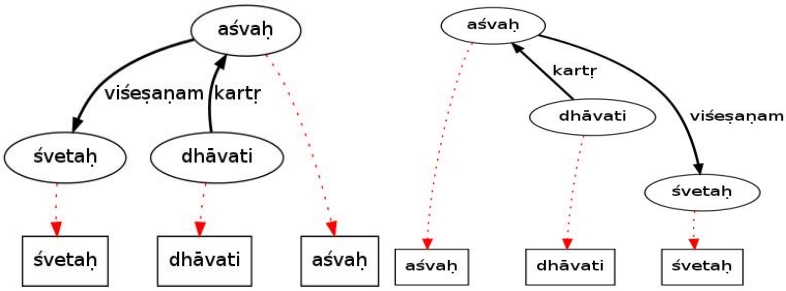


Figure 5
Projection for sentence (4)

Though both sentence (4) as well as sentences (2.2) and (2.5) show violation of the projectivity principle, there is a difference between the violation in (4) and the violation in (2.2) and (2.5). The violation in (4) is an example of sannidhi violation of utthita ākāṅkṣā while (2.2) and (2.5) are examples of sannidhi violation of utthāpya ākāṅkṣā. Thus there is a need to distinguish between these two types of violation, the distinction between which is not captured by the projectivity principle.

4.3 Weak non-projectivity (planarity)

It is possible to draw the dependency graphs for (2.2) and (2.5) by rearranging the nodes to avoid the crossing of projection lines by the dependency relations. Figure 6 shows possible graphs for (2.2) and (2.5) that avoid crossing. However, there is no rearrangement of the nodes of (4) that will avoid crossing. We capture this differ-



(a) *(2.2)* **(b)** *(2.5)*

Figure 6
Projection with Rearrangement of nodes

ence by relaxing the projectivity constraint. Instead of considering the crossing between two types of relations, viz., projection and dependency, we consider only the crossing between dependency relations with a further constraint that the nodes of the dependency

structure be represented in a linear order that reflects the surface order of the words in the sentence they represent. Such a dependency graph is weakly non-projective if there are no crossing of edges. If all the edges are drawn on the same side of a sentence (either below or above), such a graph results in a planar graph.

- *Definition:* A graph is weakly non-projective or planar, if it does not have two edges $w_i \leftrightarrow w_j$ and $w_k \leftrightarrow w_l$ with $i < k < j < l$.

Thus every projective structure is weakly non-projective, but the reverse is not true.

Let us look at the sentences (2.2), (2.5) and (4) above. All these were found to be non-projective. But as we notice from Figure 7, the sentences (2.2) and (2.5) do not have crossing edges in their linear projection. Hence these are weakly non-projective or planar. But Figure 8 shows the crossing of edges making sentence (4) non-planar.

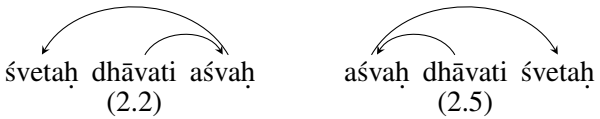


Figure 7
Planar dependency graph for sentences (2.2) and (2.5)

Planarity (or weak non-projectivity) is the precise characterisation of the sannidhi constraint. Sannidhi violation leads to non-planar graphs which correspond to the dislocation of constituents in the phrase structure tree. But the dislocation need not correspond to sannidhi violation, as we shall see below.

Consider the following sentence analyzed by Gillon (1996: 12):

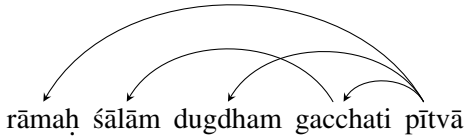


Figure 8
Planar dependency graph for sentence (4)

> *tena ca pramāṇena sādhya-dharmasya tat-mātra-anubandhaḥ khyāpyate.* (*PV* 18.1)
> And the dependence of the provable property merely on it (i.e., the proving property) is made known by that epistemic cognition.

In this example, Gillon observes that 'the third (instrumental) case noun phrase is at the clause's left periphery, having been extraposed from the passive verb "khyāpyate" ("made known"), which is at the right periphery'.

Figure 9 shows the dependency graph for this sentence which is a planar (weakly non-projective) one. Here the dislocation does not lead to a sannidhi violation.

