DIALECTOLOGICAL-DIACHRONIC GRAMMAR OF CONJUNCTION IN ARCHAIC INDO-IRANIAN

Moreno Mitrović
HUMBOLDT UNIVERSITY · ZAS, BERLIN · BLED INSTITUTE

ABSTRACT

Indo-Iranian (IIr) is an Indo-European (IE) language family which at its earliest form is constituted by Avestan and Rigvedic Sanskrit. Morphosyntactic evidence from expressions of conjunction suggests that the two languages, which are perhaps more accurately considered dialects of early IIr, reflect two distinct stages of morphosyntactic change. Early IE languages operated a double system of coordination, whereby there existed two types constructions. In one, the coordinator is non-clitic and thus occupies the medial (or first/1P) surface position, in another construction, the coordinator is enclitic and occupies the second (or final) surface position (2P). The present paper presents evidence from archaic IIr to support the view that the two dialects of early IIr reflect two distinct morphosyntactic mechanisms of expressing conjunction, one being more archaic than the other. Both Avestan and Vedic expressed 1P conjunction with *uta* and the 2P conjunction with *ca* particles, albeit to different extents, demonstrating the degree to which the grammar dis/allows movement (which is argued to be the explicans for the 1P/2P alternation). Subsequent developments within the dialects, leading to developments of classical or less archaic Iranian and Indic dialects, show two stabilised types of conjunction grammar. Novel evidence from and method for diachronic semantics is also presented, culminating in the view that syntactic-semantic change was not concomitant in one branch. Using statistical, and philological methods, this paper provides a dialectal analysis of IIr morphosyntax of conjunction, showing that Avestan represents a more retentive grammar. Aside from providing a detailed a morphosyntactic and morphosemantic analysis, it will show that, in cases of relatively stable directional historical change, that historical dialects may be identified, as well as relatively dated and cyclically interpreted, based on both the syntactic and the semantic-compositional properties reflecting the diachronic precursors and successors, under the working assumptions that various historical languages of the IIr family could and should be viewed in terms of dialectal continua.

This manuscript is submitted for first-phase internal consideration with the *Journal of Historical Syntax* to appear in the special issue on the topic historical dialectological syntax. This manuscript is the result of the paper presented at the DGfS workshop on *Towards a comparative historical dialectology: evidence from morphology and syntax* on February 24, 2022, held online and organised by the University of Tübingen.

moreno@mitrovic.co http://mitrovic.co

2022 / SUBMITTED MANUSCRIPT / v·1

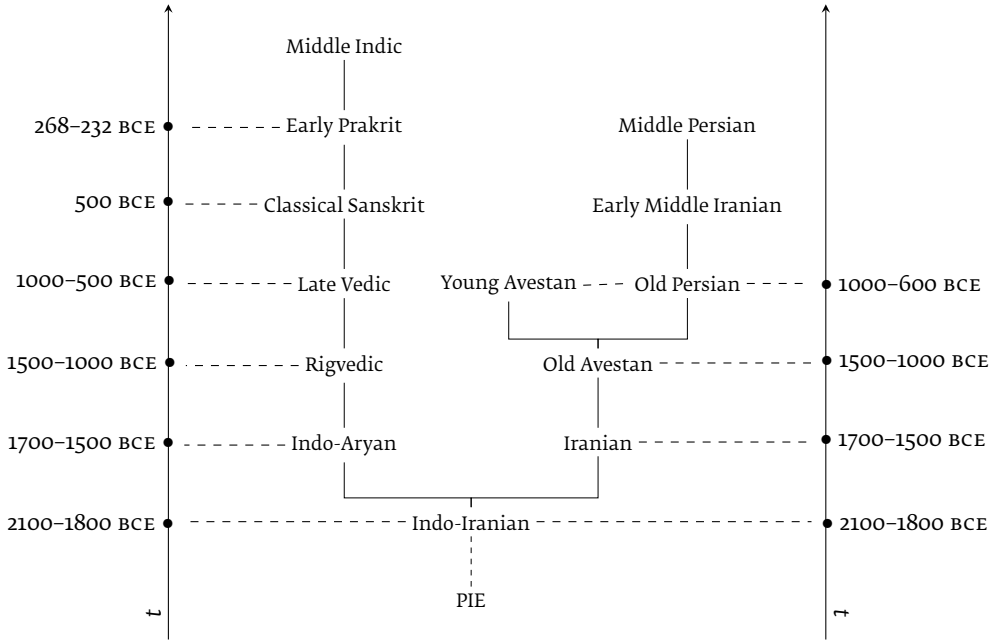


FIGURE 1: A relative chronology of IIr, based on Villalobos (2019, 117, tab. 1) and Windfuhr (2013, Chap. 3, tab. 2.1).

1 INTRODUCTION

Indo-Iranian (IIr) is an Indo-European (IE) language family which, at its earliest form, is constituted by Avestan (Av) and Rigvedic Sanskrit (RV), each representative of the two IIr-internal branches: Iranian (Ir) and Indo-Aryan (IA), respectively. While considered early IIr dialects, Avestan and Rigvedic reflect two distinct stages of morphosyntactic change, insofar as the focus of the present paper is concerned: the structure and interpretation of conjunction and conjunction-marked quantificational expressions. Furthermore, Young Avestan (YAv) is considered, at least on philological grounds, to be closer to Old Persian (OP) that followed it, rather than Old Avestan (OAv) from which it allegedly developed. This, too, I will contend does not align with the novel morphosyntactic and morphosemantic evidence from conjunction marking.

Instead, I will present evidence to support the view that the two dialects of early IIr reflect two distinct (grammars characterised by two) morphosyntactic mechanisms of expressing conjunction, one being more archaic than the other. Furthermore, I will show the diachronic morphosyntax of conjunction makes us reconsider how the branch-internal changes proceeded, at least in terms of a single descriptive parameter – that of conjunction expression strategy – and the theoretical consequences that come with it.

Ṛgveda (ṚV), on the Indic side, is the oldest religious text of the Indo-Aryans and is dated to around the second millennium BCE. The ṚV, written in Rigvedic (RV), reflects a more archaic subcorpus – the so-called family books (being books 2–7) – and the newer (books 1, 8–10). Similarly, the *Avesta*, or

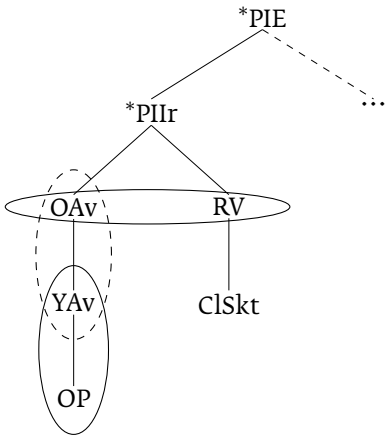


FIGURE 2: A dialect-like clustering of early IIr.

at least its final form, can be dated to the similar time of the middle of the second millennium BCE. The *Avesta*, however, falls into two chronological layers, which I analyse as historical dialects, namely Old Avestan (OAv), being grammatically very close to the language of the ṚV, and Young Avestan (YAv), being grammatically rather close to Old Persian (OP), itself a language of the second half of the first millennium BCE.[1]

The idea of dialectal clusters of early IIr, at least according to the received philological view, is that OAv and RV are more alike in that they both reflect the reconstructable Proto-IIr (*PIIr) language from which they developed. Likewise, for the later branch-internal developments, YAv and OP are considered more closely related, as Fig. 2 shows.

The evidence I present in this paper is in favour of a historical dialectological morphosyntax and semantics that suggest a different and differently motivated clustering since, as it may not be surprising, philological tradition has not concerned itself, nor could it, with detailed morpho-syntactic and -semantic analysis that the modern linguistic era affords us.

The analysis I present here may well be considered methodological for not (only) identifying historical dialects, based on formalised grammatical evidence, but rather how these dialects are reflective of a differential rate of morphosyntactic change of conjunction structure. In this way, and conjoined with the idea of a directed change (which I motivate), the languages (or dialects) may be seen as instantiating different segments of the change that IIr has undergone since its archaic stage. What is more, this method allows for an independent relative means of measurement of the branch-internal rate of retention insofar as the morphosyntax of conjunction is concerned. I sketch this in Fig. 3.

1 For details on dating and historical context, consult Witzel (1997), and extensive citations therein, for early Indo-Aryan and Skjærvø (2006), and the rich collection of those he cites, for the history of early Iranian.

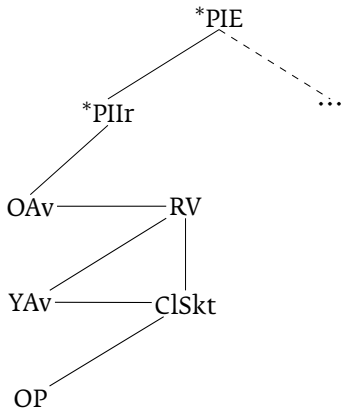


FIGURE 3: Early IIr dialects as reflections of a directed change in the morphosyntax of expression of conjunction.

THE METHOD, ARGUMENT, AND STRUCTURE OF THE PAPER

The argument starts with RV and IE: it shows, using a 'majority' reasoning, that the double system of conjunction is the most likely reconstructable candidate for PIE. This alone may be considered methodological since a detailed synchronic analysis I entertain makes available the discrete parameters surrounding the change that I allege is reflective in other IIr languages and, with it, provides a diachronic explicans, reducible to or at least in line with a third-factor narrative (in the sense of Chomsky 2005).

Therefore, the structure of the paper is as follows: before proceeding to the main empirical and anaytical part in Section 3, I present in Sec. 2 the set of assumptions and theoretical devices I will be employing in my analysis, which amounts to my motivating a rich, or richer than standard, syntactic structure for conjunction that extends to coordination more generally as well as other semantically distinct conjunction-marked expressions) that is motivated on both theoretical and empirical grounds. Section 3 thus starts out by deriving the RV facts from the rich structure. The analysis is then synchronically contrasted with the facts in the Iranian branch in Section Sec. 4. The last section (§5) moves the discussion towards a comparative diachronic-dialectal view of directed change.

2 PRELIMINARIES: CONJUNCTION STRUCTURE

Before proceeding to presenting and analysing the data, let me expound on the theoretical preliminaries concerning the structure for expressions of conjunction and related theoretical ingredients. There are three loose sets of assumptions couched within the minimalist programme (Chomsky, 2001) that I make. The first regards linearisation within the antisymmetric model of syntax (Kayne, 1994), while the second pertains to the general mechanics of cliticisation where I essentially adopt the defective goal approach developed in Roberts (2010). The third, and culminating, set of preliminaries con-

cern the syntactic template for coordinate construction, which motivates a general phrase-structure for 'junction' and couples it with the former two

2.1 ANTISYMMETRY

I will take coordinate complexes to be linearised in compliance with the Linear Correspondence Axiom (LCA), assuming conjunctions are headed and endocentric, as formulated by Kayne (1994) and given in (1).[2]

(1) $d(A)$ is a linear ordering of T,
where A is a set of non-terminals and T the set of terminals.

A prediction that stems from LCA dictates that SPEC ⟩ HEAD ⟩ COMPLEMENT order is a universal (underlying) linear order, where movement provides the only way in which word-oder differences can emerge. One core empirical prediction (1) makes in regards to how it standardly applies to coordinate structures is that HEAD ⟩ SPEC ⟩ COMPLEMENT strings are underivable or derivationally blocked as impossible structures, a prediction that is, indeed, borne out as no such type of conjunction expression is typologically attested.

There is one instance of movement resembling, in part, the latter type of movement involving the displacement of the conjunction head. The analysis I develop hinges on the crucial claim that the latter movement operation is of incorporation (terminal-to-terminal) type. In the following part, I outline the assumptions regarding clisis and incorporation that is pivotal to the claim I will make.

2.2 CLITICISATION & DEFECTIVE GOALHOOD

Roberts (2010) argues that cliticisation is an instantiation of head movement, which is part of narrow syntax and that it applies where the goal of an `Agree` relation is defective. This idea, as he shows, has empirical support from a variety of domains and is conceptually natural to the extent that movement is a special case of merger. In general, we do not and cannot prevent external merge from applying to terminals; similarly we should not prevent internal merge from applying to terminals. (Roberts, 2010, 3). The cliticisation mechanism that Roberts (2010) proposes is dubbed and defined in (2).

(2) **Defective Goal Condition for Head-Movement**:
If α probes β and iff the set of features specified on β (goal) are a (proper) subset of features specified on α (probe), then α triggers head movement of β, i.e. β undergoes incorporation.

Equipped with (2), I propose that the coordinator is in an `Agree` relation with a head of its coordinand, which is defective by virtue of the conjunction

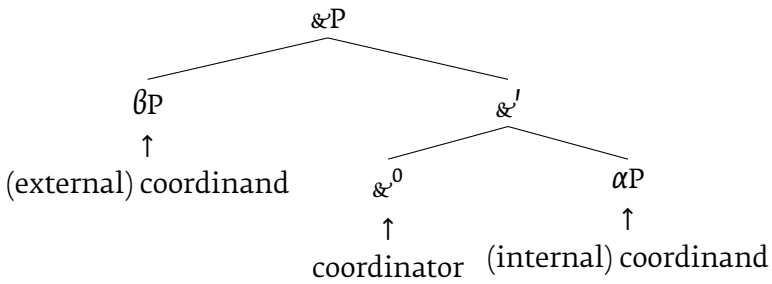
2 For a more detailed formalisation, see, for instance, Nilsen (2003, 19)

head's absorption of its complement's (head's) features. Roberts's (2010) theory will also facilitate us with a syntactic system, which captures the second-position (2P) clitic placement, which has traditionally been relegated as a postsyntactic phenomenon. The model in which narrow-syntactically defective goals surface as clitics, coupled with the assumption that the coordinate head – (5) below – absorbs the categorial features of the argument(s) it coordinates, will be instrumental in my morphosyntactic analysis of 2P conjunction clitics

2.3 STRUCTURING CON/JUNCTION

Following Kayne (1994) and Zhang (2010), *inter alia*,[3] consider the idea that conjunctions are heads projecting something a relatively traditional and default phrase-structurally compliant structure for coordination, as shown in (3).

(3)



Zhang (2010) submits some conclusions regarding the derivation of coordinate construction, namely, that the derivation of coordinate construction does not create any special syntactic configuration, other than the general binary complement and specifier/adjunct configuration, that it does not resort to any special syntactic category, that it is not subject to any special constraint on syntactic operations, and lastly that it does not require any special type of syntactic operations, other than Merge and the step-by-step, one-tail-one-head chains of Move.

