Hedging in diachrony: the case of Vedic Sanskrit *iva*

Erica Biagetti
Department of Humanities
Section of Linguistics
University of Pavia
erica.biagetti@unipv.it

Oliver Hellwig
Institute for Linguistics,
H.-Heine-Univ. Düsseldorf
Dep. of Comp. Lang. Science,
University of Zurich
Oliver.Hellwig@hhu.de

Sven Sellmer
Institute for Linguistics,
H.-Heine-Univ. Düsseldorf
Institute of Oriental Studies,
A. Mickiewicz Univ. Poznań
sellmer@hhu.de

Abstract

The rhetoric strategy of hedging serves to attenuate speech acts and their semantic content, as in English 'kind of' or 'somehow'. While hedging has recently met with increasing interest in linguistic research, most studies deal with modern languages, preferably English, and take a synchronic approach. This paper complements this research by tracing the diachronic syntactic flexibilization of the Vedic Sanskrit particle *iva* from a marker of comparison ('like') to a full-fledged adaptor. We discuss the outcomes of a diachronic Bayesian framework applied to *iva* constructions in a Universal Dependencies treebank, and supplement these results with a qualitative discussion of relevant text passages.

1 Introduction

Hedging is a rhetorical strategy by which a speaker can attenuate either the full semantic membership of an expression, as in example (1a) (propositional hedging), or the force of a speech act, as in (1b) (speech act hedging; Fraser, 2010, 22).

(1) a. *The pool has* ***sort of*** *a L-shaped design.*

b. ***I guess*** *I should leave now.*

Until recently,[1] hedges were considered to be marginal items that contribute little to communication, but their crucial role both in spoken and in written speech is now generally acknowledged by various linguistic disciplines (Kaltenböck et al., 2010, 1).

From the point of view of grammaticalization and/or pragmaticalization studies, hedges are interesting because they have been proven to emerge from different sources both intra- and cross-linguistically (Mihatsch, 2010). Furthermore, although the distinction between propositional and speech act hedges is commonly accepted, we often witness the emergence of speech act hedging as implicature of propositional hedging and vice versa (Mihatsch, 2010, 94; Kaltenböck et al., 2010).

Despite the abundance of studies on hedging in modern languages, the phenomenon has still been little studied in ancient languages, one exception being an in-depth analysis of the use of the approximation marker *hōs épos eipeîn* 'so to say' in Plato's *Gorgias* (Caffi, 2010). However, ancient languages that enjoy centuries of attestation and that have been handed down to us in sufficiently large corpora provide a privileged point of view for the study of hedging, because they allow us to trace the emergence of new hedges as well as the successive development of new functions. With centuries of attestation and an extant corpus of over three million tokens, Vedic Sanskrit (henceforth Vedic) is one such language. In this paper, we investigate the development of the particle *iva*, which in Vedic functions both as a comparison and an approximation marker. After summarizing the grammaticalization process that in the earliest texts lead to the development of the approximative function from the comparative one, we perform a quantitative analysis of the occurrences of *iva* in Vedic texts, in order to assess whether the particle underwent further syntactic and pragmatic developments.

Given the pragmatic nature of the phenomenon, corpus-based approaches are most suited for the study of hedging, because they allow to investigate genuinely attested language data rather than just invented sample sentences. In the case of an ancient language like Vedic, for which we can only make use of the available texts, a corpus-based approach to the study of *iva*'s approximating function is still a desideratum, as the only existing study (Brereton, 1982) is based on a handful of passages. In our study we employ a corpus

[1] Lakoff (1972) is the first study on hedges in English.

Proceedings of the 21st International Workshop on Treebanks and Linguistic Theories (TLT, GURT/SyntaxFest 2023), pages 21 - 31
March 9-12, 2023

with manually validated syntactic annotations, the Vedic Treebank (VTB, see Hellwig et al., 2020), which allows us to investigate the syntactic functions taken by the particle and to detect changes in the syntactic contexts in which it occurs. Since the occurrences of approximating *iva* are sparse in the VTB, we extend our data set with silver annotations obtained by an unsupervised parse of all Vedic texts contained in the Digital Corpus of Sanskrit (DCS, Hellwig, 2010–2022).

Section 2 of this paper gives a summary of previous research on approximators, including an overview of the current understanding in Vedic Studies. Sections 3 and 4 introduce the data set and describe the probabilistic model used to assess diachronic trends in the data, with a special focus on how to use data obtained in an unsupervised manner (Sec. 4.2). A more detailed qualitative evaluation is presented in Sec. 5. – Data and scripts are available at `https://github.com/OliverHellwig/sanskrit/tree/master/papers/2023tlt`.

2 The diachronic development of approximation markers

2.1 Cross-linguistic evidence

Adaptors are propositional hedges (approximators in Prince et al., 1982) that trigger loose readings of a lexical expression: in other words, they signal a loose correspondence between the referents or intended concepts and the lexemes employed, as in example (2) (see also *somewhat*, *some*, *a little bit*, etc.).

(2) a. *He's* ***sort of*** *nice.*

b. *He's really* ***like a geek***.

As explained in Mihatsch (2010), a well-attested source for adaptors are markers of similative constructions that serve to compare two entities either globally, as in (3a), or with respect to some quality, as in 3b (Haspelmath and Buchholz, 1998, 278).

(3) a. *She is* ***like*** *her grandmother.*

b. *He sings* ***like*** *a nightingale.*

Semantically, the passage from markers of similative constructions to adaptors is triggered by the very idea of similative comparison which, unlike equative comparison of quantity, always implies an approximation; compare the above similatives with the equative *Robert is* ***as tall as*** *Maria* (Haspelmath and Buchholz, 1998, 278).

