



BRILL

OLD WORLD: JOURNAL OF ANCIENT AFRICA AND EURASIA
(2021) 1-32



Evaluating Syntactic Annotation of Ancient Languages

Lessons from the Vedic Treebank

Erica Biagetti

University of Pavia / University of Bergamo, Department of Linguistics,
Pavia, Italy

erica.biagetti01@universitadipavia.it

Oliver Hellwig

University of Zurich, Department of Comparative Linguistics, Center for the
Interdisciplinary Study of Language Evolution / Heinrich Heine University
Düsseldorf, Institute for Language and Information, Zürich, Switzerland

oliver.hellwig@uzh.ch

Salvatore Scarlata

University of Zurich, Department of Comparative Linguistics, Center for the
Interdisciplinary Study of Language Evolution, Zürich, Switzerland

salvatore.scarlata@uzh.ch

Elia Ackermann

University of Zurich, Deutsches Seminar, Zürich, Switzerland

elia.ackermann@uzh.ch

Paul Widmer

University of Zurich, Department of Comparative Linguistics, Center for the
Interdisciplinary Study of Language Evolution, Zürich, Switzerland

paul.widmer@uzh.ch

Abstract

In this paper we introduce an extended version of the Vedic Treebank (VTB, Hellwig et al. 2020) which comes along with revisited and extended annotation guidelines. In order to assess the quality of our annotations as well as the usability and limits of the guidelines we performed an inter-annotator agreement test. The results show that agreement between annotators is hampered by various factors, most prominently by insufficient understanding of the content because of the cultural and temporal gap and incomplete knowledge of Vedic grammar. An in-depth discussion of disagreeing annotations demonstrates that the setup of the workflow, too, has a major influence on inter-annotator agreement. We suggest some measures that can help increase the transparency and annotation consistency according to current knowledge of the language when annotating Vedic Sanskrit, or ancient language varieties in general.

Keywords

Vedic Sanskrit – Treebank – evaluation – Inter-Annotator Agreement (IAA)

1 Introduction

Over the past decades, treebanks have become indispensable tools for studying syntactic and morphological phenomena and for enhancing Natural Language Processing (NLP). While earlier endeavors in annotating syntactic structure were largely confined to modern languages (e.g. the Penn treebank), an increasing number of treebanks of ancient languages has been published in recent years. Our paper follows in the wake of other contributions concerned with the process of building linguistic resources for ancient languages, such as the PROIEL treebanks¹ of early Indo-European languages (Eckhoff et al. 2018a,b), the ITTB² (Passarotti 2019), the Ancient Greek and Latin Dependency Treebank³ (AGLDT 2.0; Bamman & Crane 2011) or, outside of the Indo-European domain, the treebank of Old Chinese⁴ (Yasuoka 2019), and with the potential that annotated corpora have for the study of ancient languages (Eckhoff et al. 2018b).

1 <https://proiel.github.io>.

2 <https://itreebank.marginalia.it>.

3 https://perseusdl.github.io/treebank_data/.

4 <https://www.cs.brandeis.edu/~clp/ctb/>.

The Vedic Treebank⁵ (VTB; Hellwig et al. 2020) is a recent addition to this growing body of syntactically annotated corpora. Vedic Sanskrit (henceforth Vedic) is an ancient Indo-Aryan language, transmitted by a large corpus of poetry and prose texts. The creation of a treebank for an otherwise under-resourced language such as Vedic is highly welcome for various reasons. First, the Vedic corpus plays an important role in historical linguistics because of its extensive and early attestation. Also, Vedic is at the root of a long and influential tradition of linguistic description and analysis. This tradition dates back to the first millennium BCE with scholars as famous as Pāṇini, whose *Aṣṭādhyāyī* can be considered the first comprehensive grammar of the Vedic language. In addition, Vedic is crucial for understanding the early cultural, social, and religious history of South Asia up to the middle of the first millennium BCE, as archaeological and external historical evidence is either largely missing for this period or has not been studied so far (Rau 1983; Witzel 1997).

The goal of this paper is twofold. First, we introduce and describe the VTB version 2.0 and annotation guidelines for Vedic. Second, we illustrate to which degree linguistic, temporal, and cultural distance may impact on the quality of syntactic annotation by presenting the results obtained from performing an inter-annotator agreement test, i.e. a quality control step performed while annotating texts. Based on the outcome of this test, we discuss and suggest measures that may be useful for improving the annotation consistency and transparency when annotating Vedic and ancient languages more generally.

The paper is structured as follows: Since the VTB is annotated according to the Universal Dependencies scheme, in Section 2, we summarize the main assumptions on which this scheme stands. In Section 3, we first address general issues encountered when enriching corpora of ancient languages with syntactic information. In Section 4 we introduce the VTB and describe the annotation process, the size of the annotated corpus, and the genre of the texts it contains. In Section 5 we provide an evaluation of the inter-annotator agreement test performed by three annotators on a random sample of 96 text lines. Section 6, the main part of this paper, presents a qualitative evaluation of the major sources of disagreement. Section 7 contains the conclusion.

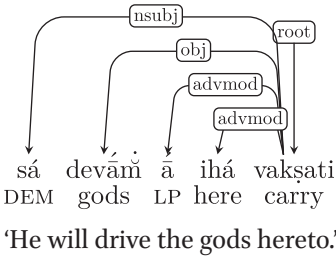
5 https://universaldependencies.org/treebanks/sa_vedic/index.html.

2 Universal Dependencies

The annotation of the VTB is based on the Universal Dependencies 2.0 (UD) annotation scheme. UD is a project that is developing cross-linguistically consistent treebank annotation for many languages (Nivre et al. 2016).⁶

Syntactic annotation in the UD scheme consists of typed dependency relations (*deprel*) between words. The basic representation forms a tree, where exactly one word is the head of the sentence depending on a conventional *root* and all the other word depend on exactly one word, as shown by example (1). An inventory of 40 items contains all possible relations,⁷ to which sub-relations of the form *relation:subrelation* may be added in order to account for language-specific constructions.

(1) RV 1.1.2C



The following principles are observed in the annotation in order to maximize parallelism while accounting for cross-linguistic differences.

Dependency relations hold primarily between content words, rather than being mediated by function words. Thus, adpositions and clitic case markers are treated as dependents of the nouns they attach to, rather than govern it. The same approach is adopted in coordinative constructions, where the leftmost conjunct constitutes the head, while other conjuncts as well as the coordinating conjunction depend on it. Finally, auxiliary verbs – with the copula as a particular case – depend on the lexical predicate, rather than being the head of the clause.

In UD, the treatment of central dependency relations between content words is based on the distinction between core arguments (subjects, objects, clausal complements) and other obliques. Furthermore, the scheme treats nominal phrases, clauses headed by a predicate, and other kinds of modifier

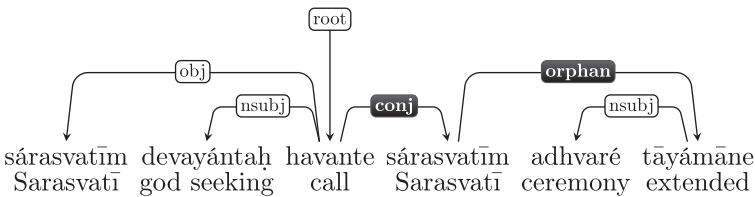
6 The latest version (2.8, released May 15, 2021) consulted during the preparation of this paper includes 202 treebanks of 114 languages.

7 For which see: <https://universaldependencies.org/u/dep/index.html>.

words differently: for example, subjects can take the relation *nsubj* or *csubj* depending on whether they are nominal or clausal subjects.

As a consequence of its lexicalist approach, UD does not make use of empty nodes in order to represent ellipsis. Instead, it marks all kinds of ellipsis by promoting a member of the elliptical clause to the head position on the base of a “coreness” hierarchy.⁸ The promoted member takes the syntactic relation that the elided element would otherwise bear, whereas all non-promoted dependents receive the relation *orphan*. Cf. example (2), which represents the treatment of ellipsis in coordination: because of the elision of the verb *havante* ‘they call’ in the second clause, the object *sárasvatīm* ‘Sarasvatī’ is promoted to the head position of the coordinate clause (*conj*), while the adjunct *tāyámāne* depends on it via the relation *orphan*.

