The Bartangi clause: Configurationality, affixes and clitics

Oleg Belyaev

Lomonosov Moscow State University, Institute of Linguistics RAS

Proceedings of the LFG'25 Conference

Miriam Butt, Jamie Y. Findlay and Ida Toivonen (Editors)

2025

PubliKon

lfg-proceedings.org

Abstract

In this paper, I provide an analysis of the clause structure of Bartangi, an Eastern Iranian language spoken in the Pamirs, which has a configurational clause structure and varying positions of agreement markers (second-position clitics and affixes) depending on tense. I show how an LrFG account can generalize the apparent variability from a regular, largely invariable c-structure. I also propose a modification to the mechanism of spanning in LrFG to account for irregular forms instantiating nonadjacent nodes.

1 Introduction

The typology of clause structure and configurationality is one of the central topics discussed in the LFG literature (see Snijders 2015) due to the framework's analytic versatility in the choice of clausal projections and structural positions for grammatical and discourse functions.[†] In this paper, I provide a detailed analysis of the simple clause structure of Bartangi, an Iranian Pamir language. The paper aims to be both an empirical contribution to the typology of configurationality in LFG and an attempt at providing a novel, non-lexicalist take on the traditional LFG notion of morphology competing with syntax (see i.a. Austin & Bresnan 1996, Bresnan 1998), by using LrFG (Asudeh & Siddiqi 2022) as a model of the morphology-syntax interface. Instead of a strict separation between morphology and syntax as distinct levels of representation, LrFG allows defining the same c-structure node as having either "syntactic" (clitics or independent words) or "morphological" (affixal) exponents depending on its environment. I argue that Bartangi combines a configurational core with a clause-initial topic position. A crucial feature of Bartangi is tense-dependent expression of agreement: by verbal affixes in the present and by second-position clitics in the past and perfect. In this analysis, I propose that both types of markers are associated with an Agr node that immediately follows the clause-initial topic; the difference between two modes of expression is reduced to the distinction between the vocabulary items (VIs) instantiating this node. This directly captures the complementary distribution between two marker types. I also discuss the syntax of suppletive and "zero-stem" forms and how it can be captured in LrFG in nonlocal structural configurations.

The paper has the following structure. In section 2, I provide general information on the Bartangi language. In section 3, I describe the clause syntax of Bartangi from a descriptive perspective. Finally, in section 4, I propose an LrFG analyis of Bartangi clause structure and agreement.

[†]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 participants of the LrFG lab at the University of Rochester, in particular Ash Asudeh, Nadeem Siddiqi, Frances Dowle, Bronwyn Bjorkman and Danil Alekseev, for their in-depth discussion of an earlier draft of this paper, as well as to the audience of LFG25 in Barcelona for an informative discussion. I would also like to thank two anonymous reviewers for their helpful and engaging comments, and Miriam Butt, Jamie Findlay and Ida Toivonen for their editorial work.

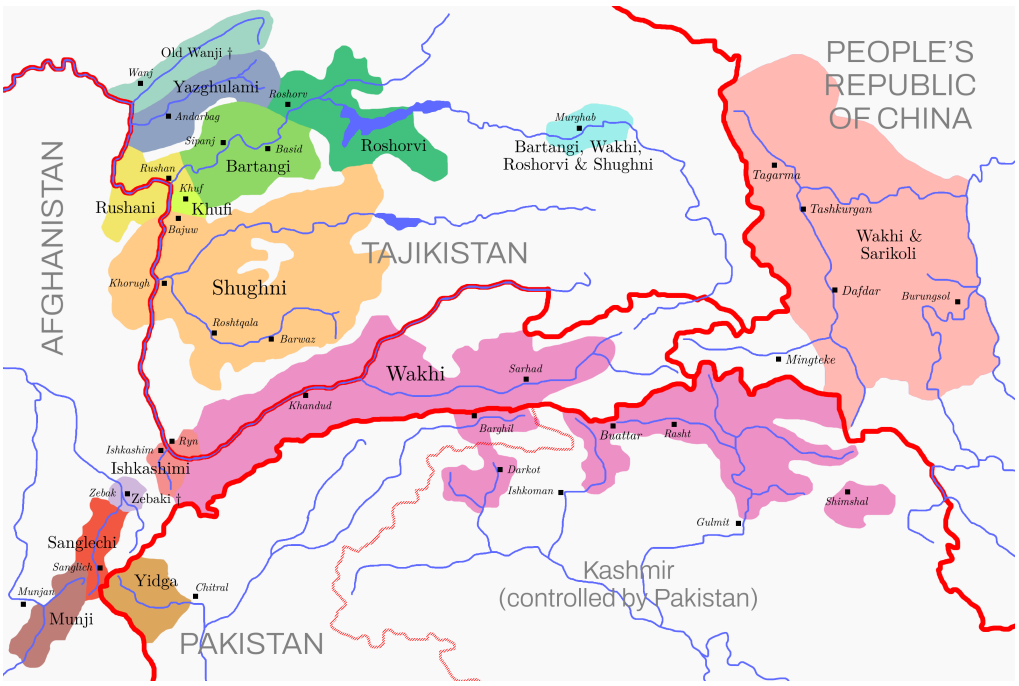

Figure 1: The Pamir languages, map © Yuri Koryakov and Maksim Melenchenko

2 General information on Bartangi

Bartangi is an Eastern Iranian language belonging to the Shughni-Rushani group, spoken in the valley of the river Bartang in the Gorno-Badakhshan Autonomous Oblast (GBAO) of Tajikistan by a few thousand speakers. Bartangi belongs to the so-called Pamir languages, an areal (Wendtland 2009) group of Iranian varieties spoken in the GBAO and in the neighbouring Afghan province of Badakhshan, and as such is characterized by many features common to this linguistic area, including SOV word order, three degrees of deixis, a two-case system on pronouns, a binary semantic gender distinction, optionality of number marking on nouns, a complex system of verbal stems, and preverbal subordinators. The area where Bartangi is spoken can be seen in the map in Figure 1.