Syntactically, similative markers that turn into adaptors lose their function of situating the object of comparison in relation to a standard and become modifiers of noun phrases, signaling their semantically loose use (Mihatsch, 2010). See examples (4a) and (4b) from French (Mihatsch, 2009, 72):[2]

(4) a. *[Q]ui a fait passer* ***<u>quelque chose</u> comme un frisson*** *dans le dos des supporters français*
'Who sent **<u>something</u> like a shiver** down the back of the French supporters'

b. *[I]l a eu comme une étrange secousse,* ***Ø comme un frisson***...
'He had like a strange spasm, **like a shiver**...'

Adaptors often develop new functions. For instance, they can be employed to signal figurative speech, as in example (5) from Italian (Mihatsch, 2010, 111); this function of adaptors derives from the fact that metaphors, like similative constructions, are also based on similarity, although across two conceptual domains.

(5) *[I] francesi hanno voluto* ***come*** *pagare un debito verso il loro poverissimo ciclismo*
'**It was as if** the French wanted to pay a debt toward their poor cyclism.' (Lit. 'The French wanted to **like** pay a debt toward their poor cyclism.')

As mentioned in the introduction, speech act hedging often arises as implicature of propositional hedging. For instance, adaptors may be used as shields for pragmatic mitigation as in French *Y'a* ***comme*** *un problème* 'there is like a problem' (Mihatsch, 2009, 84). The employment of adaptors as pragmatic shields leads to their syntactic flexibilization, allowing them to occur with parts of speech other than nouns. For instance, in languages such as Spanish and Portuguese, the same

[2] For similar developments in other Romance languages as well as Germanic languages, see Mihatsch (2009) and Mihatsch (2010); on languages outside of Europe, see Ziv (1998) and Fleischman (1999).

adaptors that have developed shield functions are also employed as rounders, i.e. as expressions that indicate imprecise numerical values (e.g., *Peter's house is **almost** 100 feet wide*; see also Spanish and Portuguese *como* 'like'; Mihatsch, 2010, 112).

2.2 The Vedic approximation marker *iva*

The Vedic corpus, whose texts cover a period ranging from the 2nd millenium BCE to around 500-300 BCE (Witzel, 1997, 2009), provides further evidence for the development of comparative markers into adaptors. In the *Rigveda* (= RV), the oldest layer of Vedic literature, the particle *iva* primarily functions as a marker of similative constructions, as in example (6). In such constructions, *iva* always follows the standard of comparison or, when this standard is a complex noun phrase, the first element of the standard (see *pitā iva sūnave* 'like a father for a son' in 6):

(6) *saḥ* *naḥ* *pitā* ***iva*** *sūnave* *agne* *sūpāyanaḥ* *bhava*
3SG.NOM 1PL.DAT father:NOM like son:DAT Agni:VOC of-easy-approach:NOM be:IMPV.2SG
'**Like** a father for a son, be of easy approach for us, o Agni.' (RV 1.1.9ab; trans. Jamison and Brereton, 2014)[3]

In more recent layers of the Vedic corpus, besides retaining its function of marking similative comparison, *iva* performs other functions that correspond to those attested cross-linguistically for adaptors, as in example (7):

(7) a. *saḥ* *avet* *pāpmānam* *vā* *asr̥kṣi* *yasmai* *me* *sasr̥jānāya* ***tamaḥ*** ***iva*** *abhūd*
3SG.NOM know:IMPF.3SG evil:ACC PTC cast:AOR.1SG REL.DAT 1SG.DAT create:ABS darkness:NOM APPROX come-to-be:AOR.3SG
'He knew, "Verily, I have created evil for myself since, after creating (the Asuras), there has come to be **a kind of darkness** for me."' (*Śatapatha-Brāhmaṇa* [M] 11.1.6.9; trans. adapted from Eggeling, 1900)

b. *tasmāt* *api* *etarhi* ***bhūyān*** ***iva*** *naktam* *saḥ* ***yāvat*** ***mātram*** ***iva*** *apakramya* *bibheti*
therefore even today big.NOM APPROX at-night 3SG.NOM as-far-as just APPROX travel:ABS be-afraid:3SG
'Therefore, even today, (although) **quite big**, he who travels even a **quite short** distance at night becomes afraid.' (*Gopatha-Brāhmaṇa* 2.5.1; trans. Brereton, 1982)

Brereton (1982) describes the different functions performed by *iva* in Vedic prose, but he does not engage in a diachronic analysis of the particle nor does he address the relation between its comparative and approximating functions. The fact that *iva*'s approximative function is already attested in some Rigvedic passages led Pinault (2004) to hypothesize that this was the original function of the particle, which only later developed a comparative function. Based on comparative and textual evidence, Biagetti (2022) makes a case for the opposite development of *iva*, namely from a marker of similative constructions into an adaptor. Through a manual scrutiny of some Rigvedic passages listed in Pinault (2004), Biagetti identifies different ambiguous contexts that may have led to the emergence of the new function and to its progressive conventionalization. In particular, *iva*'s adaptor function seems to have emerged from similative constructions whose object of comparison consists in (a) null referential argument(s), as in example (8).[4] Among such cases, those in which neither the linguistic context nor the discourse universe provide referents for a null comparee (as likely in the first half of example 9) trigger a reanalysis of the standard of comparison as the argument of the verb and of *iva* as its modifier.