(2) RV 10.17.7: Annotation scheme for verb ellipsis



‘Sarasvatī do those seeking the gods invoke, Sarasvatī. when the ceremony is being extended.’ (Jamison & Brereton 2014)

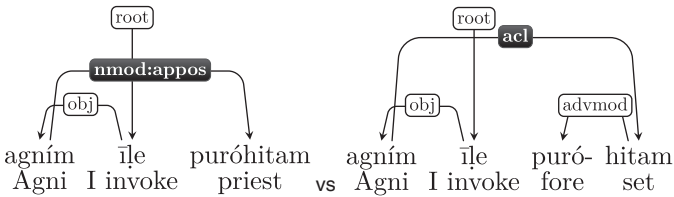
3 Syntactic Annotation of Ancient Languages

Annotating corpora of ancient languages raises several challenges that are usually not encountered when developing treebanks of modern languages. First and foremost, we cannot rely on the intuition of native speakers in order to understand complex passages or rare constructions. With chronologically remote languages such as Vedic, the lack of linguistic intuition has to be substituted by acquired competence, i.e. by makeshift knowledge based on an ongoing individual first-hand confrontation with the texts and on the collective experience of forerunners. In fact, the lead-in in the vast field of Vedic literature – which consists of texts encompassing a span of (estimated) five to seven centuries – is crucial since it surely determines the view a scholar adopts

8 Orphaned dependents are considered for promotion in the following order: *nsubj* > *obj* > *iobj* > *obl* > *advmod* > *csubj* > *xcomp* > *ccomp* > *advcl* > *dislocated* > *vocative* (cf. <https://universaldependencies.org/u/overview/specific-syntax.html#ellipsis-in-clauses>).

in handling other texts. Often, the first immersion into Vedic is the study of the Ṛgveda (RV), a primordial text which undoubtedly irradiates the bulk of the remaining texts. In spite of this evident influence, other Vedic texts do or may belong to different chronological and diatopic strata, to different schools, and to different genres (Witzel 1989, 1995a,b 1997). As a consequence, it is inaccurate and deceptive to simply impose Ṛgvedic grammar on other Vedic texts, and vice-versa. For example, the semantics of *puróhita-*, which in the RV may also be rendered etymologically as ‘set (*hitá-*) ahead (*purás*)’ shifts to ‘priest’ already in the RV and it is sometimes hard to tell which meaning is the intended one in a given passage. The choice made has an immediate bearing on the syntactic annotation, as becomes apparent when annotating the opening line of the RV:

(3) RV 1.1.1a



‘Den Priester Agni preise ich.’ (Grassmann, 1876–77, vol. 11: 1) vs. ‘Agni do I invoke – the one placed to the fore.’ (Jamison & Brereton 2014, vol. 1: 89) vs. ‘Agni berufe ich als Bevollmächtigten.’ (Geldner 1951, vol. 1: 1)

The two trees represent two different interpretations of the three words. The tree on the left, which is based on Grassmann’s translation, interprets *puróhita-* as a nominal apposition of the theonym Agni, whereas the tree on the right, based on Geldner’s translation, interprets it as a secondary predicate of Agni. Since, however, a) *purás* also denotes the cardinal point ‘east’, b) this verse opens the whole RV, and c) Agni, the god of fire, is the deity mentioned first in this text, the message intended by the poet could also have been: ‘I invoke Agni as he has just been installed in the east’ or ‘I (now) invoke Agni because he comes first’. In this situation, the annotators are faced with the ingrate task of choosing but one single option based on their individual understanding of the linguistic and socio-cultural context, being, at the same time, aware that the poet might have fancied to allude to all options reported above.

Partly linked to the temporal gap is the cultural distance, i.e. the considerable amount of uncertainty a modern scholar is confronted with when trying to interpret texts belonging to ancient cultures. This problem again gets more severe the more the cultural backgrounds of composer and annotator differ

and the less is known about the social, cultural and religious background of the text. This lack of knowledge is most pronounced when we deal with the earliest Vedic testimonies which often contain – intentionally or not – very concise ritualistic instructions, philosophical speculations, mythological intimations, and intertextual allusions. The Ṛgvedic hymn 1.164, for example, was probably designed as a riddle hymn that “makes both implicit and explicit reference to Vedic ritual” (Jamison & Brereton 2014, vol. 1: 349), and is therefore “a continuous challenge to students of the Veda” (Houben 2000:499).

Over the centuries, the study of Vedic texts has given rise to a vast number of scholarly publications: Louis Renou’s bibliography already contained more than 300 pages (Renou 1931) and the Vedic Bibliography (Dandekar 1946–2004) amounts to 6000 pages. Nevertheless, it is fair to say that a substantial amount of uncertainty with respect to the content and context of Vedic texts will persist for a long time and impair any linguistic analysis, including syntactic annotation.

4 Description of the VTB 2.0

There exist several web repositories of digitized Vedic texts, most notably the comprehensive TITUS corpus⁹ which provides the texts only, and the Digital Corpus of Sanskrit (DCS).¹⁰ For the VTB, texts were selected from the DCS because this corpus provides *sandhi* splits and morphosyntactic annotations alongside the raw source texts. We selected text passages based on their socio-linguistic importance, considering, at the same time, criteria of chronology and genre as well as the availability of preliminary work such as translations and scholarly commentaries. The annotation was performed directly in the web interface of the DCS which features a supportive, trainable machine learning classifier (see Hellwig et al. 2020 for details). At the moment, four scholars are actively involved in the extension of the VTB. The VTB comes along with detailed annotation guidelines that document the decisions made by the annotators and suggest best-practice solutions.

Our annotations comprise extracts from the following six Vedic texts (also see Table 1):

1. RV: Ṛgveda-Saṃhitā – The RV, a collection of 1028 hymns that mainly address the gods of the Vedic Pantheon, represents the most archaic form of the Vedic language and literature. The text is divided into ten books. While Books I-IX

9 <https://titus.uni-frankfurt.de/indexe.htm>.

10 <http://www.sanskrit-linguistics.org/dcs/index.php>.

Table 1 Annotated passages of the VTB with percentage of annotated text (where available), numbers of tokens and numbers of root nodes (as a rough approximation for the number of sentences)

text	%	# tokens		
		metrical	prose	# root
RV	10.5	17893	0	2360
AV(Ś)	13.5	7243	2306	1373
MS	–	0	3453	634
AB	–	463	10381	1508
ŚB(M)	–	0	1455	213
BĀU(K)	20	0	3814	533
		25599	21409	6639

are generally considered to represent an older stratum of Vedic Sanskrit, book x shows many linguistic and cultural affinities to the presumably younger Atharvaveda-Saṃhitā (see below). See Witzel & Gotō (2007: 427–466) and Jamison & Brereton (2014: 1–83) for detailed descriptions.

2. AV: Atharvaveda-Saṃhitā – The AV which has been transmitted in two recensions (Śaunakīyā and Paippalādā) consists of 730 hymns divided into twenty books. Contrary to the RV, the hymns of the AV focus on warding off immediate dangers in daily life and were therefore often labelled as “magic” in previous research. Two books of the AV (xv and xvi) are composed in a prose akin to that of the Brāhmaṇas. Book xv, which is included in the VTB, contains the earliest known description of the Vṛātyas, a sodality worshipping the Vedic god Rudra (Falk 1986). For some introductory notes to the AV see Bloomfield (1899: 1–15); on the AVP see the Paippalādā recension of the Atharvaveda, Online Edition (beta) (Zehnder et al. 2020-).
3. MS: Maitrāyaṇi-Saṃhitā – The MS is the oldest text of the Black Yajurveda tradition. It contains metrical hymns to be recited during rituals, and their prose explanations, the so called Brāhmaṇas (Amano 2009). The VTB comprises MS 2.5.1–11, a discussion of sacrifices performed in order to obtain “what is wished for” (*kāmyeṣṭi-*; see e.g. Caland 1910).
4. AB: Aitareya- Brāhmaṇa – The AB is one of the two Brāhmaṇas of the Ṛgveda (see e.g. Keith 1920). While it belongs to the oldest layer of the Brāhmaṇas, its prose is presumably slightly younger than that of the MS (Witzel 1995a: 113). The annotation covers AB 1.1–30 and AB 2.1–19, where the performance of the Soma ritual is described as well as the tale of

- Śunaḥśepa (AB 7.13–18), which belongs to a younger layer of Brāhmaṇas prose and is interspersed with metrical parts (*ślokas*).
5. ŚB: Śatapatha-Brāhmaṇa – The ŚB is the most voluminous Brāhmaṇa. It belongs to the school of the White Yajurveda and is passed down in two variants: the Mādhyandina (M) and the Kaṇva (K) recension. The VTB contains the narrative passage ŚB (M) 1.8.1 which relates the tale of Manu and the fish (a flood saga). The excerpt includes dialogues, and thus represents a different style of Vedic prose than the more formal, exegetical style found in the prose of the MS or in the first extract from the AB.
 6. BĀU: Bṛhad-Āraṇyaka-Upaniṣad – The BĀU is an integral part of the ŚB (q.v.) and ranks among the oldest Upaniṣads. It discusses metaphysical aspects (e.g. the Ātman) of the topics dealt with in the ŚB (see Deussen 1897; Olivelle 1998). The VTB comprises the first book of the BĀU in the Kāṇva recension.

5 Quantitative Evaluation

In order to evaluate the annotation process and the consistency of the annotated data we evaluated the inter-annotator agreement. For this task, 96 text lines with 1,886 tokens were randomly drawn from yet unannotated parts of the six texts currently contained in the VTB. These text lines were presented to three experts who are also involved in setting up the annotation process, developing the guidelines and annotating Vedic texts. The annotators could choose whatever linguistic and philological resources they considered necessary. As Vedic Sanskrit has no (unambiguous) clause markers (see the discussion in Section 5.1 below), a text line, i.e. a sequence of words terminated by two vertical strokes (*daṇḍas*), can contain less or more than one sentence, which explains the comparatively high number of words in this sample. Although all sampled texts deal, at the most general level, with the religious sphere and especially the all-important sacrifice, there are marked thematic differences between the oldest testimonies (RV, AVŚ) which often address individual deities, and, for instance, the samples from the much later BĀU, which rather revolves around the philosophical interpretation of the Vedic religion.

For all evaluations, we report the values for the complete annotation ('raw') and those for a post-processed data set ('cleaned'). For the cleaned setting we removed all records which were not annotated by all three annotators. This was necessary because the subdivision of the text into lines separated by *daṇḍas* sometimes returned truncated sentences that could not be annotated.

In addition, some passages could not be interpreted by some or even all (e.g. RV 10.106.8) annotators and were therefore left unannotated.