There has been little work on Bartangi since the 1970s; two major sources on the languages are the text collection and dictionary by Sokolova (1960) and the grammar by Karamxudoev (1973). In this paper, I mostly rely on data collected during my own fieldwork in 2023–2025 in Basid, GBAO. I am grateful to all native speakers I have worked with during these field trips, in particular, to Fayzmamad Nazariyev, Anoyatsho Soibnazarov, Jonali Pirmamadov, Pulod Kuchakshoev, Kuchaksho Soibnazarov, Asad Niyozov, Khushqadam Bogbonov and Sabzpari Makhramova. Needless to say, all errors are mine.

3 Clause syntax

3.1 Basic word order

Like other Iranian languages, Bartangi has basic SOV word order, as illustrated by the following examples:

(1) *pulod rūpc wīnt*
P. fox see.PST
'Pulod saw a fox.'

(2) *āz =um az wī ɣāc wīnt*
I.NOM=1SG OBJ D3.M.OBL boy see.PST
'I saw the boy.'

However, the order is not strictly SOV and is in fact freer than described for closely related languages such as Shughni and Roshani. Sokolova (1960: 7) connected this relative freedom with the fact that Bartangi extensively uses differential object marking by means of the preposition *az*; and indeed, *az*-marked arguments seem more likely to be located in a noncanonical position (3). However, unmarked objects can also be displaced, especially postverbally (4).

(3) *az portugāliya grūziya ayod*
OBJ P. G. carry.PST
'Georgia defeated Portugal.'

(4) *pulod wīnt rūpc*
P. see.PST fox
'Pulod saw a fox.'

But Bartangi cannot be characterized as a "free word order" language, because SOV is strongly preferred in all contexts, regardless of information-structure properties and other factors that may influence word order. This sets Bartangi apart from so-called discourse configurational languages, such as Ossetic (Belyaev 2022) or Hungarian (Payne & Chisarik 2000, Laczkó 2014), which have highly grammaticalized topic and focus positions. This is illustrated by (5)–(6), which differ in whether the subject is the topic and the object the focus (5) or vice versa (6), but which maintain the same word order; that is, the order of constituents is based on their grammtical function rather than their discourse function.

(5) *pulod az čī wīnt? — yā az asad wīnt*
P. OBJ who see.PST D3.SG.NOM OBJ A. see.PST
'Who did Pulod see? — He saw Asad.'

(6) *čī az asad wīnt? — pulod az asad wīnt*
who OBJ A. see.PST P. OBJ A. see.PST
'Who saw Asad? — Pulod saw Asad.'

3.2 NP structure

3.2.1 Case marking

Like many Iranian languages, Bartangi has a two-case system that contrasts nominative (used for subjects)[1] with oblique (used in all other positions). Overt case marking only appears on personal and demonstrative pronouns, as seen in the following examples:

(7) ***āz*** *=um az* ***tā*** *wīn-t*
I.NOM=1SG OBJ thou.OBL see-PST
'I saw you.'

(8) ***yā*** *ɣaδā az* ***um*** *ɣāc wīn-t*
D3.SG.NOM boy OBJ D3.F.OBL girl see-PST
'The boy saw the girl.'

In (8), we can see that case distinctions in demonstratives are also observed when these demonstratives occur as modifiers with nouns (independent and adnominal forms of demonstratives are identical). In Bartangi and other Pamir languages, adnominal demonstratives are used as definite articles (Edelman 1990: 174); more specifically, as anaphoric articles, as described for Shughni in Parker (2023: 142). Hence, while case distinctions are only retained in a limited morphological class of elements, NPs in naturally occurring texts quite often have case distinctions due to the presence of adnominal demonstratives.

3.2.2 Differential object marking

Differential object marking (DOM) in Bartangi seems to be based on NP structure rather than semantics. It is more grammaticalized than in other Pamir languages, except Roshorvi and Sarikoli. The following NP types are obligatorily marked by the object-marking preposition *az*:[2]

- human proper nouns (5), and optionally, nonhuman proper nouns, as in (3);
- pronouns (7);
- DPs with demonstrative determiners (8).

Other NPs are generally left unmarked, regardless of their semantics. This is especially well-illustrated by the behaviour of possessive NPs, which, although clearly semantically definite in the sense of identifying a unique referent, are not marked by *az* if they have no determiner. This can be seen in (9), where the possessor is the unmarked proper noun *ǰonalī* 'Jonali', and in (10), where the adnominal demonstrative *um* is interpreted as the possessor due to mismatching gender; in the general case, demonstrative determiners in oblique case and possessor are homonymous, such that *um mod* (D3.F.OBL mother) can

[1] Historically, Bartangi used to mark past-tense transitive subjects by the oblique case, in a split ergative pattern. This system has all but disappeared in the modern language and is partly preserved only by some of the older speakers (Sergienko 2025).

[2] The object-marking preposition is formally identical to the preposition *az* 'from'; I gloss the former as OBJ and the latter as ABL to avoid confusion.

mead both ‘her mother’ and ‘the/that mother’ (obl.), and *um um mod* (D3.F.OBL D3.F.OBL mother) means ‘that mother of hers’.

(9) *āz* *=um* (*ʔ**az***) ***ǰonalī*** ***pid*** *wīnt*
I.NOM=1SG OBJ J. father see.PST
‘I saw Jonali’s father.’

(10) *daδ* ***um*** ***murδā*** *um-ri* *divīs-t* *=at…*
then D3.F.OBL dead_body D3.F.OBL-DAT show.PRS-3SG =CONJ
‘Then he shows her its dead body and…’ (Sokolova 1960: 15)

The same applies to other unique but non-anaphoric definites: these trigger neither *az* nor the demonstrative, as seen in (11) for a superlative description.