[4]In example 8, the subscripts *i* and *j* indicate that *indraḥ* 'Indra' and *rathāya* 'for (his) chariot' can be interpreted as referents respectively of the null subject ($Ø_i$) and null object ($Ø_j$) of *unoti* 'urges' in the following sentence

(8)

indraḥ$_i$	*rathāya*$_j$	*pravatam*
Indra:NOM	chariot:DAT	easy-slope:ACC

kr̥ṇoti	. . .	*yūthā*	*iva*
make:3SG	...	flock:ACC.PL	like/APPROX

paśavaḥ	$\emptyset_i$	$\emptyset_j$	*vi*	*unoti*
livestock:GEN	Ø	Ø	PTC	urge:3SG

gopāḥ	*ariṣṭaḥ*
herdsman:NOM	invulnerable:NOM

yāti	*prathamaḥ*	*siṣāsan*
drive:3SG	first:NOM	win:DES.PTCP.NOM

1. Comparative reading: 'Indra makes an easy slope for his chariot [. . .]. Like a herdsman the flocks of livestock, he (Indra, *indraḥ* in *pāda* a) urges (his chariot, *rathāya* in *pāda* a). Invulnerable, he drives as the first to seek winnings.' (RV 5.31.1a-c; trans. adapted from Jamison and Brereton, 2014)
2. Approximative reading: (*pāda* c) 'The herdsman urges the flocks of livestock, as it were.'

(9)

cittiḥ	*apām*	*dame*
bright:NOM	water:GEN.PL	house:LOC

viśvāyuḥ	***sádma***	***iva***
whole-lifetime	seat:ACC	like/APPROX

dhīrāḥ	*sammāya*	*cakruḥ*
clever:NOM.PL	measure:ABS	make:PF.3PL

1. Comparative reading (unlikely): '(He is) the bright apparition in the house of the waters through his whole lifetime. Like clever men an abode, the wise have made a seat (for him), having measured it out completely.' (RV 1.67.10ab; trans. Jamison and Brereton, 2014)
2. Approximative reading: 'The clever ones made (for him, Agni) **some kind of seat** by building together.' (trans. Pinault, 2004)

Since in similative constructions *iva* always follows a noun (phrase), the adaptor function must first have developed with nouns (see example 7a) and then have spread to other parts of speech (see, e.g., example 7b with adjectives). In the following sections, we aim to trace this syntactic flexibilization of *iva* throughout different diachronic layers of Vedic literature.

3 Data

Our diachronic analysis of *iva* as an approximator is based on the dependency annotations collected in the Vedic Treebank (Hellwig et al., 2020), which is annotated using Universal Dependencies.[5] As is shown in Table 1, *iva* occurs in several syntactic functions; `discourse`, the label on which we focus in this paper, is only the third most frequent annotation of this particle. The two most frequent labels, `case` and `mark`, are employed when *iva* functions as a marker of similative comparison; in particular, the particle takes the relation `case` when it introduces a single standard of comparison (e.g. ***gaúḥ iva*** *śākináḥ* 'strong like an ox'), whereas it is labeled as `mark` when it introduces a complex standard resulting in a gapping construction (e.g. *tam tvā vayam* ***sudughām iva goduhaḥ*** *juhūmasi śravasyavaḥ* 'we call to you, as milkers [call] on a cow who gives good milk').

The alternation between the main functions of *iva* becomes much clearer when we add a chronological component to the data. Dating Vedic texts is notoriously difficult because text-internal and external chronological clues are largely missing (see e.g. Witzel, 1995). The VTB therefore assigns each Vedic text to one of five successive chronological layers, based on a general consensus in Vedic studies (details in Hellwig and Sellmer, 2021): the oldest part of the Rigveda (1-RV, ca. 15th–11th c. BCE), the metrical texts of the Mantra period (2-MA, 10th–8th c. BCE), old (3-PO ca. 8th–7th c. BCE) and late prose (4-P, ca. 7th–6th c. BCE), and the prose texts of the Sūtra period (5-SU, ca. 4th c. BCE – 3th c. CE). Rows 2ff. of Table 1 show how the syntactic functions of *iva* are distributed over these five chronological layers. We observe a clear break in the usage of *iva* between the two early metrical layers (1-RV, 2-MA) and the later prose layers: in the former, the `case` and `mark` relations are frequent, while `discourse` is virtually unattested; on the contrary, `case` and `mark` decrease in later prose layers, while `discourse` becomes more frequent. A reason for the high frequency of comparative *iva* may be found in the fact that layers 1-RV and 2-MA include metrical texts composed

[5]The current version of the VTB, which is available at `https://github.com/OliverHellwig/sanskrit/tree/master/papers/2020lrec/treebank`, contains 140,442 words in 18,958 sentences.

Time	case	mark	discourse	other
Global	397	132	126	5
1-RV	135	53	1	1
2-MA	183	72	7	1
3-PO	31	2	49	0
4-PL	33	3	65	3
5-SU	15	2	4	0

Table 1: Gold labels for *iva*. First row: global counts; following rows: counts per time slot in the VTB. See p. 4 for the chronological labels in the first column.

in a highly formulaic diction which is characterized, among other figures of speech, by the extensive use of similes introduced by *iva*.[6] Moreover, `discourse` is employed so rarely in 1-RV and 2-MA because at this diachronic stage *iva*'s approximative function has not yet fully developed (see Sect. 2.2 and below). At this point it should be added that the ambiguous nature of the function of *iva* in examples such as (8) is not explicitly reflected in our data, as the annotation software used did not allow to assign two or more alternative labels. As far as the annotation guidelines are concerned, they did not contain any specific rules as to the treatment of these two functions of *iva*.