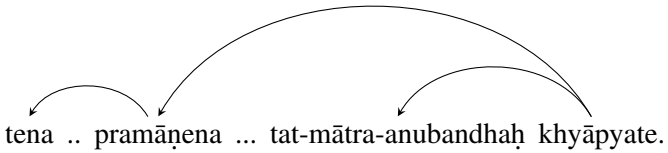


Figure 9
Dislocation without sannidhi violation

5 Empirical evaluation

Now that we have a precise mathematical criterion for evaluating the sannidhi violation, we test it on a real corpus consisting of both prose as well as verse. For prose, we take the same corpus as studied by Gillon. Since sannidhi violation always leads to dislocation, we need not look at the sentences which do not have dislocation. We study only the cases of dislocation discussed by Gillon. For the verse, we chose the *Bhagavadgītā* (*BhG.*).

5.1 Cases of sannidhi violation from Gillon's data

Gillon's corpus consists of about a thousand sentences, approximately half of them from the Pramāṇvārttika by the Buddhist philosopher Dharmakīrti and the rest from Apte's (1885) broadly representative selection of examples of classical Sanskrit literature. Gillon observed dislocation in about 160 sentences which he categorised into three classes:

1. extraposition from subject position,
2. extraposition from verb complement position, and
3. verb complement topicalisation.

We noted earlier that not all dislocation leads to sannidhi violation. Only those cases which have crossing edges in their planar dependency graph are cases of sannidhi violation. Almost 75% of the examples discussed by Gillon 1996 do not involve crossing edges, hence do not lead to sannidhi violation. Among the examples of dislocation that involve sannidhi violation, we noted that mainly two relations, viz., those of the adjective and the genitive, are involved in crossing.

5.1.1 Dislocation of a genitive

We give below two examples of dislocation of a genitive discussed by Gillon (1996).

(7) *tayoḥ baddhayoḥ kim-nimittaḥ ayam mokṣaḥ* (*ASG.* 14.1.2)
What basis does the release of the two prisoners have?

A copulative verb *asti* is supplied in the dependency stucture. The two Indian schools Vyākaraṇa (grammar) and Nyāya (logic) differ in the analysis of sentences with missing copulative verbs. The Vyākaraṇa school supplies the missing copulative verb *asti*[5] and then establishes relations between the substantive and its predicative adjective through this verb, while the Nyāya school establishes the relation between the substantive and the predicative adjective directly. We follow the Vyākaraṇa school. The non-planar and planar dependency graphs are shown in Figure 10. As one can see, the crossing links involve the verb *asti*, and the position of this verb is crucial in the planarity of the dependency graph. If the verb *asti* is placed in juxtaposition with *kim-nimittaḥ* then there is no crossing. The verb *asti* in this sentence has only two arguments. In example (8) (Figure 11) below, where there are more than two arguments, the linear representation of the sentence has crossing of links no matter where the copulative verb is placed.

(8) *sarvatra audarikasya abhyavahāryam eva viṣayaḥ.* (*ASG.* 1.1.2)
In every case, a glutton's object is only food.

[5] *astīr bhavantīparaḥ prathamapuruṣo 'prayujyamāno 'pyasti.* (A.2.3.1 vt. 11)