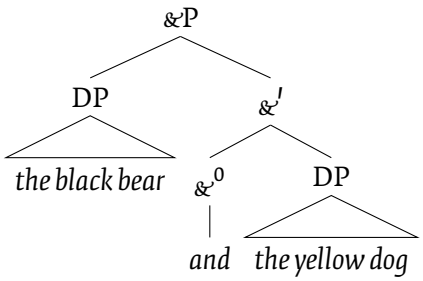
The structurer in (3) are compliant with the minimalist tenets and invoke no special devices, configurations, operations, or categories.[4] Whereas Kayne (1994) and Zhang (2010) stand by the Spec-X^0-Compl configuration underlying coordinate structure, Munn (1993), for instance, proposes a structure whereby the first conjunct (external coordinand in (3) is adjoined to a Boolean Phrase (BP), which is headed by a Boolean head (B^0), which instantiates the

3 A structure like the one in (3) has been argued for by Blümel (1914), Bloomfield (1933), Bach (1964), Chomsky (1965), Dik (1968), Dougherty (1969), Gazdar et al. (1985), Goodall (1987) and Muadz (1991), among many others.

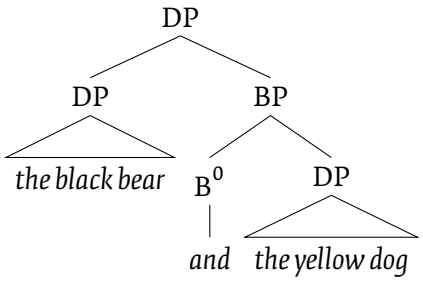
4 Other approaches to coordinate phrase structure, such as those by Munn (1993) and Velde (2005), suppose a minimally different structure, as far my analysis is concerned. Alternative binary, and mono-dimensional (*contra* Progovac 1998a, 1998b, *i. a.*), approaches generally differ with respect to the 'mode' of merger of the external/first coordinand with the coordinand+internal/second coordinand complex.

coordinator and whose complement is the second conjunct (i.e. internal coordinand). Velde (2005), on the other hand, proposes a counterintuitive structure, which is more reminiscent of sub- than co-ordination structure. These three binary and mono-dimensional approaches to coordinate structure of a DP complex like 'the black bear and the yellow dog' are sketched in (4).

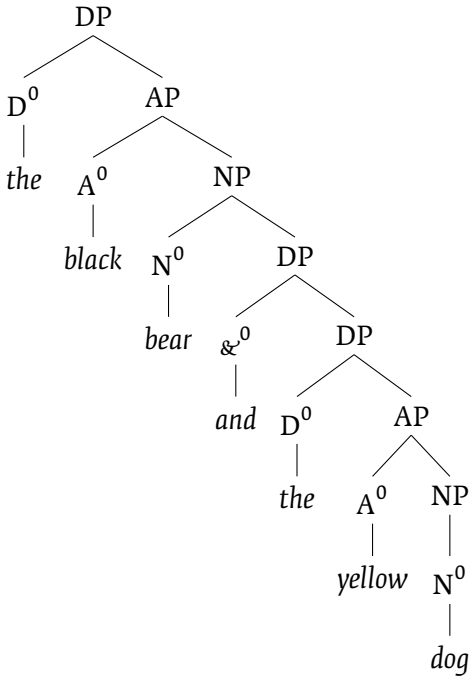
(4) a. Spec-X^0-Compl structure (Kayne 1994; Zhang 2010)



b. BP+adjunction structure (Munn, 1993)



c. Asymmetrically total structure (Velde, 2005)

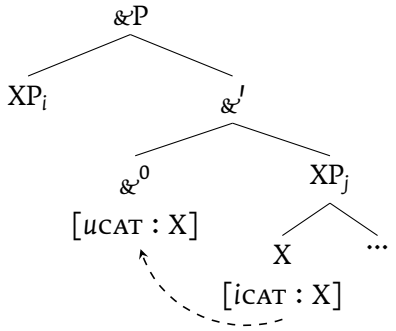


In terms of Munn's (1993) derivation, it is not clear how the adjunction of

the Boolean Phrase to the first coordinand is theoretically-conceptually motivated. In the case of Velde's (2005) model, it is even more difficult to discern the mechanics, let alone the theoretical motivations, underlying (i) the adjunction of the coordinator to the the internal coordinand, as well as (ii) the complementation structure of the higher NP, where N^0 freely takes a DP, with the adjoined $\&^0$, as complement (where the external adjunction of a minimal category to a maximal one goes against the minimalist tenets we aim to maintain). For these fundamental technical and conceptual reasons, I adopt the model that is consistent with the current theoretical assumptions of Minimalist syntax and which requires the least amount of stipulation, namely the simple Spec-X^0-Compl structure in (4-a).

The core intuition in theorising about the general coordinate structures is as follows. A coordinate complex of two DPs should itself be a DP and, along the same lines, a complex of two propositions should itself be a proposition, since we would like to restrain from positing and invoking *ad hoc* categories, such as $\&^0$. In our theory of coordinate syntax, there are two desiderata: we want to derive the coordination so that the $\&^0$ inherits the category (i.e. categorial features) from its coordinand/s, while still maintaining that $\&^0$ carries a primitive concatenating feature, which semantically functions like a connective operator. Any syntactic theory of coordination should adhere to this intuition and satisfy the percolation of the categorial makeup of coordinands. A way of implementing this condition on the overall categoricity of coordinate complexes within a Minimalist framework (Chomsky, 1995) is to posit an uninterpretable categorial feature [*u*CAT] on $\&^0$, which is checked under `Agree`, as per (5).

(5) Categorising conjunction by categorial absorption:



EMPIRICAL MOTIVATION FOR A (CON)JUNCTION SUPERSTRUCTURE

In the final step, let me reproduce the arguments for upgrading the standard coordination structure in light of the cross-linguistic data such as the following that align even with the English 'long conjunction' expressions ('*both* Bilbo *and* Gandalf') which do not fit into the standard structure we have been entertaining, as the following set of evidence from genetically varied languages show.

(6) *i* *Mujo* *a* *i* *Haso*
AND/ALSO= μ NAME AND/BUT AND/ALSO= μ NAME
"both Mujo and Haso" (Czech) or
"not only Mujo but also Haso" (Ser-Bo-Croatian)

(7) *i* *Mujo* *i* *i* *Haso*
AND/ALSO= μ NAME AND AND/ALSO= μ NAME
"both Mujo and Haso" (Macedonian)

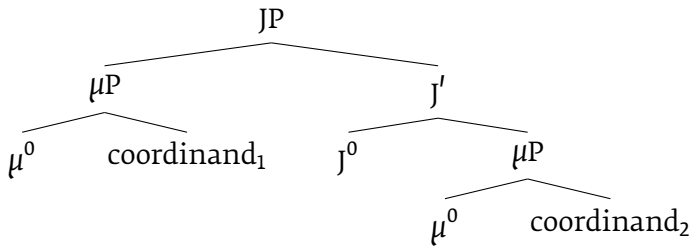
(8) *Mujo* *-is* *és* *Haso* *-is*
NAME AND= μ AND NAME AND= μ
"(both) Mujo and Haso" (Hungarian)

(9) *Mujo* *-gi* *va* *Haso* *-gi*
NAME AND/ALSO= μ AND NAME AND/ALSO= μ
"(both) Mujo and Haso" (Avar)

(10) *Mujo* *-ts* *da* *Haso* *-ts*
NAME AND/ALSO= μ AND NAME AND/ALSO= μ
"(both) Mujo and Haso" (Georgian)

To accommodate such strings under and within a common structure for conjunction, I adopt the Junction Phrase (JP) structure, building on Slade (2011), in which both (what we have been labelling as) the internal and the external coordinand positions are headed by a 'lower' or 'light' conjunction head, dubbed μ^0, culminating in a rich conjunction structure such as the one given in (11).[5]

(11) Rich JP-conjunction structure (Mitrović, 2014; Mitrović, 2021):



While the medially placed conjunction marker, taken to instantiate the J head, in the examples above may well fit into the standard phrase-structre, the additional conjunction markers cannot. The upgraded Junction structure, joining two μPs, can – without any additional stipulation. Since this may well be a universal structure for conjunction (Mitrović, 2021; Mitrović & Sauerland, 2016), where the amount of pronounced structure is cross-linguistically variable, I will consider it to be the underlying structure in IIr given the typological arguments developed elsewhere in the literature. Furthermore, and based on this empirical motivation for a richer underlying structurer for conjunction expressions, I will show that it is the pronounce-

5 For details and extensive discussion, see Mitrović (2021, Ch. 2) and citations there.

ment of this superstructure that is diachronically and dialectally variable across IIr, to which I now turn.

3 CONJUNCTION SYNTAX IN ARCHAIC INDO-ARYAN

This section provides a synchronic analysis the syntax of conjunction in the history of IA, starting with RV.

3.1 RIGVEDIC

RV Sanskrit, along with a majority of early Indo-European (IE) languages, operated – what I dub here – the double system of coordination, whereby coordinate constructions are two types. In the first type, the coordinator (*utá*) occupies a medial (or final in case of short nexus) surface position with respect to its coordinating arguments (coordinands) (12). In regard to categories it coordinates and placement, the coordinator *utá* behaves very much like English *and* in terms of the position it occupies in the coordinate configuration.

(12) *mā́ no mahā́ntam utá mā́ no arbhakám*
NEG US great.ACC UTA NEG US small.ACC
'[O Rudra, harm] not either great or small of us' (RV: ṚV, 1.114.07[a])

The second type of coordinate construction is headed by an enclitic or postpositive coordinator like *ca*, which is restricted to second-position (2P) in the coordinate complex. In case of simplex coordination, the 2P is simultaneously a superficially final-position, as shown in (13). This second type of coordinator is unlike *and*, both in regard to the categories with which *ca* tends to combine or the configurational 2P status.

(13) *bhāsā śrávobhiś ca*
radience.INSTR fame.INSTR and
'with (thy) radiance and with (thy) fame.' (RV: ṚV, 6.1.11[ab])

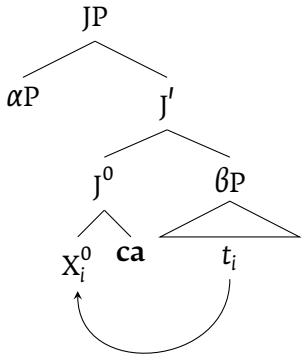
As Klein (1985a, 88) observes, *ca* in *Ṛgveda* normally functions as a coordinator signalling tighter nexus between shorter units, while *utá* serves as a higher level concatenator conjoining longer stretches of discourse. An LCA-compliant approach to phrase structure allows us to view the differential surface placements of the coordinator in the coordinate allo-sentences – *utá*/*ca* as in (12) and (13) – as underlyingly occupying a single position and therefore to derive from different featural makeup of, prima facie, the two seeming types of coordinating heads. As seen in (14), the two conjunction markers generally accord with Klein's description, which we explore at greater length below.

(14) *yásmin víśvāś carṣaṇáya utá cyautnā́*
upon.whom.M all men UTA achievements.PL.NOM

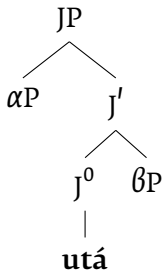
jráyāṃsi$_j$ *ca*
regions.N.PL.NOM CA
'He upon whom all men depend, [and] all regions, [and] all achievements, [he takes pleasure in our wealthy chiefs.]' (RV: ṚV, 8.2.33^{ab})

The analysis I develop derives both coordinate allo-structures (i.e., *utá*- and *ca*-type) from a single default structural template. The derivation will generally follow along the following lines, stemming from the basic assumptions I started with. Assuming a rich conjunction, resting on a relatively default syntactic template for coordination (Kayne 1994; Zhang 2010), coupled with suppositions of universal antisymmetry operating in narrow syntax (Kayne 1994; Biberauer et al. 2010), I am led to maintain that all (or both) coordinate configurations departing from head-initial configuration are derived through movement. As the *utá*-type of coordinate exprerssion (12) is consistent with the tenets of the LCA that all underlying configurations are head-initial, we consider that no movement is involved. For the *ca*-type coordination, which includes second position (2P) placement and encliticisation of the coordinator, we may posit a movement operation as *ca*-type configurations departs from the Spec-X^0-Compl linear base. In line with this preliminary idea, let me tentatively submit a cursory analysis along these lines whereby one conjunction head, realising as a 2P conjunctive marker *ca*, triggers head movement of, or from, its complement/internal coordinand, which additionally and phonologically feeds cliticisation, in line with Roberts's (2010) Defective Goalhood model, while the other type of conjunction head, realising as *utá*, does not trigger such movement. I sketch the two types in (15) below.

(15) a. The syntax of a *ca*-type configuration:



b. The syntax of an *utá*-type configuration:



The mechanics of movement signalled in (15-a) derived from the absorption of the categorial feature by *ca* (5), turning the closest minimal category of its argument, generally belonging to that category, to constitute a defective goal, as per (2).

One important empirical prediction that the spirit of the narrow-syntactic anlysis laid out above in (15) concerns the subjacency of the possibly movement-triggering conjunction head and its complement. This concerns the first of the two signature properties of the double system that I now turn to addressing, briefly reproducing the arguments put forth in Mitrović (2013).

There exists a distributional asymmetry between *ca* and *utá* conjunctions in RV in regard to the type of the category their coordinands belong to. The following table, drawing on counting from Klein (1985a,b), summarises the distribution of the two conjunction markers.

conjunction marker	overall distribution		clausal		subclausal	
	#	%	#	%	#	%
utá	705	47.64	364	51.66	341	48.34
ca	775	52.56	59	7.61	714	92.39

TABLE 1: The overall and categorial distribution of coordinands in ṚV, based on Klein (1985a,b).

Mitrović (2013) shows that the difference between the grammatical behaviour and the empirical distribution of the two conjunction is rather better understood in terms of the availability of syntactic objects which the enclitic conjunction head *ca* may probe. The crux of the analysis is the inability of overt clausal heads to incorporate into a conjunction head due to their strongly[6] phasal (π) status.

Aside from the category-selecting tendencies, the two types of conjunction markers also differ in terms of their morphological structure and complexity. The particles *uta* and *ca* are taken to be phasally-conditioned allomorphs: when *ca* cannot probe for a host from within its complement, *u* is realised to satisfy *ca*'s 2P requirement, surfacing as a word-internal Wackernagel effect. The bimorphemic *utá* is analysed to reflect the overt pronunce-

6 By *strongly* phasal, I refer here to those phases whose edge excludes the minimal category, refering back to Chomsky's (2001) original formulation of the Phase Impenetrability Condition (PIC), distinguishing between strong and weak Phases; but see Richards (2007) and those he cites for the relevant discussion.