(11) *anoyatǩo* *ar* *māš* *maktab* (*ʔ**az***) *sarbašānd* *talabā* *ǩoyond*
A. LOC.DOWN we school OBJ best student teach.PST
‘Anoyatsho taught the best student at our school.’

3.2.3 NP structure

Based on the facts described above, I propose the structure in Figure 2 for the Bartangi noun phrase and the object marker *az*. Since the latter predominantly occurs on NPs with determiners, or on elements from specific lexical classes which can be viewed as inherently involving a determiner (pronouns, proper nouns), I assume that Bartangi has a DP and *az* is licensed by the object being a DP. The possessor is located within NP because of both its position closer to the head than the determiner and the fact that the presence of a possessor does not by itself license *az*.

The structure in Figure 2 has some properties that are unfamiliar from the perspective of traditional LFG. In LrFG, the terminal nodes of the c-structure tree itself are not morphological words or morphemes but the category nodes (P, D, $\sqrt{}$ and n) supplied with their f-descriptions just as words are in standard LFG.[3] These terminal nodes are mapped via the projection function ν to exponents – formalized in LrFG as v-structures (Asudeh, Bögel & Siddiqi 2023) – that are defined by vocabulary items (VIs) such as one in (22) below that couple a v-structure with a pair consisting of a list of c-structure categories and an f-description. As per the Elsewhere Principle, the realization for a given terminal node is the most specific match for this node’s category and f-description. A VI can define a list of c-structure categories; in this case, it instantiates all these categories at once if they are adjacent, similarly to lexical sharing (Wescoat 2002); the resulting

[3] Work in LrFG has been somewhat ambiguous on the status of these terminal nodes and their associated f-descriptions, and on whether these f-descriptions have to be reified in any sense. The current line of thinking (Ash Asudeh, p.c.) is that the category nodes together with their f-descriptions are stored in a “syntactic lexicon” – the so-called syntacticon – that acts as a language-specific repository of *exponenda*, things that can be realized by the morphology. Note that any realizational theory needs such an enumeration of exponenda – for example, a similar role is played by the content paradigm in PFM2 (Stewart & Stump 2007). The nodes stored in the syntacticon are inserted into the c-structure as per standard LFG mechanisms, but the f-description for each syntacticon entry is also interpreted by the function Φ that takes an f-structure and returns the (possibly empty) set of f-structures that it describes. It is against these sets of f-structures that f-descriptions in VIs are compared. A more explicit formalization of this mechanism is to be published in future work.

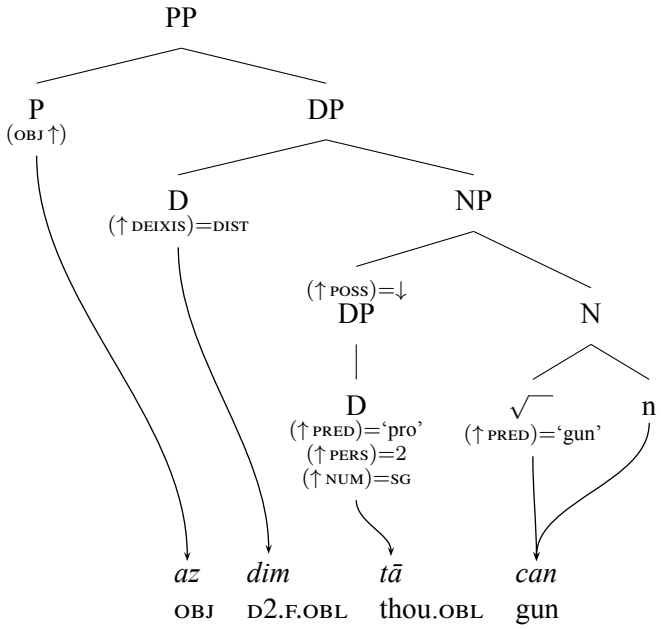

Figure 2: The general structure of Bartangi DPs

realization is called a *lexical span*. Lexical spanning models cumulation and suppletion and, if available, is always preferred to separate realization; an example of a lexical span in Figure 2 is the realization of $\sqrt{}$ and n by *can*.[4] If a node has no suitable realization, so-called *Pac-Man spanning* can link it to an adjacent node that maps to the same c-structure. For a more detailed and explicit discussion of LrFG's formal mechanisms, see Asudeh & Siddiqi (2022).

Bartangi (and Pamir languages in general) is one of the few Iranian languages that has retained a gender system, albeit in a highly modified form. The gender distinction itself is binary (masculine vs. feminine), but only based on sex for higher animals; lower animals and inanimates are assigned to various genders based on complex, poorly understood semantic factors; see Karamšoev (1978). In terms of marking, nouns do not have productive formal gender marking while agreement targets distinguish gender by means of vowel and stem change. Gender is marked on demonstratives, some verb forms and a few adjectives (cf. m. *ʒul* vs. f. *ʒil* 'small', although not all speakers use this distinction in the modern language).

Since demonstratives figure prominently in the examples in this paper, I provide the demonstrative paradigm in Table 1. Demonstratives distinguish between three degrees of deixis (proximal, medial and distal), number, case, and gender (in the oblique singular only). I do not use a CASE feature, because for modern Bartangi, a simple inside-out subject/nonsubject distinction is sufficient.

There are other complications in the NP/DP structure and object marking that I will not go into here, as the purpose of this paper is to describe the clause structure of Bar-

[4]Nodes like n are called *categorizers* and are a convention that LrFG inherits from DM that models cross-categorial roots. For the purposes of describing Bartangi, all stems (except perhaps non-verbal components of complex verbs) can be viewed as root + categorizer lexical spans.

