

Clause-edge markers in Bartangi: An LrFG analysis

Oleg Belyaev 

Lomonosov Moscow State University, Institute of Linguistics RAS

Proceedings of the 32nd International Conference on
Head-Driven Phrase Structure Grammar

Centro de Linguística da Universidade de Lisboa, Lisbon, Portugal

Stefan Müller, Shûichi Yatabe (Editors)

2025

Berlin: Berlin Universities Publishing

pages 197–220



Keywords: LFG, LrFG, Distributed Morphology, subordination, particles, Iranian languages, Pamir languages, Bartangi

Belyaev, Oleg. 2025. Clause-edge markers in Bartangi: An LrFG analysis. In Stefan Müller & Shûichi Yatabe (eds.), *Proceedings of the 32nd International Conference on Head-Driven Phrase Structure Grammar, Centro de Linguística da Universidade de Lisboa, Lisbon, Portugal*, 197–220. Berlin: Berlin Universities Publishing. DOI: 10.83067/hpsg.2025.10.



4.0

Abstract

In this paper, I provide an analysis of the major clause-peripheral functional elements (particles and conjunctions) in Bartangi (Iranian > Indo-European) from the perspective of Lexical Realizational Functional Grammar (LrFG), a non-transformational framework that blends elements of Lexical Functional Grammar (LFG) and Distributed Morphology (DM). The analysis reveals that at least three distinct phrase structure positions for such elements – initial, second and final – are required. The coexistence of such opposite types of directionality is probably due to the interaction of two distinct contact influences: the right-branching (clause-initial and clause-second) type, as found in Persian/Tajik and most other modern Iranian languages, and the left-branching (clause-final) type, as found in languages of the adjacent South Asian linguistic area. This view on the emergence of functional layers of the clause is consistent with the LFG approach to constituent structure as a surface-based, language-specific phenomenon.

1 Introduction

While the issue of phrase structure and configurationality has received considerable attention in the literature on Lexical Functional Grammar (see Snijders 2015), there has been comparatively little research on the possible extent of dedicated functional positions in the clause. A version of X' Theory is usually assumed, with the lexical projection VP extending to the functional projections IP and CP (Bresnan et al. 2016: 102–103), which are not viewed as universal. A language, then, may lack CP or IP (which is common in so-called nonconfigurational languages, see Austin & Bresnan 1996), but these two projections are normally enough to capture all word order and constituency facts within the clause.

However, what has generally been overlooked in the LFG literature on X' theory is that languages may possess a large number of clause-level markers (conjunctions, particles) with fixed but distinct linear slots, which may require more than two functional clausal layers if these slots are identified on structural rather than prosodic grounds. In this paper, I consider the data of Bartangi, an Iranian (> Indo-European) Pamir language with a very large number of such grammaticalized functional elements, from the point of view of Lexical Realizational Functional Grammar (LrFG) – a version of LFG that abandons a strict version of lexicalism in favour of a realiza-

[†]This research has been supported by the Russian Science Foundation, project no. 24-18-00199 “Clause structure and positional phenomena in SOV languages”. I am grateful to the audiences of the Workshop on Minor Sentence Types at the 32nd International Conference on Head-Driven Phrase Structure Grammar (Lisbon, 2–3 September 2025), the 30th International Lexical Functional Grammar Conference (Barcelona, 22–24 July 2025) and the International conference on Clause Linkage in Iranian Languages and Beyond (Uppsala, 21–23 August 2025), where parts of this research were presented, for their helpful comments and suggestions. I am especially grateful to Ash Asudeh, Nadeem Siddiqi, Adam Przepiórkowski, Jaroslava Obrtelová, Vladimír Plungian and Jakob Maché, as well as to all native speakers of Bartangi. All errors are mine.

tional model that associates terminal c-structure nodes with their morphonological representations.¹

Crucially for this analysis, in LrFG the same phrase structure position can be associated with different linearizations (e.g. an affix vs. a clitic) depending on the vocabulary item (VI) that realizes this position. In other versions of LFG this is either impossible or possible to a very limited degree, cf. the analysis of Pashto endoclefts as prosodic inversion at the phonological level in Bögel (2015: 210ff.). This flexibility of LrFG allows the same c-structure to be mapped to distinct realizations, but the analysis still requires at least three different phrase structure positions for functional elements in Bartangi: initial, second and final. This means that L(r)FG must allow for a larger inventory of functional projections than traditionally assumed, and that their exact makeup and directionality are language-specific outcomes of particular diachronic processes. This generally agrees with the functional-typological approach to clause structure, and LrFG can be said to provide a more formalized version thereof.

The paper is structured as follows. In Section 2, I describe general facts on Bartangi and its clause structure, drawing on the LrFG analysis of Belyaev (2025a). In Section 3, I focus on descriptive generalizations on the behaviour of conjunctions and clause-level particles in Bartangi, which are then analyzed in Section 4. In Section 5, which concludes the paper, I offer some discussion of what the analysis means for the treatment of conjunctions, particles and other clause-peripheral elements in L(r)FG.

2 Bartangi

2.1 General information and data sources

Bartangi is a Shughni-Roshani (Iranian > Indo-European) language spoken in the lower part of the Bartang river valley in the Mountainous Badakhshan Autonomous Region (henceforth MBO), Tajikistan. Shughni-Roshani is part of the Pamir linguistic area (Edelman 1980), whose languages share a number of structural and lexical similarities due to long-term language contact but do not form their own phylogenetic group within Iranian (Wendtland 2009).

Available descriptions of Bartangi are very limited, even compared to its larger relative Shughni. The first data on Bartangi come from Zarubin (1937), reflecting the author's fieldwork in 1915–1916, which includes folk poetry and a glossary; these texts are of limited use for the purposes of this paper, as they include few instances of clause linkage markers. The only grammatical description of Bartangi is Karamxudoev (1973), which, however, has almost nothing to say on syntax; it also includes several transcribed and translated texts. Sokolova (1960) is a collection of texts and vocabulary, which nevertheless contains much valuable grammatical information in

¹Due to space constraints, I cannot provide a detailed introduction to LrFG in this paper. A reasonably detailed exposition can be found in Asudeh & Siddiqi (2023). The only part of the analysis that strongly depends on the specifics of LrFG is the treatment of the preverbal subordinators *ca* and *di_{IN}* and the interrogative particle *=o*; the relevant features of the framework will be explained in the relevant sections below.

the introduction and within lexical entries – in particular, the entries for coordinating and subordinating conjunctions. A few Bartangi texts have also been published in the collection of Pamir folklore edited by Šakarmamadov (2005).

In 2023–2025, new materials on Bartangi have been collected by the author and his colleagues (Artyom Badeev, Alexander Sergienko) during field trips to Basid (Rushan district, Mountainous Badakhshan AO, Tajikistan). These materials include both elicited texts and spoken data recorded from native speakers. The volume of newer transcribed and annotated texts is still much smaller than that of the earlier materials; therefore, I will be mainly relying on the texts published by Sokolova, Karamkhudoev and Shakarmamadov and my own elicitation. As far as clause combining is concerned, the grammar of Bartangi does not seem to have changed much between the earlier texts and the present.

2.2 Fundamentals of clause structure

Bartangi clause structure has been described and analyzed from the perspective of LrFG in Belyaev (2025a). The general organization of the simple clause is illustrated in Figure 1 for a sentence meaning ‘I drank water yesterday’. According to this representation, the clause includes a configurational VP core that includes the direct object, a TP layer where the subject is attached as a specifier, and a dedicated structural topic position immediately in front of the clitic agreement marker. Following standard LFG principles (see Bresnan et al. 2016: 90), the subject and object c-structure positions are optional, such that any argument may be displaced and appear in the topic slot before Agr. The pre-Agr topic slot, however, is not optional, which is especially evident when the agreement marker is an enclitic, such as in Figure 1.

In this example, the core arguments are found in their configurational positions and the topic position is occupied by an adverb; but more commonly, in the absence of a topical adverb, the initial position will be occupied by the subject, and the Spec, TP position will be empty, as in (1).