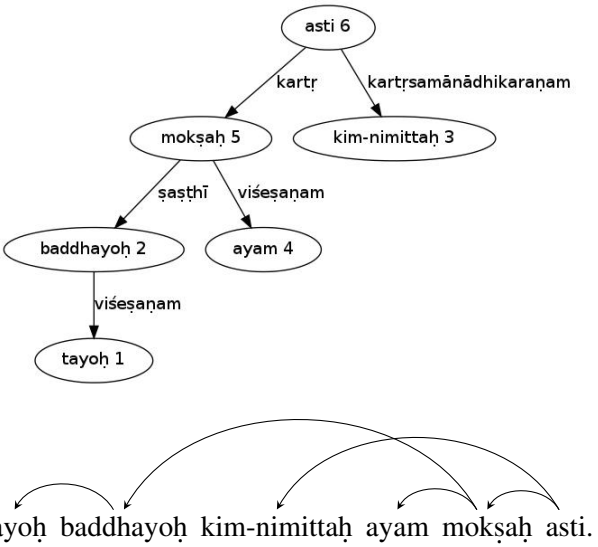


Figure 10
Dislocation of a genitive

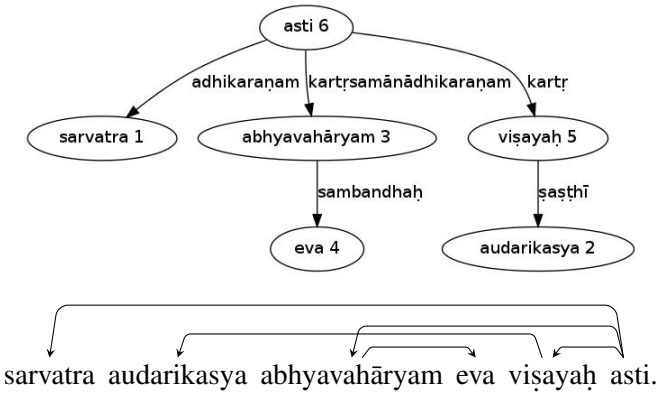


Figure 11
Dislocation of a genitive

5.1.2 Dislocation of a viśeṣaṇa

Example (9) (Figure 12) shows the evidence of sannidhi violation due to dislocation of an adjective. We supply the missing copulative verb *bhavanti* following the Vyākaraṇa school.

(9) *ete hi hr̥daya-marma-bhidaḥ saṁsāra-bhāvāḥ* (*ASG.* 8.1.3)
For, these worldly things are heart breakers.

Here the position of the missing copulative verb *bhavanti* dictates the crossing of links. Certain positions of *bhavanti* lead to crossing of links; others do not.

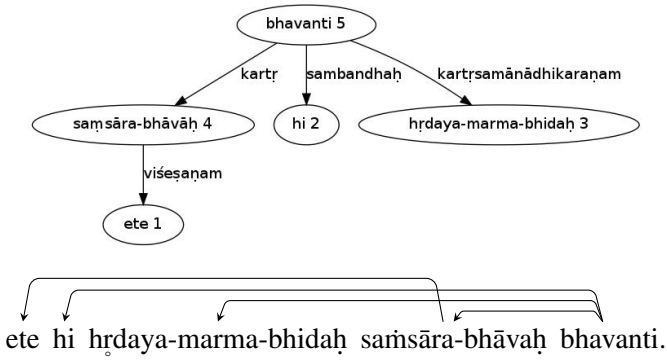


Figure 12
Dislocation of a viśeṣaṇa

5.1.3 Other relations

Example (10) is a case where the relations involved in sannidhi violation are other than viśeṣaṇa and genitive (Figure 13).

(10) *aham manda-autsukyaḥ asmi nagara-gamanam prati* (*ASG.* 3.1.3)
I am (one who is) little eager about going to the city.

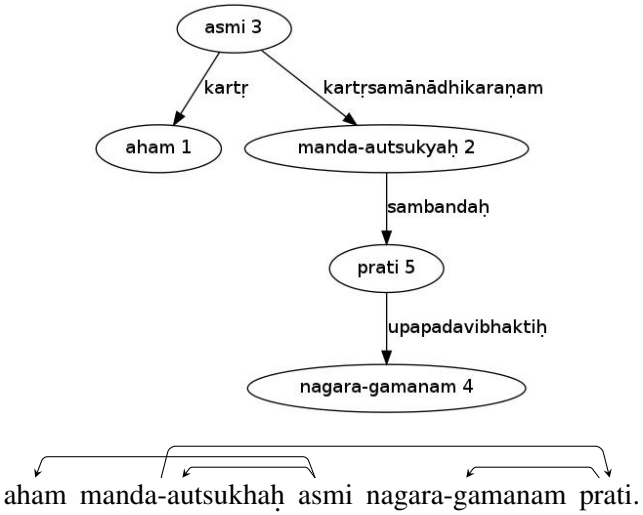


Figure 13
Dislocation of an argument

The edge marking the non-kāraka relation between the words *manda-autsukyaḥ* and *prati* crosses the edge between the kartṛ relation between *aham* and *asmi*.

5.2 Sannidhi violation in the *Bhagavadgītā*

The word order in verse, though free, is constrained by metrical considerations. In order to study the effect of metrical considerations on sannidhi, we analysed all the verses of the *BhG*. We did not consider the ślokas which have either conjunctive or disjunctive particles. Among the remaining 344 ślokas, 300 instances of sannidhi violations were found. Out of these, 193 cases involved either an adjectival or genitive relation. The remaining cases had other non-kāraka relations such as negation, vocative, precedence,

simultaneity, etc. A total of 9 cases of anomalous behaviour were found. Of these, in 7 cases both of the relations involved in crossing are kāraka relations. In two cases there is crossing of an adverb with a kāraka relation. We discuss all these cases below.

5.2.1 Sannidhi violation involving kartṛ and karman

Five among the nine anomalous cases had the kartṛ relation of one verb crossing with the karman relation of another verb. These instances are examined below.