Table 1: Demonstrative forms

	D1			D2			D3		
CASE	M	F	PL	M	F	PL	M	F	PL
NOM	*yim*		*māδ*	*yid*		*dāδ*	*yā*		*wāδ*
OBL	*mī*	*mim*	*mif*	*dī*	*dim*	*dif*	*wī*	*um*	*uf*

Table 2: Diagnostic forms of the verbs *sifedow* 'go up', *čegow* 'do' and *wāɣdow* 'shout'

PRS	PRS.3SG	PST.M	PST.F/PL	PF.M	PF.PL	PF.F	INF
sifān-	*sifen-t*	*sifö-d*	*sifā-d*	*sifö-ǰ*	*sifā-ǰ*	*sife-c*	*sife-d(-ow)*
kin-	*kiǐt*	*čū-g*		*čūǰ*			*če-g(-ow)*
wāɣ-	*wāɣ-d*			*wāɣ-ǰ*			*wāɣ-d(-ow)*

tangi.

3.3 Verb morphology

3.3.1 Stems

Like in all other modern Iranian languages, Bartangi verb inflection is based on a system of stems whose formation involves a number of irregularities. In Bartangi, verbs maximally distinguish between eight different basic forms: the present stem;[5] a special irregular form for 3rd person singular present; the masculine and feminine/plural past-tense forms; the masculine, plural, and feminine perfect forms; and the infinitive. In practice, most verbs do not have all eight forms; in particular, gender distinctions are only found in a subset of intransitive verbs. Minimally, verbs have to distinguish between the present (unmarked), the past+infinitive and the perfect, the latter two of which use distinct suffixes. In Table 2, I show the simplest case (*wāɣdow* 'to shout' with just one stem), the most complex case (*sifedow* 'to go up' with all eight distinct forms) and a more typical "middle ground", *čegow* 'to do'.

This system of gender marking via ablaut is quite unusual and has been the object of several explanation attempts, both synchronic (Muravyeva 1975, Makarov & Plungian 2023, Makarov 2024, Stump & Hippisley 2011) and diachronic (Edelman & Jusufbekov 1999). A concise overview of the historical development of verb stems in Shughni is found in Parker (2023).

While it may be possible to provide a synchronic account of gender agreement by treating vowels as some kind of autosegmental morphemes, the value of such an approach is doubtful, considering that the total number of verbs that have gender agreement of any kind in the dictionary of Sokolova (1960) is only 29 (out of a total of 131). Therefore, I consider treating all stems as listed items to be justified.

[5] I use the term "stem" to designate the part of the verb form which precedes the tense or infinitive suffix: e.g., for the past-tense form *ǎxovd* 'slept', the stem is *ǎxov-* while *-d* is the tense suffix. This is different from common usage in the Iranological tradition, where the whole form *ǎxovd* is commonly treated as a past stem.

However, this does not mean that one cannot capture any regularities involved in Bartangi verb inflection. The past-tense and infinitive always use the same suffix, which is *-d/t* for all verbs, with the exception of five verbs – *če-**g**-ow* ‘do’, *me-**g**-ow* ‘die’, *ve-**g**-ow* ‘bring’, *xe-**g**-ow*, *zide-**g**-ow* ‘sweep’ – which use *-g*.[6] The perfect suffix is *-č/j*; *-c* is used in the feminine with those verbs that display a gender distinction in the perfect. Further, stems can be organized into two hiearchies: PRS > PST > PRF and PRS > PST > INF, on which each distinct stem must occupy a continuous area. There are other, more complex interactions involving gender (e.g. if there is a distinct feminine perfect form, then there is also a distinct infinitive stem, and these two forms share a vowel), but gender agreement is a separate issue that is beyond the scope of this paper.

3.3.2 Person-number marking

Finite verbs in Bartangi always agree in person and number with the subject. The expression of this agreement differs in the present tense on the one hand, and in the past and perfect, on the other. In the present tense, person and number are expressed by affixes that attach to the verb, regardless of its linear position (12). In the past and perfect forms, second-position clitics are used instead; the clitic typically stands after the first NP or adverb in the clause (13), but, if the verb is clause-initial, it follows the verb (14).

(12) *āz* *az* *pulod* *wīn**-um***
I.NOM OBJ P. see.PRS-1SG
‘I see Pulod.’

(13) *āz**=um*** *az* *pulod* *wīnt*
I.NOM=1SG OBJ P. see.PST
‘I saw Pulod.’

(14) *wīnt**=um*** *az-ī*
see.PST=1SG OBJ-D3.M.OBL
‘See him, I did.’

The affixes and clitics are mutually exclusive within a single clause. Their paradigms have many things in common but are generally distinct, as seen in Tables 3 and 4. In the affix paradigm, the 2nd person singular is unmarked, and the two allomorphs in 3rd person singular are phonologically distributed. In the clitic paradigm, it is the 3rd person singular that is unmarked, while the 3rd person plural has two variants: the general *=an* and the specialized *=af*, which only occurs with transitive verbs (and some unergatives; see Chistyakova 2025) and has an additional restriction: it is only used when the subject is pro-dropped (15). The only exception from this generalization is the combination *uf=af* (16), which combines the oblique form of the 3rd person plural distal demonstrative (a relic of the earlier split ergativity) with *=af*. In all other forms, pro-drop is optional, as seen from the use of *=an* in both sentences in (15).

[6] This historically goes back to the stem-final change **rd* → *g*; cf. Persian *ka**rd**-an* ‘do’, *mo**rd**-an* ‘die’, *bo**rd**-an* ‘bring’, *xo**rd**-an* ‘eat’. However, in Bartangi this is not a synchronic process anymore as it does not occur, for example, in the present tense: *xer-d* ‘eats’, but not **xeg*.

Table 3: Person-number affixes

	SG	PL
1	*-um*	*-an*
2	∅	*-af*/*-at*
3	*-d*/*-t*	*-an*

Table 4: Person-number clitics

	SG	PL
1	*=um*	*=an*
2	*=at*	*=af*/*=at*
3	∅	*=an*, *=af* (tr.)