- (1) āz =um ʃac biroz-d
 I.NOM=1SG water drink-PST
 ‘I drank water’.

As mentioned above, Bartangi has mobile agreement markers. This means that the agreement marker is a second-position enclitic in the past and perfect tenses, but an affix in the present. The affix and the clitic are thus found in different linear positions, but they are in complementary distribution. In the LrFG analysis of Belyaev (2025a), this is modeled by VIs for both types of markers realizing the same features in the Agr node, but the affixes constrained as appearing in the present tense, and the clitics, in the past and perfect. The affix will “hop” onto the verb as per the head identification mechanism described in Asudeh et al. (2023).

In spite of certain formal complications stemming from the underdevelopment of some aspects of LrFG (discussed in the original paper), this analysis adequately models key facts with respect to Bartangi basic word order. However, the position of other

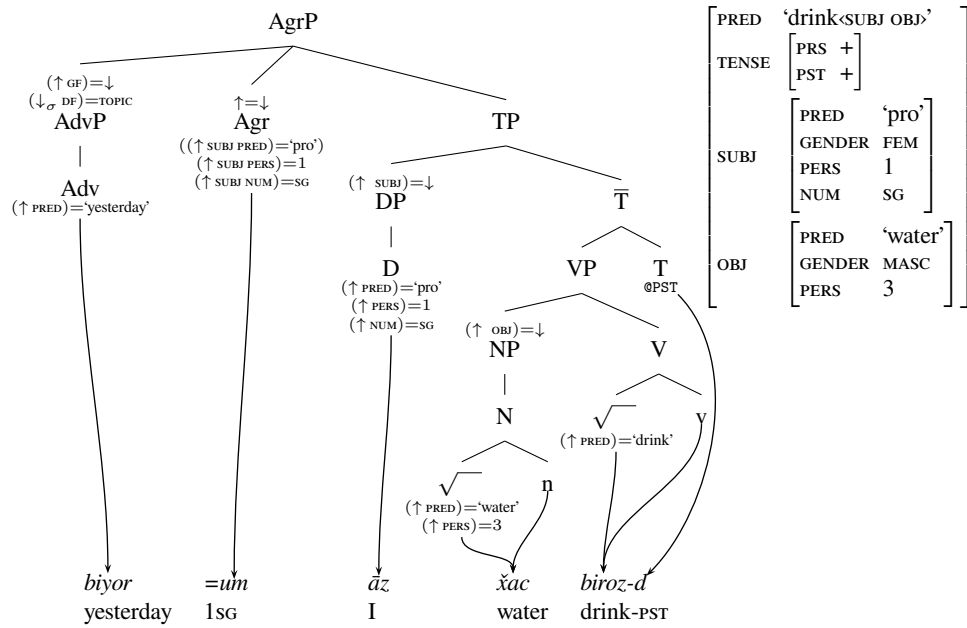


Figure 1: Clause structure with clitic person-number marker (Belyaev 2025a)

elements, traditionally viewed as subordinators or conjunctions, has remained outside the scope of the paper. Below, I will demonstrate that, contrary to what can be expected from an Iranian language, Bartangi has developed a full-fledged sentence-final area that requires postulating a separate clause-final position for particles or subordinators in addition to the clause-initial one.

3 The clause-peripheral area in Bartangi

3.1 “Initial” markers

According to traditional descriptions, Pamir languages, including Bartangi, have two types of subordination markers: clause-initial and clause-internal (preverbal). The most common clause-initial marker, under the traditional analysis, is the polyfunctional *di*, which in its range of meanings is akin to Tajik *ki* (= Persian *ke*), as will be discussed in Section 3.2 below. Other initial subordinators are mainly Tajik loanwords, of which two stand out in terms of frequency: *agar* ‘if’ and *to* ‘until, etc.’; a native word with the same linear distribution is *lāk* ‘let’, although it is perhaps better treated as a jussive rather than a subordinator. These latter markers seem to alternate between clause-initial, second, and preverbal positions (which is somewhat akin to Ossetic “floating” subordinators, see Belyaev 2014, Belyaev & Haug 2014); occasionally, even postverbal positions are encountered (cf. Windfuhr & Perry 2009: 519 on final *agar* in spoken Tajik):

- (2) a. initial
agar lap-dör yoc=af *ca* čũĵ, mun =um tamāš-ri mis
 if much-CMPR fire 3PL.TR SUBD do.PRF I.OBL=1SG you-DAT also
 vorĵ vũĵ
 horse bring.PRF
 ‘If you had made a bigger fire, I would have brought you a horse too.’
 (Sokolova 1960: 8)
- b. second
 yā *agar* ĵāl-dör *ca* tezd, pa rošorv nūr=aθ
 D3.SG.NOM if soon-CMPR SUBD go.PRS:3SG LOC.UP R. today=EXH
 firāp-t
 reach.PRS-3SG
 ‘If he goes out earlier, he will reach Roshorv today.’
 (Karamxudoev 1973: 273)
- c. preverbal
 az sabo tar-um *agar* *ca* vārδī-m, bašānd-dör
 from tomorrow LOC.EQ-DIST.LOC if SUBD be.able.PRS-1SG good-CMPR
 kin-um
 do.PRS-1SG
 ‘Starting from tomorrow, if I am able to, I will work better.’
 (Karamxudoev 1973: 208)
- d. postverbal
 ar mun qāp δīd *agar*, vār-um
 LOC.DOWN I.OBL handful fall.PRS:3SG if bring.textscprs-1sg
 tā-ri
 thou.OBL-DAT
 ‘If (it) falls into my hands, I will bring (it) to you.’
 (Karamxudoev 1973: 245)

Significantly, in all of these examples, *agar* occurs together with the native subordinator *ca*. Even though this is not strictly necessary (and not as frequent for *to* ‘until’), it underscores the peripheral status of such elements in the system. Nevertheless, when *agar* is clause-initial, it cannot host clitics (2a), which means that it has to be viewed as occupying a dedicated structural position. In other positions, *agar* and other elements of this class can be treated as a special kind of adverbial.

3.2 “Initial” *di*_{EX}

As stated above, traditional grammars treat *di* (var. *idi*) as the native clause-initial subordinator in Bartangi (*di*^l in Sokolova 1960: 101–102). In order to differentiate it from another, formally and historically identical but synchronically distinct clause-

internal subordinator *di* (*di*² in Sokolova 1960: 102),² I will label the clause-external marker as *di*_{EX} and the clause-internal marker as *di*_{IN}, following Belyaev (2025c,d).

As suggested in Belyaev (2025c,d), the analysis of *di*_{EX} as an initial conjunction is problematic from several points of view. Like Persian *ke*, prosodically it encliticizes to the linearly preceding constituent (which belongs to the “main” clause). It is said to encode a wide range of functions (3)–(6), but in all of these, it is optional, and the meaning is sometimes doubled by a separate marker, such as the jussive *lāk* in (4). There is no evidence to suggest that the second clause in these constructions is embedded in any sense: the two clauses cannot be swapped, the “subordinate” clause cannot be put inside the “main” one, etc. At best, these constructions are in a very loose dependency relationship.

(3) complement

yā fikri čūg *di*, asal pa diraxt
D3.SG.NOM thought do.PST LNK honey LOC.SUPER tree
‘He thought that honey was on the tree.’

(4) purpose

āz=um az xu bajoy čūg *di* (lāk) az mun mā wīn-an
I.NOM=1SG OBJ self hidden do.PST LNK JUS OBJ I.OBL PROH see.PRS-3PL
‘I hid so that they wouldn’t see me.’