(11) चञ्चलं हि मनः कृष्ण प्रमाथि बलवद्दृढम्।
तस्याहं निग्रहं मन्ये वायोरिव सुदुष्करम्॥
cañcalam hi manaḥ kṛṣṇa pramāthi balavat dṛḍham;
tasya aham nigraham manye vāyoḥ iva suduṣkaram. (*BhG.* 6.34)
O kṛṣṇa, the mind is fickle, turbulent, obstinate and strong, hence I think it is as difficult as to control the wind.

In the second line of this verse the main verb is *manye* whose kartṛ is *aham*. The karman of the verbal noun *nigraham* is the pronominal *tasya*, which refers to *manaḥ* in the first sentence. Thus the word sequence *tasya aham nigraham manye* produces two crossing edges involving the relations of kartṛ and karman (Figure 14).

(12) धूमो रात्रिस्तथा कृष्णः षण्मासा दक्षिणायनम्।
तत्र चान्द्रमसं ज्योतिर्योगी प्राप्य निवर्तते॥
dhūmaḥ rātriḥ tathā kṛṣṇaḥ ṣaṇmāsāḥ dakṣiṇāyanam;
tatra cāndramasam jyotiḥ yogī prāpya nivartate. (*BhG.* 8.25)
The yogī on travelling the path of smoke, night, the dark half of the month and the six months of the southern path (of the sun) after death, obtains the lunar light and returns (to this world).

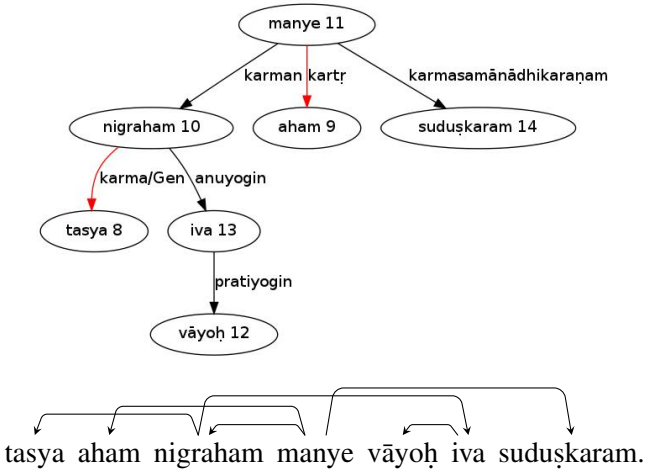


Figure 14
Analysis of BhG. 6.34

The second example with similar crossing of kartṛ and karman was found in verse *BhG.* 8.25. In this verse, in the word sequence *jyotiḥ yogī prāpya nivartate*, *jyotiḥ* is the karman of the absolutive verb *prāpya* and *yogī* is the kartṛ of the verb *nivartate* resulting in the crossing (Figure 15).

(13) वक्तुम् अर्हस्यशेषेण दिव्या हि आत्मविभूतयः।
याभिर्विभूतिभिर्लोकान् इमां त्वं व्याप्य तिष्ठसि॥
vaktum arhasi aśeṣeṇa divyāḥ hi ātmavibhūtayaḥ;
yābhiḥ vibhūtibhiḥ lokān imān tvam vyāpya tiṣṭhasi. (*BhG.* 10.16)
You could tell me about your divine manifestations by which you exist pervading these worlds.

The third example is from the verse *BhG.* 10.16. Here the word sequence *lokān imān tvam vyāpya tiṣṭhasi* results in crossing edges

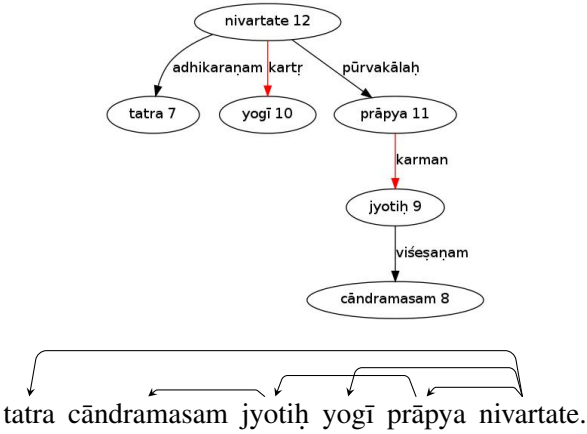


Figure 15
Analysis of BhG. 8.25

with *lokān* being the karman of *vyāpya* and *tvam* being the kartṛ of *tiṣṭhasi* (Figure 16).