$X_{\in CP}$ + *ca* combination	distribution of clausal *ca*-conjunction	general [±CP] distribution
[Spec, $CP_{\pi}^{0\star}$] + *ca*	77.97% (N =46)	5.94%
$C_{\pi}^{0\star}$ + *ca*	22.03% (N =13)	1.68% (p <0.001)

TABLE 2: Distribution of clausal *ca* conjunction: head *vs* Spec hosts. (N = 59))

ment of both the J and the μ heads that feature in the derivation of the conjunction expression. Mitrović (2013; 2014; 2021) shows that the fact that nearly all archaic IE languages that have a clausal 1P conjunction marker are bimorphemic reflects this aspect of inherited conjunction structure.

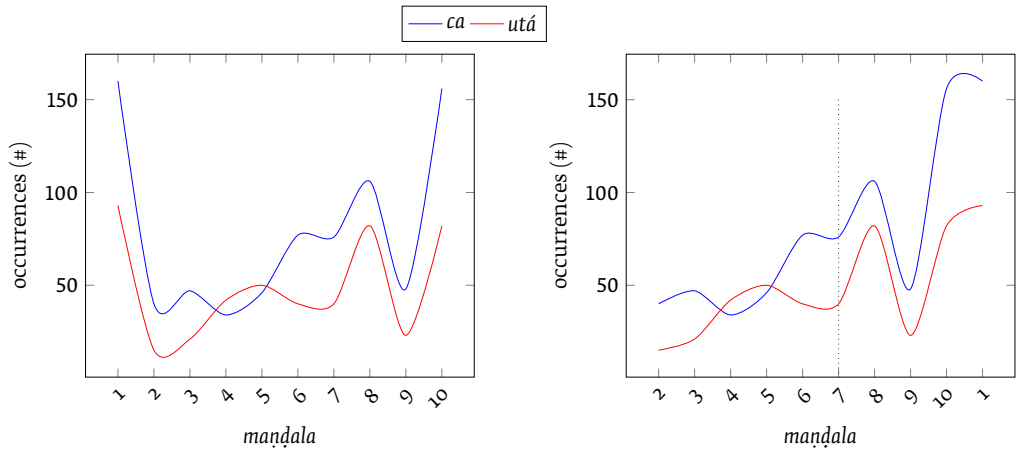
While the enclitic coordinator *ca* is considered to derive from *-$k^{w}e$, *utá* cannot be considered to derive from a *single* reconstructed form. In fact, *utá* is pleonastic, i.e. compound of two coordinators, reconstructable as a word-level particle compound $\langle ^{\star}h_2u + {}^{\star}$-$te\rangle$, comprising of an orthotone and an enclitic part, as shown in (16-a). Dunkel (1982) also recongnises the pleonasticity of many IE orthotone coordinators listed in (16-b)–(16-d), from which I extrapolate a generalised form for IE in (16-e).

(16) a. Vedic *utá*, Greek *aute*, Latin *aut* $= {}^{\star}h_2u + {}^{\star}$-$te$
b. Ved. *u ca*, Goth. *uh* $= {}^{\star}h_2u + {}^{\star}$-$k^{w}e$
c. Goth. *jau* $= {}^{\star}yó + {}^{\star}$-$h_2u$
d. Hit. *takku*, OIr. *toch* $= {}^{\star}tó + {}^{\star}$-$k^{w}e$
∴
e. **Freestanding/1P coordinator in IE** $= {}^{\star}J^{0} + {}^{\star}\mu^{0}_{[2P]}$

The head-initial and configurationally medial coordinators across archaic IE, as Dunkel (1982) notes, are therefore not single heads but in fact pleonastic forms, comprising of an an orthotone and an enclitic half.

These two morphemic halves are underlyingly taken to be reflective of two functional elements J and μ. The analysis in which two morphemes are analysed as verbalisation of the two functional heads makes another prediction. The JP structure is bicyclic in nature insofar as its analysis predicts that the lower μP cycle is independent of the higher JP, *ceteris paribus*. In fact, the μ-markers should feature independently in logically related expressions, such as those of additivity and quantification – what I call here 'monadic conjunction' since they combine with a single argument. If *ca* is indeed μ, then *ca* should express non-conjunctive meanings, to the exclusion of 1P/non-2P conjunction makers like *utá*. This is borne out in full format as the following example from ṚV demonstrates (this is valid throughout early IE).

(17) *(prát)īdám vīśvam modate yát kim -ca prthivyāmádhi*
DEM world exults REL WHAT CA= μ] world.F.ACC-upon
"This whole world exults **whatever**$_{\mu P}$ is upon the earth." (RV: ṚV 5.83.9^{c})



(A) Sequential arrangement (B) Historical arrangement

FIGURE 4: The distribution of occurrences of the 2P and 1P conjunction markers *ca* and *utá*, respectively, across the ten books of the Rigveda corpus, given in (A) sequential and (B) diachronic formats with the more archaic (family) books (2–7) placed on the left.

The core synchronic analysis of the double system of conjunction here supposes that there was a competition between *ca* triggering incorporation of a 2P-host and realising a last-resort J-host structurally from above to satisfy its Wackernagel condition. The change in the loss of the double system, therefore, reflects the shift in the grammar that in RV was economy-based.

(18) **Diachronic Inflation**:
Make more expensive what used to be cheaper

The main change supposes that there is a generalised competition between a derivational strategy in which the relevant head (μ^0) searches for the closest and most eligible minimal category, to be probed and moved, versus a more blind last-resort reliance on the structurally higher head (J^0) to act as host satisfying the probing head's (μ^0) Wackernagel requirements (where the morphosyntactially cheaper/more economical options bleeds semantic multi-functionality of μ).. While I will show that the Iranian branch, along with the other IE language families, uniformly opted for the latter, post-RV Sanskrit shows the opposite trend.

There is reason to suppose that there were RV-internal precursors to the start of the principle of Diachronic Inflation. Consider first the basic distribution of *ca*- and *utá*-marked conjunction in ṚV, according to the traditional sequencing of the maṇḍalas (books), shown in Fig. 4.

The oldest part of the RV, maṇḍalas 2 through 7, the so-called family books (see Witzel 1997, 262 and references therein), seem to correlate statistically although a significant effect is not detected.[7] In the Late Vedic (LV) period, comprising the dates of composition of the other vedic texts, a decline of

7 The χ^2 statistic with Yates correction: $\chi^2(1, N = 1278) = 0.4735$, the p-value is 0.491371, not significant at $p < 0.05$.

text	time			*ca*		*uta*		Σ	corpus size
	earliest	latest	mean	#	%	#	%		
Ṛgveda	-1700	-1100	-1400	1092	59.32%	749	40.68%	1841	170930
Sāmaveda	-1200	-800	-1000	139	79.89%	35	20.11%	174	41266
Kṛṣṇayajurveda	-1200	-800	-1000	117	83.57%	23	16.43%	140	19565
Atharvaveda	-1200	-800	-1000	1194	82.23%	258	17.77%	1452	71259
Taittirīyabrāhmaṇa	-400	-300	-350	120	96.00%	5	4.00%	125	14416
Mahābhārata	-300	300	0	48421	99.19%	393	0.81%	48814	1145905

TABLE 3: Distribution of 2P *ca* and 1P *uta* conjunction markers in the history of Sanskrit, from early Vedic to classical. NB: the negative time-points refer to periods BCE.

the 1P conjunction maker is evident and, being symptomatic of, the gradual loss of, what I have dubbed, the double system of coordination. This is demonstrated in Fig. 5, based on the statistical data from Tab. 3:

3.2 CLASSICAL SANSKRIT

In Classical Sanskrit, namely the post-LV period of *Mahābhārata*, the 1P strategy of expression conjunction is nearly non-existent, and the 2P *ca*-based means of expressions completely overtakes the system.

This overtaking of *ca* does not seem to have structurally changed since its Vedic stage at all, since the monadic conjunction structures expressing quantificational meanings is likewise expressible with *ca*:

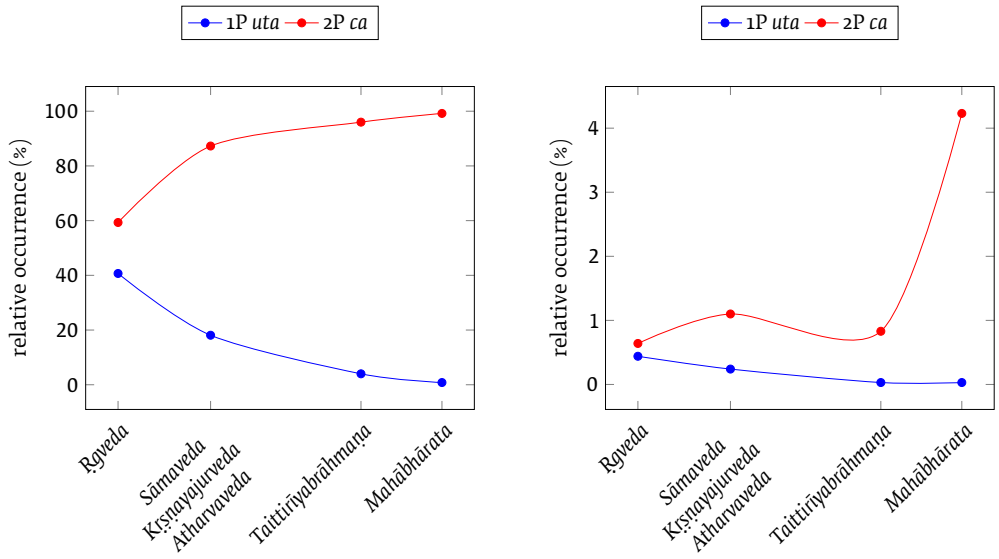
(19) *na yasya kaś -ca tititarti māyā?*
NEG whom.GEN who.M.SG CA.μ able to overcome illusions.PL
"No one [=not **anyone**] can overcome that (=the Supreme Personality of Godhead's) illusory energy." (CLSkt: BP, 8.5.30)

I return to the diachronic semantic details of the change in post-Classical Sanskrit in Sec. 5.2.1.

3.3 MIDDLE INDO-ARYAN

3.3.1 EARLY MIDDLE INDO-ARYAN: AŚOKAN PRAKRIT

While I generally try to provide a temporal view of IIr diachronic varieties as dialects, I turn in this subsection to the more truly dialectal varieties of IA with an areal distribution. Let me, therefore, finally discuss the conjunction grammar of early Aśokan Prakrit (AP) Māghadī, as reflected by the Edicts of Aśoka, a collection of over thirty multilingual inscriptions on pillars, boulders and cave walls – Tab. 4 shows the languages in which the edicts were inscribed. These inscriptions, dating to the Mauryan rule of Aśoka between 268 BCE and 232 BCE, are generally divided into four categories, according to size (minor versus major) and medium (rock versus pillar). Diachronically, the rock edicts predate the pillar inscriptions.



(A) Proportional distribution.

(B) Normalised distribution.

FIGURE 5: Plotting the relative occurrence of 1/2P conjunction systems in all of the core vedic texts by their respective mean time of composition, spanning circa a millennium, up to the classical period: *Ṛgveda*, *Sāmaveda*, *Kṛṣṇayajurveda (Taittirīyasaṃhitā)*, *Atharvaveda*, *Taittirīyabrāhmaṇa*, and *Mahābhārata*, as per Tab. 3.

	ROCK	PILLAR
MAJOR	Prakrit, Greek	Prakrit
MINOR	Prakrit, Greek, Aramaic	Prakrit

TABLE 4: The languages of inscriptions on the Edicts of Aśoka (269–233 BCE)

The Minor Rock Edicts, being the earliest, were inscribed in the tenth year of Aśoka's reign, i.e. dated to cca. 259 BCE. The Minor Pillar Edicts are contemporaneous with the Major Rock Edicts and date to the twelfth year of Aśoka's reign, i.e. 257 BCE. Since the chronological difference between the two periods is slight, barely any valid diachronic analysis can be attempted, hence my focus here is solely on the areal-dialectal features.

In Māghadī AP of the third century BCE, we see exclusively the 2P *ca* conjunctions just like, and inherited from, ClSkt. As was the case in ClSkt, Aśokan Prakrit does show a single use of the 1P conjunction *utá*. See Oberlies (2003) and those he cites for context and details on Aśokan Prakrit. What may be interesting is the following conjunction featuring a *wh*-pronoun and *ca* and which is not interpreted quantificationally, i.e. not as monocyclic μP, but rather as CP conjunction:

(20) *ki ca iminā katavyataraṃ yathā svagāradhi*
what CA more desirable than this
"And what is more desirable than this, viz. the attainment of heaven?"
(AP – G: GKSh, IX^{L})

While the syntax of conjunction in AP does not show any change from the the contemporaneous ClSkt, there is evidence to be found in the inscriptions for a semantic change. I return to the details of this in Section. 5.2.

4 CONJUNCTION SYNTAX IN ARCHAIC IRANIAN

4.1 OLD AVESTAN

Old Avestan (OAv) shows the distribution of *ca*, both in its morphosyntactic and semantic profiles, identical to that of RV, including its independent productivity to express quantificational monadic conjunctions:

(21) a. at ahurā huuō mainiiūm zaraϑuštrō vərəṇtē mazdā
"Thus, he there, Zarathustra, O Ahura, prefers (your) inspiration"
b. *yas-tē ciš-cā*
REL.NOM.SG.-DAT./GEN.SG.ENCL who.NOM-AND
spəništō
most-holy.NOM.SG.M
"**whichever**$_{\mu P}$, O Mazdā, (is) your most life-giving" (Y 43.16)

(22) a. *yōi mōi ahmāi*
REL.NOM.PL.M 1.SG.DAT.GEN.SG.ENCL DEM.DAT.SG.M
səraošəm dąn caiias-cā
NAME.ACC.SG.M GIVE.3.PL.AOR.INJ.ACT WHO.NOM./ACC.N-AND
"**Whosoever**$_{\mu P}$ shall give readiness to listen to this one of mine,"
b. upā.jimən hauruuātā amərətātā
"shall come to wholeness (and) immortality" (Y 45.5).

(23) a. *yā.* *zī* *cī-cā* *vahištā*
REL.INSTR.SG.M/N PRT/indeed what-AND best.INSTR.SG.N
"For **whatever** best (things)"
b. hanarəə θβahmāt zaošāt drəguuā baxšaitī.
"the one possessed by the Lie shall give out, (it is) without (thereby obtaining) your pleasure,"
c. ahiiā šiiaoθanāiš akāt āšiiąs manaŋhō.
"(because of) dwelling—on account of his (own) actions—on the side of bad thought" (3.47.5).