(5) result

xil-ak=an pa wī nāst *di*, yā mis piḏo
many-DIM=1PL LOC.UP D3.M.OBL sit.PST.PL LNK D3.SG.NOM ADD visible
sut
go.PST.M
‘We waited for a while, and he appeared.’ (Karamxudoev 1973: 276)

(6) manner

ik-dond azob az tā kin-um *idi*, tū mir-i
EMPH-SO.much pain OBJ thou.OBL do.PRS-1SG LNK thou.NOM die.PRS-2SG
‘I will beat you up so much that you will die.’ (Sipanj, Sokolova 1960: 102)

There is a frequent class of examples where *di*_{EX} is even less amenable to interpretation as a subordinating conjunction. In such cases, illustrated by (7)–(8), *di*_{EX} introduces an unexpected turn of events in the discourse, with both clauses constituting new information and no backgrounding characteristic of subordinating constructions.

²In the closely related Shughni, the clause-external marker is *idi* while the clause-internal marker is *di* (Parker 2023: 428, 432, 435).

- (7) daδ az ĵingāl tūyĵ xu yoδ-d yi ĵoy *idi*,
 then ABL forest go_away.PRF.M and come.PRS-3SG one place LNK
 wī-ri paryend dučor saw-d
 D3.M.OBL-DAT ovring encountered go.PRS-3SG
 ‘He moved away from the forest, comes to a place, and he encounters an ovring
 (narrow mountain path).’ (Sokolova 1960: 59)

- (8) yā yač *idi*, wī šögbuc nist
 D3.SG.NOM come.M.PRF LNK D3.M.OBL lamb NEG.EXST
 ‘He came, and his lamb is not there.’ {brt_txt_tribrata}

In yet other examples, it is not even clear to what extent *di* can be assigned any particular interpretation:

- (9) a. δaw kud yi ĵondor inĵuvd *di*... (500ms)
 two dog one animal catch.PST LNK
 ‘Two dogs caught one animal (and)...’
 b. wī-tör=af vo wī nusk-tör=af wilčak
 D3.M.OBL-SUPER=3PL.TR again D3.M.OBL muzzle-SUPER=3PL.TR measure
 inĵuvd
 put.PST
 ‘...they put a measure on its muzzle.’ {jonali1_200623}
- (10) a. sarpiro=an di tar ĵīw rawon ca sad *di*... (1 s 130ms)
 before=3PL SUBD LOC.EQ hunt going SUBD go.PST.PL LNK
 b. e... yi daxo vud
 HES one prayer be.PST
 ‘When in the old times we went hunting... there was one prayer [before
 the hunt].’ {jonali1_200623}

Especially in the latter example, the first clause already has the subordinator *ca*; therefore, *di*_{EX} cannot be a clause-initial complementizer here, because technically it introduces the MAIN clause.

Quite often *di*_{EX} seems to behave merely as a hesitation marker, or perhaps a discourse deictic particle, without any obvious clause linkage functions (11). For this function, cf. the Ossetic hesitation marker *jed* (Digor *idi*), which possibly comes from the same demonstrative stem **aita* (Abaev 1958: 410, Rastorgueva & Edelman 2000: 132).

- (11) a yim *idi*... tupxona
 and D1.SG.NOM LNK fortress
 ‘And this is... a fortress.’ {fortress_180623}

Finally, *di*_{EX} can be used in exclamative clauses without any clause linking context whatsoever:

Table 1: Frequency of functions of di_{EX} in the texts

<i>Function</i>	<i>#</i>
Quotative	88 (49.4%)
Other complement	21 (11,8%)
“Consecutive”	21 (11,8%)
“Hesitation”	13 (7.3%)
Result	5 (2.8%)
Purpose	3 (1.7%)
Relative (‘such’), manner / degree (‘so... that’)	17 (9.6%)
Temporal (<i>waxto di...</i> etc.)	8 (4.5%)
Unclear interpretation	2 (1.1%)
<i>Total</i>	<i>178</i>

- (12) *yi lap naxčir umder di(dī)!*
 one many goat D3.LOC:IN LNK
 ‘There are so many mountain goats there!’ (Karamxudoev 1973: 277)

All of this appears to suggest that di_{EX} is not really a clause-initial subordinator in a large share of its functions. This claim appears controversial given that di_{EX} is generally assumed to be a universal subordinator akin to Persian *ke*, Tajik *ki* and similar markers in other Iranian languages. Therefore, in order to assess the degree to which the meanings listed above occur in natural discourse, I have annotated 178 examples of di_{EX} (and other clause linkage markers) in natural texts collected in 2023–2024 and published in Sokolova (1960). The results are displayed in Table 1. A very large share of these are sentences where di_{EX} is used to introduce direct speech. If these are excluded, consecutive and hesitation examples make up 38% of all observations, which demonstrates that such uses are by no means marginal.

Moreover, if only texts collected in 2023–2024 are considered, of the total 77 examples, 37 are quotative (mostly in fossilized form with the verb: *luvd-i* ‘s/he said’), 11 are consecutive and 13 are hesitation-like. This means that such “discourse-structuring” uses make up 60% of all uses of di_{EX} outside of direct speech marking. Perhaps the higher percentage of such uses can be explained by hesitations and other unclear uses having been removed from the texts in Sokolova (1960) during editing.

Such uses of ostensibly “subordinating” markers are widespread in neighbouring languages as well. For example, Wakhi has very similar uses of “clause-initial” *ki*, which Obrtelová (2019) describes as introducing post-nuclear temporal clauses. In such constructions, the subordinate clause is “highlighted (foregrounded) while the main clause is backgrounded” (Obrtelová 2019: 265). This also applies to Bartangi from a pragmatic point of view. However, it is unclear whether such examples can be treated as temporal clauses, as there is no evidence of syntactic embedding, and functionally, they form part of a consecutive narrative. Thus Bashir (2025: 255–256), in her analysis of Khowar (Dardic), which belongs to the same broader Hindu Kush-

Karakorum linguistic area as Bartangi and Wakhi (Liljegren 2020), treats sentences with similar uses of *ki* as instances of left-branching with final subordinators:

- (13) Khowar (Dardic, Bashir 2025: 256)
 [te-ɣár múɬi bakas-ó bil-ó ayh a-ré-r ki] tshap
 REM.SG-ABL below box-OBL lid-OBL up AUG-DO.PST.D-3SG when full.of
 ašrafí
 gold.coins
 ‘[When he lifted the lid of the box below that one], (he saw that) it was full of gold coins.’

While a fully left-branching analysis along these lines seems too strong for Bartangi, viewing *di*_{EX} as a kind of clause final marker at least in some contexts is more plausible than viewing it as a clause-initial subordinator, and I will adopt it here.

3.3 “Clause-internal” *ca* and *di*_{IN}

The second class of subordinators distinguished in traditional grammars includes two markers that generally appear within the dependent clause: the general subordinator *ca*, which is immediately preverbal, separable only by negation (14)–(15), and the temporal-conditional-causal *di*_{IN}, which fluctuates between second and preverbal positions (16).

- (14) wāð=an *ca* yat, āz=um az uf qīwd
 D3.PL.NOM=3PL SUBD come.PST I.NOM=1SG OBJ D3.PL.OBL call.PST
 ‘When they came, I called them.’
- (15) ibod dars ðod, tonac pulod *ca* na yač
 I. lesson give.PST while P. SUBD NEG come.PRF
 ‘Ibod was doing lessons until Pulod came.’
- (16) (**di*) tū=t (^{OK}*di*) māš-ri (**di*) bazaygarak (^{OK}*di*)
 SUBD thou.NOM=2SG SUBD we-DAT SUBD toy SUBD
 dāčūg, māš=ti xoš sān
 give.PST we=FUT happy go.PRS:1SG
 ‘When you give us toys, we are happy.’ (Belyaev 2025c: 35)

While *ca* or *di*_{IN} is preverbal, the verb to which it attaches can be freely positioned in the clause:

- (17) ibod dars ðod, tonac *ca* na yač pulod
 I. lesson give.PST while SUBD NEG come.PRF P.
 ‘Ibod was doing lessons until Pulod came.’