4 Quantitative evaluation

4.1 Gold data from the VTB

In this section, we focus on the data in column four of Table 1. For this study, the counts of these words are further split by the word class of the head of *iva*. This view of the gold annotations is presented in the first two rows of Table 2, along with the proportion of noun constructions in each time slot. The first three rows of Table 2 suggest that the diachronic distribution of *iva* with nouns is influenced not only by chronology but also by the register of the texts (metrical vs. prose). First, only one construction of this type is found in the first two layers of the VTB which contain the early metrical texts (1-RV, 2-MA) although there are seven cases in which *iva* is labeled as discourse marker here.[7] Second, the proportion of this construction (see row 3 of Table 2) decreases in the three layers that contain middle and late Vedic prose texts (3-PO, 4-PL, 5-SU). The two factors of time and register are not easy to disentangle because the metrical texts constitute all of the two oldest strata. In order to test how these factors influence the frequency of *iva* with nouns, we fit a binomial logistic regression to the gold data in the upper half of Table 2. Such a model generates the observed counts of *iva* with nouns in a time slot given the total number of instances in this slot and the values of the covariates (predictors). As the data set is small, we use a Bayesian approach that restricts the values of the inferred coefficients. We develop models that test the plausibility of the following three scenarios:

1 Time alone is responsible for the distribution in Table 2. Let t_i denote the time slot, scaled to the range $[-1, +1]$,[8] n_i the number of cases in which the head of *iva* is a noun in time slot i (row 1 of Table 2), N_i the total number of occurrences of *iva* in slot i (sum of rows 1 and 2 of Table 2), and $\sigma(\ldots)$ the logistic link function. After placing standard Normal priors on the coefficients a, b, the observed frequencies of *iva* (n_i) are generated in the following way:

$$n_i \sim \text{Binomial}(N_i, \sigma(a + bt_i)) \quad (1)$$

2 The distribution in Table 2 is solely caused by register, i.e. the opposition between (early) metrical and (late) prose texts. The link function in Eq. 1 changes to $\sigma(a + cr_i)$, with r_i denoting the register of layer i encoded as a binary factor.

3 Each row in Table 2 is generated by jointly considering register and time. If $t_i \in (1, 2)$, p_i is generated as in model 2; else as in Eq. 1.

We implement all models in `RStan` (Stan Development Team, 2022) and compare them using the

[6]On the formulaic nature of Rigvedic similes, see Pinault (1985) and Pinault (1997), among others.

[7]This may partly be due to the fact that, as explained in Sect. 5, in the first two layers *iva*'s adaptor function has not yet become conventionalized. While cases where *iva* follows another part of speech are easy for annotators to interpret, some cases where *iva* follows a noun can be ambiguous between comparative and adaptor reading (see example 8): since the former reading is by far the most frequent in the RV and mantra language, annotators in these cases are likely to chose the label `case` or `mark`, and therefore approximative *iva* with a noun as head may be slightly under-represented in these two layers.

[8]This implies that an ordinal variable is transformed into a scalar. Such an approach is problematic (see e.g. McKelvey and Zavoina 1975 for the case of ordinal predicted variables), and it would be more meaningful to model time either with an ordered factor or to estimate at least the widths of the time slots before performing the transformation. The data set studied here is, however, not large enough to obtain reliable estimates of the additional parameters.

		1-RV	2-MA	3-PO	4-PL	5-SU
VTB (gold)	noun	0	1	24	28	0
	other	1	6	25	37	4
	Prop.	0	14.3	49	43.1	0
DCS (silver)	noun	1	3	84	128	4
	other	3	12	91	240	23
	Prop.	25	20	48	34.8	14.8
	correct/wrong	0/0	1/1	11/1	19/2	2/0

Table 2: POS tag of the syntactic head of *iva* used as discourse marker, conditioned on the time slot (columns). Gold data in the upper half is from the VTB. Rows with 'noun': The head of *iva* is a noun; 'other': The head has any other POS. For the silver data in the lower half of the table, refer to Sec. 4.2 of this paper.

expected log pointwise predictive density (elpd) in a leave-one-out setting (Vehtari et al., 2017). Each model is trained for 5,000 iterations and with four parallel chains. Model diagnostics ($\hat{R}$, ESS) show no problems in the sampling process.

The results in Table 3 show that the elpd of model 1, which only considers time, is more than one standard error (column 'SE') lower than that of the two other models which include the register split. This outcome suggests that time alone cannot explain the distribution of *iva* with nouns, and register information is relevant for modeling the data in Table 2. This conclusion finds further support by a posterior predictive check the results of which are given in the column labeled 'β' in Table 3. To calculate β, we sample values of n_i (counts of *iva* with nouns) from the posterior distributions of the three models at each post-burn-in iteration of the sampler. The five sampled values n' (one for each time slot) are compared with the observed distribution of n (row 1 of Table 2) using the exact Fisher test for quantifying the goodness of fit. β in Table 3 is the proportion of these tests in which the null hypothesis (n and n' come from the same distribution, i.e. the model generates "naturally looking" samples) was rejected at an error level of 5%. β can thus be interpreted as an approximation of the type II error of wrongly accepting the null hypothesis. The chance of making such an error is clearly higher (0.0192, i.e. 1.9%) for the model that only considers time than for those that integrate register as well (0.63% and 0.44%). As the values of β and elpd show, the best fit of the data is achieved by the third model which combines time and register. This outcome is not surprising. While model 2 (register only) adapts to the observed counts using two estimates that remain constant over slots 1-2 and 3-5, model 3 has the chance to capture the temporal dynamics in the three later slots of Vedic and thus achieves a better fit.