Despite the full-fledged grammar of *ca*, in both its conjunctional and non-conjunctional profiles, it unclear how this contrast with the allegedly closest dialect, that of RV, is to be understood in light of the relative absence of *uta* in OAv. In the OAv, two occurrences of *uta* are found:

(24) Y 35.6 (YH): "*As thus both man or woman* knows (the duty), both thoroughly and truly, so let him, or her, declare it and fulfil it, and inculcate it upon those who may perform it as it is. 7. We would be deeply mindful of Your sacrifice and homage, Yours, O Ahura Mazda! and the best, (and we would be mindful) of the nurture of the Kine. And that let us inculcate, and perform for You according as we may; and (for) such (praisers as we are)."
a. *yaθā āṯ* *utā* *nā* *vā nāirī* *vā*
like EMPH.PTC UTA man or woman or
b. vaēdā haiθīm
c. aθā haṯ vohū
d. taṯ əə-ād-ū vərəziiōtūcā īṯ ahmāi
e. fracā vātōiiōtū īṯ aēibiiō
f. yōi īṯ aθā vərəziiąn
g. yaθā īṯ astī

(25) Y 40.4 (YH): "So let there be a kinsman lord for us, with the laborers of the village, and so likewise let there be the clients (or the peers). And by the help of those may we arise. *So may we be to You, O Mazda Ahura!* holy and true, and with free giving of our gifts."
a. aθā xvaētūš
b. aθā vərəzənā
c. aθā haxəmąm xiiāṯ
d. yāiš hišcamaidē
e. *aθā* *və* *utā* *xiiāmā*
so/thus/likewise to-you.DAT UTA may-we-be.1.PL.PRES.OPT.ACT
mazdā ahurā
NAME
f. aṣauuanō ərəšiiā ištəm rāitī

It also seems reasonable to me to consider one of the two diachronic the-

ories light of this fact alone: either RV is more retentive and therefore OAv lost the double system of conjunction marking by the time of its composition. The other theory is inverse: perhaps OAv is more retentive and a single enclitic 2P morphosyntax of conjunction, in both its dyadic/standard and monadic/quantificational uses (structurally allosemic) is the original strategy for expressing conjunction. To get closer to answering questions such as this one, let me bring into the discussion evidence from Young Avestan in the next subsection.

4.2 YOUNG AVESTAN

The *ca*-based marking remains the most productive strategy of conjunction in Young Avestan (YAv), just like in OAv, which is in stark contrast to Old Persian (OP) with which YAv is alleged to be relatively contemporaneous and more closely related. In my analysis, I look at two chronological layers of YAv: the early YAv (EYAv) and Late YAv (LYAv). Nonetheless, there are some novel trends discernible at that stage of language change compared to OAv. Let me discuss two: the demonstrably gradual appearance of 1P *utā* marking as well as the semantic decline of the monadic conjunctions – these two facts testify to an onset of a change of which we later see the effects in OP and ultimately in Middle Persian (MP).

EARLY YOUNG AVESTAN: THE NOVEL APPEARANCE OF UTĀ

The distribution of the novel *uta* in YAv is contained only to three hāitis ("sections tied together", from *hā-*, "to bind, tie"): 9 through 11, as plotted in Fig. ?? – the *Hom Yašt*, being a later liturgical text, itself presumably reflecting late YAv. Compared to the even distribution of *ca* and *utá* across the Rigvedic texts, consider the uneven distribution of the *utá* in YAv, as given in Fig. ??, which can be taken as evidence of novelty at that stage of the language development.

The late YAv found in the Yašt texts, comprising 21 hymns, on the other hand, shows a much more even distribution of the conjunction marker *uta*. Given that Yt is written in late YAv, then the discernibly more even distribution of *uta* may be taken to suggest a more canonical place of *utā* in the late YAv grammar.

As noted, the 1P conjunction maker *uta* is found only in YAv, in fact only 8 times in books nine through eleven, already presumably belonging to the younger canon.

(26) The eight occurences of *uta* in EYAv (Hom Yt):

a. 9.22 (1×)
b. 10.4 (2×)
c. 10.7 (1×)
d. 11.1 (2×)
e. 11.3 (2×)

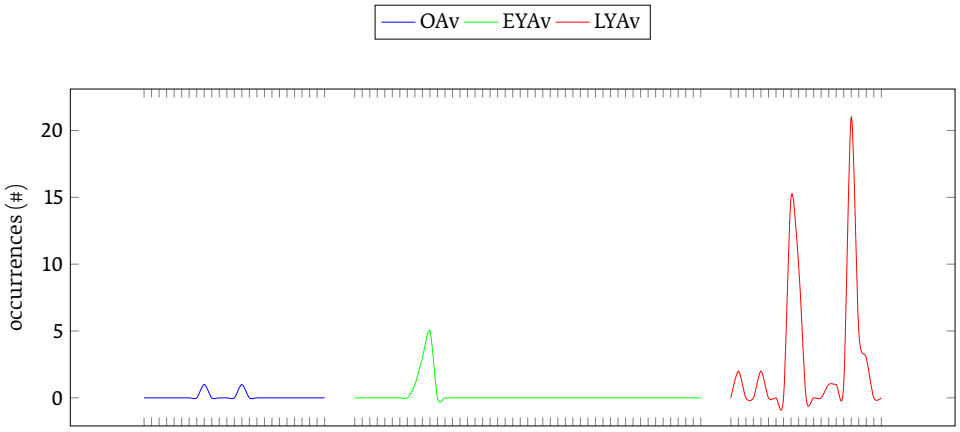


FIGURE 6: The distribution of occurrences of the 1P conjunction marker *uta* in OAv, Early YAv (EYAv) and Late YAv (LYAv) across Y and Yt corpora, respectively.

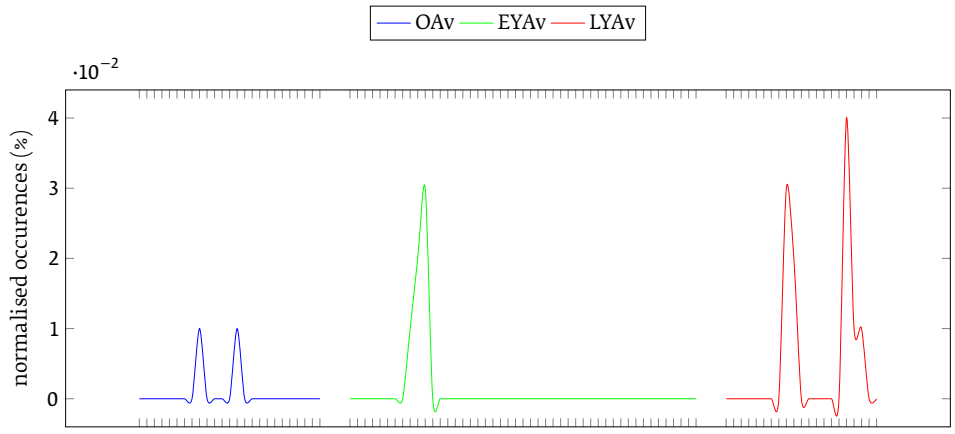


FIGURE 7: The normalised distribution of occurrences of the 1P conjunction marker *uta* in OAv, Early YAv (EYAv) and Late YAv (LYAv) across Y and Yt corpora, respectively.

Despite being relatively small in size, the pool of data with the eight occurrences suggest that, while the novel morphosyntax of *uta* is uniform, in that it consistently occupies the first and never the second position, its semantics is far from being stably conjunctive.

Out of the eight examples, let me exemplify two from Chapters nine and ten which showcase the *uta* particle in its early conjunctive roles, in both cases at the level of nominal argument, presumably at the DP-level.

(27) 9.22 (Hom Yt): "Haoma grants to racers who would run a course with span both speed and bottom (in their horses). Haoma grants to women come to bed with child a brilliant offspring and a righteous line. Haoma grants to those (how many!) who have long sat searching books, more knowledge and more wisdom."

a. haomō aēibiš yōi auruuaṇtō
b. hita taxšəṇti arənāum
c. zāuuarə aojåsca baxšaiti
d. haomō āzīzanāitibiš
e. daδāiti xšaētō puθrīm
f. *uta aṣauuafrazaiṇtīm*
UTA righteous-line.ACC.SG.F
g. haomō taēcit yōi kataiiō
h. naskō frasåŋhō åŋhəṇte
i. spānō mastīmca baxšaiti

Just as in (27), the other example also shows a nominal-level conjunction in the last line:

(28) 10.7 (Hom Yt): "Wasting doth vanish from that house, and-with it foulness, whither in verity they bear thee, and where thy praise in truth is sung, the drink of Haoma, famed, health-bringing (as thou art) to his village and abode they bear him."

a. nasiieiti haθra frākərəsta
b. ahmat haca nmānāt āhitiš
c. yaθra bāδa upāzaiti
d. yaθra bāδa upastaoiti
e. haomahe baēšaziiehe
f. ciθrəm dasuuarə baēšazəm
g. *ahe vīse uta maēθanəm*
here.GEN.SG.N village.DAT.SG.F. UTA dwelling.ACC.SG.N.

In later Av historical dialects of Ir, *uta/utā* sees a steady increase, as shown in Figs. 6 and 7, nearly completely overtaking the grammar of conjunction by the time of OP. Let me comment on some other changes that occur in the later Av period.

THE NOVEL DISAPPEARANCE OF MONADIC *CA*

Another novelty regards the semantics of *ca*-based marking in monadic structural settings. Skjærvø (2003) notes that the *ca*-based indefinites, themselves a hallmark of monadic conjunction, are less common in YAv. My corpus search shows that there are nearly none: out of 144 occurrences of *čiˇ*, corresponding to who.NOM (compared to 949 occurences in OAv), no monadic *čiˇ-ca* forms are found in the YAv texts. While textitci occurs twice (what.NOM), only one occurrence of *ci-ca* is found, in the *Vidēvdād*[8] with unclear, reduplicated and distorted use of the universal quantificational contribution of *ca*, while the long (Free-Choice) indefinite-based formula seems to reflect only relativisation, void of quantificational force expected from a fully grammaticised *ca* associated with the μ^0 status.[9]

(29) *spaiieite vīspa tā šiiaoθna yā ci -ca vərəziieiti*
takes away all DEM act.SG.N REL what CA work
"it takes away any sin that may be sinned." (Vd 3.41)

I return to a more in depth discussion of the diachronic semantics of the Iranian monadic *ca* in Sec. 5.2.2, so let me now turn to OP, a later stage of the language showing only remnants of the previously predominant *ca*-expressions of conjunction.

4.3 OLD PERSIAN

Old Persian (OP) is attested in the royal inscriptions of the Achaemenian Kings who left extensive cuneiform inscriptions dating roughly between 600 BCE and 300 BCE.

Klein (1988) showed that the 2P *ca* (*cā*) conjunction is relatively rare in OP, although a closer inspection of the OP corpus shows a productive morphosyntax of OP *ca*, nonetheless, the productivity of which I take to reflect an inherited feature, where the closest comparandum from which such an inheritance may be best be modelled as having taken place is YAv.

Let me start with the oldest OP corpus, namely the Darius's inscriptions from Behishtan. As was the case with the early YAv texts, the distribution of the relevantly rare conjunction markers – in the case of early YAv this was *uta* and in the case of OP it is *cā* – is not even across the relevant texts. For this reason, I give in Tab. 5 a specific entry for the inscriptions on the first column, as also plotted in Fig. 8.

Let me now turn to a later set of inscriptions, the so-called "Daiva Inscription" of Xerxes. The following evidence shows the productive co-occurence of *cā* and *utā*, where the former carries the additive meaning (see Mitrović 2021 for discussion of how additive meanings behind conjunction markers

8 Note the repeated verses in Vd 3.41j , 8.29j.
9 See Mitrović (2021); Mitrović & Sauerland (2016) for semantic criteria on μ-particle status cross-linguistically, which all relevant archaic IE languages pass.

	conjunction		
	cā	*utā*	Σ
(overall) DB I–V	9.62% (5)	90.38% (47)	(52)
(specific) Db I	28.57% (4)	71.43% (10)	(14)

TABLE 5: Distribution of 2P *cā* and 1P *utā* conjunction markers in DB (522–486 BC).

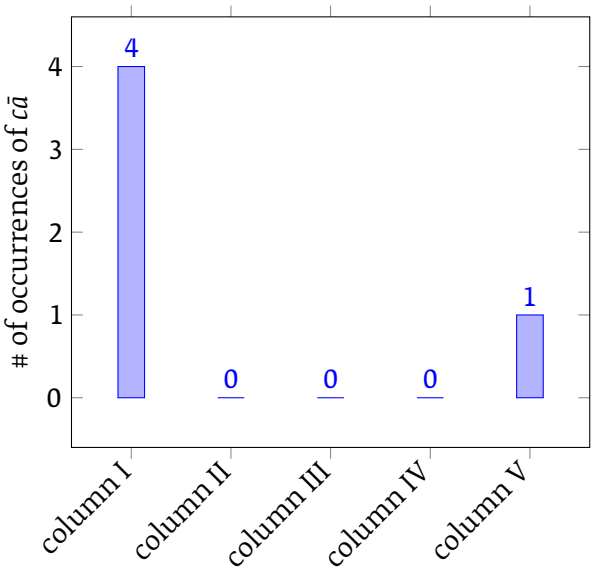


FIGURE 8: The distribution of occurrences of the 2P conjunction marker *cā* in DB.

	conjunction	
	cā	*utā*
absolutely	3	9
relatively	25%	75%

TABLE 6: Distribution of 2P *cā* and 1P *utā* conjunction markers in XPh (486–465 BCE).

	conjunction	
	cā	*utā*
absolutely	11	118
relatively	8.53%	91.47%

TABLE 7: Distribution of 2P *cā* and 1P *utā* conjunction markers in the entire OP corpus of inscriptions, containing the OP inscription texts as given in the list at the end.

such as OP *cā* demonstrate the lower-conjunction status of logical markers).