However, when *ca* is used in a clause that has no overt verb, it has to be final:

- (18) a. pul tā-yā ca, dī-qatī čöd viröxt vārδī
 money thou.OBL-POSS SUBD D2.M.OBL-INS house build.INF can.PRS[2SG]
 ‘With the money you have, you can build a house.’
 b. *pul ca tā-yā, dī-qatī čöd viröxt vārδī

This could be due to possessive constructions, or copular clauses in general, disallowing postverbal or post-predicate constituents. But possessive constructions can have an overt copula, and in this case, there can be postverbal constituents just like in clauses with verbal predicates:

- (19) yi nāw-δūs kitob yal yast mun-ā
 one nine-ten book still COP I.OBL-POSS
 ‘I still have around nine or ten books.’ (Karamxudoev 1973: 335)

Hence, this “freezing” of word order in the absence of an overt verb seems to be due to some inherent properties of *ca* rather than the copular construction itself. The most straightforward way to account for this behaviour is to assume that *ca* is fundamentally a clause-final subordinator, and its preverbal position is due to additional constraints on its position. An LrFG implementation of this positional variability will be proposed in the next section.

The preverbal *di*_{IN} is similar to *ca* in that it is also a proclitic to the verb. However, it does not show the behaviour characteristic of *ca* in copular constructions, i.e. it cannot terminate a clause on its own. Thus there is no evidence that it is fundamentally clause-final, unlike *ca*.

As noted above, *di*_{IN} can also appear in second position. In this case, it always follows the person-number clitics: *tū=t di yat* (thou.NOM=2SG SUBD come.PST), but not **tū di=t yat*. But there is an important difference: if the clause consists of only a single verb, PN clitics appear postverbally, but this position is not available for *di*_{IN}:

- (20) a. di yat =um, qīw-um az tā
 SUBD come.PST=1SG call.PRS-1SG OBJ thou.OBL
 ‘When I will come, I will call you.’
 b. *yat=um di, qīwum az tā

Hence, unlike *ca*, there is no evidence that any of the positions of *di*_{IN} is inherently postverbal. This will be important for the analysis below.

3.4 “Coordinators” *xu* and *at*

Bartangi is described as having two coordinating conjunctions: *xu*, which connects clauses forming a temporal sequence (21), and *at*, which is a more general coordinator that also conjoins NPs and other constituents; in the function of clausal coordination, it is mostly used in simultaneity contexts, but it is not limited to any particular temporal relation (22).

- (21) bād yač pa xu mallā =xu... (440ms) aynan
 then come.PRF.M LOC.UP REFL house =SEQ same
 k-az-um xu bobokalon-ā k-az wī piltā-yī
 EMPH-OBJ-D3.F.OBL REFL grandfather-POSS EMPH-OBJ D3.M.OBL fuse-ADJ
 can qadimā zoz-d =xu sawd
 gun old take.PRS-3SG =SEQ go.PRS:3SG
 ‘Then he came to his house and... nevertheless he takes his grandfather’s old
 matchlock rifle and goes.’ {moustache_man_260624}

- (22) ...podžo-duxtar tar toq-aθ xu kāl zuwosč-at, az
 king-daughter LOC.EQ window-EXH REFL head take.out.PST-CONJ OBJ
 wī čāst
 D3.M.OBL look.PST
 ‘...the princess took out her head and looked at him.’ (Sokolova 1960: 16)

Both markers are encliticized to preceding words and at least *xu* has syntactically left-branching properties. In (23), the clause with final *xu* is positioned after the other clause, but the temporal order of the basic configuration ($A=xu\ B \rightarrow A$ precedes B) is preserved (23); however, in other contexts of this type the temporal relation is not as clear, see (24).

- (23) zaxmī sad yā čabūd, žer-ac =um az um δod
 wounded go.PST.M D3.SG.NOM pigeon stone-TERM=1SG OBJ D3.F.OBL hit.PST
 =xu
 =SEQ
 ‘I hit the pigeon with a stone, and it became wounded.’
 (Karamxudoev 1973: 258)

- (24) wazīr-puc sud az um daraxt sont =at, yā yurč
 vizier-son go.PST.M OBJ D3.F.OBL tree raise.PST =CONJ D3.SG.NOM bear
 zibud, quloγ mis wī-r na luvd =xu
 run_away.PST.M thanks ADD D3.M.OBL-DAT NEG say.PST =SEQ
 ‘The vizier’s son lifted the tree and the bear ran away, and he didn’t even say
 thank you.’ (spoken text)

The conjunction *xu* has several other left-branching properties that demonstrate that it has drifted towards subordination, as detailed in Belyaev (n.d.). For example, it can be used without a second clause in answers to questions, or as a reaction to another speaker’s utterance:

- (25) mun puc yoδd =xu
 I.OBL son come.PRS:3SG =SEQ
 (‘Let’s go.’ —) ‘When my son comes.’ (‘My son comes and.’)

By itself, *at* has no such left-branching properties, apart from its prosody. However, when *at* appears with *to* ‘until’ in the first clause, the construction acquires the seman-

tics of immediate precedence (26a), as described in Belyaev (2025b). Importantly, this construction acquires subordinating properties, as seen in (26b) where the *to...at* clause is postposed.

- (26) a. šer to k-az wī um wayd-ow řud =at, az
 lion until EMPH-OBJ D3.M.OBL D3.F.OBL cry.INF-INF hear.PST=and from
 řoř zibud
 fear jump.PST.M
 ‘As soon as the lion heard her [the donkey’s] cry, he ran away in fear.’
 (Sokolova 1960: 32)
- b. šer az řoř zibud, um sado to řud =at
 lion from fear jump.PST.M D3.F.OBL sound until hear.PST=and
 ‘The lion ran away as soon as he heard her cry.’ (constructed example)

Thus in this case we are forced to recognize =at as a clause-final subordination marker. As for *to*, its position is variable, similarly to *agar* discussed above, but with more restrictions: in addition to the second position illustrated in (26a), it can also occupy initial (27) and preverbal (28) positions.

- (27) to wāδ =an yat =at, wī vo az wī vōst
 until D3.PL.NOM=3PL come.PST=and D3.M.OBL again OBJ D3.M.OBL bind.PST
 ‘As soon as they came back, he bound him again.’ (Sokolova 1960: 35)
- (28) yā potřo-razen pay newd sad, um yořk ar
 D3.SG.NOM king-daughter PROSP cry.INF go.PST.F D3.F.OBL tear LOC.SUB
 wī pec to čekt =at, yā agā sud
 D3.M.OBL face until drip.PST=and D3.SG.NOM awake go.PST.M
 ‘The princess started crying, and as soon as her tear fell on his face, he woke up.’
 (Sokolova 1960: 38)

Observe that in (27), the second-position agreement clitic does not attach to *to*, but to the following pronoun *wāδ* ‘those’, which means that it should be assigned to the same clause-initial position as *agar* and other Tajik loans.

3.5 Question particle =o

In addition to the elements identified above, Bartangi has one unambiguously clause-final marker, the question particle =o. This particle marks yes-no questions and normally occurs immediately after the verb when the usual SOV order is used (29a). However, if the clause is not verb-final, two position options are available: sentence-final (29b) and immediately postverbal (29c).

- (29) a. pulod az zarifā wīnt =o?
 P. OBJ Z. see.PST=Q
 ‘Did Pulod see Zarifa?’

Table 2: Positions of functional elements in the Bartangi clause

		initial	second	verb-adjacent	final
<i>di</i> _{EX}	general linker	? ⁱ	—	—	+
<i>di</i> _{IN}	temporal-conditional	—	+	+	—
<i>ca</i>	general subordinator	—	—	+	+ ⁱⁱ
<i>xu</i>	sequential	—	—	—	+
<i>at</i> ⁱⁱⁱ	temporal	—	—	—	+
<i>to</i>	temporal	+	+	+	—
<i>agar</i>	conditional	+	+	+	+
<i>=o</i>	question particle	—	—	+ ^{iv}	+

ⁱ Traditionally analyzed as initial, but see Section 3.2.

ⁱⁱ Only final when no verb is present in the clause, i.e. in present-tense copular clauses.

ⁱⁱⁱ In its function as a subordinator (together with *to*) only.