5.1 *Sentence Segmentation and Boundaries*

Vedic Sanskrit lacks unambiguous grammatical sentence and clause boundary markers (e.g. strict verb position at the periphery, obligatory conjunctions) and a sharp-cut distinction between main and subordinate clauses. In addition, the indigenous tradition partitions the texts with vertical strokes (|, *daṇḍa*) only at higher levels of compositional complexity (books, chapters, paragraphs in prose; stanzas and hemistiches in metrical texts), but does not feature a punctuation system that structures utterances, clauses, sentences and their constituents. For these reasons, sentence-segmentation must be performed manually as part of the annotation process. The uncertainty with respect to segmentation translates into a substantial amount of variation in root assignment and subsequently in dependency and label annotation. Notably, co- and subordination syntax has been dealt with in several voluminous publications, e.g. Minard (1949–1956); Klein (1978b, 1985); Hettrich (1988); Klein (1992); Viti (2007, 2008); Kulikov (2017). The bulk of this work, however, focuses on a single text, the RV, the syntax of which is comparatively idiosyncratic, and much more work is needed to achieve a full understanding of co- and subordination syntax in Vedic.

When evaluating the agreement of the segmentations, we therefore focus on perfect matches between sentences using the Jaccard coefficient as a measure of agreement. Let s_i, s_j denote the sentence boundaries that emerge from the annotations of annotators i and j , and assume that the identity of $s_a \in s_i$ and $s_b \in s_j$ can be ascertained by comparing their start and end indices in the complete data set. The measure of agreement is now defined as:

$$a_{ij} = \frac{|s_i \cap s_j|}{|s_i \cup s_j|}$$

a_{ij} ranges from 0 (no agreement) to 1 (full agreement). The results in Table 2 show a remarkably low agreement in this fundamental task, especially in the overall agreement between all three annotators which oscillates around the value of 0.5. While the qualitative evaluation in Section 5.2 will provide detailed insights into some linguistic structures responsible for the observed low scores, one may hypothesize that the annotation agreement is closely correlated with the lengths of the annotated text lines, because longer lines have a larger number of possible continuous partitions. In order to test this hypothesis, we find those sentences that are demarcated in the same way by all annotators and collect the lengths of text lines containing these sentences in a set

Table 2 Jaccard coefficients for the agreement of sentence segmentations. Columns labelled 1–2, 1–3, 2–3 show the pairwise agreement scores between Annotators 1 and 2, between Annotators 1 and 3, and between Annotators 2 and 3. The column ‘all’ reports the agreement between all three annotators.

	1–2	1–3	2–3	all
raw	0.601	0.595	0.618	0.472
cleaned	0.641	0.661	0.74	0.556

l^+ ; the same is repeated for all sentences not demarcated in the same way by all annotators, resulting in the set of lengths l^- . When comparing the lengths in l^-, l^+ using a t-test (null hypothesis: the numbers in l^- are larger than those in l^+), we obtain a highly significant p-value of 0.00090 in the raw and of 0.00001 in the cleaned setting. This p-value suggests that the lack of a punctuation system and of grammatical markers for sentence boundaries indeed complicates the segmentation of long text lines to a major degree.

The lengths of the text lines are coupled with the metrical structure of the texts and the editorial decisions made when publishing the texts. Table 3 therefore splits the (dis-)agreement about sentence segmentation (columns) by the annotated texts (rows). Column 2 of this table (‘agree’) reports the number of sentences that are segmented in the same way by all three annotators, while column 3 contains the remaining cases. The table shows that we reached a significantly higher agreement¹¹ for the two metrical texts (especially for

Table 3 Numbers of cases in which all three annotators agree (column 2), in which at least one annotator disagrees (column 3) about sentence boundaries, split by the six texts from which samples were drawn.

	Agree	Disagree
AB	28	29
AVŚ	26	20
BĀU	56	51
MS	12	24
RV	29	4
ŚB	43	76

¹¹ We collapse Table 3 into a 2×2 count table with the metrical structure in its two rows (RV, AVŚ → metrical; all others → prose). A χ^2 -test of this table yields a highly significant p-value of $p < 0.0001$.

the RV) than for the four prose texts. This result does not come as a surprise: boundaries of stanzas or hemistiches in the metrical texts often coincide with sentence boundaries since enjambement is rare. Sentences in prose texts are usually longer, more variegated and their boundaries are not systematically constrained by syllable count.

As sentence segmentation turned out to be a source of considerable disagreement, we report a third setting ‘cleaned-sameSeg’ for the evaluation of the actual syntactic annotation in the following Section 5.2. In this setting we only consider those sentences that are demarcated in the same way by all three annotators.

5.2 Syntactic Annotation

Considering that sentence segmentation already poses notable problems, it is not surprising to see that the agreement scores for the actual syntactic annotation are comparatively low. The upper compartment of Table 4 reports Fleiss’ Kappa (Fleiss & Cohen 1973) for the label-only agreement (LOA) where only the labels of tokens, but not their heads are taken into consideration (see e.g. Ragheb & Dickinson 2013). As in Table 2, we report scores for pairs of annotators (‘1–2’ etc.) as well as the overall agreement (‘all’). Using the categorization proposed in Landis & Koch (1977:165), we observe what they call a substantial

Table 4 Label-only (LOA), unlabelled (UAA) and labelled attachment agreement (LAA) for the test annotation.

Setting	1–2	1–3	2–3	“all”
	LOA			
all	0.653	0.646	0.684	0.661
cleaned	0.695	0.722	0.746	0.721
cleaned-sameSeg	0.764	0.757	0.789	0.77
	UAA			
all	0.695	0.709	0.693	0.601
cleaned	0.754	0.807	0.803	0.696
cleaned-sameSeg	0.823	0.823	0.852	0.76
	LAA			
all	0.607	0.606	0.618	0.498
cleaned	0.663	0.692	0.716	0.578
cleaned-sameSeg	0.732	0.722	0.755	0.638

agreement in all settings and for all combinations of annotators, although the actual scores range from 60% to almost 80%.

The following two compartments of Table 4 show the scores for the unlabelled (UAA) and labelled (LAA) attachment agreement (see e.g. Kübler et al. 2009) for which we report the percentual agreement. The overall and pairwise scores in these two settings are significantly lower than those reported for other ancient languages. While we obtain 76.0% of UAA and 63.8% LAA, Bamman et al. (2010), for example, reach 87.4% UAA and 80.6% LAA for ancient Greek, their annotators mostly being graduate students. Zeldes & Abrams (2018) even report 96% UAA on Coptic texts for their expert annotators after inserting punctuation, a score surpassing ours by a margin of 20%.

Having given a quantitative overview of the issues we encountered in the test annotation, we present more detailed descriptions of the major sources of disagreement in the following section.

6 Qualitative Evaluation

The constructions dealt with in this section were chosen by assigning a flag to all words with at least one diverging edge or label annotation. Each of these flags denotes one recurrent type of disagreement (see Table 5, and Table 6 in the Appendix for a detailed description of the flags). From among the most frequent types of disagreement we discuss examples that, in our opinion, are the most representative ones. Note that the selected examples can address multiple types of disagreement. Furthermore, we do not discuss cases of disagreement tagged with the flag 'Inattention' which subsumes annotation errors influenced by non-literal translations or types of disagreement which could not be assigned to any of the other flags.

6.1 Quotes

Exegetical prose texts such as the Brāhmaṇas frequently refer to the four metrical Saṃhitās when discussing aspects of the Vedic ritual. Such quotes are marked with the particle *iti* 'thus, so' which is also used for marking direct speech. In many cases such references indicate which hymn or stanza has to be recited while performing a ritual act. They mostly consist of truncated verbatim quotes, often reporting only the first few words of the cited passage and thus lacking a coherent syntax (see e.g. Apte 1939). In these cases, the citation serves as a mere proxy of a single stanza, a group of stanzas or even a complete hymn. Just like direct speech, the quote may be an integral part of the sentence it is embedded in (*comp*) or it may be syntactically isolated. In the following

Table 5 Flags denoting types of disagreement with numbers of instances

	Flag	#
1	Nominal Sequence	152
2	Quote	140
3	Guidelines	77
4	Coordination	71
5	Ellipsis	69
6	Particle	59
7	Inattention	47
8	Pronoun	47
9	Secondary Predication	42
10	Case	36
11	Subordination	31
12	Core-Oblique	22
13	Reported Speech	20
14	Copula	19
15	Compound	16
16	Root	10
17	Morphology	9
18	Relative Clause	9
19	Local Particle	8

prose passage from a manual of the domestic ritual (Kauśikasūtra 5.8.11), for instance, the word *nissālām* followed by the quotative particle *iti* indicates that the Atharvanic hymn AVŚ 2.14, which begins with the word *niḥsālām*, has to be recited while performing the acts described in the rest of the text line.¹²

- (4) *nissālām ity ulmukena triḥ prasavyaṃ pariharaty anabhipariharan ātmānam*
 ‘(Reciting the mantra beginning with) “*nissālām* (etc.)” he (the sacrificer) carries a fire brand three times (around the sacrificial animal, keeping it) to his left, without carrying it around himself.’