(30) *artā- cā brazmaniya utā aniyaš -ca*
truth.INST.SG.N CA reverent.NOM.SG.M UTA other.NOM.SG.N CA
"(being reverent, I worshipped Ahura Mazda) and Truth And there was yet another thing" (XPh 41)

It seems non-trivial, and in fact insightful, to plot this statistical insight against a timeline. The overall statistical inspection of the entire corpus of OP inscriptions, while low-resolution in nature and not relying on parsed texts, still shows a non-negligible grammatical presence of *cā*, as given in Tab. 7

In regard to the monadic conjunction, which we take to be a hallmark of a fully operative double system of conjunction with a fully fledge JP structure and its μP substructure, it may not be surprising that the *ca*-marked quantificational meanings are not found. I hypothesised a start of this loss for YAv based on the low-level distribution of *ca*-based indefinites and I take the negative facts in OP as evidence of this.

Nonetheless, Middle Persian (MP), while void of *ca*-based marking altogether, shows a sign for a rebirth of the cycle, namely the re-appearance of the rich JP structure with the novel μP substructure.

4.4 MIDDLE PERSIAN

In Middle Persian (MP), or Pahlavi, the literary language of the Sasanian Empire (224–651 CE), no 2P *cā*-descendant conjunction marker is found as 1P *ud* (< *utā*) is the only type of surviving conjunction.

What is additionally interesting is the appearance of a new particle, *ham* that fills the position that the archaic IIr *ca* occupied: a focus particle used to

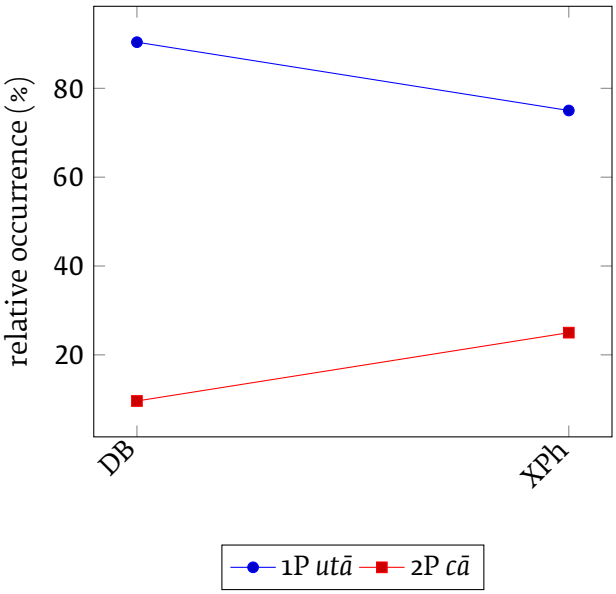


FIGURE 9: Plotting the relative occurrence of 1/2P conjunction systems in two OP texts: Xerxes' "Daiva Inscription" from Persepolis (XPh) 486–465 BCE, averaged at 476 BCE, and the Darius inscriptions from Behistun (DB) from cca. 520 BCE.

reinforce the conjunction with *ud* and mark the distributivity of the entire conjunction expression:

(31) *ham abar ahlawān ud ham abar druwandān*
HAM TO.PV righteous-PL UD HAM TO.PV unrighteous-PL
"both for the righteous and for the unrighteous" (MP: AW 52.12)

The rise of this 'new' particle is in line with the general semantic predictions for the underlying conjunction structure, both in its syntactic and semantic profiles, of the analysis put forth in Mitrović (2021); Mitrović & Sauerland (2016). While I leave the details of the theory of cyclical change in the domain of logical vocabulary in Iranian (as well as in Indic in regard to Prakrits) for the future, I take the appearance of the novel particle *ham* in MP as potential evidence for the renewed pronouncement of the rich JP structure I have alleged for archaic IIr. I return to this briefly in the Discussion.

While MP particle *ham* features in JP-headed long conjunction, functioning as an emphatic (or focus) particle, it also shows the signature μ semantics we otherwise find across other IE languages, as well as cross-linguistically. In MP, *ham* in monadic (non-conjunctive) contexts expresses additivity (cite Pahlavi dictionary, p. 39) and also features in building universal quantificational terms. (MacKenzie, 1971, 39–40)

(32) a. *ham* [*hm*], 'also': found both in Manichaean Middle Persian and (early) New Persian
b. *ham-āg* [*hm'k'*], 'all' (*h'm'g* in Manichaean Middle Persian, *hama* in

(early) New Persian), where *-āg* is a nominaliser-like formative in MP

c. *ham-bāstag* [*hmb'stk'*], 'all'

While I could not discern any combinations of wh-terms with *ham*, qualifying the MP particle *ham* to be analysed as a re- or up-cycled *μ* logical particle, we can see its multifunctional *μ*-status not only retained in Modern Persian (functioning as a marker of additivity, long conjunction, and universal quantification) but possibly extended to build indefinites in Negative Polar contexts (Negative Polarity Items, NPIs). Note that the internal structure of both the universal and existential quantifier terms involve no wh-term, which is typologically (and diachronically) less common for logical expressions of this kind.

(33) Modern Persian *ham* as *μ*:

a. *ham sib (-o) ham berenj xarid-am*
HAM=CONJ apple O=AND HAM=CONJ rice buy.PST-1.SG
"I bought (both) apples and rice."
(CONJUNCTION; Ghomeshi 2020, 69n21a)

b. *Ali Bahar-rā ham be Sara mo'arefi kard.*
NAME NAME-OBJ HAM=ALSO TO NAME INTRODUCTION DO.PST.3.SG
"Ali also introduced Bahar to Sara."
(ADDITIVE MARKER; Balogh & Kazemian 2021, n1)

c. *Man (hattâ) ye ketâb (-ham) na-xarid-am.*
I EVEN A BOOK HAM=ALSO NEG-buy.PST-1.SG
"I didn't buy any books."
(EXISTENTIAL QUANTIFIER/NPI; Toosarvandani & Nasser 2017, 673n22a)

d. *Hame=ye yax âb shod=e.*
all.√HAM-EZ ice water become.PTCP=be.PRES.3.SG
"All the ice melted."
(ARGUMENTAL UNIVERSAL QUANTIFIER; Toosarvandani & Nasser 2017, 683n57)

e. *Sohrâb hamishe qabl az xâb dandun-â-sh-o mesvâk*
NAME always.√HAM before from sleep tooth-PL-3SG-ACC brush
mi-zan-e.
IMPF-hit.PRES-3SG
"Sohrab always brushes his teeth before bed."
(ADVERBIAL UNIVERSAL QUANTIFIER; Toosarvandani & Nasser 2017, 685n62)

The productivity of *ham* in Modern Persian universal quantifiers seems to appear in fossilised form, both in argumental and adverbial quantification contexts in (33-d) and (33-e) above, with the marker displaying additive meaning (33-b) being homophonous with the conjunctive (33-a) and the NPI marker (33-c).

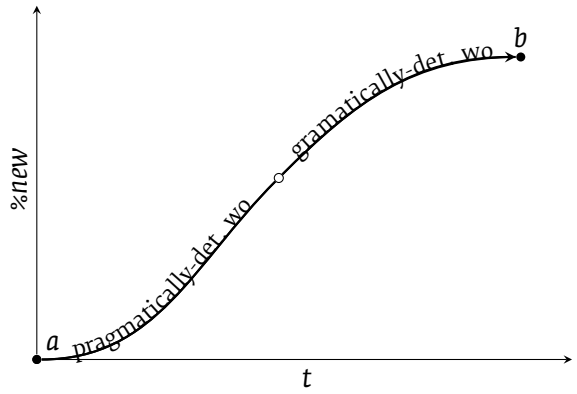


FIGURE 10: A sigmoid curve idealising word order (wo) change from the pragmatically determined to the grammatically determined, as per Denison's (2003) model.

The decline, or rather weakening of the quantificational force in monadic μ-contexts, that I conjecture for the history from Middle to (early) modern Persian is perhaps best paralleled by two independent phenomena. Firstly, the development of the negative-polarity sensitivity of μ-markers in the history of Japonic, and, secondly, by the inherent quantificational split of μ-markers across the whole of IE. I have discussed the former in Mitrović (2021, Ch. 4) and the latter in Mitrović (2019).[10]

5 A DIACHRONIC ANALYSIS

5.1 DIACHRONIC SYNTAX: ROTATIONAL-PARAMETRIC CHANGE

This section lays out and reproduces the general argument for the directed loss of 2P conjunction marking in RV that can also be found across the IE family. In doing so, I predominantly rely on Mitrović (2014; 2018; 2021) and work cited therein. The general backbone of this section relies on the idea that the diachrony of IE syntax arose from depragmaticalisation.

In this vein, Ledgeway (2015) convincingly shows that the rotational change is tied to the more question, or super macroparameter deciding, whether the word-order is dictated grammatically or pragmatically. This observational hypothesis can be traced back, as Ledgeway (2015, 35–36) notes, to Meillet (1908, 330) who held that "word order had an expressive, and not a syntactic, value" (which in itself can further be traced back to Weil (1844), but also see Ledgeway 2015, 35fn11). This notion of 'expressive', rather that 'syntactic value' finds an obverse in modern generative theory as information-structural rather than argument-structural, respectively, as I will contend. We can plot this generalised trend of change in determination of word-order as a sigmoid function as given in Fig. 10.

10 For a comparative discussion on Old Persian indefinites, see Villalobos (2019, 2022).

In theoretically more informed terms, the aetiology of the pragmatic determination of word order can be relegated to the Ā-processes that probe elements within the sentential core and trigger fronting to the clausal edge: targeting maximal categories yields the signature of configurational syntax, while discontinuous expressions result from fronting of minimal categories, a hallmark of non-configurational syntax (see Ledgeway 2015, 68ff). This shift in configurationality is also clearly reflected in the grammatical domain of conjunction. Across all branches of the IE, the 2P Wackernagel conjunction marker, such as **kwe*, sadly, uniformly, and directedly died all. It was subsequently replaced with an orthotonic 1P marking, most commonly bimorphemic in nature – a sign I have taken to support the view that IE reflected a riche JP superstructure for conjunction at word-level.

This decline in 2P conjunciton marking, detectably not only via configurational differences with the competing 1P conjunction strategy, but also independently via semantics: only monomorphemic 2P markers had the power of 'monadic conjunction': seemingly conjunctive marking of non-conjunctive meanings (indefinites, quantifiers, additives, for instance). The decline of 2P conjunction and its ultimate disappearance can, as I have suggested int this paper, help us ploy the historical trajectories of the morphosyntactic change – also at the level of dialects.

The evidence I brought forth from IIr was used to demonstrate a variable state of the underlyingly same grammatical system of conjunction marking. I have suggested that not only does the Ir branch in its archaic form parallel that of archaic IA, but its development is best understood if considered more retentive that the Indic. I have entertained the cursory evidence from MP as signalling a renewal of the cycle that rearticulates teh rich JP structure.

Given that the majority of IE behaved like RV, and unlike OAv, one may consider arguing for a reconstruction of the double system for the IE family as a whole. To square off

With these considerations, let be submit a cyclic theory that explains the dialectal differences in IIr by supposing a differential speed change, or a differential onset of the start of change. Plotted in Fig. 12 is an idealised version of the facts presented in Fig. 11.

The differential S-curves can be understood in Kauhanen & Walkden's (2017) formalisation as involving a different intercept parameter, what they dub the *k*-parameter: "[t]he *k* parameter serves to translate the curve along the time axis, indicating the point of greatest growth, or the *tipping point*," Kauhanen & Walkden (2017, 485) while the change itself proceeds at a constant rate as first proposed and demonstrated by Kroch (1989) (see Kauhanen & Walkden (2017) for further references and arguments).

The tangential evidence from the brief history of logical marking in Iranian suggested a reanalysis of a novel marker that seemingly replaces the extinct *ca*, namely the *ham* that appears in MP and survives seemingly nearly intact, and extends to cover the marking of negative polar indefinites, in

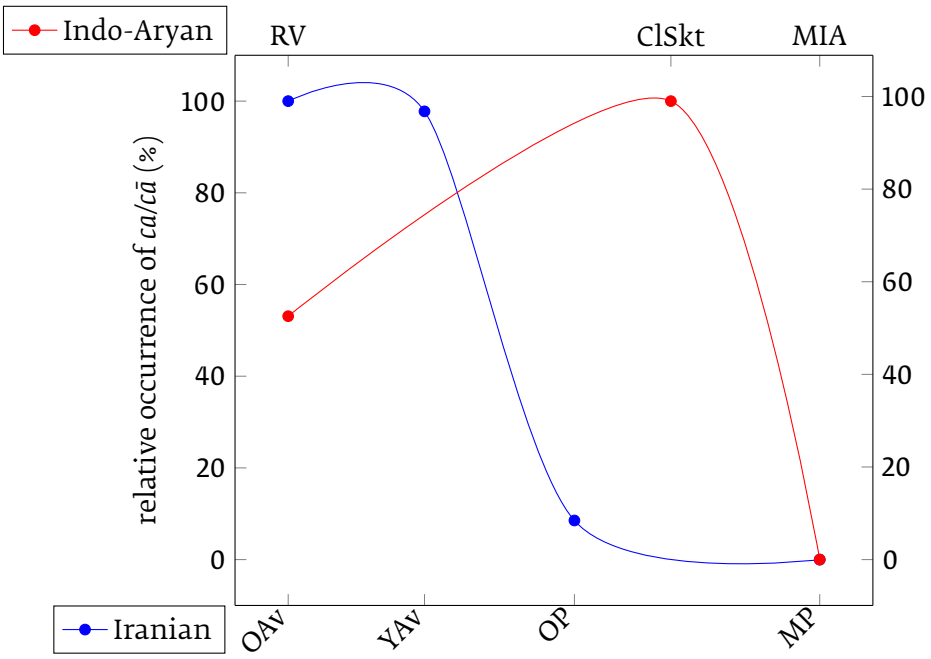


FIGURE 11: Plotting the relative occurrence of 2P conjunction marker *ca/cā* across the two branches of IIr with the Indo-Aryan historical texts plotted above and the Iranian historical texts plotted below.

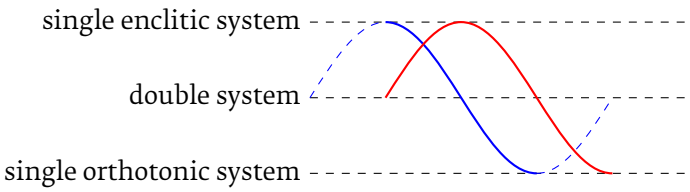


FIGURE 12: A sine wave idealisation of the cyclicity of S-curved change for the Indic (red) and Iranian (blue) branches supposing a differential onset of the start of the changes (cycles).