^{iv} Immediately postverbal rather than preverbal.

b. pulod wīnt az zarifā=*yo*?

c. pulod wīnt=*o* az zarifā?

This distribution is reminiscent of that of *ca* discussed above: the question particle is essentially sentence-final but may “hop” onto the verb. In the case of *ca*, this “hopping” is obligatory, while for *=o* it is optional.

3.6 Summary

A summary of the positions of the elements discussed above is provided in Table 2. The column “verb-adjacent” refers to the immediately postverbal position for *=o* and the immediately preverbal position for the other markers. The second position of *di*_{IN} and *to* is equivalent to that of clitic agreement markers; the subordinators always follow the agreement clitics.

These data demonstrate that in Bartangi we have to assume a rather large range of positions for conjunctions and particles: (i) *initial*: *to*, *agar* and other Persian/Tajik loans; (ii) *second*: *di*_{IN}, *to*; (iii) *preverbal*: *ca*, *di*_{IN}; (iv) *postverbal*: *=o*; (v) *final*: *xu*, *at* (with *to*...), *di*_{EX}, *ca* (without overt verb), *=o*. In the following section, I will discuss how this seeming complexity can be at least partially reduced in the analysis.

4 Analysis

4.1 General outline

Of the five descriptive types discussed in the preceding section, the preverbal and postverbal can be eliminated in the analysis, because both markers that appear in this position also have alternative positional options (final for *ca*, second for *di*_{IN}); moreover, the degree of their proximity to the verb is so high that they can be treated

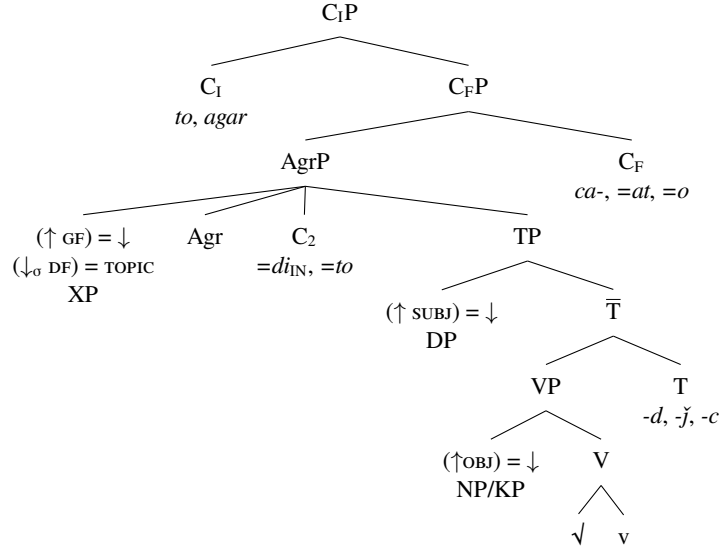


Figure 2: Bartangi clause structure with additional positions

as affixes rather than independent clitics (see formalization in Section 4.3 below). The other positions, however, are hardly reducible to one another, as their distributional properties are too distinct.

We can thus augment the structure in Figure 1 with additional structural positions as shown in Figure 2: C_I for clause-initial markers, C_F for clause-final and “preverbal” markers and C₂ for second-position markers. In what follows I will discuss each of these additional positions in more detail.

4.2 Clause-initial markers

Clause-initial markers are independent words that bear their own stress, thus their VIs are quite straightforward, as illustrated in (30) for *agar*. A VI in LrFG defines the relationship between terminal c-structure nodes providing syntactic and semantic information with their realizations. The left-hand side of a VI (before the right arrow) defines the *exponendum* – the information to be realized – which is a pair consisting of a list of c-structure categories (more than one category indicates a span) and a union of three types of information: f-structure constraints, Glue Semantics (Asudeh 2022) meaning constructors and information-structure-related definitions. In this specific case, I assume that markers such as *agar* ‘if’ are not associated with any f-structure features, thus the VI provided in (30) only realize the relevant semantic information, which is here represented by $\llbracket \text{if} \rrbracket$.³ As for the exponent in the right-hand side of the VI,

³Since this paper is concerned with the structural properties of these items rather than their semantics, I will not include the meanings in this analysis below. This does not affect the morphosyntactic discussion in any way.

here it is defined simply as the string *agar*, which is a shorthand for the corresponding phonological form of the word. We will see shortly that the exponent may be more complex.

$$(30) \quad \langle [C_I], [\text{if}] \rangle \xrightarrow{\nu} \text{agar}$$

4.3 Clause-final markers

The clause-final elements are less homogeneous. The conjunction *=at* in its subordinating function is a simple clitic in the sense of Zwicky & Pullum (1983): an element that does not form a separate prosodic word yet which syntactically behaves like a complete word occupying its respective position in the c-structure. In terms of LrFG, this amounts to the clitic having a DEP feature (in this case, LEFT as it is an enclitic) and a restricted prosodic frame, as in (31).

$$(31) \quad \langle [C_F], (\uparrow \text{CONJ}) = \text{AND} \rangle \xrightarrow{\nu} \begin{bmatrix} \text{PHON.REP} & /at/ \\ \text{PFRAME} & (\dots (\cdot)_{\sigma})_{\omega} \\ \text{DEP} & \text{LEFT} \end{bmatrix}$$

Note that *at* can only appear in its subordinating role in the presence of *to* in clause-initial or clause-internal position; since outside of this construction it is a general coordinating marker, I analyze it as having no semantics of its own in this context; its realization is instead linked to the f-structure feature CONJ, defining the specific conjunction used in this context.⁴

In (31), we can see a more complex exponent than in (30) above: the right-hand side of the VI is a feature structure (a *v-structure*) that includes not only the phonological form of the morpheme, but also additional information that determines a number of additional properties: prosodic status and context, linear position, inflectional class, constraints on the host etc.; see Asudeh et al. 2023 for an in-depth discussion of the handling of exponents in LrFG. In (31), PFRAME is intended to capture the fact that *at* is a clitic that appears at the end of the prosodic word. The value of the feature DEP similarly reflects its being hosted by the exponent to its left.

The status of *=o* is different: it can behave either as a “leaner”, like *=at*, or it can “hop” onto the verb. The latter process in LrFG is formalized via the notion of host identification: if a v-structure has the feature [HOST [IDENT +]], its host is determined not through c-structure-based adjacency but through c-command as the structurally *closest* realization. A complete definition of this notion is given in Asudeh et al. (2023). Informally, for a v-structure v_1 that realizes node n_1 , $\text{closest}(v_1)$ defines the v-structure v_2 that realizes node n_2 such that both n_1 and n_2 map to the same f-structure f , n_1 c-commands n_2 , and there is no intervening node n_3 that c-commands n_2 while also mapping to f . This is illustrated in Figure 3: in (a), v_2 can serve as the host for v_1 because there are no intervening c-commanding nodes; in (b), v_2 can only

⁴For a similar “semantically empty” treatment of a coordinating conjunction serving as a complementizer, see Belyaev (2015: 314). The difference between Ossetic and Bartangi is that in Ossetic, the conjunction still has the c-structure of a coordination marker, while in Bartangi, it has been reanalyzed as a final complementizer.

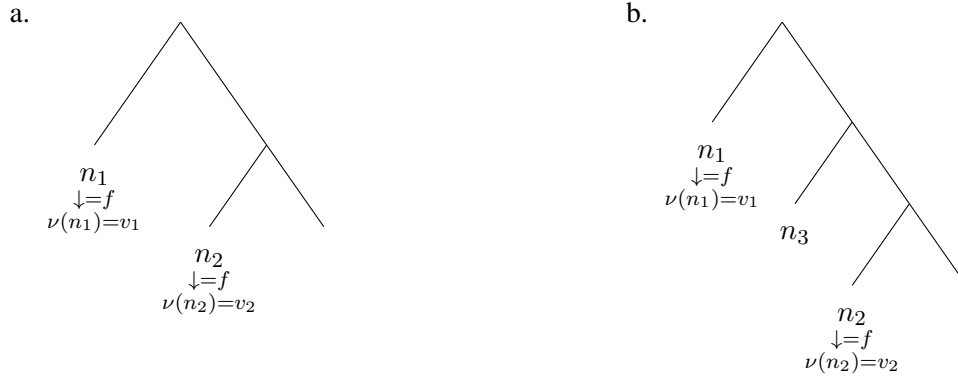


Figure 3: Illustrations for host identification in LrFG

be host to v_1 if the intervening n_3 does not map to the same f -structure f ; otherwise, *closest* is violated.