It should be noted that the quoted part (*nissālām*) does not depend on any other element in the sentence: the citation is thus neither syntactically nor semantically connected with the rest of Kauśikasūtra 5.8.11, but just a placeholder for

¹² In our translations, supplements or explanatory notes are given in parentheses (...), textual context in square brackets [...].

a speech act. Our guidelines therefore suggest to systematically label dependencies within quotes of metrical passages as *flat*. However, identifying such quotes is not always easy. Bloomfield's extensive collection of Vedic quotes (Bloomfield 1906; Franceschini & Bloomfield 2007) is an excellent point of reference, but even this list is not exhaustive and a number of passages, which are likely references to mantras cannot be attributed to a source. In Kausikasūtra 5.8.14, for example, the source of the quoted text is unknown:

- (5) *sam asyai tanvā bhavety anyataram darbhā avāsyati*
 '(Saying) 'come together with her body' he throws one of the darbhā blades below (the cow).'

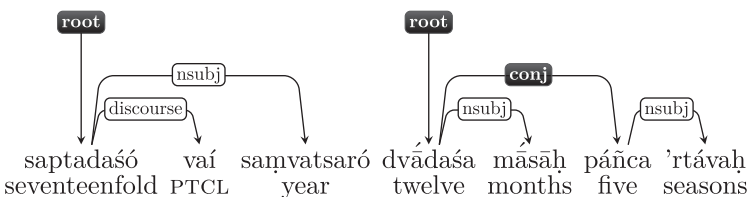
Considering these difficulties, the guidelines leave it open to the annotator to decide whether passages marked with *iti* should be annotated as a case of direct speech or as citations with all internal relations being labelled *flat*.

6.2 Sentence Segmentation and Coordination

In this section, we discuss cases of disagreement subsumed under the flags 'Coordination', 'Subordination', 'Reported Speech', 'Relative Clause', and 'Nominal Sequence' (see the introduction to Section 5), which all eventually lead to differences in sentence segmentation.

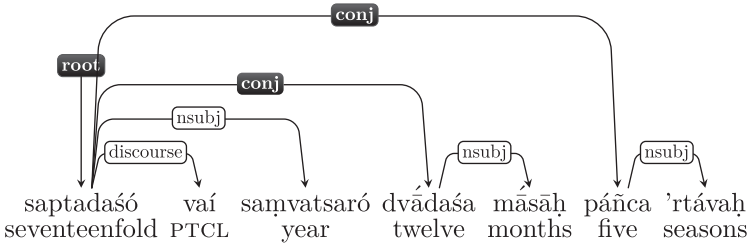
In Vedic, as in many other ancient and modern Indo-European languages, asyndetic coordination is more usual than explicit coordination (Viti 2008: 37). The lack of explicit markers for coordination is one of the most common reasons for disagreement in sentence segmentation, as it is often impossible to decide whether two clauses should be regarded as coordinated asyndetically or as self-standing sentences. This issue is illustrated in example (6):¹³

- (6) ŚB 6.2.2.8



Annot. 2 and 3: 'The year consists of seventeen parts. Twelve are the months and five the seasons.'

13 Divergent labels and edges in the illustrating dependency trees are signaled by a dark box. Occasionally other elements or structures up for discussion are highlighted; these are declared ad loc.

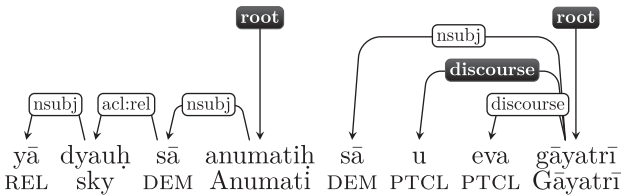


Annot. 1: ‘The year is seventeenfold, the months are twelve, the seasons are five.’

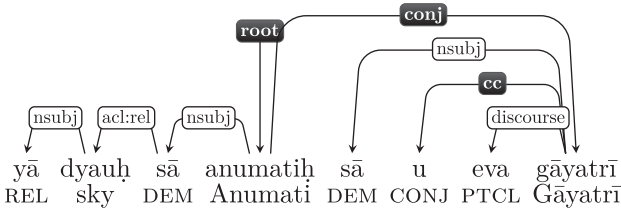
In example (6), the second clause reformulates the first one by making the reason for considering the year seventeen-fold explicit. Since the logical relation between the two clauses is not one of addition, nor is it disjunctive or adversative as is typical of coordinated clauses (Mauri 2008), annotators may easily disagree. Annotators 2 and 3 prefer a two-sentence interpretation, whereas Annotator 1 chooses an asyndetic relation using the label *conj*. The latter is a sensible choice, too, since causal relations between clauses are often expressed with juxtaposition (Viti 2008: 45).

A slightly different source of disagreement is the particle *u*, which can serve as a conjunction or as a discourse marker in Vedic (see Klein 1978a, Klein 1978b, Dunkel 1982–1983:179ff., Dunkel 2014: 335, 882). The ambiguity is mirrored in the annotation of example (7): Annotator 1 understands *u* as introducing the second sentence, whereas Annotators 2 and 3 take it as a coordinating conjunction. The same issue is reported by Eckhoff et al. (2018a) who exemplify it with the Old Church Slavonic and Old East Slavic particle *i* ‘and’.

(7) AB 3.48.3



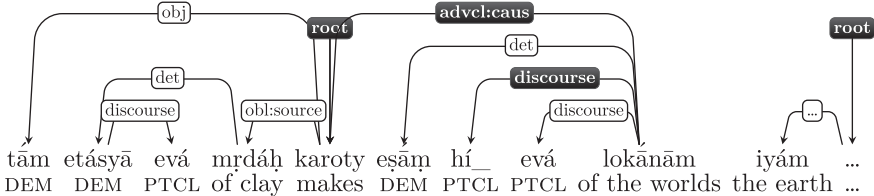
Annot. 1: ‘[To the sky he should offer a pap.] The sky is Anumati. She (Anumati) is also the Gāyatrī (metre).’



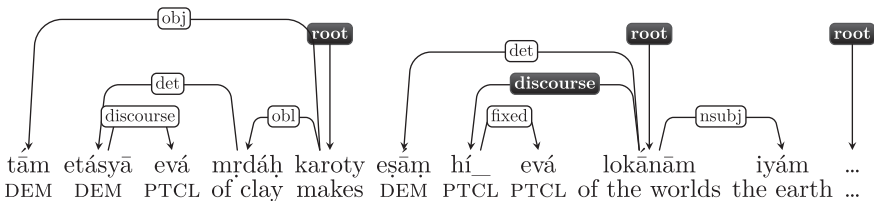
Annot. 2 and 3: ‘The sky is Anumati, and she (Anumati) is also the Gāyatrī.’

The same problems encountered in distinguishing (asyndetic) coordination from independent clauses are sometimes found in the distinction between subordinated and main clauses. While subordinate clauses introduced by relativizers (such as *yād*, *yāthā*, or relative pronouns) are easily detectable, some clauses introduced by particles such as *hī* (causal) and *kuvid* (final) hold an ambiguous status. The ambiguity is caused by the fact that their verb is accented, as in ordinary subordinate clauses, but they can retain illocutionary force like independent ones (see Hettrich 1988, Viti 2007: 134ff., 161ff., Viti 2008:41ff.). As a consequence, annotators can analyze the segments either as independent sentences or as subordinate clauses. This is shown in example (8), which is part of a discussion on how and by whom bricks for building up the fire altar are to be manufactured:

(8) ŚB 6.5.3.1



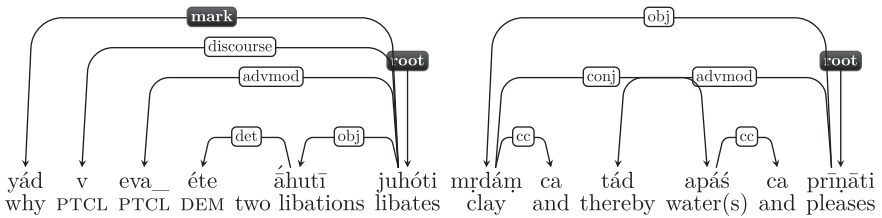
Annot. 1: [‘Of that same (clay) she (the queen) forms the first, the “invincible” (brick); for the invincible one (Aṣḍhā) is this earth, and this earth was created first of these worlds.’] (Eggeling 1982) She forms it of that same clay, for (this earth is one) of these worlds. This earth ...’



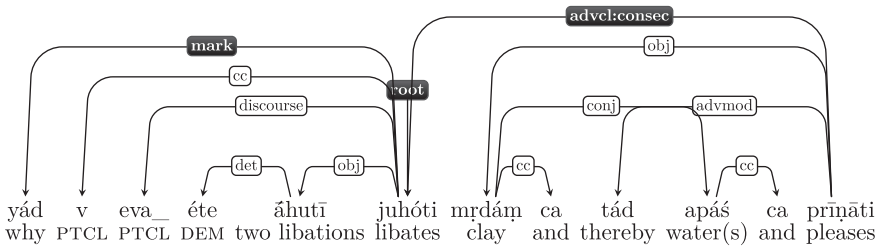
Annot. 2 and 3: ‘She forms it of that same clay. For this earth is (one) of these worlds ...’

The following example (9) presents another kind of asyndetic construction which is especially frequent in exegetical texts such as the ŚB (see Delbrück 1888:581ff., §288) and which shows that relativizers might not always introduce subordinate clauses. The constructions in point consist of two parts: the first is rather short and functions as a sort of introduction or title to a section, typically separated by a *daṇḍa* from a following longer exegesis (Minard 1949–1956, vol. I: 32, §84):

(9) ŚB 6.3.3.17



Annot. 3 (and 1): ‘(The reason) why he offers these two libations – he thereby gratifies both the clay and the waters.’