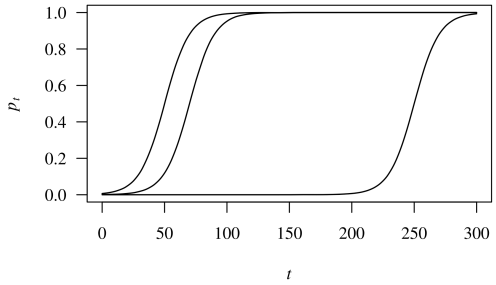


FIGURE 13: Three logistic curves with identical *s* ('slope') parameters but differing *k* ('intercept') parameters, from Kauhanen & Walkden (2017, 486, fig. 2)

modern Persian. The theoretically stronger claim underlying the sine-functional analysis of cyclicity may suggest that a single wave-lenth correspond to a change of the overt form of the corresponding marker of logical meaning. The primacy of the semantic makeup of marking versus the morphosyn-

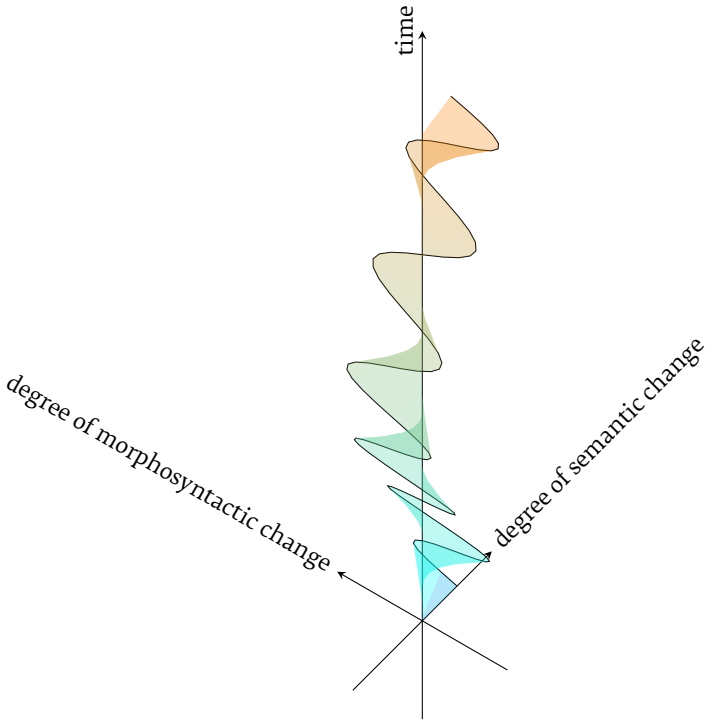
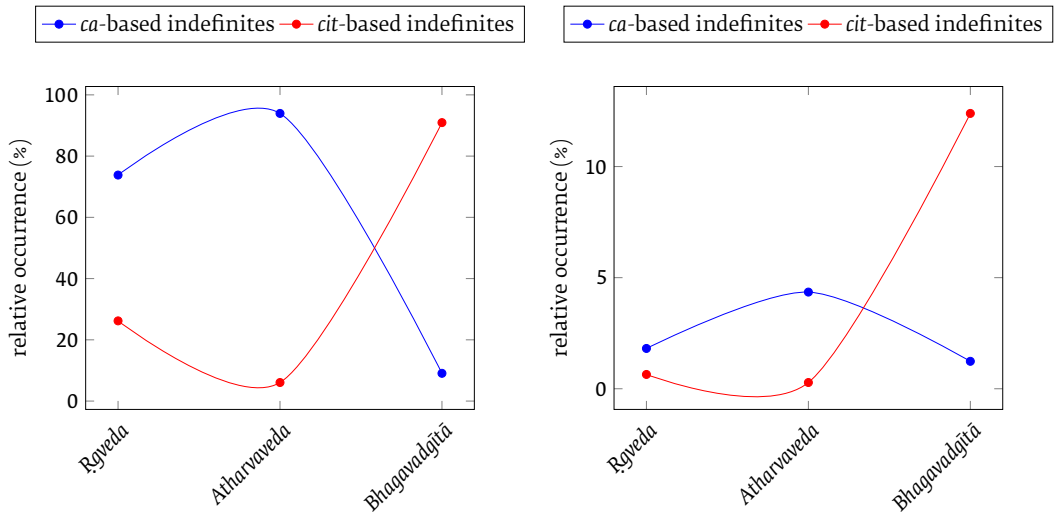


FIGURE 14: A three-dimensional metaphor for the morphosyntacto-semantic cyclicity of change.

tactic featural composite/signature of the maker – perhaps a double, or multidimensional, wave models would capture best both the independence of morphosyntax and semantics as well as its interlock in the sense that one pushes and pulls the other. The next subsection, accordingly, looks at the diachronic semantics of μ-markers in Indo-Iranian.

5.2 DIACHRONIC SEMANTICS: COMPOSITIONAL CHANGE

This subsection examines the rate and extent of semantic change associated with the fine-grained conjunction structure. I rest my historical semantic analysis of IIr dialects on the argument regarding the monadic conjunction expressions, i.e. quantificational *ca*-based terms: the relative share of quantificational expressions compared to the overall employment of the *ca* particle and its subtle overtaking by a competing quantificational particle *cit*/*ciṯ*/*cī ṯ* in later dialects of two two branches of IIr. This argument is demonstrated for the Indo-Aryan and the Iranian branches in turn in Sections 5.2.1 and 5.2.2. For each branch, I give the statistical evidence for two types of context: the first displays the relative competition between two particles (CA- versus CIT-based quantifier terms), and the second looks at the relative presence of the particle in the entire corpus (a normalised distribution).



(A) Proportional distribution.

(B) Normalised distribution.

FIGURE 15: Distribution of *ca*-based and *cit*-based indefinites in the history of Sanskrit, given in (A) raw proportional distribution reflecting the relative share of the competing markers, and (B) with the distribution normalised to corpus size (per 10k tokens).

5.2.1 SEMANTIC CHANGE IN EARLY INDO-ARYAN

While the syntax of *ca*, as a marker of conjunction, remains stably unchanged throughout the history of Sanskrit, the sematic profile of *ca* as a *μ* superparticle (in the sense of Mitrović 2021) can be seen to have changed since the archaic and post-Vedic stages of Sanskrit. In the following analysis, a sample of three texts from three historical stages is gathered: *Ṛgveda*, *Atharvaveda*, and *Bhagavadgītā* with the mean dates used as per Tab. 3 (where the dating of *Bhagavadgītā* is set to 600 BCE).

As Fig. 15 shows, the monadic semantics of *ca* can be characterised by an overall decline and substitution by the *cit*-based indefinite terms by the middle of the first millennium BCE, as the analysis of *Bhagavadgītā* shows. Prior to this decline, we can see the rise of the *ca*-built quantifiers at the very end of the second millennium BCE in *Atharvaveda*.

Note that in both (A) the proportional-competitive and (B) the normalised graphs in Fig. 15, the archaic RV grammar shows a more equal distribution of CA- and CIT-based quantifier terms, which in late Vedic becomes more unequal, showing a development trend which ends up being reversed in the late Classical period.

The declining trend of monadic *ca* and its overtaking by *cit* in Sanskrit is paralleled in early Prakrit, also. The corpus of the fourteen rock inscriptions of the Rock Edicts in AP shows only one instance of a *ca*-based. The piece of relevant evidence is given below in parallel format and involving the three AP Māghadī dialects of the edicts: the dialect of Girnār (G), Kālsī (K), and Shāhbāzgaṛhī (S). While G and S feature a *ci(t)*-based particle to build their

free-choice indefinite quantifier, K employs a *c(h)a* particle.

(34) Pillar XII^{D}: "For whosoever praises his own sect …"

(G) *yo hi ko -ci ātpa-pāsaṃḍaṃ pūjayati*
COREL PTC who CI(T) ≠ μ his-sect praises

(K) *ye [h]i ke -cha [a]ta-pāśaḍa punāti*
COREL PTC who C(H)A = μ his-sect praises

(S) *yo hi ka -ci ata-praśaḍam pujeti*
COREL PTC who CI(T) ≠ μ his-sect praises

It is serendipitous to find the single instance monadic *ca* in AP, which reveals that the semantics of early prakrit *ca* is unlike its counterpart in early Sanskrit. The quantificational role of free-choice indefinites in AP is almost exclusively handled by the *-ci/-ti* particle. Compared to *Bhagavadgītā*, being composed several centuries earlier and constituting a different dialect or sociolect, the relative disappearance of quantificational *ca* is confirmed also statistically as weakly significant compared to mid-first millenium Sanskrit (*Bhagavadgītā*) with $\chi^2(1, N = 345) = 2.6165$ ($p = 0.105755$), and strongly significant compared to its 1st millennium Sanskrit precursor (*Atharvaveda*) with $\chi^2(1, N = 80) = 68.0937$ ($p < 0.00001$).

While considered a single remnant the monadic conjunction function of *ca*, the aspirated spelling of the monadic *c(h)a* in K (34) may also be taken as an allomorph (and alloseme, in the sense of Mitrović 2021, building on Marantz 2013) of *ca*, that also has in K a freely varying allomorph *cā*, as a conjunction snippet in (35) from pillar XII shows.[11]

(35) *baha-śutā cā kayānāgā ca*
very/much-learned CA possessed of good scriptures CA
"(that all sects) should be full of learning, and should be pure in doctrine [possessed of good scriptures; Charpentier 1931, 319fn1]." (AP – K: GKSh XII^{J})

The allomoprhy and allosemy of CA which shows remnants of the inherited semantic profile of a μ logical particle CA in K dialect of AP can thus be captured as in (36): the contextual allosemy of the surface logical interpretation of ⟨CA⟩ presumably encoded on the μ category will yield conjunctive interpretation as well as the allomorph *ca* with freely varying vowel length when featuring in the context immediately local to the commanding J^{0}; independently from the J formative, ⟨CA⟩ realises as (non-conjunctive) *cha*, with an aspirate reflex in the phonological domain.

(36) Logical allosemy and free allomorphy of CA in K:

$$\langle \text{CA} \rangle_{\mu^0} \Leftrightarrow \begin{cases} \{/\text{ca}/, /\text{cā}/\} & \text{if } [_{\text{JP}} \_\_ \\ /\text{cha}/ & \text{otherwise} \end{cases}$$

11 The *cā* particle variant appearing 49 times, compared to 27 instances of *ca*, in the K Rock Edits. No contextual environment that would go against my analysis of free allomorphic variation is discernible.

text →	Y						Yt			
dialect →	OAv			YAv						
subdialect →				EYAv			LYAv			Σ
INDEFINITE BASE	#	%PROP	%NORM	#	%PROP	%NORM	#	%PROP	%NORM	
CA	5	41.67%	6.50%	29	70.73%	16.00%	5	11.11%	1.00%	39
CIT	7	58.33%	9.10%	12	29.27%	6.62%	40	88.89%	8.01%	59
Σ	12			41			45			98

TABLE 8: A distribution of the particle-based indeterminate quantification in diachronically continuous dialects of Avestan, given in raw proportional ($\%_{PROP}$) and normalised ($\%_{NORM}$) formats (to 100 token given the corpus size).

Let me now turn to the Iranian branch, where a similar diachronic semantic trend can be discerned.

5.2.2 SEMANTIC CHANGE IN EARLY IRANIAN

A parallel diachronic decline of monadic *ca*-based quantifiers can be observed in the Iranian branch, also. Just like in the IA branch, the monadic *ca* sees a bump in early YAv before its relative decline in late YAv as given in Fig. 16. Unlike the relative overtaking by the competing marker *ciṯ/cīṯ̃*, the history of Avestan dialects shows more clearly the decline of *ca*-based quantification. In light of the background theory of *μ*-particles and the rich conjunction structures they feature in, both archaic Indic and archaic Iranian show a semantic change associated with the semantic profile of *ca/cā*. Fig. 16 graphically summarised Tab. 8.

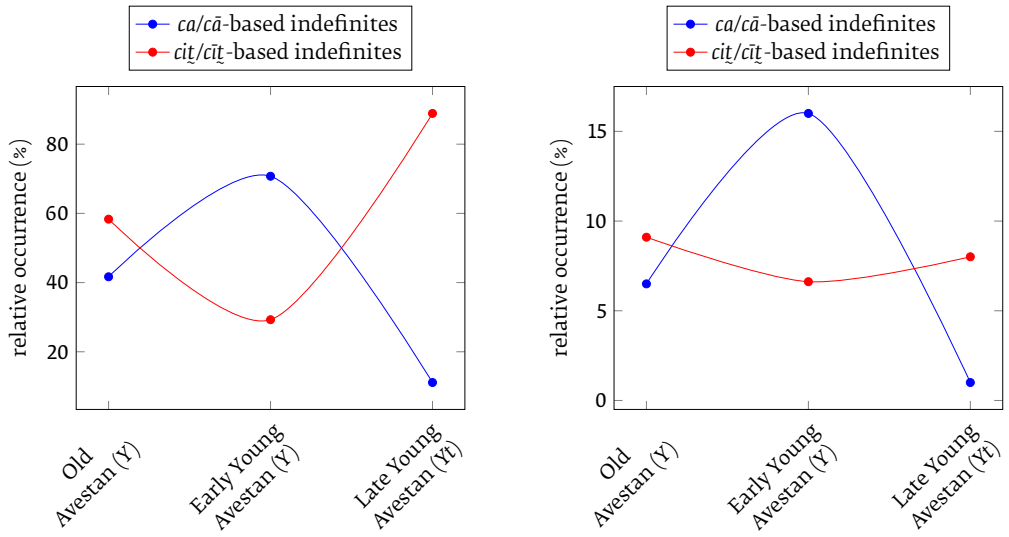
While a mild statistical significance is detectable for the quantification strategies in OAv versus non-OAv historical dialects, the effect is statistically far more significant when late YAv is compared to its predecessors. This is summarised in Tab. 9.

	OAv ~ non-OAv	LYAv ~ non-LYAv
ca/cā ~ *ciṯ/cīṯ̃*-based indefinites	0.0301. (p = 862296)	26.4051 (p < 0.00001)

TABLE 9: An analysis of statistical significance of particle-based indeterminate quantification in diachronically continuous dialects of Avestan; values for $\chi^2(1, N = 98)$.

AN EXCURSUS ON THE PARALLEL PHONOLOGICAL MIRROR OF HISTORICAL DIALECTS OF AVESTAN

The three stages of Avestan semantics are also mirrored in the morphophonology of the 'shape' of the relevant particle, pertaining to both *ca* and *ciṯ*. Let me take a brief excursus to demonstrate it. Phonological differences are



(A) Proportional distribution.

(B) Normalised distribution.