(15) a. *wāδ =an / *=af az-ī kud wīnt*
D3.PL.NOM=3PL =3PL.TR OBJ-D3.M.OBL dog(M) see.PST
'They saw the dog.'

b. *az-ī kud =af/ =an wīnt*
OBJ-D3.M.OBL dog(M)=3PL.TR =3PL see.PST
id.

(16) *uf =af az-ī kud wīnt*
D3.PL.OBL=3PL.TR OBJ-D3.M.OBL dog(M) see.PST
'They saw the dog.'

There are no idiosyncrasies or phonological constraints involved in the attachment of clitics to their hosts, with the exception of the optional elision of the initial vowel after the final vowel of the host (*tū=at* ~ *tū=t*) and equally optional *y*-epenthesis (*zarifā=(y)um*). But in the affix paradigm, there are three types of irregularities. First, many verbs have irregular 3rd person singular present forms, such as the verb 'do' shown in Table 2, which has the present stem *kin-* but the 3rd person *kiǐt*. Second, a few frequently used forms have contracted present forms, e.g. *sām* 'I go', *sām* 'we go' for *sāw-um*, *sāw-an*. Third, while the imperative and the 2nd person present are generally not distinguished in Bartangi, some verbs drop the final consonant in the imperative (with compensatory lengthening) but not in the present: *sāw* 'thou goest', *sǎ̄* 'go'; *pataw* 'thou throwest', *patā* 'throw' (Karamxudoev 1973: 175).

3.4 The second-position constraint

Cross-linguistically, second-position clitics can be licensed prosodically, syntactically, or via a mixed approach. In Bartangi, purely syntactic licensing seems to be sufficient. The pre-clitic material always forms a single constituent; in particular, NPs cannot be split by clitics, as seen in (17), where angle brackets indicate two potential positions for the clitics, only one of which is grammatical.[7]

(17) *pulod <* =um > čöd < =um > wīnt*
P. =1SG house =1SG see.PST
'I saw Pulod's house.'

[7] An exception is when the possessor is marked by the special case/adposition *-ā* and – in most instances – doubled by a possessive pronoun, such as *pulod-ā* ***=yum*** *wī čöd wīnt* (P.-POSS =1SG D3.M.OBL house see.PST) 'I saw Pulod's house', lit. 'to Pulod his house I saw'. But the possessor in such constructions seems to be external, attaching at clause level; cf. Serdobolskaya & Belyaev (2022) for the analysis of a similar construction in Ossetic.

Furthermore, the pre-clitic constituent has a special information-structure status. As seen in (5)–(6) above, the subject can occupy the initial position regardless of its information-structure status.[8] But when another argument occupies this position, it must be topical: in (18), the answer is illicit because Asad is treated as topical, but would have been felicitous if the question was 'Who saw Asad?'.

(18) *tū=t* *az* *čī* *wīnt?* — *#az* *asad=um* *āz* *wīnt*
thou.NOM=2SG OBJ who see.PST OBJ A.=1SG I.NOM see.PST
'Whom did you see? — # I saw Asad.'

The initial position is especially common for topical, stage-setting adverbs. As seen in (19), in this case the rest of the arguments follow in their standard configurational order. In terms of clause structure, this seems to suggest an initial topical position marked by the clitic followed by the configurational core of the clause.

(19) *biyor* *=um* *āz* *ǰac* *birozd*
yesterday=1SG I.NOM water drink.PST
'Yesterday I drank water.'

4 Analysis

The overall structure I propose for the Bartangi clause is displayed in Figure 3, which illustrates (19). I postulate that the Bartangi clause has three layers: AgrP, TP and VP. I will now provide a motivation and description for each of these levels in turn.

4.1 AgrP: The finite clause

AgrP is the layer where the person-number markers are introduced. As we saw above, the pre-clitic position must be occupied either by the subject or by a topic constituent.[9] This is reflected in the phrase structure rule in (20). I assume a flat structure for AgrP because I see no good evidence for treating the clitic as forming a constituent with the following TP.

(20) AgrP → XP $\left\{ \begin{array}{l|l} (\uparrow \text{GF}) = \downarrow & (\uparrow \text{SUBJ}) = \downarrow \\ (\uparrow_{\sigma} \text{DF}) = \text{TOPIC} & \end{array} \right\}$ Agr TP $\uparrow = \downarrow$

Both the initial XP and Agr are obligatory. The Agr is a terminal, and is thus provided with no f-description in the rule; the f-description is inserted from the syntacticon, which provides all combinations of person-number features of the subject, plus an optional PRED. Agr is realized by a clitic or an affix depending on the conditions for the corresponding vocabulary items (VIs). For this reason, the affix/clitic properties of the marker can be associated with its VI directly, as described in detail in section 4.4 below.

[8] An anonymous reviewer suggests that subjects appear in this position more often than other constituents because they are default topics. However, this is contradicted by examples like (6), where the subject is clearly in focus.

[9] This analysis does not currently handle verb fronting: if V is allowed to appear in the initial position, the tense affix will be unable to find a valid host. It is possible that verb-initial clauses should be given a prosodic treatment; more research on this topic is needed.