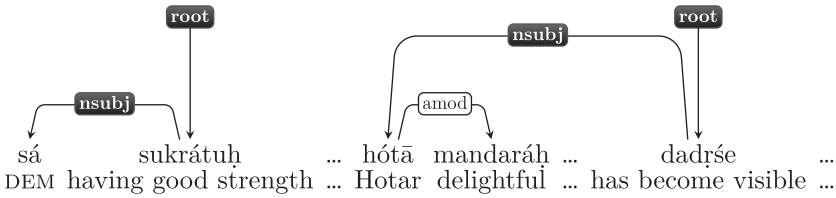


Annot. 2: ‘Why (/that) he then offers these two libations (is that) he thereby gratifies both the clay and the waters.’

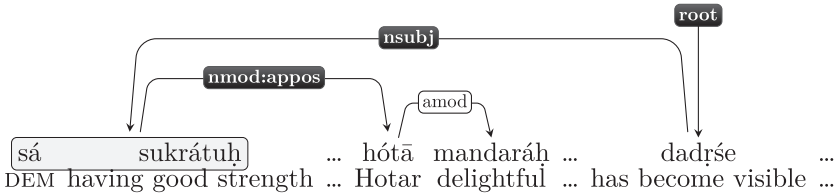
The annotation suggested by Annotator 3 mirrors the original Vedic text and evokes two independent clauses, whereas most translators consider it necessary to connect the clauses by a dash, thus ascertaining a nexus between the two clauses. This nexus is overtly recognized by Annotator 2 who links the two clauses with the relation *advcl*. The relation between the two linked statements is clearly causal; however, since the content of the first clause is not the cause for the content of the second but rather its consequence, Annotator 2 chooses the sublabel *:consec* for the dependency and posits an anticipated or fronted consecutive clause.

Finally, since the copula is not mandatory in Vedic, nominal sequences can also cause disagreement during the segmentation task, as example (10) makes clear:

(10) RV 7.9.2



Annot. 1: 'He is of strong resolve The delighting Hotar has become visible ...'



Annot. 2 and 3: 'He of strong resolve ..., the delighting Hotar, has become visible ...'

Both annotations presented in example (10) are viable. The addressee of the stanza is in any case Agni, the god of fire. Annotators 2 and 3 choose to link all words denoting the god to the first available representative of the subject in the stanza (boxed, gray) – *sá* 'he' (Annotator 2) or *sukrátuḥ* 'of strong resolve' (Annotator 3) –, whereas Annotator 1 splits the sequence into two independent sentences.

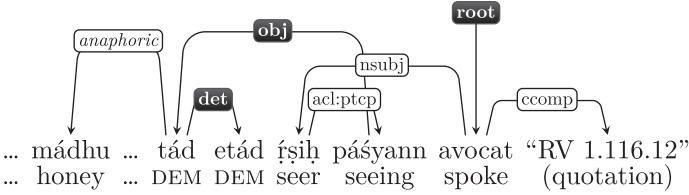
6.3 Edges

Annotators often disagree when there is more than one possible parent for a dependent and neither grammar nor context provide evident clues for which to select. A case in point is disagreement in the annotation of second-position pronominal clitics, corresponding to the flag 'Pronoun' in Table 5. Such clitics can depend on any noun in the clause or on the verb, and alternative interpretations of the text lead to alternative dependencies, all of which are acceptable from the point of view of Vedic grammar. Irrespective of the host, the dependency type may remain unchanged, as it happens for instance with certain discourse markers which can take different scopes.

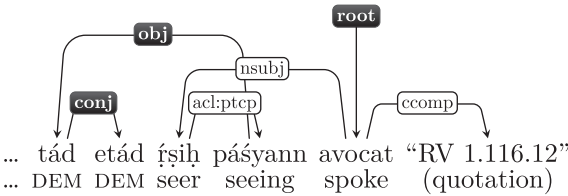
Other multi-functional elements such as demonstrative pronouns are especially subject to various interpretations. In example (11), two neuter forms of the pronouns *tád* and *etád* occur next to each other in sentence-initial position. The former can be analyzed as the anaphoric direct object and the latter as reinforcing *tád* via the relation *det* (Annotator 3). Alternatively, *tád* may be regarded as an adverb 'there, thus' and *etád* as cataphorically referring to the

following quote from the RV (Annotator 2). Finally, the two pronouns may simply be coordinated (Annotator 1).

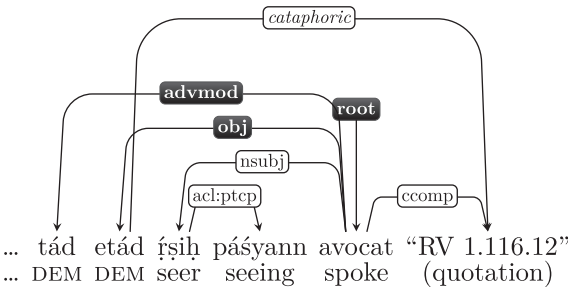
(11) BĀU 2.5.16



Annot. 3: '(This is the same) honey (as Dadhyañc Ātharvaṇa told the Aśvins). Seeing this here, the seer declared: ...'



Annot. 1: 'Seeing this and this here, the seer declared: ...'



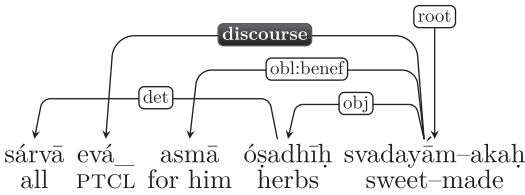
Annot. 2: 'There, seeing (that) the seer spoke this (/as follows): ...'

Disagreement on the scope of particles ('Particle' in Table 5) also belongs in this section. In example (12), Annotators 1 and 2 agree on the relation type (*discourse*) of the particle *evá*, but not on its parent. Annotator 1 makes the particle depend on the main predicate *svadayám akah*¹⁴ '(he) made edible' thus attributing to it an explicative meaning which refers to the preceding subordinate clause ('since he puts in one log, he therefore/for this reason (*evá*) ...'). Annotator 2, instead, attributes an emphatic function to the particle and

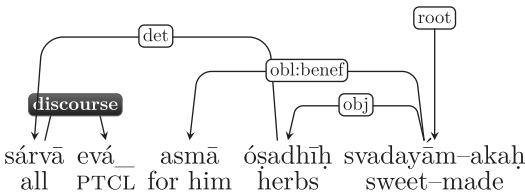
14 The texts report an unintelligible form *svadhayám akah*.

attaches it to the preceding quantifier *sārvāḥ* ‘all’. Finally, Annotator 3 agrees with Annotator 1 on the parent of *evā*, but prefers an adverbial reading for it (*advmod*).

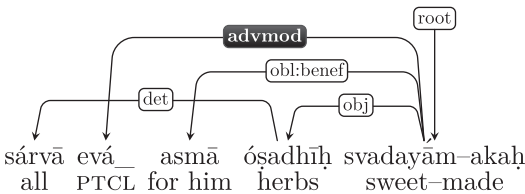
(12) MS 1.8.4.68



Annot. 1: ‘Thereby, he (the priest) has made all the herbs edible for him (the patron).’



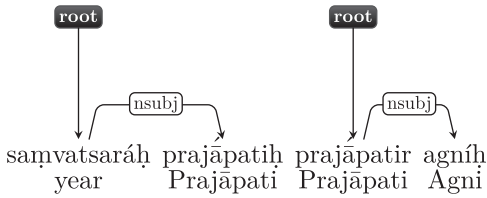
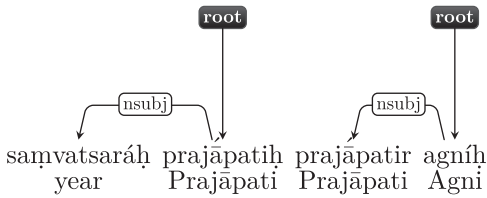
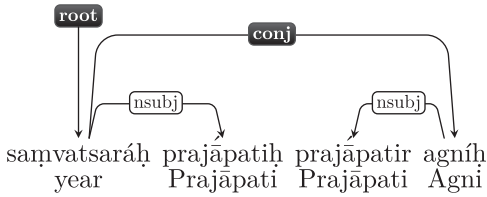
Annot. 2: ‘He has made *all* the herbs edible for him.’



Annot. 3: ‘He has *thus* made all the herbs edible for him.’

Given that in copular clauses the copula can be omitted and the linear order of subject and nominal predicate is not fully constrained (an issue already encountered in example (6)), the sequence in example (13) can be highly ambiguous if taken out of context (see ‘Nominal Sequence’ in Table 5). Annotator 1 interprets *saṃvatsarāḥ* ‘year’ as the predicate of the first instance of the theonym *prajāpatiḥ* lit. ‘lord of offspring’, and then follows the same order of predicate and subject in the remainder of the sequence. Annotator 2, instead, annotates the inverted order throughout. Finally, Annotator 3 has a fixed subject *prajāpatiḥ*, and opts for an asyndetic coordination of the two predications (see Section 6.2). Even with more context at hand such identifying equations remain challenging to interpretation and analysis (see Witzel 1996 and Gren-Eklund 1978).

(13) ŚB 6.2.2.8

**Annot. 1:** 'Praṇāpati is the year. Agni is Praṇāpati.'**Annot. 2:** 'The year is Praṇāpati. Praṇāpati is Agni.'**Annot. 3:** 'Praṇāpati is the year, and Praṇāpati is Agni.'