FIGURE 16: Distribution of *ca*/*cā*-based and *cit̰*/*cīt̰*-based indefinites in the history of Avestan, given in (A) raw proportional distribution reflecting the relative share of the competing markers, and (B) with the distribution normalised to corpus size (per 100 tokens).

clear in the diachronic dialects of Iranian insofar as the vowel length in each of the two particles is concerned, as Tab. 10 shows.

	text →	Y				Yt	
	dialect →	OAv		YAv			
	subdialect →			EYAv		LYAv	
		#	%	#	%	#	%
CA	*ca*	21	4.23%	1585	84.49%	3499	78.37%
	cā	476	95.77%	291	15.51%	966	21.63%
	total	497	100.00%	1876	100.00%	4465	100.00%
CIT	*cit̰*	0	0.00%	40	48.19%	228	68.06%
	cīt̰	39	100.00%	43	51.81%	107	31.94%
	total	39	100.00%	83	100.00%	335	100.00%

TABLE 10: Allomorphs of *ca* ~ *cā* and *cit̰* ~ *cīt̰* in the historical dialects of Avestan.

Both the old versus non-old dialects of Avestan show a significant difference, as does the contingency χ^2 analysis (with Yates correction) of late-young versus non-late-young Avestan, as shown in Tab. 11, where the significance is set at $p < .05$.

Likewise, the overtaking *cit*-marker itself also ends up dying from the Iranian branch by the time of OP; only two instances of *kašciy* found in DB (Col-

	OAv ~ non-OAv	LYAv ~ non-LYAv
ca ~ *cā*	1401.127 $\binom{*}{*}$	92.9695 $\binom{*}{*}$
cit̰ ~ *cīt̰*	54.668 $\binom{*}{*}$	44.4352 $\binom{*}{*}$

TABLE 11: An analysis of statistical significance of phonological differences in diachronically continuous dialects of Avestan; values for $\chi^2(1, N = 7295)$, where $\binom{*}{*}$ stands for $p < 0.00001$.

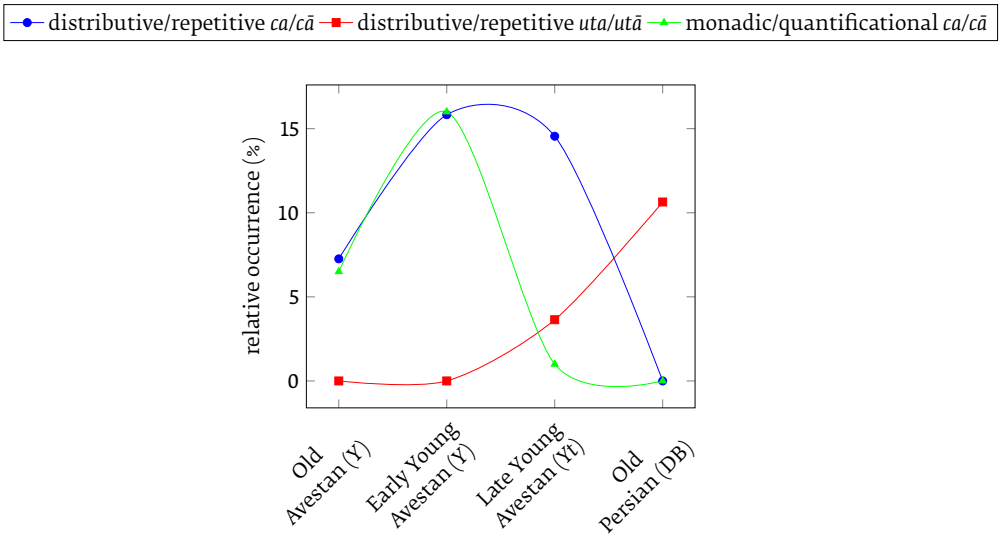


FIGURE 17: Three types of semantic changes in the history of Ir. (normalised to respective corpora size).

umn I: lines 49, 53).[12]

The semantic change can also be observed outside the monadic conjunction meanings, namely in the distributive and long-form conjunctions. By the time of OP, *ca*-marking nearly completely disappears and *utā* takes over as connector of also smaller units, such as DPs. The repetitive (or long) form of *utā* with distributive meaning increases accordingly as the sole marker of distributive conjunction (which started out at null in OAv).

6 DISCUSSION & CONCLUSION

SUMMARY

This paper tried meeting several desiderata: empirically to introduce novel data and consider them under and within a more general theory of logical marking, namely, the JP superstructure coding for coordination which derives from a logical core encoded on the μ head, itself the predominant realisation locus of 'superparticles', in the sense of Mitrović (2021). Using a refined structure for the syntax and semantics of conjunction structures, I

12 But see also Villalobos 2019 for a historical-philological and comparative discussion of facts.

also tried providing a continuous diachronic and comparative theory of the IIr languages and sketching their relation as a dialectal one. This afforded me a synchronic, comparative, and ultimately diachronic analysis of the IIr historical dialects. The diachronic syntactic evidence suggested a delayed or *k*-parametric development and ultimately loss of the double system of coordination in IIr, involving two types of conjunction markers (the 1P *uta* and the 2P *ca* particle), culminating in Fig. 11. In terms of compositional semantic variation change, I presented novel evidence and means for a comparative and diachronic semantic analysis of μ-conjunctive and -logical expressions under the working assumptions that various historical languages of IIr could and should be viewed in terms of dialectal continua.

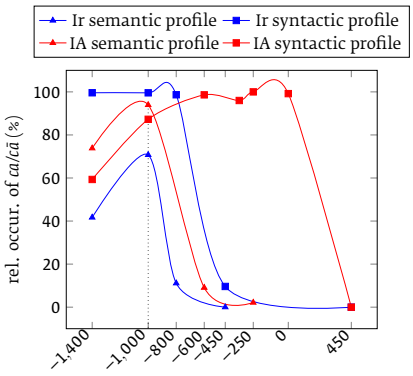
COMPOSITIONAL CHANGE & SEMANTIC RECYCLING

The rise of CIT-marked quantificational terms, and the inverse relative decline of CA-marking, can be explained by the morpho-semantic specification. While ca-terms can be considered ambiguous between its various semantic profiles, CIT-marking is not. In time, the lexical entry of the original *ca*, or indeed the $^{\star}k^{w}e$ and $^{\star}k^{w}e$-like particles across archaic IE, is lost and reshaped from the multifunctional one to a more specific – or featurally more specified and restricted – one, resulting in its subsequent inability to function as a marker of monadic conjunction, a term I used here for quantifier-building. This is confirmed by the novel facts, statistically summarised and diachronically reported in Figs. 15 and 16. What is more, the single instance of a CA-based quantificational term in the K dialect of AP Māghadī may be seen as a serendipitous remnant of the arguably archaic superparticle system.

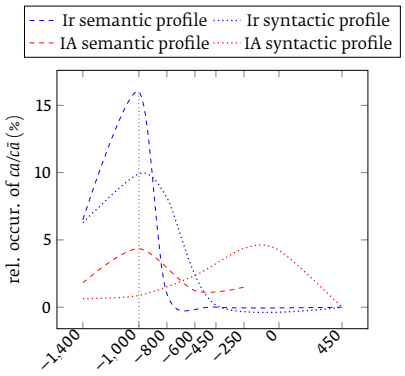
Therefore both of the IIr branches independently yet in a synchronised manner strengthened the μ-grammatical particle CA before losing it to its competitor CIT, which is sketched in Fig. 18 for the monadic use of the particle CA historically across the two branches.

The diachronic 'tipping point' – marked with a thin dotted vertical line in Fig. 18 – in both branches is concomitant at around 1,000 BCE, that is the presumed time of Late Vedic (LV) in the Indic and the time of YAv in the Iranian branch. Both the proportional and the normalised distribution show a relative decline of the monadic employment of the CA-marker in subsequent dialects, signifying a period of semantic compositional change. While the *k*-parametric intercept of the semantic change is different for the two historical-dialectal continua, the nature of change is presumably the same, assuming the original inherited lexical entry for the CA-particle is that of a μ-superparticle. In IA the CE period involve later later MIA in which 2P *ca* is lost and, with it naturally, any superparticle semantics associated with it.

Another relevant observation one can make in regard to the culminating diachronic facts presented in Fig. 18 concerns the synchronisation of change



(A) Proportional distribution.



(B) Normalised distribution.

FIGURE 18: Plotting the relative occurrence of the quantificational use of the 2P conjunction marker *ca/cā* across the two branches of IIr: blue for the Ir branch, red for the IA branch; triangle (▲) for the semantic profile (or dashed for normalised values (B)) of the conjunction/logical marker, square (■) for the syntactic profile (or dotted for normalised values (B)) of the conjunction marker.

in the semantic and the morphosyntactic profiles of the 2P *ca* marker. While Iranian shows a concomitant change in both the structural and the interpretational signature of CA, namely its Wackernagel effects in syntax and monadic meaning-building in semantics, the IA branch shows a syntactic inheritance of archaic CA and a semantic loss of its original superparticle meaning.

My analysis of the general[13] diachronic analysis of a directed semantic-compositional change proceeds from a functionally stronger to a functionally weaker, where weakness is understood in terms of the powerset of meanings the relevant μ-particle can generate at the given stage it is detectable.

I thus far managed to avoid the conceptual and technical details that underlie the superparticle semantics. Very briefly, I sketch here the semantic backbones of the analysis I am submitting: the superparticle μ is allosemic depending on the moprho-syntactic and -semantic context in which it features. Both the local and global logical properties of the context matter: if the μ-superparticle combines with an indefinite stem ($\exists_{[-\text{DEF}]}$), depending on the global context, the resulting meaning can be that of a Polarity-Sensitive (PS) indefinite (when restricted to negative ($\neg$) or antitonic or downward-entailing (DE) global contexts) a Free-choice (FC) indefinite (when confounded to modalised global contexts licensing fluctuation), or a universal quantifier term (when no global restrictions are at play). When the μ-host is not an indefinite, but a definite DP, the result can be that of an additive expression given that the relevant exhaustification operator (EXH, which can be considered as a covert version of 'only' in English) is parametrically al-

13 For a detailed makeup of parametric changes associated with the μ particle, based on the evidence from and applying to the diachronic semantics of Indo-European and Japonic, see Mitrović (2021, Chap. 4, Sec. 3).

lowed to apply iteratively (or twice). If this parameter is absent, additive, Free-choice and universal terms are predicted to be absent from the grammar. A very cursory sketch of the semantics of μ is given in (37), where I use the symbol $\neg$ for negation, symbol $\prec$ for structural embedding that conditions the contextual allosemy (reducible to or paralleled by syntactic command: so $x \prec y$ would read as 'y commands/may proble x'), the symbol $\diamond$ for existential modal (which licenses FC inferences), notation $\textsc{exh}^2$ for recursive (or twice applying) exhaustification, and symbols $\exists_{[-\textsc{def}]}$ and $x_{[+\textsc{def}]}$ for indefinite (existentials), like *wh*-terms, and definite DPs which can host the μ particle.

(37) $$[\![\ \mu\ ]\!](p_{\ni\mu}) \Rightarrow \textsc{exh}(p) = \begin{cases} \textsc{p} & \text{PS indef.} & \text{if } \mu \prec \neg \\ \textsc{f} & \text{FC indef.} & \text{if } \textsc{exh}^2 \text{ and } \mu \prec \diamond \\ \forall & \text{univer. quant.} & \text{if } \textsc{exh}^2 \text{ and } \mu(\exists_{[-\textsc{def}]}) \\ + & \text{additive} & \text{if } \textsc{exh}^2 \text{ and } \mu(x_{[+\textsc{def}]}) \\ \wedge & \text{conjunctive} & \text{if additive and } J^0 \\ & \bot & \text{otherwise} \end{cases}$$

The effect of diachronic change is therefore observable in the meanings of μ-collocations and interpretations that are subsets of the overall meanings generated by μ in (37) that cannot be parametrically captured without losing the lexical entry of or the rule for composition with the superparticle μ. As discussed in Mitrović (2019); Mitrović (2021), the structure of meanings generated by μ is a logical one and in absence of some non-conjunctive expressions featuring μ, no superparticle meaning can be obtained. Hence, the later Ir and IA evidence suggest, on that analysis, a demonstrable loss of the original, or sufficiently fully specified, superparticle meaning of a μ particle like CA took place by the time of the fifth century in the Iranian and by the time of the third century in the Indic branch, as per line e of (38).[14]

(38) decreasingly nested meaning subsets generated by μ_n (where n is a function of time)

		∧	+	∀	F	P
a.	$[\![\ \mu_1\ ]\!](\phi)$ = (37)	✓	✓	✓	✓	✓
b.	$[\![\ \mu_2\ ]\!](\phi) : \textsc{exh}(\phi) \subset \phi = \textsc{exh}(\phi)$	✓	✓	✓	✓	
c.	$[\![\ \mu_4^{[u\pm\textsc{def}]}\ ]\!](\phi) = \textsc{exh}^2(\phi)$	✓	✓	✓		
d.	$[\![\ \mu_3^{[u+\textsc{def}]}\ ]\!](\phi) = \textsc{exh}^2(\phi)$	✓	✓			
	⋮					
e.	$[\![\ \mu_5\ ]\!](\phi,\psi) = \phi \wedge \psi$	✓				

The loss of non-conjunctive μ-meanings in later IIr is consistent only with a lexical entry for μ which is void of its inherited alternative-based semantic profile, such as the e-example in (38), signifying and signalling a semantic change.