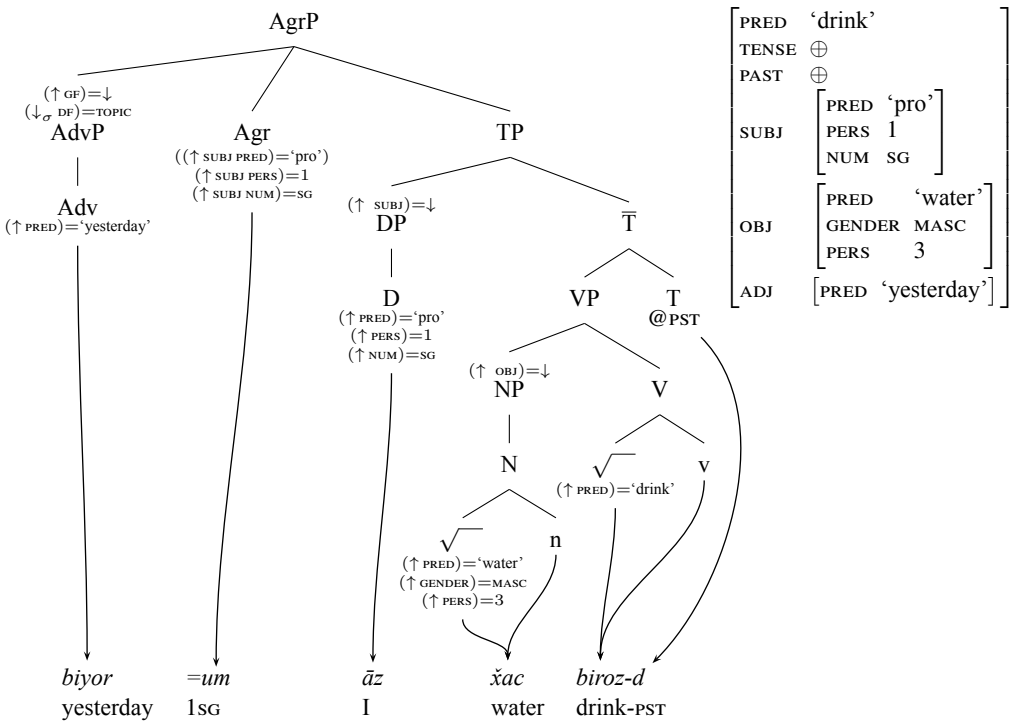

Figure 3: Clause structure with clitic person-number marker

4.2 TP: The clausal core

TP is headed by the T that is instantiated by the tense suffixes on the verb and includes the subject. The reason for positioning T relatively high in the tree and placing the subject in SpecTP rather than VP is the behaviour of infinitival clauses. So-called "infinitives" in Bartangi are rather nominalizations, mixed categories that combine internal verbal and external nominal syntax. The direct object in Bartangi infinitives can be expressed in the same way as in independent sentences (21) or by a possessive phrase, while the subject can only be expressed by a possessor, which is seen from the ungrammaticality of using the nominative pronoun. From this, I conclude that Bartangi infinitives construct NP on top of VP, and the infinitive marker realizes the N head.[10] A sample structure for an infinitival clause is shown in Figure 4.[11]

[10]This analysis implies that the identity of the infinival marker *-d/t* and the past stem marker *-d/t* is due to homonymy. But note that *-d/t* is also used as the 3rd person singular present ending, which is certainly not to be considered the same morpheme. As for the affinity between the infinitival and past stems, of the 131 simplex verbs in Sokolova (1960), only 15 have the same vowel in the infinitive and past stems that is different from the vowel in the present stem, and the majority of these verbs use the vowel *e*, the same vowel that is always used in infinitives having their own distinct stem. Overall, it is unclear how a nominalization and a past tense could be connected in the syntax from the synchronic point of view.

[11]The complex category N_V is used following Asudeh & Siddiqi (2022) to capture the selectivity of this N with respect to verbal complements. A possible alternative is to handle such constraints at v-structure, i.e. in the morphological projection.

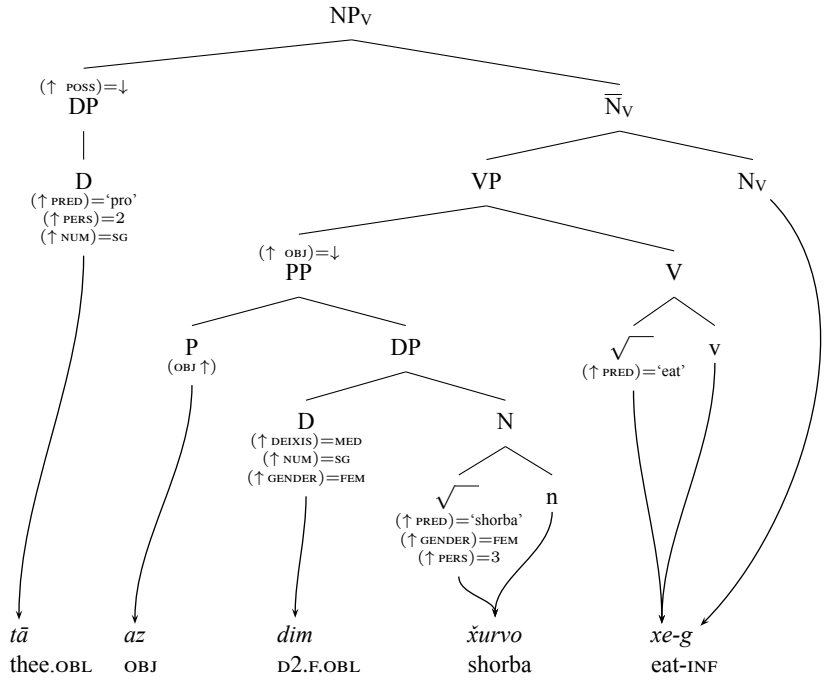

Figure 4: Structure of an infinitival clause (up to NP)

(21) *tā* / **tū* ***az*** *dim* *x̌urvo* *xe-g-qatī* *māš* *sān*
thou.OBL thou.NOM OBJ D2.F.OBL shorba eat-PST-INS we go.PRS:1PL
'We will go after you eat this shorba.' (lit. "with your eating this shorba")

The VIs for the tense markers that realize T are provided in (22); curly braces around f-descriptions are included for clarity. The present-tense feature has no associated VI, while the past and perfect have their own consistent exponents, with the exception of the *-g* past stem,[12] feminine perfect *-c*, which has to be morphologically conditioned to attach only to stems that are defined as having a feminine distinction in the perfect. The templates for the tense features themselves will be explained in the next section.