(14) अश्रद्दधानाः पुरुषा धर्मस्यास्य परन्तप।
अप्राप्य मां निवर्तन्ते मृत्युसंसारवर्त्मनि॥
aśraddadhānāḥ purūṣāḥ dharmasya asya parantapa;
aprāpya mām nivartante mṛtyusaṁsāravartmani. (*BhG.* 9.3)
O Parantapa, those who have no faith in this dharma return to the circle of death and rebirth without attaining me.

The fourth example is from *BhG.* 9.3, where the dependency arrow corresponding to the karman of the kṛdanta *a-śraddadhānāḥ* crosses the arrow corresponding to the kartṛ of the main verb *nivartante* (Figure 17).

(15) तान् समीक्ष्य सः कौन्तेयः सर्वान् बन्धूनवस्थितान्।
कृपया परयाविष्टो विषीदन्निदमब्रवीत्॥

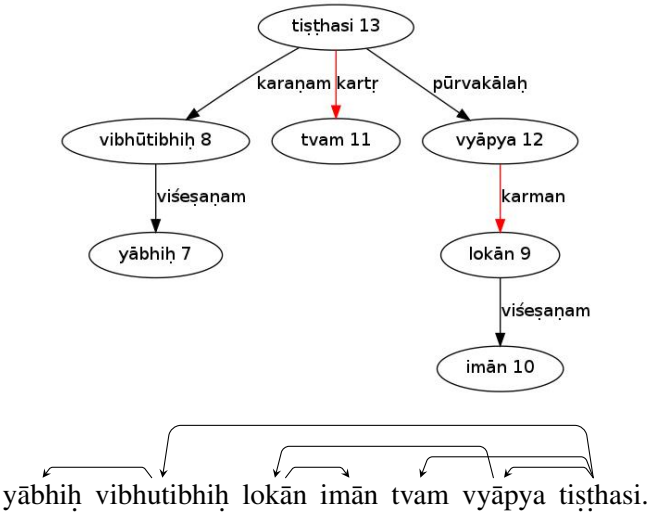


Figure 16
Analysis of BhG. 10.16

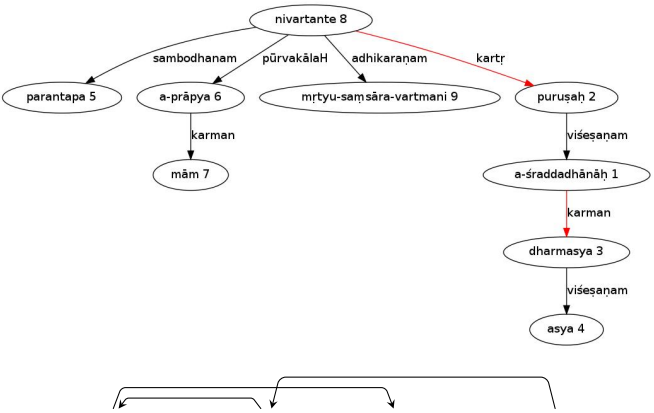


Figure 17
Analysis of BhG. 9.3

tān samīkṣya saḥ kaunteyaḥ sarvān bandhūn avasthitān. (*BhG.* 1.27cd)
kṛpayā parayā āviṣṭaḥ viṣīdan idam abravīt; (*BhG.* 1.28ab)
Seeing all these relatives present there (on the battle-field) Kaunteya filled with compassion uttered these words in dejection.

In *BhG.* 1.27–28 also, we find the crossing of the dependency arrows pointing to a kartṛ and a karman. *Bandhūn* is the karman of the absolutive *samīkṣya* and *kaunteya* is the kartṛ of the verb *abravīt*, and the edges marking these relations cross (Figure 18). In

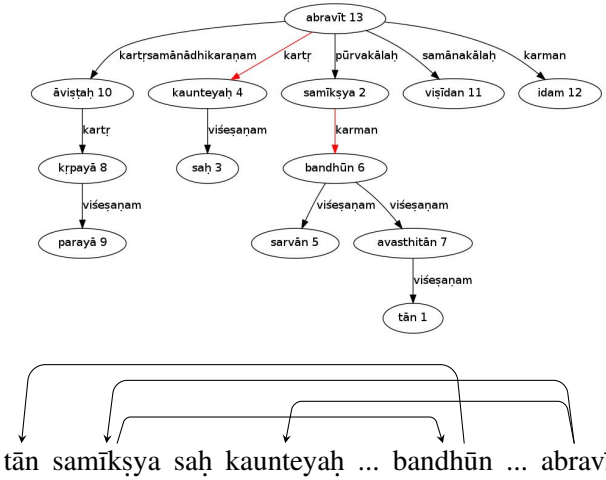


Figure 18
Analysis of BhG. 1.27

this example, if we choose *tān*, which is an adjective of *bandhūn* as the karman of *samīkṣya*, then this crossing vanishes! This, in fact, happens to be an example of cataphora, and would be analysed as in Figure 19.