To avoid affixes attaching to clitics or other affixes with inverse direction of attachment (suffixes to prefixes, prefixes to suffixes), we can add an additional constraint that host identification as described above only applies to other affixes with the same directionality or to roots.⁵ Since this paper is not dedicated to formal issues of LrFG, I will leave a more developed definition of host identification for future work.

Given this definition, the VI for $=o$ is given in (32). The optionality of the [HOST [IDENT +]] specification means that it can attach either to the verb or to the element that appears immediately to its left according to c-structure-based linearization. When attaching to the verb, $=o$ can “skip” intervening nodes per the definition above if they do not map to the same f -structure; thus, in (29c), the question marker can attach to the verb while ignoring the postverbal direct object NP *az zarifā*.

$$(32) \quad \langle [C_F], [\text{yes-no-q}] \rangle \xrightarrow{\nu} \left[\begin{array}{l} \text{PHON.REP} \quad /o:/ \\ \text{PFRAME} \quad ((\cdot)_{\sigma} \dots)_{\omega} \\ \text{DEP} \quad \text{LEFT} \\ \left(\text{HOST} \quad [\text{IDENT} \quad +] \right) \end{array} \right]$$

The status of *ca* is similar, with the difference that when *ca* appears in final position, it is a fully stressed word, and it only does so when there is no valid host for it to attach to (i.e. no overt verb). I propose modeling this via the VI in (33), where not only host identity, but also the DEP and PFRAME features are optional. This means that here, the choice is not between an affix and a clitic, but between an independent word and an affix with a prosodic constraint.

⁵I am grateful to Ash Asudeh, Nadeem Siddiqi, Bronwyn Bjorkman and other participants of the LrFG lab for discussion of this issue and the proposed solution.

$$(33) \quad \langle [C_F], \llbracket ca \rrbracket \rangle \xrightarrow{\nu} \left[\begin{array}{l} \text{PHON.REP } /t^h a/ \\ \left(\begin{array}{l} \text{PFRAME } ((\cdot)_{\sigma} \dots)_{\omega} \\ \text{DEP } \quad \text{RIGHT} \\ \text{HOST } \quad [\text{IDENT } +] \end{array} \right) \end{array} \right]$$

Note that in both cases, LrFG’s MostSpecific constraint (Asudeh et al. 2023: 32–33) should ensure that the more constrained realization is chosen, i.e. the affix. This is the right prediction in the case of *ca*, because in Bartangi it cannot occur clause-finally if there is an overt verb for it to attach to. But for *=o*, both options are available, so we must assume that there is either a significant prosodic difference between two variants, such that one v-structure does not subsume the other, or an information-structure difference (such that the clitic variant is associated with higher emphasis, for example, like English *will* vs. *’ll*), or both.⁶ This is an empirical question raised by the analysis that has to be investigated during further research. If it turns out that no such differences are to be found, then postulating two different c-structure positions for the realization of *=o* seems unavoidable within the current framework.

4.4 Second position

All particles and conjunction that appear in second position are always adjacent to the agreement clitics and follow it. Since there is currently no evidence in favour of the second-position conjunctions forming a constituent with any adjacent nodes, the slot for such conjunctions can be located immediately after Agr and can be labeled as *C*₂. Since these markers are not roots or affixes, they do not interfere with agreement suffix placement in any way. A VI for the second-position variant of *to* is given in (34).⁷

$$(34) \quad \langle [C_2], \llbracket \text{until} \rrbracket \rangle \xrightarrow{\nu} to$$

Since *di*_{IN} has a prefixal variant in addition to the second-position one, we can treat it similarly to *ca*, as in (35); the curly braces stand for disjunction between two options. In this case free variation between second and preverbal positions for *di*_{IN} is predicted by neither of the VI being more specific than the other, as they are associated with different prosodic constraints.

⁶Note that in Wakhi, another Pamir language with a system of conjunctions and particles very similar to that of Bartangi, the counterpart to *ca*, *cə*, has a stressed variant *cəy* that usually appears sentence-finally when there is no overt verb, as in Bartangi. But, unlike Bartangi, it can also appear together with a verb (Grjunberg & Steblin-Kamenskij 1976: 654). However, according to Jaroslava Obrtelová (p. c.), in the latter case *cəy* “usually emphasises a relation that would normally be expressed with just *cə*”. Presumably a similar explanation can be provided for Bartangi sentence-final vs. postverbal *=o*.

⁷An alternative to postulating two positions for *to* is to use a prosodic inversion rule (Halpern 1995), such that *to* is optionally a clitic that moves after the first prosodic word if it ends up in clause-initial position, see Bögel (2015) for an LFG implementation. However, there is no evidence that the placement of *to* obeys prosodic rules or that it is an enclitic; in fact, /o/ in Bartangi is phonologically a long vowel, and no known clitics in Bartangi end in long vowels. Furthermore, a dedicated second position for conjunctions seems necessary to account for *di*_{IN}, which has no clause-final or clause-initial variants.

$$(35) \quad \langle [C_2], [\text{di}] \rangle \xrightarrow{\nu} \left[\begin{array}{c} \text{PHON.REP } /di/ \\ \left(\left[\begin{array}{cc} \text{PFRAME} & (\dots (\cdot)_{\sigma})_{\omega} \\ \text{DEP} & \text{LEFT} \end{array} \right] \vee \left[\begin{array}{cc} \text{PFRAME} & ((\cdot)_{\sigma} \dots)_{\omega} \\ \text{DEP} & \text{RIGHT} \\ \text{HOST} & [\text{IDENT } +] \end{array} \right] \right) \end{array} \right]$$

4.5 Worked examples

In Figures 4 and 5, I provide examples of a subordinate clause with the initial and final subordinators *to...at* ‘as soon as’ and a clause with preverbal placement of *ca*, respectively. The implementational details of the f-structure features need not concern us here; semantic labels are supposed to stand for the corresponding meaning constructors. In Figure 4, I treat the postposition *xīz* as a relational noun, which corresponds to the descriptive tradition (Karamxudoev 1973: 263) and its morphosyntactic properties – fully stressed, governs NP in oblique case, etc. – although a more careful study of the properties of Bartangi adpositions is yet to be undertaken. The analysis of verbal morphology follows Belyaev (2025a); specifically, the possibility of vocabulary spanning of the nodes Agr, T, v and $\sqrt{}$, which are adjacent in this example (*sawd* ‘goes’ being an intransitive verb), but which are non-adjacent if the direct object, an adverbial phrase or another element appears between Agr and the verb. An open question is how the collocation of *to...at* is to be established; for the time being, I assume that the syntactic lexical entry⁸ for the meaning ‘as soon as’ in C_I contains a constraint requiring a specific value for the feature *CONJ*, which can only be supplied by the conjunction *=at*. When *to* realizes the meaning ‘until’ or ‘while’, it will not be associated with this constraint.⁹

In Figure 5, the subordinator *ca* realizes the final position C_F but appears preverbally due to the more specific variant in (33) being chosen; in LrFG terms, it is a prefix. Were the clause to lack a verb, as it does in nonverbal predication (18), it would be an independent word, i.e. the features that appear in parentheses in (33) would not be used. I omit a more explicit discussion of such examples in this paper, as they would require a proper analysis of Bartangi nonverbal predication. Presumably, examples without an overt verb ‘to be’ have a different c-structure configuration that lacks T and V nodes, because LrFG does not allow zero root realizations. The important point is that in such a structure, the subordinator will be located at the very end of its clause, predicting that no clausal constituents may appear to its right.