(22) a. past
$\langle [\mathrm{T}], \{ @\textsc{pst} \} \rangle \xrightarrow{\nu}$ *-t*

b. past in *-g*

$$\langle [\mathrm{T}], \{ @\textsc{pst} \} \rangle \xrightarrow{\nu} \begin{bmatrix} \textsc{phonrep} & /\mathrm{g}/ \\ \textsc{dep} & \textsc{lt} \\ \textsc{host} & \begin{bmatrix} \textsc{class} & \textsc{g-pst} \end{bmatrix} \end{bmatrix}$$

c. perfect
$\langle [\mathrm{T}], \{ @\textsc{prf} \} \rangle \xrightarrow{\nu}$ *-č*

[12] Alternatively, since there are only five verbs with this stem ending, the whole words *čūg*, *vūg* etc. can be treated as vocabulary spans.

d. feminine perfect

$$\langle [\mathrm{T}], \left\{ \begin{array}{c} @\textsc{prf} \\ (\uparrow \textsc{subj gender}) = \textsc{fem} \end{array} \right\} \rangle \xrightarrow{\nu} \begin{bmatrix} \textsc{phonrep} & /\mathrm{ts}/ \\ \textsc{dep} & \textsc{lt} \\ \textsc{host} & \left[\textsc{class} \quad \textsc{prf-f}\right] \end{bmatrix}$$

4.3 VP: The verbal stem and direct object

Above I motivated the positioning of the direct object inside VP, as opposed to the higher subject in TP. Whether other verbal complements, such as recipients, or obliques, are also positioned as VP complements, or lack their own positions and are adjoined at different levels in the clause, is a question for further research.

Turning now to the verbal stems, under this analysis they are stored as vocabulary spans specified for tense features. According to the description in section 3.3.1, I use a template inheritance hierarchy for tenses shown in (23), in the spirit of Asudeh et al.'s (2024) analysis of Latin case. I will also define templates to be used in defining the recurrent irregular forms of verbal stems in (24). Double angle brackets are an LrFG convention for indicating constraining equations; VIs only instantiate defining equations, and constraining equations within VIs provide additional context that must be satisfied for the VI to be a valid realization.

(23) a. $@\textsc{prs} := (\uparrow \textsc{tense}) = \oplus$

b. $@\textsc{pst} := @\textsc{prs}$
$(\uparrow \textsc{past}) = \oplus$

c. $@\textsc{prf} := @\textsc{pst}$
$(\uparrow \textsc{perf}) = \oplus$

(24) a. template for past/perfect masculine stem
$@\textsc{pst-m} := @\textsc{pst}$
$\langle\langle (\uparrow \textsc{subj gender}) = \textsc{masc} \rangle\rangle$

b. template for perfect feminine stem
$@\textsc{prf-f} := @\textsc{prf}$
$\langle\langle (\uparrow \textsc{subj gender}) = \textsc{fem} \rangle\rangle$

This hierarchy[13] ensures that the absence of a VI for a stem lower in the hierarchy will lead to the realization for the less-marked stem being used in its stead. Thus, for instance, the verb *wāɣdow* 'shout' from Figure 2 will have a single stem which is not marked for tense at all, as it is also used in the infinitive. The VI in (25) will thus be compatible with features from all three templates in (23). The verb *birexťow* 'to drink' (26) will have a

[13]The original version of this paper defined @PRS as having the feature PRS = +, which meant that the past tense was defined as having the features PRS and PST simultaneously. An anonymous reviewer observed that this looks self-contradictory, thus I changed the PRS feature to TENSE: this captures the intuition that the present is the least marked "tensed" form, from which all other tenses inherit. I retained the label PERF for the perfect because, while today it is more of a past evidential than a perfect (Melenchenko 2023), its full range of uses is too complex for any simple label to be an exact fit. The ⊕ sign is a notational convention to indicate that these are privative features.

present stem, a past stem (that is used as both past and perfect) and an infinitive stem,[14] which I treat as unspecified for tense. Thus among the forms in (26), the unmarked form will only surface if the TENSE feature is not present, because there are more marked forms available; similarly, the present stem will only be used in the present tense, because there is a specialized past stem. Finally, the maximally distinctive verb *sifedow* 'to go up' (27) has an infinitive stem, a present stem, specialized stems for masculine past/perfect and feminine perfect,[15] and a general past stem which emerges as the unmarked in all other combinations of gender and number.

(25) $\langle [\sqrt{\ }, \text{v}], \{(\uparrow \textsc{pred}) = \text{'shout'}\} \rangle \xrightarrow{\nu} wāɣd$

(26) a. infinitive stem (inf. *bireǩ-t(-ow)*)
$\langle [\sqrt{\ }, \text{v}], \{(\uparrow \textsc{pred}) = \text{'drink'}\} \rangle \xrightarrow{\nu} bireǩ$

b. present stem (prs. *biroz-*)
$\langle [\sqrt{\ }, \text{v}], \left\{ \begin{array}{c} (\uparrow \textsc{pred}) = \text{'drink'} \\ \langle\!\langle @\textsc{prs} \rangle\!\rangle \end{array} \right\} \rangle \xrightarrow{\nu} biroz$

c. past + perfect stem (pst. *biroǩ-t*, prf. *biroǩ-č*)
$\langle [\sqrt{\ }, \text{v}], \left\{ \begin{array}{c} (\uparrow \textsc{pred}) = \text{'drink'} \\ \langle\!\langle @\textsc{pst} \rangle\!\rangle \end{array} \right\} \rangle \xrightarrow{\nu} biroǩ$