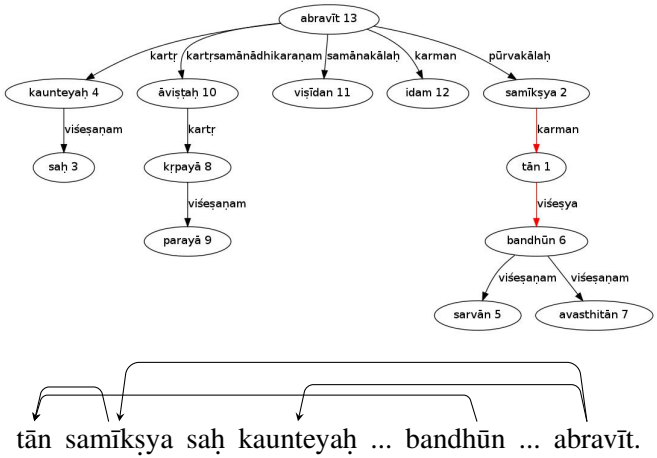


Figure 19
Modified Analysis of BhG. 1.27

5.2.2 Sannidhi violation involving two kartṛ relations

Verse *BhG.* 8.19 has an instance of crossing edges involving two kartṛ relations.

(16) भूतग्रामः स एवायं भूत्वा भूत्वा प्रलीयते।
रात्र्यागमे ऽवशः पार्थ प्रभवत्यहरागमे॥
bhūtagrāmaḥ saḥ eva ayam bhūtvā bhūtvā pralīyate;
rātri-āgame avaśaḥ pārtha prabhavati ahar-āgame. (*BhG.* 8.19)
O Pārtha, these living beings merge and re-emerge at the onset of night and on the coming of the day.

Here *saḥ* is the kartṛ of the verb *pralīyate* and *bhūtagrāmaḥ* is the kartṛsamānādhikaraṇa for the absolutive *bhūtvā*, and these relations cross (Figure 20).

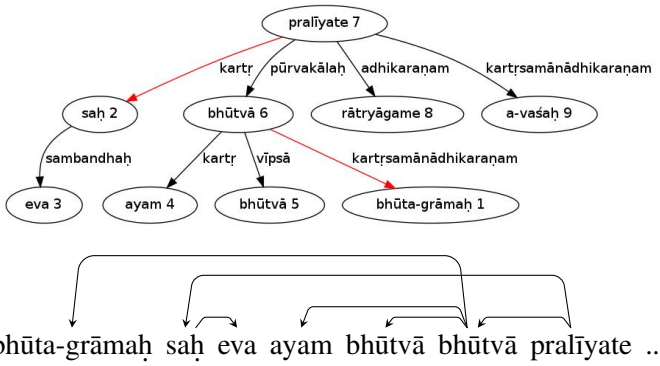


Figure 20
Analysis of BhG. 8.19

5.2.3 Sannidhi violation involving karman and kriyāviśeṣaṇa

In *BhG.* 1.37, the relation of karman of the absolutive *hatvā* is crossed by the kriyāviśeṣaṇa of the main verb *syāma* (Figure 21).

(17) तस्मान्नार्हा वयं हन्तुं धार्तराष्ट्रान्स्वबान्धवान्।
स्वजनं हि कथं हत्वा सुखिनः स्याम माधव॥
tasmāt na arhāḥ vayam hantum dhārtarāṣṭrān svabāndhavān;
svajanam hi katham hatvā sukhinaḥ syāma mādhava. (*BhG.* 1.37)
Therefore O Mādhava, it is not suitable on our part to kill our relatives, for by killing our own kinsmen, the sons of Dhṛtarāṣṭra, how can we remain happy?

A similar example is found in *BhG.* 11.32 where the relation of kartṛ crosses with that of the adjective.