⁸In LrFG, there is a distinction between the syntactic lexicon – which can be called the *syntacticon* – which supplies the annotations that appear on terminal nodes in the c-structure; and the morphological lexicon, or vocabulary, i.e. the list of VIs that serve to realize the information supplied by the syntacticon.

⁹Is is unclear whether the two words for *to* are to be treated as the same VI, or are even etymologically related. In the meaning ‘while, until’ it is clearly a Tajik loanword, but the meaning ‘as soon as’ is not attested in Tajik.

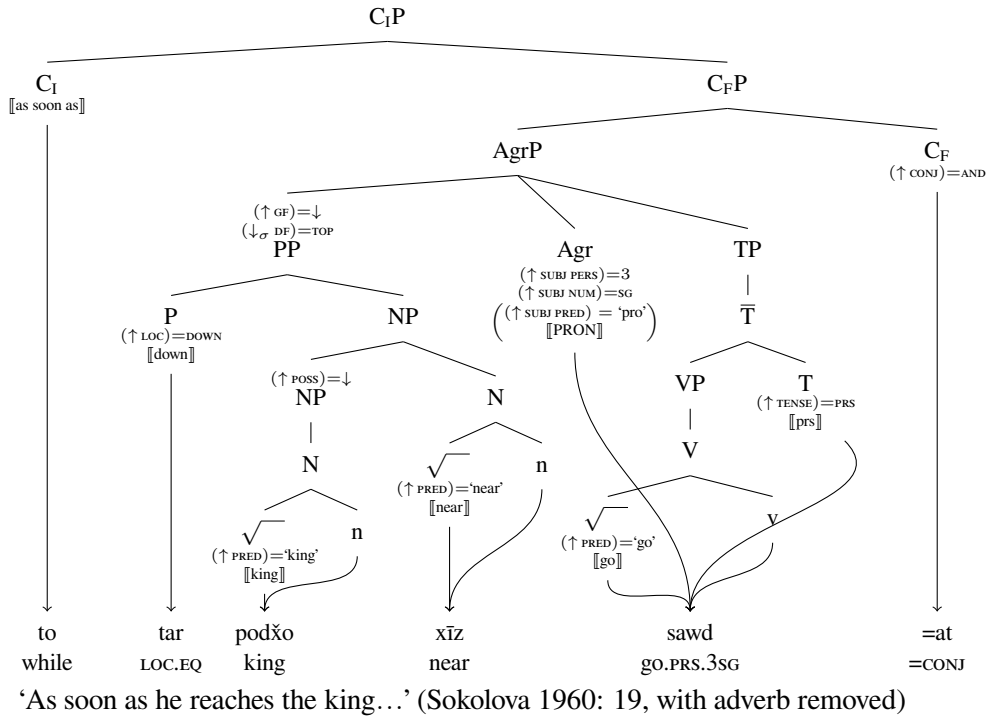


Figure 4: C- to v-structure mapping for an example with initial and final subordinators

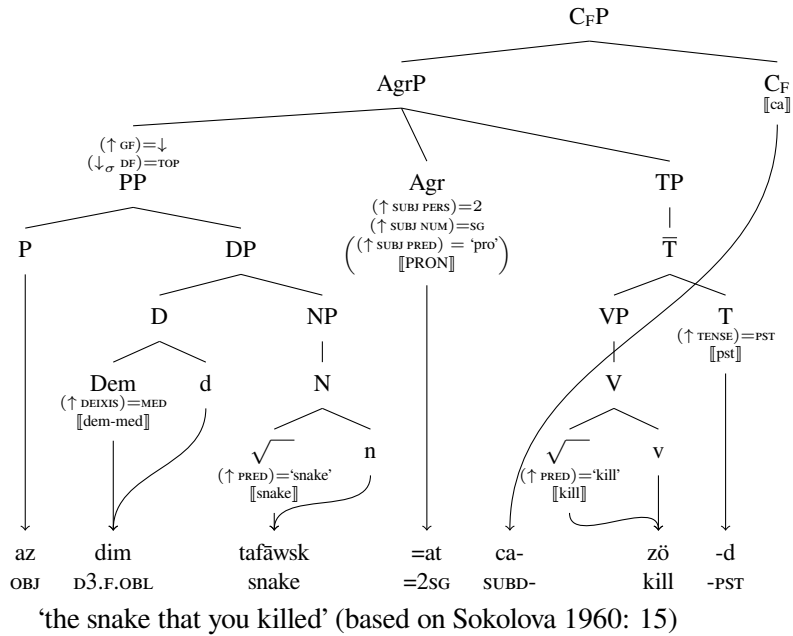


Figure 5: C- to v-structure mapping for a sentence with preverbal *ca*

5 Conclusions

In this paper, I have described aspects of the particle and conjunction system of Bartangi in terms of LrFG. I have shown that this language has three developed functional areas: clause-initial, second-position and clause-final. Preverbal subordinators, while being a prominent feature of the Pamir languages, can be described as essentially clause-final (*ca*) or second-position (*di*_{IN}) due to their distributional properties. Thus, while the initial, second and final positions are independently motivated, the preverbal placement of these subordinators is a consequence of them having affixal allomorphs. In standard LFG, the same data would have to be accounted for by either postulating a preverbal clausal position, or by including the *ca*- or *di*- prefixed verbs into the verbal paradigm, which is a different, lexicalist way of capturing the same competition between morphology and syntax.

This result is typologically significant. While sentence-final particles of various kinds seem to be very widespread cross-linguistically (cf. the papers in Hancil et al. 2015), the degree of their obligatoriness and integration into the grammatical system, as well as the range of functions they may encode, widely varies for different linguistic areas. In particular, Panov (2020) identifies an areal type of final particles typical of East and Southeast Asia, but not of Central or South Asia.¹⁰ Iranian languages specifically are not generally described as having a developed inventory of clause-final markers. For instance, all Iranian languages in the sample of Dryer (2013) are described as having clause-initial adverbial subordinators, with the exception of Wakhi (clause-internal type) and Ossetic (mixed type, i.e. clause-initial and clause-internal, see Erschler 2012, Belyaev 2014).

From this perspective, Bartangi comes across as mixed to an extreme, having three distinct structural positions under this analysis and as many as four (initial, final, second, preverbal) from a descriptive point of view. This diversity can potentially be explained by the Pamir languages being located between two linguistic areas as a kind of buffer zone (Stilo 2005): the tendency towards clause-initial markers follows the common Iranian pattern as exemplified by Tajik, while the clause-final positioning corresponds to the pattern found in neighbouring Dardic (described in Bashir 2003: 907 as mostly having left-branching subordinating structures) and Nuristani.

The LrFG analysis provided herein is compatible with this view of the development of phrase structure in Bartangi as a language-specific development rather than a reflection of any universal underlying principles of clausal organization. More broadly, LFG's conception of phrase structure and configurationality (Snijders 2015) does not seem to warrant any *theoretical* notion of "clause-final particle" or "conjunction"; such categories are more akin to cross-linguistic comparative concepts (Haspelmath 2010) along the lines of the types identified in works like Panov (2020). Despite being a formal generative framework, in this respect LFG aligns more with

¹⁰It is telling that no Iranian language in the sample of Panov (2020) has a final polar question particle, except Bukharan Tajik, which borrowed the question particle *mi* from Uzbek (Ido 2007: 7).

functional-typological approaches to language than with the universalist Chomskyan tradition.