Model	elpd	SE	β
Time	-12.4	2.43	0.0192
Register	-9.92	2.34	0.0063
Time/register	-9.92	2.6	0.0044
Time/reg., silver	-320.76	29.47	0.0297

Table 3: Summary evaluation of the models applied to the data in Table 2. 'elpd' and 'SE' quantify the predictive power and its standard error. 'β' reports the results of a posterior predictive check. Higher values of elpd and lower ones of β are better. elpd and SE of the fourth model are not comparable with the values of the other three models and are only given for reference.

4.2 Exploring silver annotations

While the results discussed so far are in favor of a diachronic scenario that explains NOUN + *iva* constructions with a combination of the Vedic register split and a chronological model, one should keep in mind that the data set on which this conclusion is built consists of only 125 observations and is therefore tiny. As the DCS, on top of which the Vedic Treebank is built, is much larger than the VTB and a parser for Vedic is available,[9] it is obvious to extend the data set with silver annotations made by this parser. We therefore extract all occurrences of *iva* from an up-to-date unsupervised parse of the DCS[10] and merge gold and silver an-

[9]This parser uses a biaffine architecture (see Dozat and Manning, 2017) with the addition of a character based CNN (see Rotman and Reichart, 2019; Zhang et al., 2015), and reaches a performance of 87.61 UAS and 81.84 LAS. For further details see Hellwig et al. (Forthcoming).

[10]The Con-LLU data are available at `https://github.com/OliverHellwig/sanskrit/tree/master/dcs/data/conllu`. Silver parses are contained in the `conllu_parsed` files.

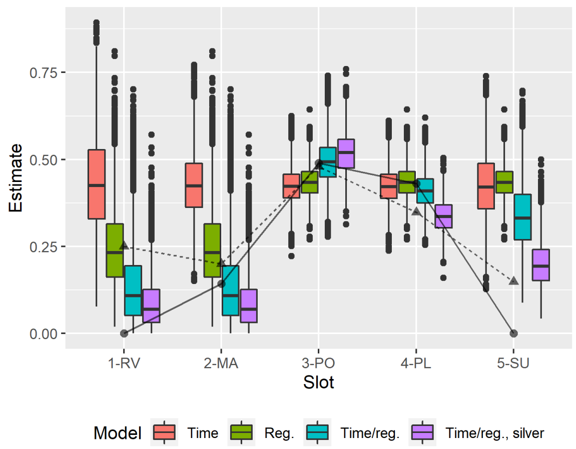

Figure 1: Temporal dynamics predicted by the four models. The y-values are parameters of binomial distributions that predict the presence of *iva* with nominal heads. Dots and lines give the observed proportions in the gold (circles; Sec. 4.1) and silver data (triangles; Sec. 4.2).

notations; statistics of the silver data are presented in the lower half of Table 2. While the gold data used in Sec. 4.1 only contain instances of *iva* labeled as discourse, the merged data set contains all gold and silver annotations of *iva* regardless of their syntactic labels, because we want to recover instances of *iva* as discourse particle that were mislabeled by the parser.

Instead of the plain GLM of Eq. 1, we now use a hierarchical model that integrates a mechanism for error detection. At the first level, this model decides if, in a given record k, the true label of *iva* is `discourse`. The corresponding binary variable z_k (1 = record k is an instance of *iva* as discourse, 0 = it is not) is only partly observed. Somewhat over-confidently, we assume that all gold annotations are labeled correctly. To get an estimate of the error level in the silver data, one author of this paper manually annotated the correct label of 100 randomly chosen silver records, marking those cases in which a wrong head was chosen for *iva*. z_k is predicted using the following covariates: time; the label; the distance between *iva* and its head; the POS and label of the head; interactions between time and position difference and time and head POS. At the second level, the model proceeds with model 3 (Time/register) from Sec. 4.1 if z_k has the value 1, i.e. is correct according to the first level of the model. The only difference is that the binomial is replaced with a Bernoulli distribution because individual records are inspected. More formally, let $\mathbf{x}_k$ denote the vector of covariates for the Bernoulli logistic model at level 1, and $\mathbf{d}$ the vector of the corresponding coefficients to be estimated. After placing standard normal priors on the coefficients, z_k is drawn from a Bernoulli distribution $\text{Bern}(\sigma(\mathbf{x}^T\mathbf{d}))$. If $z_k = 1$, model 3 from Sec. 4.1 is used for describing the diachronic distribution.

Figure 1 provides a graphical comparison of the results produced by the four models discussed in Sections 4.1 and 4.2. Here, the values on the y-axis are the estimated proportion parameters that model the occurrence of *iva* with nominal heads. In addition, Fig. 1 also shows the proportions observed in the gold and silver data as points connected with lines. As could already be deduced from the β values in Table 3, neither the time-only nor the register-only models fit the observed gold data well. The outcome is much better for the two models that combine register and time. Both predict low values for the early metrical texts (1-RV, 2-MA), and they appropriately describe the decreasing trend in the three prose levels. Note that neither 'Register' nor 'Time or register' fully capture the low frequencies in the last chronological layer (5-SU). This suggests that further, probably domain or genre specific, factors are in effect here. One possible explanation may be that the language of the late Vedic Sūtra texts differs markedly from that of the earlier Vedic literature (see e.g. Renou, 1957, 15-16).