5.2.4 Sannidhi violation involving karman and apādāna

The last example is from the last chapter of *BhG.* Here the karman of the non-finite verb *kathayataḥ* crosses the relation of apādāna

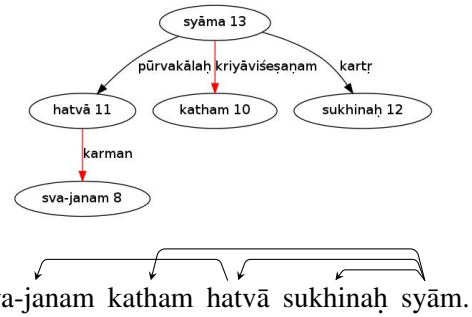


Figure 21
Analysis of BhG. 1.37

between *śrutavān* and *kr̥ṣṇāt* (Figure 22).

(18) व्यासप्रसादाच्छ्रुतवानेतद्गुह्यमहं परम्।
योगं योगेश्वरात्कृष्णात्साक्षात्कथयतः स्वयम्॥
vyāsaprasādāt śrutavān etad guhyam aham param
yogam yogeśvarāt kr̥ṣṇāt sākṣāt kathayataḥ svayam. (BhG. 18.75)
By the grace of Vyāsa I have heard this supreme secret yoga directly from the lord of yoga, Kr̥ṣṇa.

6 Conclusion

We conclude with two observations. The first observation is that the number of violations is much greater in verse than in prose. This may be attributed to metrical considerations. Out of 344 verses, 154 verses had sannidhi violation, while Gillon reported only 160 instances of violations in about 1000 prose sentences.

The second observation is with respect to the relations involved. The two major relations involved in sannidhi violation in

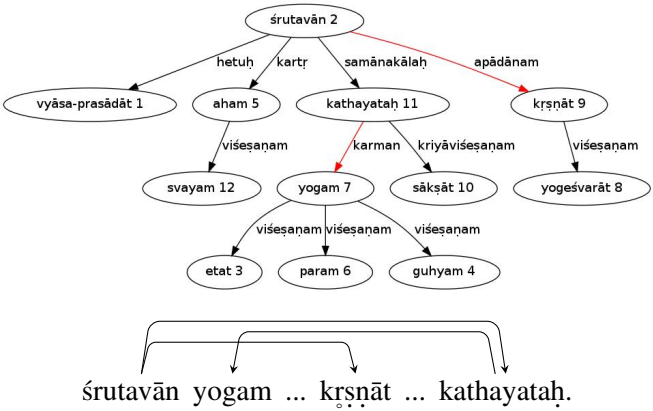


Figure 22
Analysis of BhG. 18.75

both prose and verse are the viśeṣaṇa and the genitive. Some of the other relations involved in the sannidhi violation are sambodhana (vocative), negation, precedence, and simultaneity. All these relations have unilateral expectancy and thus correspond to utthāpya ākāṅkṣā.

Empirical study thus reveals that mutual expectancies are tightly coupled, and the words that have mutual expectancy are always in close proximity (sannidhi). In contrast, a word that has unilateral expectancy (*utthāpya ākāṅkṣā*) may be moved away from its relatum and its relational path may be interrupted by one or more words unrelated to them.

Table 1
Abbreviations

A.	*Aṣṭādhyāyī*. Reference to: chapter . section . verse
ASG.	Apte 1885. Reference to: chapter . exercise set . example sentence

BhG. *Bhagavadgītā*. Reference to: chapter . verse
SK. Siddhāntakaumudī. See Govindacharya 2010. Reference to: dhātusaṅkhyā

References

Apte, Vaman Shivaram. 1885. *The student's guide to Sanskrit composition: being a treatise on Sanskrit syntax for the use of schools and colleges*. Second edition ... enlarged. Pune: H. N. Gokhale. Printed at Āryabhūṣaṇa Press.

Bharati, Akshar, Medhavi Bhatia, Vineet Chaitanya, and Rajeev Sangal. 1998. "Paninian grammar framework applied to English." *South Asian Language Review* 8.1: 1–23.

Bharati, Akshar, Vineet Chaitanya, and Rajeev Sangal. 1994. *Natural language processing: a Paninian Perspective*. New Delhi: Prentice Hall of India Pvt Ltd.

Bodirsky, M., M. Kuhlmann, and M. Möhl. 2005. "Well-nested drawings as models of syntactic structure." *Tenth Conference on Formal Grammar and Ninth Meeting on Mathematics of Language*, pp. 195–203. University of Edinburgh. Edinburgh.