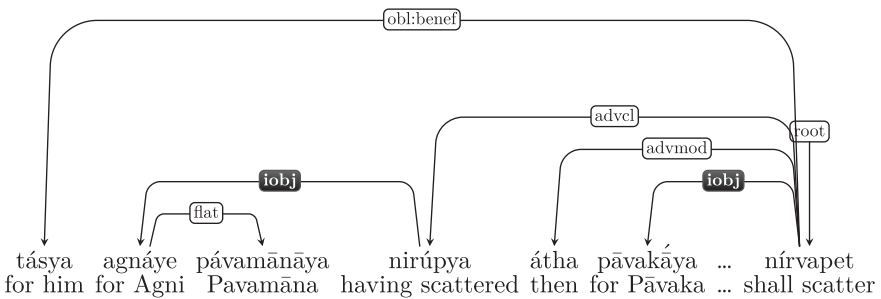
6.4 Labels

Gaps in the grammatical description of Vedic often lead to disagreement when labelling edges. For example, despite recent progress (Hettrich 2007), our knowledge of Vedic argument frames is still incomplete. Discrepancies in the annotation due to such difficulties are subsumed under the flag 'Core-Oblique' in our sample (Table 5).

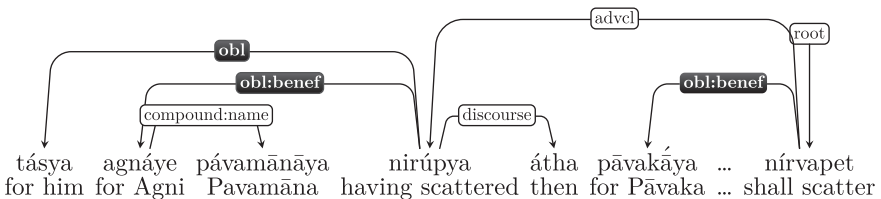
Preverb-verb combinations go through a process of grammaticalization and lexicalization in Vedic, whose stages are witnessed in the texts contained in our sample. Often, the presence of a preverb determines changes in a verb's valency and argument structure and therefore affects dependencies and relation types. For instance, the transitive verb *as-* 'throw, cast something (Acc)' gains a semantic specification of directionality when a preverb is attached and thus becomes trivalent (Casaretto & Schneider 2015:244–245). In some cases

the situation is even more complex. A case in point is the compound verb *nir-vap-* lit. ‘scatter off’, which Hettrich (2007) defines as a trivalent verb with the argument structure NomAccLoc in the RV. However, this compound verb seems to develop a less-compositional meaning ‘to offer, to distribute’ in later Vedic and to change its argument structure into NomAccDat (see Böhlingk & Roth 1855–1875 s.v. *vap + nis*), as can be observed in example (14). Depending on whether the annotators include the datives *agnāye pāvamānāya* ‘to the purifying Agni’ and *pāvakāya* ‘to the purifying one’ in the verb valency or not, they label them as indirect objects (*iobj*) or obliques beneficiaries (*obl:benef*).

(14) MS 1.6.8.14



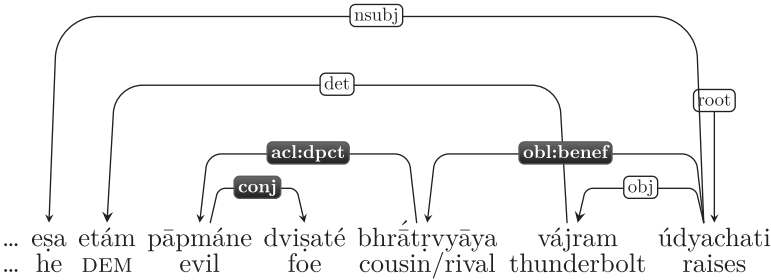
Annot. 2: ‘For the benefit of him (the patron) he (the priest) shall then, after having scattered (a sacrificial cake) for Agni Pāvamāna, scatter (two more) for (Agni) Pāvaka (and Agni Śuci).’



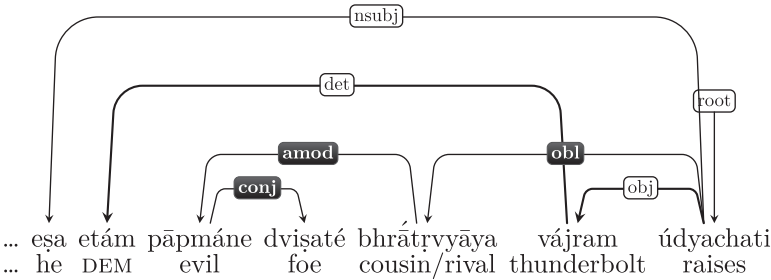
Annot. 3: ‘After having scattered (a sacrificial cake) to Agni Pāvamāna for the benefit of him (patron), he (the priest) shall then scatter to (Agni) Pāvaka (and Agni Śuci).’

Label differences also arise as a consequence of our imperfect understanding of word order constraints within nominal expressions (see Section 6.3). Take for instance example (15) in which the famous myth of Indra’s killing of the dragon Vṛtra (see e.g. Watkins 1995; Witzel 2008) is evoked to explain a ritual action:

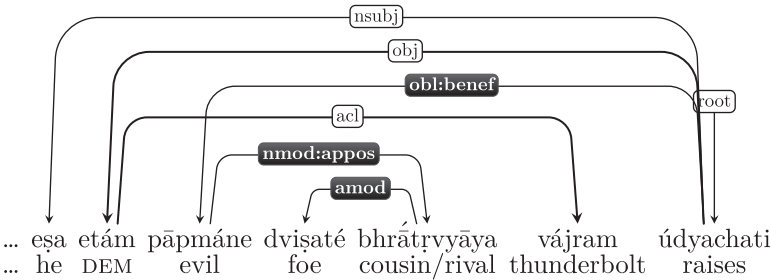
(15) ŚB 1.2.4.3



Annot. 1: ‘[Now when he takes up the wooden sword,] he raises that thunderbolt against a cousin, if he is wicked and spiteful, [just as Indra once raised the thunderbolt against Vṛtra: that is the reason why he takes the wooden sword.]’



Annot. 3: ‘[Now when he takes up the wooden sword,] he raises that thunderbolt (Vajra) against a wicked and spiteful rival, [just as Indra once raised the thunderbolt against Vṛtra: that is the reason why he takes the wooden sword.]’



Annot. 2: ‘[Why he takes up the wooden sword: just as Indra once raised the Vajra against Vṛtra, in the same way] he raises that (sword) as if it were (Indra’s) Vajra against the evil that consists in a spiteful rival. [That is the reason why he takes the wooden sword.]’

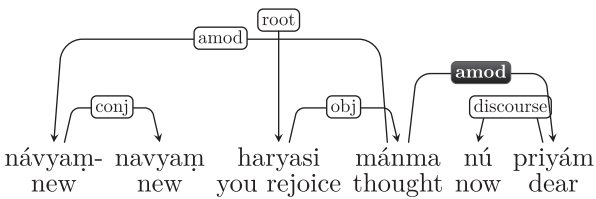
The differences in the annotations are mainly due to the differing interpretations of two of the three nominals in the dative case, namely *pāpmáne* and

dviṣaté. The lemma *pāpmán-* m. (*pāpmáne*) is often used in apposition to other nouns and is thus easily interpreted as an adjective as in the case of Annotator 3. On the other hand, the lemma *dviṣánt-*, while formally a participle ('hating, spiteful'), has been lexicalized as a substantive noun already in the RV ('enemy, foe'), a fact overlooked by all annotators, but reported e.g. by Grassmann (1873–75:652, s.v. *dviṣ*) and Lowe (2015:260ff.).

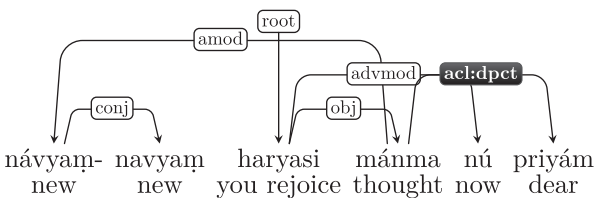
A second difference (marked with thick edges lines) emerges in the handling of the sequence *etám ... vájram*. While Annotators 1 and 3 take the pronoun *etám* as a determiner of the object *vájram*, Annotator 2 reads it as referring anaphorically to the word *sphyám* 'wooden sword' in the preceding clause, and interprets *vájram* as a depictive of this pronoun.

Finally, annotators often disagree on the function of adjectives, compounds, and participles ('Secondary Predication' in Table 5). For Vedic, unambiguous grammatical cues (e.g. agreement, word order) are lacking or not yet well understood and the context alone often does not constrain annotation options. Thus, in example (16), which is taken from a Ṛgvedic hymn addressing the god Indra, the adjective *priyám* 'dear' can be interpreted as an attribute of *mánma* 'thought' (label *amod*) or, alternatively, as a depictive secondary predicate (label *acl:dpct*; Schultze-Berndt & Himmelmann 2004; Himmelmann & Schultze-Berndt 2005; Casaretto 2020) meaning that Indra, the addressee of the hymn, does not generally rejoice at every thought, but only *when* the thoughts are dear to him.

(16) RV 10.96.11b



Annot. 1: 'You now rejoice in every new thought, if it is dear (to you).'



Annot. 2 (and 3): 'You rejoice in every new thought now dear (to you).'

7 Summary and Outlook

In this paper, we introduced and described the VTB version 2.0. In addition, we reported the results obtained from a sanity check aiming at assessing the usability of the VTB guidelines and the consistency of the annotation in order to improve the overall procedure and, ultimately, the quality of the annotation.