5 Qualitative analysis

The data presented in Section 4 confirms the syntactic flexibilization of *iva* hypothesized in Section 2: as the proportion of NOUN + *iva* constructions decreases, the particle starts occurring with other parts of speech; furthermore, the analysis in Section 4 suggests that this development is to be attributed to both register and time. In this section, we provide a more detailed qualitative evaluation of the data showing that the extension of *iva*'s scope to other parts of speech co-occurs with the development of new functions for this particle.

The grammaticalization process described in Section 2 first results in the employment of *iva* with nouns (see example 7a). Other parts of speech occurring with *iva* in layers 3-PO and 4-PL belong either to open classes, such as verbs (example 10) and adjectives (11), or to closed classes such as conjunctions (12) and particles (13). In example (10), king Janaka asks the sage Yājñavalkya about the possible substitutes for the *agnihotra*, a meal offering usually consisting of milk. The conversation comes to an end when Yājñavalkya states that, even in the absence of water, the *agnihotra* can be celebrated by offering, ‘in some way’ (*iva)*, ‘truth in faith’ (*satyam śraddhāyām*). Similarly, in (11), the author explains that, during the Soma sacrifice, the sacrificial post is anointed from its base upwards because it is for heaven that it is anointed and heaven is ‘in some way’ (*iva*) ‘upwards’ (*parāṅ*).

(10) *yat* *āpaḥ* *na* *syuḥ*
if water:NOM.PL NEG be:OPT.3SG
kena *juhuyāḥ* *iti.*
what:INST offer:OPT.2SG QUOT
saḥ *ha* *uvāca* *na* *vai* *iha*
3SG.NOM PTC say:PF.3SG NEG PTC here
tarhi *kiṃcana* *āsīt* *atha* *etat*
then nothing be:IMPF.3SG but here
u ***hūyate*** ***iva*** *satyam*
PTC offer:PASS.3SG APPROX truth:NOM
śraddhāyām *iti*
faith:ACC QUOT

‘If there would be no water, with what would you perform the offering?’ He said: ‘Then, indeed, there would be nothing at all here, and yet **there would be offered in some way** here, namely, truth in faith.’ (*Jaiminīya-Brāhmaṇa* 1.19.23.1; trans. adapted from Bodewitz, 1973)

(11) *parāñcam* *prokṣati.* ***parāṅ***
upwards:ACC anoint:3SG upwards:NOM
iva *hi* *suvargaḥ* *lokaḥ*
APPROX for heavenly:NOM world:NOM

‘He anoints (he sacrificial post) from the foot upwards, for **upwards as it were** is the world of heaven.’ (*Taittirīya-Saṃhitā* 6.3.4.1)

In examples (12) and (13), where *iva* follows the conjunctions *uta* ‘and’ and the causal expression *tasmāt vā* ‘therefore’, the particle seems to have scope not only on the preceding lexical item, but on the whole proposition. In (12), Ajātaśatru explains to Gārgya that, when one is asleep, one gathers the cognitive power of the vital functions into the space within one’s heart. The dream then consists of the perceptions that the sleeping person experiences in her heart, rather than in the external world. In this example, the sequence of *uta iva* marks the fictive nature of the events experienced in the dream.

(12) *saḥ* *yatra* *etat* *svapnyayā*
3SG.NOM wherever thus in-dream
carati ... *tat* ***uta*** ***iva***
go:3SG ... then CONJ APPROX
mahārājaḥ *bhavati* ***uta***
great-king:NOM become:3SG CONJ
iva *mahābrāhmaṇaḥ* ***uta***
APPROX great-Brahmin:NOM CONJ
iva *uccāvacam*
APPROX high-and-low(-region):ACC
nigacchati
enter:3SG

‘Wherever he may travel in his dream [...] He may appear to become a great king or an eminent Brahmin, or to visit the highest and the lowest regions.’ (*Bṛhadāraṇyaka-Upaniṣad* 2.1.18.3; trans. Olivelle, 1998)

Example (13) is concerned with explaining the creation of the universe by pointing out similarities between words. At the beginning there was nothing but seven vital airs; they were turned into seven persons and these, in turn, into body parts of Prajāpati, the ‘lord of generation’. In this process, the best part (*śrī-*) of each person was concentrated and became Prajāpati’s head (*śiras-*; note the phonetic similarity of the words *śrī* and *śiras*). In the example, the sequence ***tasmāt vā iva*** *etat śiraḥ* seems to present the preceding clause (’It was thereto that the vital airs resorted’), which involves the verb *śri-* ’rest on’, as a further possible explanation of the word *śiras* ’head’ to which, however, the author does not fully commit.

(13)

tasmin	etasmin	prāṇāḥ
3SG.LOC	DEM.LOC	vital-air:NOM.PL

aśrayanta.	**tasmāt**	**vā**	**iva**
resort:IMPF.3PL	therefore	PTC	APPROX

etat	śiraḥ
3SG.NOM	head:NOM

'{And because (in it) they concentrated the excellence (*śriyaṃ* < *śrī*), therefore it is (called) the head (*śiras*).} It was thereto (in the head) that the vital airs resorted (*aśrayanta* < *śri-*): possibly therefore it is the head (*śiras*). (*Śatapatha-Brāhmaṇa* [M] 6.1.1.4.4; trans. adapted from Eggeling, 1894)