Culotta, A. and J. Sorensen. 2004. "Dependency tree kernels for relation extraction." *42nd Annual Meeting of the Association for Computational Linguistics: proceedings of the conference*, pp. 423–29. Stroudsburg, Pennsylvania: acl.

Gillon, Brendan S. 1996. "Word order in Classical Sanskrit." *Indian Linguistics* 57.1–4: 1–36.

Govindacharya. 2010. *Vaiyākaraṇasiddhantakaumudī*. 2 vols. Varanasi: Chaukhambha surabharati prakashan.

Haghighi, A. D., A. Y. Ng, and C. D. Manning. 2005. "Robust textual inference via graph matching." *Human Language Technology Conference (HLT) and Conference on Empirical Meth-*

ods in Natural Language Processing (EMNLP), pp. 387–94. Stroudsburg, Pennsylvania: acl.

Havelka, J. 2005. "Projectivity in totally ordered rooted trees." *Prague Bulletin of Mathematical Linguistics* 84: 13–30. Prague: Charles University.

Hudson, R. 1984. *Word grammar*. Oxford: Basil Blackwell.

Jere, Atmaram Narayan. 2002. *Kārikāvalī*. Varanasi: Chowkamba Krishnadas Academy.

Joshi, K. R. 1985. *Nyāyasiddhāntamuktāvalī*. Bhandarkar Oriental Series 21. Pune: Bhandarkar Oriental Research Institute.

Kulkarni, Amba. 2013. "A deterministic dependency parser with dynamic programming for Sanskrit." *Proceedings of the Second International Conference on Dependency Linguistics (DepLing 2013)*, pp. 157–66. Prague, Czech Republic: Charles University in Prague, Matfyzpress, Prague, Czech Republic. URL: `http://www.aclweb.org/anthology/W13-3718`.

Kulkarni, Amba, Sheetal Pokar, and Devanand Shukl. 2010. "Designing a constraint based parser for Sanskrit." *Sanskrit computational linguistics: 4th International Symposium, New Delhi, India, December 2010, Proceedings*, ed. by Girish Nath Jha, pp. 70–90. Lecture Notes in Artificial Intelligence 6465.

Lin, Dekang. 2003. "Dependency-based evaluation of MINIPAR." *Treebanks: building and using parsed corpora*, ed. by Anne Abeillé, pp. 317–29. Text, speech, and language technology 20. Berlin; Heidelberg: Springer.

Marcus, S. 1967. *Algebraic linguistics: analytical models*. Mathematics in Science and Engineering 29. New York: Academic Press.

Marneffe, Marie-Catherine de, Bill MacCartney, and Christopher D. Manning. 2006. "Generating typed dependency parses from phrase structure parses." *Proceedings of LREC-06*, pp. 449–54. Paris: elra.

Melĉuk, I. 1988. *Dependency syntax: theory and practice*. Albany: State University of New York Press.

Nivre, J. 2006. "Constraints on non-projective dependency parsing." *Eleventh Conference of the European Chapter of the Association for Computational Linguistics (EACL)*, pp. 73–80. Trento, Italy: acl.

Quirk, C., A. Menezes, and C. Cherry. 2005. "Dependency treelet translation: syntactically informed phrasal SMT." *43rd Annual Meeting of the Association for Computational Linguistics: proceedings of the conference*, pp. 271–79. Stroudsburg, Pennsylvania: acl.

Raja, K. Kunjunni. 1963. *Indian theories of meaning*. The Adyar Library Series 91. Reprints, 1969, 1977. Madras: The Adyar Library and Research Center.

Sastri, Gangadhara. 1903. *Tantravārtikam*. Varanasi: Chowkamba Sanskrit Book Depot.

Segall, P., E. Hajičová, and J. Panevová. 1986. *The meaning of the sentence in its semantic and pragmatic aspects*. Berlin; Heidelberg: Springer.

Sleator, D. and D. Temperley. 1993. "Parsing English with a link grammar." *Third International Workshop on Parsing Technologies (IWPT)*, pp. 277–92. Tilburg, The Netherlands; Durbuy, Belgium.

Speijer, J. S. 1886. *Sanskrit syntax*. Leiden: E. J. Brill. [Reprint: Delhi: Motilal Banarsidass, 1973.]

Staal, J. F. 1967. *Word order in Sanskrit and universal grammar*. Foundations of Language, Supplementary Series 5. Dordrecht: D. Reidel Publishing Co.

Tesnière, L. 1959. *Éléments de syntaxe structurale*. Paris: Klinksieck.