As trivial as it may seem, the evaluation of the inter-annotator agreement test illustrates perfectly well that the overall quality of the annotation is basically driven by the expertise of the annotator, the familiarity with the text, the availability of resources (grammars, editions, translations, concordances, commentaries, detailed guidelines, etc.), and time restrictions. In addition to these merely annotator- and text-specific factors, the evaluation of the procedural set up revealed several suggestions for improvement which apply to the syntactic annotation of Vedic Sanskrit as well as of other ancient languages.

First and foremost, we have seen in Section 5.1 that an important amount of disagreement among annotators is due to segmentation issues. We therefore suggest an additional step during preprocessing in which the text is segmented in a consistent manner whenever possible. If the criteria adopted in this step are made explicit in the treebank documentation, a more consistent segmentation would also help future users to make more targeted queries.

As dealing with a random sample has shown, annotating sentences out of context leads to an increase of effort and a decrease of quality. Therefore, it is important to annotate large continuous portions of text incrementally. Furthermore, availability of additional resources has a positive effect when working on ancient languages: we recommend starting the annotation with well-studied texts for which reliable translations are available and to profit from other available resources such as dictionaries and commentaries. In fact, among the samples taken from the six texts, portions drawn from the RV are the one in which we reached the highest inter-annotator agreement score. This is certainly due to the high number of studies this text enjoys, compared to the Atharvaveda and to prose texts.

Thorough documentation is crucial, and the annotation guidelines should be as explicit as possible and draw the attention of the annotator to gaps and issues in the description of the language. For obvious reasons, this is of great help for users, who will know what to expect from the resource, as well as for developers willing to take up the torch and extend the VTB or build from scratch another treebank. In fact, multiplication of effort is well attested in this domain and in many cases a detailed documentation would have helped to forestall this.

Summing up, the acquired knowledge gained by performing the IAA-test is certainly a significant input to be taken into account in order to improve our guidelines and to work on a new version of the treebank building on firmer bases.

Finally, we hope that ours can be a useful contribution for all those who want to approach the study of Vedic grammar from a new perspective. Indeed, opting for an annotation process that takes into account the extraordinary long philological tradition Vedic studies rely on, we hope that our work will help to bridge traditional philology and the most recent methods adopted by computational linguistics.

References

- Amano, Kyoko. 2009. *Maitrāyaṇī Saṃhitā I–II. Übersetzung der Prosapartien mit Kommentar zur Lexik und Syntax der älteren vedischen Prosa*. Bremen: Hempen.
- Apte, Vaman Shivaram. 1939. The Ṛgveda Mantras in their Ritual Setting in the Grhya Sūtras. *Bulletin of the Deccan College Post-Graduate and Research Institute* 1(1):14–44.
- Bamman, David, Francesco Mambrini & Gregory Crane. 2010. An Ownership Model of Annotation: The Ancient Greek Dependency Treebank. In *Proceedings of the Eighth International Workshop on Treebanks and Linguistic Theories (TLT8)*, Marco Carlo Passarotti, Adam Przepiórkowski, Savina Raynaud, & Frank Van Eynde (eds.), Milan, Italy, https://www.researchgate.net/publication/242088885_An_Ownership_Model_of_Annotation_The_Ancient_Greek_Dependency_Treebank.
- Bamman, David & Crane, Gregory. 2011. The Ancient Greek and Latin Dependency Treebank. In *Language Technology for Cultural Heritage*, Caroline Sporleder, Antall van Den Bosch & Kalliopi Zervanou (eds), 79–89. Berlin: Springer.
- Bloomfield, Maurice. 1899. The Atharvaveda. *Grundriss der Indo-Arischen Philologie und Altertumskunde*, II. Band, 1. Heft B, Strassburg: Verlag von Karl J. Trübner.
- Bloomfield, Maurice. 1906. *A Vedic concordance, being an alphabetic index to every line of every stanza of the published Vedic literature and to the liturgical formulas thereof, that is, an index to the Vedic mantras, together with an account of their variations in the different Vedic books*. Cambridge, Mass: Harvard University Press.
- Böhtlingk, Otto & Rudolph Roth. 1855–1875. *Sanskrit-Wörterbuch*. St. Petersburg: Kaiserliche Akademie der Wissenschaften, Eggers.
- Caland, Willem. 1910. *Altindische Zauberei*. Amsterdam: Johannes Müller.
- Casaretto, Antje. 2020. On Secondary Predicates in Vedic Sanskrit – Syntax and Semantics. *International Journal of Diachronic Linguistics and Linguistic Reconstruction* 17:1–63.
- Casaretto Antje & Caroline Schneider. 2015. Vedic local particles at the syntax-semantics interface. In *Language change at the syntax-semantics interface*, Chiara Gianollo, Agnes Jäger & Doris Penka (eds), Berlin: de Gruyter, pp 223–259.
- Dandekar, Ramchandra Narayan. 1946–2004. *Vedic bibliography*. Bombay: Karnatak Publ. House.

- Delbrück, Bertold. 1888. *Altindische Syntax*. Halle: Verlag der Buchhandlung des Waisenhauses.
- Deussen, Paul. 1897. *Sechszig Upanishad's des Veda*. Leipzig: Brockhaus.
- Dunkel, George E. 1982–1983. IE conjunctions: pleonasm, ablaut, suppletion. *Zeitschrift für vergleichende Sprachforschung* 96:178–199.
- Dunkel, George E. 2014. *Lexikon der indogermanischen Partikeln und Pronominalstämme*, LIPP. Heidelberg: Winter.
- Eckhoff, Hanne Martine, Kristin Bech, Gerlof Bouma, Kristine Eide, Dag Haug, Odd Einar Haugen, & Marius Jøhndal. 2018a. The PROIEL Treebank Family: A Standard for Early Attestations of Indo-European Languages. *Language Resources and Evaluation* 52:29–65.
- Eckhoff, Hanne Martine, Silvia Luraghi & Marco Passarotti. 2018b. The added value of diachronic treebanks for historical linguistics. *Diachronica* 35:3:297–309.
- Eggeling, Julius. 1882–1900. *The Satapatha-Brāhmana*. Oxford: Clarendon.
- Falk, Harry. 1986. *Bruderschaft und Würfelspiel*. Untersuchungen zur Entwicklungsgeschichte des vedischen Opfers. Freiburg: Hedwig Falk.
- Fleiss, Joseph L & Jacob Cohen. 1973. The Equivalence of Weighted Kappa and the Intraclass Correlation Coefficient as Measures of Reliability. *Educational and Psychological Measurement* 33(3):613–619.
- Franceschini Marco, Maurice Bloomfield. 2007. *An updated Vedic concordance: Maurice Bloomfield's A Vedic concordance enhanced with new material taken from seven Vedic texts*. Cambridge, Mass.: Harvard University Press.
- Geldner, Karl Friedrich. 1951. *Der Rigveda. Aus dem Sanskrit ins Deutsche übersetzt und mit einem laufenden Kommentar versehen*. Cambridge, Mass.: Harvard University Press.
- Grassmann, Hermann. 1873–75. *Wörterbuch zum Rig-Veda*. Leipzig: Brockhaus.
- Grassmann, Hermann. 1876–77. *Rig-veda. Übersetzt und mit kritischen erläuternden Anmerkungen versehen*. Leipzig: Brockhaus.
- Gren-Eklund, Gunilla. 1978. *A Study of Nominal Sentences in the Oldest Upaniṣads*. (Stockholm), Uppsala: Almqvist and Wiksell.
- Hellwig, Oliver, Scarlata, Salvatore, Ackermann, Elia & Widmer, Paul. 2020. The Treebank of Vedic Sanskrit. In Proceedings of The 12th Language Resources and Evaluation Conference (LREC 2020), Nicoletta Calzolari, Frederic Bechet, Philippe Blache, Khalid Choukri, Christopher Cieri, Thierry Declerck, Sara Goggi et al. (eds.), 5137–5146.
- Hettrich, Heinrich. 1988. *Untersuchungen zur Hypotaxe im Vedischen*. Berlin: de Gruyter.
- Hettrich Heinrich. 2007. *Materialien zu einer Kasussyntax des Rgveda*. Universität Würzburg, Institut für Altertumswissenschaften, Lehrstuhl für Vergleichende Sprachwissenschaft, Würzburg.
- Himmelmann Nikolaus P. & Eva F. Schultze-Berndt. 2005. Issues in the syntax and semantics of participant-oriented adjuncts. In *Secondary Predication and Adverbial*