The source for the syntactic flexibilization of *iva* may also be found in the very sort of texts contained in layers 3-PO and 4-PL. In the *Brāhmaṇas*, ancillary texts providing detailed explanations of rituals, *iva* is often employed in order to point out correspondences among elements of the ritual realm, of the cosmic realm, or of daily life. In such cases, the sequence NOUN + *iva* co-occurs with a causal particle or adverb such as *hi* 'for, because' or *tasmāt* therefore': see example (14), where the phrase *vājinam iva* 'some sort of steed' is followed by the particle *hi* (see also example 11):

(14) *paryagnaye kriyamāṇāya anubrūhi iti āha adhvaryuḥ [...]*
'Recite for the carrying round of fire' the Adhvaryu (priest) says [...]'

vājī	san	pari
steed:NOM	be:PTCP.NOM	around

nīyate	iti	**vājinam**
carry:PASS.3SG	QUOT	steed:ACC

iva	**hi**	enam	santam
APPROX	for	DEM.ACC	be:PTCP.ACC

pariṇayanti
around-carry:3PL

'Being a steed he (the fire, god Agni) is carried round' (the Adhvaryu says), **for him being as it were a steed** they carry round.' (*Aitareya-Brāhmaṇa* 2.5.3.2; trans. Keith, 1920)

In example (14), for instance, *iva* does not signal a loose reading of the noun *vājinam* 'steed' alone, but rather the metaphorical nature of the correspondence between the fire (god Agni) and a steed. The frequency of structures such as (14) in the *Brāhmaṇas* may have caused an interpretation of *iva* as having scope not only on the preceding lexical item, but on the whole proposition, and may eventually have caused the emergence of sequences such as *tasmāt vā iva* in (13), where the particle directly follows the causal adverb and the disjunctive particle.

6 Summary and conclusion

Originally a marker of phrasal comparison, the Vedic particle *iva* grammaticalized into an approximation marker signaling the semantically loose use of the preceding noun (Sect. 2.2). This grammaticalization process can already be traced in the oldest texts (layers 1-RV and 2-MA) by manual scrutiny (Biagetti, 2022), but is not captured by the syntactic annotation contained in the VTB; this is because, in ambiguous contexts that may have been responsible for the reanalysis of *iva* into an adaptor, the particle was usually annotated as a marker of comparison (deprels `case` or `mark`) by the annotators, as this is by far its most common function in the RV.

In this paper we have focused on the further syntactic flexibilization of *iva* in later Vedic texts. Bayesian analysis (Sect. 4) has shown that the proportion of NOUN + *iva* constructions, in which the particle has scope on the immediately preceding lexical item, decreases in layers 3-PO, 4-PL and 5-SU and that this break is to be attributed to both time and register. Accordingly, *iva* starts occurring with parts of speech other than noun and, as shown by the qualitative analysis of Sect. 5, gradually develops new functions. First, the frequent occurrence of *iva* with other particles or conjunctions leads to an extension of its scope to the whole proposition (see examples 11 to 14); second, in some such cases *iva* seems to mark the metaphorical meaning of the expression (example 14) or seems to function as a speech-act edge, signaling lack of commitment in the statement being uttered (example 13).

Ultimately, the quantitative and qualitative analyses of *iva* in Vedic prose seems to mirror the diachrony of adaptors as attested cross-linguistically and thus provides further evidence for the development of the particle from a marker of comparison to an approximation marker, and not vice versa.

7 Acknowledgments

We would like to thank three anonymous reviewers for their remarks and suggestions.
Oliver Hellwig and Sven Sellmer were funded by the German Federal Ministry of Education and Research, FKZ 01UG2121, while doing research for this paper.

References

Erica Biagetti. 2022. From standard marker to adaptor: The case of Vedic *iva*. *Bhasha. Journal of South Asian Linguistics, Philology and Grammatical Traditions*, 1(2).

Henk Bodewitz. 1973. *Jaiminīya Brāhmaṇa I, 1-65: Translation [from the Sanskrit] and commentary*. Number v. 17 in Orientalia Rheno-Traiectina. Brill, Leiden.

Joel P. Brereton. 1982. The particle *iva* in Vedic prose. *Journal of the American Oriental Society*, 102(3):443–450.

Claudia Caffi. 2010. Weakening or strengthening? A case of enantiosemy in Plato's *Gorgias*. In Gunther Kaltenböck, Wiltrud Mihatsch, and Stefan Schneider, editors, *New Approaches to Hedging*, pages 181–202. Emerald Group Publishing Limited, Bingley.

Timothy Dozat and Christopher D. Manning. 2017. Deep biaffine attention for neural dependency parsing. In *5th International Conference on Learning Representations*, pages 1–8.

Julius Eggeling. 1894. *The Śatapatha-Brāhmaṇa. According to the text of the Mādhyandina School. Translated. Part III, Books V, VI and VII*, volume 3. Clarendon Press, Oxford.

Julius Eggeling. 1900. *The Śatapatha-Brāhmaṇa. According to the text of the Mādhyandina School. Translated. Part V, Books XI, XII, XIII andXIV*, volume 5. Clarendon Press, Oxford.

Suzanne Fleischman. 1999. Pragmatic markers in comparative perspective. In *Paper presented at PRAGMA 99, Tel Aviv, Israel*.

Bruce Fraser. 2010. Pragmatic competence: The case of hedging. In Gunther Kaltenböck, Wiltrud Mihatsch, and Stefan Schneider, editors, *New approaches to hedging*, pages 15–35. Emerald Group Publishing Limited, Bingley.

Martin Haspelmath and Olga Buchholz. 1998. Equative and similative constructions in the languages of Europe. In Johan Van der Auwera, editor, *Adverbial Constructions in the Languages of Europe*, pages 277—334. Mouton de Gruyter, Berlin.