A parallel argument for the semantic aspect of the μ-particle system that is

14 The parameters associated with each iteration of μ in (38) are given in very cursory format – consult Mitrović (2021) for details on technical and conceptual translation.

partially inherited and retained can be made on the basis of modern South Slavonic. While Slovenian and Ser-Bo-Croatian have conjunction makers that seem similar, and in fact develop from a common ancestral **i* particle, itself boasting a μ-profiel, they have radically different morphosyntactic statuses and compositional-semantic behaviours, testable on whether or not they may form 'monadic conjunctions': only Ser-Bo-Croatian *i*, and not the Slovenian *in*, can form indefinites since Slovenian *in* is pleonastic and derives from the merger of two particles, at least diachronically, only one of which had the semantic profile of μ, as noted in Sec. 3.1 generally and (16-e) specifically:

(39) a. *i -(t)ko*
AND = μ WHO
"anyone" (Ser-Bo-Croatian)

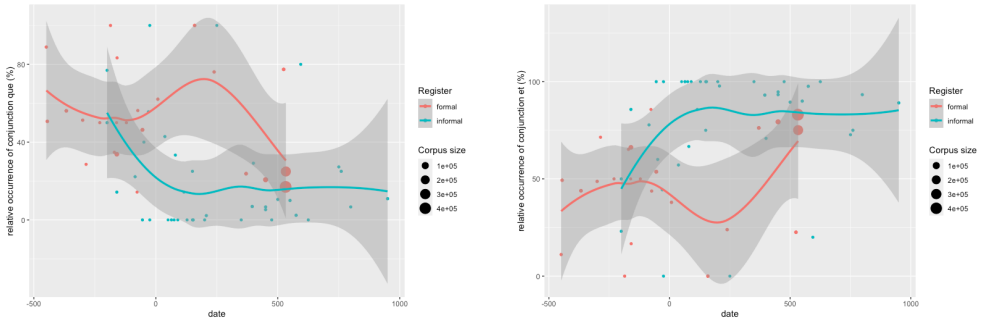
b. **in -kdo*
AND ≠ μ WHO
"anyone" (Slovenian)

REGISTER AND/AS DIALECT

One important aspect of the present dialectal analysis of historical IIr, which I have not foregrounded sufficiently, concerns register. While allowing myself to, at least terminologically if not conceptually, consider different micro- and macro-varying IIr languages as historical dialects, I have not addressed the role register plays and the the synchronic grammatical and diachronic status it has in terms of promoting retention.

Both of the most archaic languages in both branches of IIr are on a par in terms of the presumed religious and liturgical register and poetic structure. While the intermediate language phases, characterised by YAv on the one and post-Vedic ClSkt on the other hand, broadly share the poetic and religious register, the subsequent earliest dialects show the nearly completed syntactic change and the loss of the old superparticle system that the logical particle *ca* associated with in the archaic dialects. For instance, both AP in the Indic and the OP in the Iranian branch, reflect a register different to their dialectal predecessors.

The difference in register as reflecting different compositional-semantic profile of conjunction marking specifically, and logical marking of μ-meanings more generally, can also be observed in Latin. A preliminary study based on the corpus containing both formal and informal (cca. 775k tokens) texts, spanning 15 centuries confirms that the Latin formal registers show a prolonged grammatical retention of the enclitic conjunction *que*, compared to the texts with informal registers – the history and facts of which are plotted in Fig. 19 as the competition between conjunction markers *que* (which can also encode other μ-type meanings, as per §2.3) and *et*, where the scattered data is plotted as the relative occurrence of the two competing conjunction markers using local regression (the Loess smoothing algorithm).



(A) 2P marker *que*. (B) 1P marker *et*.

FIGURE 19: The distribution of the 1 and 2P conjunction markers *et* and *que* across formal and informal registers in the history of Latin.

Naturally, we are at least a decent formal theory of register short of providing an answer to such questions and a more holistic grammatical account of the drivers and vehicles of socio-historical change. Nonetheless, this paper is hopefully a step in such a direction.

LANGUAGE ABBREVIATIONS

IIr Indo-Iranian

IA Indo-Aryan

Ir Iranian

Av Avestan

OAv Old Avestan

YAv Young Avestan

EYAv Early Young Avestan

LYAv Late Young Avestan

OP Old Persian

MP Middle Persian

RV Rigvedic

ClSkt Classical Sanskrit

Skt Sanskrit

Pkt Prakrit

AP Aśokan Prakrit (Māghadī)

G Girnār (dialect of Māghadī)

K Kālsī (dialect of Māghadī)

S Shāhbāzgaṛhī (dialect of Māghadī)

MIA Middle Indo-Aryan

HISTORICAL TEXTS

Each text is prefixed with the language (abbreviation) in which it was written.

Y OAv/(E)YAv: Yasna

ṜV RV: Ṛgveda

AmH OP: Ariaramnes, Hamadan

AsH OP: Arsames, Hamadan

BP ClSkt: Bhagāvata Purāṇa

CM OP: Cyrus, Murghab (Pasargadae)

DB OP: Darius, Behishtan (5 columns). 522–486 BCE

DN OP: Darius, Naqsh-i Rustam

DS OP: Darius, Susa

DZ OP: Darius, Suez inscriptions

DE OP: Darius, Elvend

DH OP: Darius, Hamadan

GM AP: Girnār Māgadhī

XP OP: Xerxes, Persepolis

XPh OP: Xerxes' "Daiva Inscription" from Persepolis (Trilingual, on stone tablets, 2 copies). 486–465 BCE

XS OP: Xerxes, Susa

XE OP: Xerxes, Elvend

XV OP: Xerxes, Van

XH OP: Xerxes, Hamadan

A1Pa OP: Artaxerxes I, Persepolis A

A1I OP: Artaxerxes I, incerto loco

D2S OP: Darius II, Susa

A2S OP: Artaxerxes II, Susa

A2H OP: Artaxerxes II, Hamadan

A?P OP: Artaxerxes II or III, Persepolis

A3Pa OP: Artaxerxes III, Persepolis

W OP: Inscriptions on weights: Wa, Wb, Wc, Wd

Seals OP: Inscriptions on Seals

Vase OP: Vase Inscriptions

Vd YAv: Vidēvdād

AW MP: *Ardā Wīrāz*. 9-10c. CE.

GKSh AP: Girnār, Kālsī, Shāhbāzgaṛhī rock edicts (Māgadhī) (Braarvig & Nesøen, 2016)

ACKNOWLEDGEMENTS

I gratefully acknowledge financial support from the Alexander von Humboldt foundation and Mercator fellowship support (grant #ME 4125/5-1; PIs: Meyer, Willis).

REFERENCES

Bach, E. 1964. *An Introduction to Transformational Grammars*. New York: Holt, Rinehart & Winston.

Balogh, Kata & Reza Kazemian. 2021. The additive particle *ham* in Modern Persian. Paper presented at the *16th International Conference on Role and Reference Grammar*.

Biberauer, Theresa, Anders Holmberg & Ian Roberts. 2010. A Syntactic Universal and its Consequences. Ms. University of Cambridge.

Bloomfield, L. 1933. *Language*. New York: Holt, Rinehart & Winston.

Blümel, R. 1914. *Einführung in die Syntax*. Heidelberg: C. Winter.

Braarvig, Jens & Asgeir Nesøen. 2016. Bibliotheca Polyglotta: The Ashoka Library. University of Oslo, Norwegian Institute of Palaeography and Historical Philology. https://www2.hf.uio.no/polyglotta/index.php?page=library&bid=14.

Charpentier, Jarl. 1931. Antiochus, King of the Yavanas. *Bulletin of the School of Oriental and African Studies* 6(2). 303–321. DOI: 10.1017/s0041977x0009279x.

Chomsky, N. 1965. *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA: MIT Press.

Chomsky, Noam. 2001. Derivation by Phase. In M. Kenstowicz (ed.), *Ken Hale: A Life in Language*, 1–52. Cambridge, MA: MIT Press.

Chomsky, Noam. 2005. Three Factors in Language Design. *Linguistic Inquiry* 36. 1–22.

Denison, D. 2003. Log(ist)ic and Simplistic S-curves. In R. Hickey (ed.), *Motives for language change*, 54–70. Cambridge: Cambridge University Press.

Dik, S. C. 1968. *Coordination: its implications for the theory of general linguistics*. Amsterdam: North-Holland.

Dougherty, R. C. 1969. Review of *Coordination: Its Implications for the Theory of General Linguistics* by Simon C. Dik. *Language* 45. 624–636.

Dunkel, George E. 1982. IE conjunctions: pleonasm, ablaut, suppletion. *Zeitschrift für vergleichende Sprachforschung* 96(2). 178–199.

Gazdar, Gerald, Ewan Klein, Geoffrey Pullum & Ivan Sag. 1985. *Generalized Phrase Structure Grammar*. Oxford: Basil Blackwell.

Ghomeshi, Jila. 2020. The additive particle in Persian: A case of morphological homophony between syntax and pragmatics. In *Advances in Iranian Linguistics*, 57–83. Amsterdam: John Benjamins. DOI: 10.1075/cilt.351.04gho.

Goodall, G. 1987. *Parallel Structures in Syntax: Coordination, Causatives and Restructuring*. Cambridge: Cambridge University Press.

Kauhanen, Henri & George Walkden. 2017. Deriving the Constant Rate Effect. *Natural Language & Linguistic Theory* 36(2). 483–521. DOI: 10.1007/s11049-017-9380-1.

Kayne, R. 1994. *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.

Klein, Jared S. 1985a. *Toward a Discourse Grammar of the Rigveda. Part 1.*, vol. I. Heidelberg: Carl Winter Universitätsverlag.

Klein, Jared S. 1985b. *Toward a Discourse Grammar of the Rigveda. Part 2.*, vol. II. Heidelberg: Carl Winter Universitätsverlag.

Klein, Jared S. 1988. Coordinate Conjunction in Old Persian. *Journal of the American Oriental Society* 108(3). 387–417.

Kroch, Anthony S. 1989. Reflexes of grammar in patterns of language change. *Language Variation and Change* 1(3). 199–244. DOI: 10.1017/s0954394500000168.

Ledgeway, Adam. 2015. *From Latin to Romance: Morphosyntactic Typology and Change*. Oxford: Oxford University Press.

MacKenzie, David Neil. 1971. *A Concise Pahlavi Dictionary*. London: Oxford University Press.

Marantz, Alec. 2013. Locality Domains for Contextual Allosemy. In Matushansky & Alec Marantz (eds.), *Distributed Morphology Today: Morphemes for Morris Halle*, 95–115. Cambridge, MA: MIT Press. Paper presented at the Columbia Lingusitic Society.

Meillet, A. 1908. *Introduction à l'étude comparative des langues indoeuropéennes*. Paris: Hachette.

Mitrović, Moreno. 2013. Configurational change in Indo-European coordinate construction. In A. M. Martins & A. Cardoso (eds.), *Word order change*, Oxford: Oxford University Press. In press. DOI: jcks.

Mitrović, Moreno. 2014. *Morphosyntactic atoms of propositional logic: a philo-logical programme*: University of Cambridge dissertation.

Mitrović, Moreno. 2018. Configurational changes in Indo-European coordinate constructions. In Ana Maria Martins & Adriana Cardoso (eds.), *Word order change* (Oxford Studies in Diachronic & Historical Linguistics 29), chap. 2, 19–44. Oxford: Oxford University Press. DOI: jcks.

Mitrović, Moreno. 2019. Quantificational cycles and shifts. In Anne Breitbarth, Elisabeth Witzenhausen, Miriam Bouzouita & Lieven Danckaert (eds.), *Cycles in Language change* (Oxford Studies in Diachronic & Historical Linguistics 37), chap. 9, 155–176. Oxford: Oxford University Press. Forthcoming. DOI: jgcc.

Mitrović, Moreno. 2021. *Superparticles: A Microsemantic Theory, Typology, and History of Logical Atoms* (Studies in Natural Language & Linguistic Theory 98). Dordrecht: Springer. DOI: g6sn.

Mitrović, Moreno & Uli Sauerland. 2016. Two conjunctions are better than one. *Acta Linguistica Hungarica* 63(4). 471–494. DOI: g6sj.

Muadz, H. 1991. *Coordinate structure: a planar representation*: University of Arizona dissertation.

Munn, A. 1993. *Topics in the Syntax and Semantics of Coordinate Structures*: University of Maryland dissertation.

Nilsen, O. 2003. *Eliminating Positions*: University of Utrecht dissertation.

Oberlies, Thomas. 2003. Aśokan Prakrit and Pali. In George Cardona & Dhanesh Jain (eds.), *The Indo-Aryan Languages* Routledge Language Family Series, 179–224. London: Taylor & Francis.

Progovac, Liljana. 1998a. Structure for Coordination: Part I. *Glot International* 3(7). 3–6.

Progovac, Liljana. 1998b. Structure for Coordination: Part II. *Glot International* 3(8). 3–9.

Richards, Marc. 2007. On phases, phase heads, and functional categories. *Nanzan Linguistics* 1(1). 105–127.

Roberts, I.G. 2010. *Agreement and Head Movement: Clitics, Incorporation, and Defective Goals* Linguistic Inquiry Monographs. Cambridge, MA: MIT Press.

Skjærvø, Prods Oktor. 2003. An Introduction to Young Avestan. Unpublished monograph. Harvard University.

Skjærvø, Prods Oktor. 2006. Iran vi.: Iranian Languages and Scripts. *Encyclopaedia Iranica* 13(3). 344–377. http://www.iranicaonline.org/articles/iran-vi-iranian-languages-and-scripts.

Slade, Benjamin Martin. 2011. *Formal and philological inquiries into the nature of interrogatives, indefinites, disjunction, and focus in Sinhala and other languages*: University of Illinois at Urbana-Champaign dissertation.

Toosarvandani, Maziar & Hayedeh Nasser. 2017. Quantification in Persian. In *Handbook of Quantifiers in Natural Language: Volume II*, 665–696. Springer International Publishing. DOI: 10.1007/978-3-319-44330-0_13.

Velde, John R. te. 2005. *Deriving Coordinate Symmetries: A phase-based approach integrating Select, Merge, Copy and Match*. Amsterdam: Benjamins.

Villalobos, Juan E. Briceño. 2019. *Negation, indefinites, and polarity in early Greek and IndoIranian: a typological and comparative approach*: Universidad Complutense de Madrid dissertation.

Villalobos, Juan E. Briceño. 2022. Achaemenid Elamite and Old Persian Indefinites: A Comparative View. In Michele Bianconi, Marta Capano, Domenica Romagno & Francesco Rovai (eds.), *Ancient Indo-European Languages between Linguistics and Philology: Contact, Variation, and Reconstruction*, vol. 18 Brill's Studies in Historical Linguistics, chap. 2, 48–87. Leiden: Brill. DOI: 10.1163/9789004508828_004.

Weil, Henri. 1844. *The Order of Words in the Ancient Languages Compared with That ofthe Modern Languages*. Boston: Ginn & Co.

Windfuhr, Gernot (ed.). 2013. *The Iranian Languages*. Routledge. DOI: 10.4324/9780203641736.

Witzel, Michael. 1997. The Development of the Vedic Canon and its Schools: The Social and Political Milieu. In Michael Witzel (ed.), *Inside the Texts, Beyond the Texts: New Approaches to the Study of the Vedas*, vol. 2 Harvard Oriental Series, Opera Minora, 257–348. Cambridge: Harvard University Press.

Zhang, Niina Ning. 2010. *Coordination in Syntax* Cambridge Studies in Linguistics. Cambridge: Cambridge University Press. DOI: 10.1017/CBO9780511770746.

Moreno Mitrović
Humboldt University
Unter den Linden
DE-10117 Berlin