- Modification*, Nikolaus P. Himmelmann & Eva F. Schultze-Berndt (eds). Oxford & New York: Oxford University, pp 1–67.
- Houben, Jan E. M. 2000. The Ritual Pragmatics of a Vedic Hymn: The “Riddle Hymn” and the Pravargya Ritual. *Journal of the American Oriental Society* 120.4:499–536.
- Jamison Stephanie W. & Joel P. Brereton. 2014. *The Rigveda*. New York: Oxford University Press.
- Keith, Arthur Berriedale. 1920. *Rigveda Brahmanas: The Aitareya and Kauṣītaki Brāhmaṇas of the Rigveda*. Cambridge, Mass.: Harvard University Press.
- Klein, Jared. 1978a. The Diachronic Syntax of the Particle *u* in the Rigveda. *Journal of the American Oriental Society* 98:266–276.
- Klein, Jared. 1978b. *The particle “u” in the Rigveda*. Göttingen: Vandenhoeck & Ruprecht.
- Klein, Jared. 1985. *Toward a Discourse Grammar of the Rigveda*, vol 1 and 11. Heidelberg: Carl Winter Universitätsverlag.
- Klein, Jared. 1992. Some Indo-European Systems of Conjunction: Rigveda, Old Persian, Homer. *Harvard Studies in Classical Philology* 94:1–51.
- Kübler, Sandra, Ryan McDonal & Joakim Nivre. 2009. *Dependency Parsing*. Morgan & Claypool Publishers.
- Kulikov, Leonid. 2017. The syntax of Indic. In *Handbook of Comparative and Historical Indo-European Linguistics*, Jared Klein, Brian Johseph & Matthias Fritz (eds), no. 1 in Handbücher zur Sprach- und Kommunikationswissenschaft, HSK 41 Berlin: de Gruyter, pp 377–409.
- Landis, J. Richard & Gary G. Koch. 1977. *The Measurement of Observer Agreement for Categorical Data*. *Biometrics* 33(1):159–174.
- Lowe, John J. 2015. *Participles in Rgvedic Sanskrit*. Oxford: Oxford University Press.
- Mauri, Caterina. 2008. *Coordination relations in the languages of Europe and beyond*. Cambridge, Mass: de Gruyter.
- Minard, Armand. 1949–1956. *Trois énigmes sur les Cent chemins*. Paris: Belles-lettres.
- Nivre, Joakim, Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Jan Hajič, Christopher D. Manning, Ryan McDonald, Slav Petrov, Sampo Pyysalo, Natalia Silveira, Reut Tsarfaty & Daniel Zeman. 2016. Universal dependencies vi: A multilingual treebank collection. In *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC 2016)*.
- Olivelle, Patrik. 1998. *The early Upaniṣads: annotated text and translation*. Oxford: Oxford University Press.
- Passarotti, Marco. 2019. The Project of the Index Thomisticus Treebank. In *Digital Classical Philology. Ancient Greek and Latin in the Digital Revolution*, Monica Berti (eds.), no. 10 in Grundfragen der Informationsgesellschaft. Berlin: de Gruyter, pp 299–319.
- Ragheb, Marwa & Markus Dickinson. 2013. Inter-annotator Agreement for Dependency Annotation of Learner Language. In *Proceedings of the Eighth Workshop on Innovative Use of NLP for Building Educational Applications*, Joel Tetreault, Jill Burstein & Claudia Leacock (eds.), pp 169–179.

- Rau, Wilhelm. 1983. *Zur vedischen Altertumskunde*. Wiesbaden: Steiner.
- Renou, Louis. 1931. *Bibliographie Védique*. Paris: Librairie d'Amérique et d'Orient.
- Schultze-Berndt Eva F. & Nikolaus P. Himmelmann 2004. Depictive secondary predicates in crosslinguistic perspective. *Linguistic typology* 8(1):59–131.
- Viti, Carlotta. 2007. *Strategies of Subordination in Vedic*. Milano: Angeli.
- Viti, Carlotta. 2008. The meanings of coordination in the early Indo-European languages. *Revue de sémantique et pragmatique* 24:35–64.
- Watkins, Calvert. 1995. *How to Kill a Dragon*. New York: Oxford University Press.
- Witzel, Michael. 1989. Tracing the Vedic dialects. In *Dialectes dans les littératures indoaryennes*, Michael Witzel & Colette Caillat (eds). Paris: Collège de France, pp 97–265.
- Witzel, Michael. 1995a. Early Indian history. In *The Indo-Aryans of Ancient South Asia*, George Erdosy (ed.). Berlin: de Gruyter, pp 85–125.
- Witzel, Michael. 1995b. Ṛgvedic history. In *The Indo-Aryans of Ancient South Asia*, George Erdosy (ed.). Berlin: de Gruyter, pp 307–352.
- Witzel, Michael. 1996. How To Enter The Vedic Mind? Strategies In Translating A Brāhmaṇa Text. In *Translating, Translations, Translators from India to the West*, Enrica Garzilli (ed). Harvard Oriental Series, Opera Minora 1, Cambridge, Mass: Harvard University Press, pp 163–176, URL http://crossasia-repository.ub.uni-heidelberg.de/109/1/How_to_Enter_1996.pdf.
- Witzel, Michael. 1997. The Development of the Vedic Canon and its Schools. In *Inside the Texts, Beyond the Texts*, South Asia Books, Michael Witzel (ed). Columbia: MO, pp 257–345.
- Witzel, Michael. 2008. Slaying the Dragon in Eurasia. In *In Hot Pursuit of Language in Prehistory. Essays in the Four Fields of Anthropology*, John D. Bengtson (ed.). Amsterdam: John Benjamins Publishing Company, pp 263–286.
- Witzel, Michael & Toshifumi Gotō. 2007. *Rig-Veda. Das heilige Wissen. Erster und zweiter Liederkreis*, vol 1. Frankfurt am Main/Leipzig: Verlag der Weltreligionen.
- Yasuoka, Koichi. 2019. Universal Dependencies Treebank of the Four Books in Classical Chinese. *10th International Conference of Digital Archives and Digital Humanities* pp 20–28.
- Zehnder, Thomas, Robert Leach, Oliver Hellwig, Aneglrika Malinar & Paul Widmer. 2020–. *Paippalāda Recension of the Atharvaveda*. URL <http://www.atharvaveda-online.uzh.ch>.
- Zeldes, Amir & Mitchell Abrams. 2018. The Coptic Universal Dependency Treebank. In *Proceedings of the Second Workshop on Universal Dependencies (UDW 2018)*, Marie-Catherine de Marneffe, Teresa Lynn & Sebastian Schuster (eds.). Brussels: Association for Computational Linguistics, pp 192–201. DOI 10.18653/v1/W18-6022, URL <https://www.aclweb.org/anthology/W18-6022>.

Appendix

Table 6 Disagreement types reported in Section 6

Flag	Description
1 Nominal Sequence	The annotators disagree on the choice of the head for a sequence of nominals.
2 Quote	The annotators disagree on the annotation of quotes: e.g. use of <i>flat</i> and <i>parataxis</i> .
3 Guidelines	The annotators give the same interpretation to a passage but annotate it differently due to: <ol style="list-style-type: none"> a. gaps in the guidelines, b. guidelines not properly followed.
4 Coordination	The annotators segment the passage differently due to fuzzy boundaries between independent elements and asyndetic coordination (both at clause and at phrase level).
5 Ellipsis	The annotators deal differently with ellipsis due to: <ol style="list-style-type: none"> a. problems in determining the element that should be promoted; b. disagreement on the annotation of extra-clausal elements such as vocatives; c. counterintuitive annotation of leftward gapping, which is nevertheless linguistically realistic.
6 Particle	The annotators disagree on the function of particles: e.g. <i>discourse</i> vs. <i>advmod</i> . Sometimes related to the issue of sentence boundaries: independent clauses vs. asyndetic coordination (e.g. <i>u</i>), independent clauses and subordination (e.g. <i>hî</i>).
7 Inattention	Comprises errors influenced by non-literal translations and types of disagreement which could not be subsumed under any of the other flags.
8 Pronoun	Annotators disagree on the function of pronouns (e.g. <i>det</i> vs. head of the noun phrase, with noun attached as <i>appos</i> or <i>nmod:appos</i>).
9 Secondary Predication	Annotators disagree on the interpretation of a noun, adjective, compound, or participle as an attributive modifier or a secondary predicate: e.g. <i>amod</i> vs. <i>acl:attr</i> vs. <i>acl:dpct</i> , etc.

Table 6 Disagreement types reported in Section 5

Flag	Description
10 Case	Annotators disagree a) on the case of a form (e.g. clitic personal pronouns with syncretic genitive, dative, and accusative forms), b) on which function to attribute to a given case form (e.g. dative as <i>advcl:fin</i> or <i>obl:benef</i>); not to confound with Core-Oblique (see below).
11 Subordination	Annotators disagree on the function of a subordinate clause (e.g. <i>ccomp</i> , <i>xcomp</i> , <i>advcl</i>). In some cases, this flag is a counterpart, at clause level, of the Core-Oblique flag (<i>ccomp</i> vs. <i>advcl</i>); fuzzy boundaries between independent sentences and subordination (e.g. <i>hí</i> -clauses) also belong here.
12 Core-Oblique	Annotators disagree on whether a dependent is a core or oblique argument of the verb.
13 Reported Speech	Annotators disagree on the annotation of reported speech; not to confound with Quote, since reported speech and quotes from other texts are treated differently in the VTb.
14 Copula	Annotators disagree on whether the verb <i>bhū-</i> 'be, become' should be treated as copula (<i>cop</i>) or as a lexical verb (<i>root</i>).
15 Compound	Annotators disagree by interpreting compound syntax differently: e.g. the first element can be <i>nmod</i> , <i>obj</i> , <i>obl</i> etc. depending on the nature of the second element (note that, differently from UD, relations between compound elements are analyzed in the VTb).
16 Root	Annotators disagree on the choice of the head; in case of nominal sentences the appropriate flag is Nominal Sequence.
17 Morphology	Only one or two of the annotators noticed an error in the morphological annotation and decided to leave the token unannotated.
18 Relative Clause	Annotators disagree either on the function of the relative pronoun within the relative clause or on the function of the relative clause within the sentence; can be assimilated to Subordination, but also includes Pronouns.
19 Local Particle	Annotators disagree on the syntactic function of a local particle and annotate it either as <i>advmod</i> depending on a verb or as <i>case</i> depending on a noun.