References

- Abaev, Vasilij I. 1958. *Istoriko-ètimologičeskij slovar' osetinskogo jazyka*. [Historical etymological dictionary of Ossetic]. Vol. 1: A-k'. Moscow & Leningrad: Izdatel'stvo AN SSSR.
- Asudeh, Ash. 2022. Glue semantics. *Annual Review of Linguistics* 8. 321–341.
- Asudeh, Ash, Tina Bögel & Daniel Siddiqi. 2023. Modelling exponents. In Miriam Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *Proceedings of the LFG'23 conference*, 23–44. Konstanz: PubliKon. <https://lfg-proceedings.org/lfg/index.php/main/article/view/38> (2 October, 2025).
- Asudeh, Ash & Daniel Siddiqi. 2023. Morphology in LFG. In Mary Dalrymple (ed.), *The handbook of Lexical Functional Grammar*, 855–901. Berlin: Language Science Press. DOI: 10.5281/zenodo.10185976.
- Austin, Peter & Joan Bresnan. 1996. Non-configurationality in Australian Aboriginal languages. *Natural Language and Linguistic Theory* 14(2). 215–268. DOI: 10.1007/BF00133684.
- Bashir, Elena. 2003. Dardic. In George Cardona & Dhanesh Jain (eds.), *The Indo-Aryan languages*, 905–990. London: Routledge.
- Bashir, Elena. 2025. *A grammar of Khowar: Descriptive and comparative analysis*. London: UCL Press.
- Belyaev, Oleg. 2014. Anaphora in Ossetic correlatives and the typology of clause combining. In Pirkko Suihkonen & Lindsay Whaley (eds.), *On diversity and complexity of languages spoken in Europe and North and Central Asia*, 275–310. Amsterdam & Philadelphia: John Benjamins.
- Belyaev, Oleg. 2015. Systematic mismatches: Coordination and subordination at three levels of grammar. *Journal of Linguistics* 51(2). 267–326. DOI: 10.1017/S0022226714000450.
- Belyaev, Oleg. 2025a. The Bartangi clause: Configurationality, affixes and clitics. In Miriam Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *Proceedings of the LFG'25 Conference*, 41–62.
- Belyaev, Oleg. N.d. From coordinating conjunctions to clause-final subordinators in Bartangi and Ossetic. Submitted.
- Belyaev, Oleg & Dag Haug. 2014. Pronominal coreference in Ossetic correlatives and the syntax-semantics interface. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG'14 conference*, 89–109. Stanford, CA: CSLI Publications. <https://web.stanford.edu/group/cslipublications/cslipublications/LFG/19/lfg14.html> (13 September, 2020).
- Belyaev, Oleg I. 2025b. Nekompozicional'nyj sostavnoj konnektor *to...* at v bartangskom jazyke. In Maria B. Panich, Natalia V. Serdobolskaya, Sergei G. Tatevosov & Olga V. Fedorova (eds.), *Diskurs. Predloženie. Slovo (sbornik statej k*

- jubileju Iriny Mixajlovny Kobozevoj*) [Discourse. Sentence. Word (a collection of papers for Irina Kobozeva's anniversary)], 223–236. Moscow: Buki Vedi.
- Belyaev, Oleg I. 2025c. Polifunkcional'nye podčinitel'nye konstrukcii v bartangskom jazyke: K edinomu analizu. [Polyfunctional subordinating constructions in Bartangi: Towards a unified analysis]. *Voprosy jazykoznanija* 3. 30–61.
- Belyaev, Oleg I. 2025d. Pozicionnye klassy sojuzov v bartangskom jazyke [Position classes of conjunctions in Bartangi]. *Lomonosov Philology Journal* 2. 70–87. DOI: 10.55959/MSU0130-0075-9-2025-48-02-6.
- Bögel, Tina. 2015. *The syntax-prosody interface in Lexical Functional Grammar*. Universität Konstanz. (Doctoral dissertation).
- Bresnan, Joan, Ash Asudeh, Ida Toivonen & Stephen Wechsler. 2016. *Lexical-functional syntax*. 2nd edn. Oxford: Wiley Blackwell.
- Dryer, Matthew S. 2013. Order of relative clause and noun. In *The World Atlas of Language Structures Online*. Leipzig: Max Planck Institute for Evolutionary Anthropology. <http://wals.info/chapter/90> (12 April, 2019).
- Edelman, Joy I. 1980. K substratnomu naslediju Central'noaziatskogo jazykovogo sojuza. [Towards the substratum heritage of the Central Asian Sprachbund]. *Voprosy jazykoznanija* 5. 21–32.
- Erschler, David. 2012. From preverbal focus to preverbal “left periphery”: The ossetic clause architecture in areal and diachronic perspective. *Lingua* 122(6). 673–699.
- Grjunberg, Aleksandr L. & Ivan M. Steblin-Kamenskij. 1976. *Vaxanskij jazyk: Teksty, slovar'. grammatičeskij očerk*. Moscow: Glavnaja redakcija vostočnoj literatury.
- Halpern, Aaron. 1995. *On the placement and morphology of clitics*. Stanford, CA: CSLI Publications.
- Hancil, Sylvie, Alexander Haselow & Margje Post (eds.). 2015. *Final particles*. Berlin: De Gruyter Mouton.
- Haspelmath, Martin. 2010. Comparative concepts and descriptive concepts in crosslinguistic studies. *Language* 86(3). 663–687.
- Ido, Shinji. 2007. *Bukharan Tajik*. München: LINCOM EUROPA.
- Karamxudoev, Nodir. 1973. *Bartangskij jazyk (fonetika i morfologija)* [Bartangi language (phonetics and morphology)]. Dushanbe: Doniš.
- Liljegren, Henrik. 2020. The Hindu Kush–Karakorum and linguistic areality. *Journal of South Asian Languages and Linguistics* 7(2). 239–285. DOI: 10.1515/jsall-2021-2027.
- Obrtelová, Jaroslava. 2019. *From oral to written: A text-linguistic study of Wakhi narratives*. Uppsala: Acta Universitatis Upsaliensis.
- Panov, Vladimir. 2020. Final particles in Asia. *Linguistic Typology* 24(1). 13–70.
- Parker, Clinton. 2023. *A grammar of the Shughni language*. Montréal: McGill University. (Doctoral dissertation). <https://escholarship.mcgill.ca/concern/theses/0v8385834> (22 December, 2025).
- Rastorgueva, Vera S. & Joy I. Edelman. 2000. *Ètimologičeskij slovar' iranskix jazykov*. [Etymological dictionary of the Iranian languages]. Vol. 1: A–ā. Moscow: Izdatel'skaja firma «Vostočnaja literatura» RAN.

- Šakarmamadov, Nisormamad (ed.). 2005. *Fol'klor Pamira*. Vol. 2: *Mify, legendy i skazanija*. [Folklore of the Pamirs: Myths, legends and tales]. Dushanbe: Imperial-grupp.
- Snijders, Liselotte. 2015. *The nature of configurationality in LFG*. University of Oxford. (Doctoral dissertation). <https://ora.ox.ac.uk/objects/uuid:1310f160-283e-411e-a8d7-20ab4b3380c2> (2 October, 2025).
- Sokolova, Valentina S. 1960. *Bartangskie teksty i slovar'* [Bartangi texts and dictionary]. Moscow: Izdatel'stvo Akademii nauk SSSR.
- Stilo, Donald. 2005. Iranian as buffer zone between the universal typologies of Turkic and Semitic. In Éva Ágnes Csató, Bo Isaksson & Carina Jahani (eds.), *Linguistic convergence and areal diffusion: Case studies from Iranian, Semitic and Turkic*, 35–63. London: Routledge.
- Wendtland, Antje. 2009. The position of the Pamir languages within East Iranian. *Orientalia Suecana* LVIII. 172–188. <http://uu.diva-portal.org/smash/get/diva2:305500/FULLTEXT02.pdf> (2 October, 2025).
- Windfuhr, Gernot & John R. Perry. 2009. Persian and Tajik. In Gernot Windfuhr (ed.), *The Iranian languages*, 416–544. London: Routledge.
- Zarubin, Ivan I. 1937. *Bartangskie i rušanskije teksty i slovar'* [Bartangi and Rushani texts and dictionary]. Moscow & Leningrad: Izdatel'stvo Akademii nauk SSSR.
- Zwicky, Arnold M. & Geoffrey K. Pullum. 1983. Cliticization vs. inflection: English n't. *Language* 59(3). 502–513.