Oliver Hellwig. 2010–2022. DCS - The Digital Corpus of Sanskrit.

Oliver Hellwig, Sebastian Nehrdich, and Sven Sellmer. Forthcoming. Data-driven dependency parsing of Vedic Sanskrit. *Language Resources and Evaluation*.

Oliver Hellwig, Salvatore Scarlata, Elia Ackermann, and Paul Widmer. 2020. The treebank of Vedic Sanskrit. In *Proceedings of the 12th Language Resources and Evaluation Conference*, pages 5137–5146, Marseille, France. European Language Resources Association.

Oliver Hellwig and Sven Sellmer. 2021. The Vedic treebank. In Erica Biagetti, Chiara Zanchi, and Silvia Luraghi, editors, *Building New Resources for Historical Linguistics*, pages 31–40. Pavia University Press, Pavia.

Stephanie W. Jamison and Joel P. Brereton. 2014. *The Rigveda: the earliest religious poetry of India. 3 volumes*. Oxford University Press, New York.

Gunther Kaltenböck, Wiltrud Mihatsch, and Stefan Schneider. 2010. Introduction. In Gunther Kaltenböck, Wiltrud Mihatsch, and Stefan Schneider, editors, *New approaches to hedging*, pages 1–14. Emerald Group Publishing Limited, Bingley.

Arthur Berriedale Keith. 1920. *Rigveda Brahmanas: The Aitareya and Kausītaki Brāhmanas of the Rigveda*, volume 25 of *The Harvard Oriental Series*. Harvard University Press, Cambridge, Massachusetts.

George Lakoff. 1972. Hedges: A study in meaning criteria and the logic of fuzzy concepts. In J.N. Levi Deborah James, Paul M. Peranteau and G.C. Phares, editors, *Papers from the Eighth Regional Meeting of the Chicago Linguistic Society*, pages 183–228. Pavia University Press, Chicago.

Richard D McKelvey and William Zavoina. 1975. A statistical model for the analysis of ordinal level dependent variables. *Journal of Mathematical Sociology*, 4(1):103–120.

Wiltrud Mihatsch. 2009. The approximators French comme, Italian come, Portuguese como and Spanish como from a grammaticalization perspective. In Corinne Rossari, Claudia Ricci, and Adriana Spiridon, editors, *Grammaticalization and pragmatics: facts, approaches, theoretical issues*, pages 65–91. Brill, Leiden.

Wiltrud Mihatsch. 2010. The diachrony of rounders and adaptors: approximation and unidirectional change. In Gunther Kaltenböck, Wiltrud Mihatsch, and Stefan Schneider, editors, *New approaches to hedging*, pages 93–122. Emerald Group Publishing Limited, Bingley.

Patrick Olivelle. 1998. *The Early Upaniṣads. Annotated text and translation*. Oxford University Press, Oxford.

Georges Pinault. 1985. Négation et comparaison en védique. *Bulletin de la Société de linguistique de Paris*, 80(1):103–144.

Georges Pinault. 1997. Distribution des particules comparatives dans la *Rik-Samhitā*. *Bulletin d'Études Indiennes*, 13-14:307–367.

Georges Pinault. 2004. On the usages of the particle *iva* in the Rigvedic hymns. In *The Vedas. Texts, languages and ritual. Proceedings of the Third International Vedic Workshop (Leiden, May 29-June 2, 2002)*, pages 285–306, Groningen. Groningen Oriental Studies.

Ellen F. Prince, Charles L. Bosk, and Joel E. Frader. 1982. On hedging in physician-physician discourse. In Robert J. di Pietro, editor, *Linguistics and the professions*, pages 83–97. Ablex, Norwood, NJ.

Louis Renou. 1957. *Altindische Grammatik, Introduction Générale*. Vandenhoeck & Ruprecht, Göttingen.

Guy Rotman and Roi Reichart. 2019. Deep contextualized self-training for low resource dependency parsing. *Transactions of the Association for Computational Linguistics*, 7:695–713.

Stan Development Team. 2022. RStan: the R interface to Stan. R package version 2.21.5.

Aki Vehtari, Andrew Gelman, and Jonah Gabry. 2017. Practical Bayesian model evaluation using leave-one-out cross-validation and WAIC. *Statistics and Computing*, 27(5):1413–1432.

Michael Witzel. 1995. Early Indian history: Linguistic and textual parametres. In George Erdosy, editor, *The Indo-Aryans of Ancient South Asia. Language, material Culture and ethnicity*, volume 1, pages 85–125. Walter de Gruyter, Berlin, New York.

Michael Witzel. 1997. The development of the Vedic canon and its schools: the social and political milieu (Materials on Vedic Sakhas, 8). In Michael Witzel, editor, *Inside the texts, beyond the texts. New approaches to the study of the Vedas*, pages 258–348. Harvard Oriental Series, Cambridge.

Michael Witzel. 2009. Moving targets? Texts, language, archaeology and history in the late Vedic and early Buddhist periods. *Indo-Iranian Journal*, 52(2/3):287–310.

Xiang Zhang, Junbo Zhao, and Yann LeCun. 2015. Character-level convolutional networks for text classification. In *Advances in Neural Information Processing Systems*, volume 28, pages 649–657. Curran Associates, Inc.

Yael Ziv. 1998. Hebrew *kaze* as a discourse marker and lexical hedge: conceptual and procedural properties. In Andreas H. Jucker and Yael Ziv, editors, *Discourse markers: descriptions and theory*, pages 203–221. Benjamins, Leiden.