(27) a. infinitive stem (inf. *sife-d-ow*)
$\langle [\sqrt{\ }, \text{v}], \{(\uparrow \textsc{pred}) = \text{'go up'}\} \rangle \xrightarrow{\nu} sife$

b. present stem (prs. *sifān-*)
$\langle [\sqrt{\ }, \text{v}], \left\{ \begin{array}{c} (\uparrow \textsc{pred}) = \text{'go up'} \\ \langle\!\langle @\textsc{prs} \rangle\!\rangle \end{array} \right\} \rangle \xrightarrow{\nu} sifān$

c. past/perfect masculine stem (pst. m. *sifö-d*, prf. m. *sifö-ǰ*)
$\langle [\sqrt{\ }, \text{v}], \left\{ \begin{array}{c} (\uparrow \textsc{pred}) = \text{'go up'} \\ \langle\!\langle @\textsc{pst-m} \rangle\!\rangle \end{array} \right\} \rangle \xrightarrow{\nu} sifö$

d. perfect feminine (prf. f. *sife-c*)
$\langle [\sqrt{\ }, \text{v}], \left\{ \begin{array}{c} (\uparrow \textsc{pred}) = \text{'go up'} \\ \langle\!\langle @\textsc{prf-f} \rangle\!\rangle \end{array} \right\} \rangle \xrightarrow{\nu} \begin{bmatrix} \textsc{phonrep} & /\textsc{sife}\text{ː}/ \\ \textsc{class} & \textsc{prf-f} \end{bmatrix}$

e. past feminine / plural and perfect plural
(unmarked past stem: pst. f./pl. *sifā-d*, pst. pl. *sifā-ǰ*)
$\langle [\sqrt{\ }, \text{v}], \left\{ \begin{array}{c} (\uparrow \textsc{pred}) = \text{'go up'} \\ \langle\!\langle @\textsc{pst} \rangle\!\rangle \end{array} \right\} \rangle \xrightarrow{\nu} sifā$

4.4 Person-number markers

The distribution of the affix and clitic markers is captured in LrFG by defining them as such: the former, as an affix, through the feature [HOST IDENT +], while the latter, through

[14] It might be that the consonantal alternation in the past and infinitive (*z*+*t* → *ǩt*) is morphophonological. If so, this verb can be described as having just two stems: present and infinitive.

[15] The equivalence of the vowel in the infinitive and perfect feminine is a recurring feature in all Bartangi verbs. However, I have to view these forms as homonymous because there is no natural class where these two forms would fit; neither can they be viewed as the unmarked case. In the feminine perfect, I also make use of the CLASS feature to ensure that such forms always use the special feminine perfect affix *-c*.

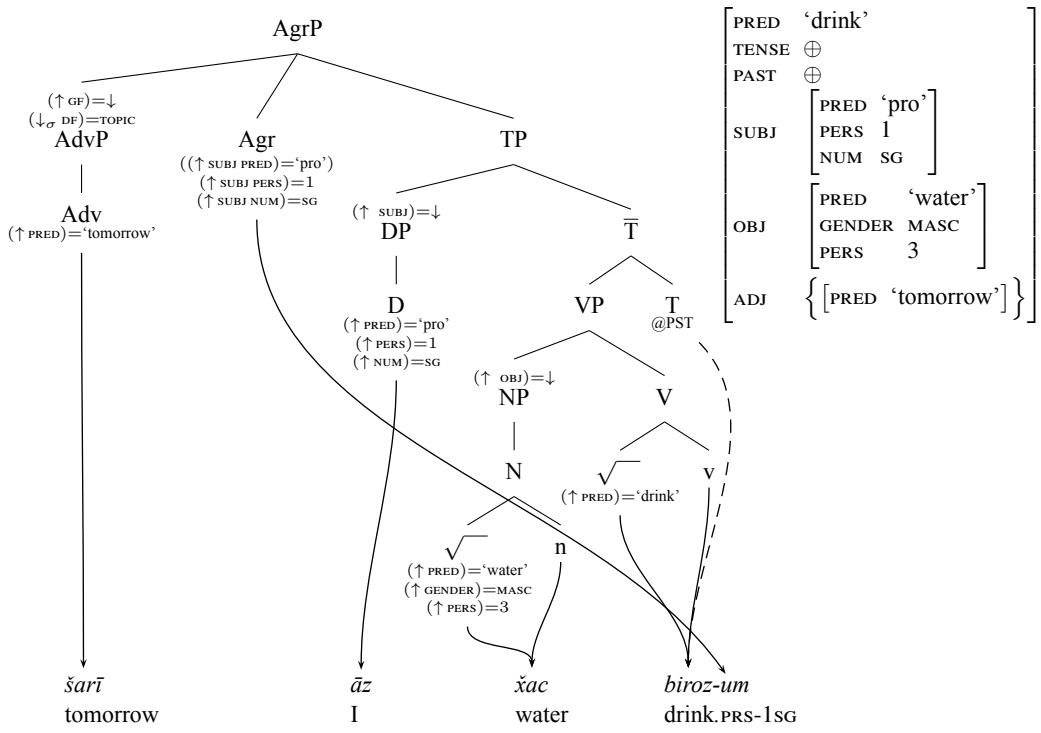

Figure 5: Clause structure with affixal person-number marker

left-dependence and prosodic constraints. The clitic will be linearized in its c-structure position, while the affix will attach to the closest c-structure node it c-commands that maps to the same f-structure (Asudeh, Bögel & Siddiqi 2023). The VIs for a sample affix and clitic are given in (28); the feature hierarchy in (23) ensures that the clitics defined as @PST will also be used in the perfect. The PFRAMES given in the v-structures are not meant as fully formalized representations; they merely serve to indicate that the exponent must stand at the end of a phonological word. A sample affixal realization, for the sentence in (29), is shown in Figure 5. Dashed lines indicate so-called Pac-Man spanning, where a VI co-instantiates an adjacent node if the latter has no licit realization.

(28) a. $\langle$ [Agr], $\left\{ \begin{array}{c} (\uparrow \text{SUBJ PERS}) = 1 \\ (\uparrow \text{SUBJ NUM}) = \text{SG} \\ \langle\langle @\text{PRS} \rangle\rangle \end{array} \right\} \rangle \xrightarrow{\nu} \left[\begin{array}{ll} \text{PHON.REP} & /\text{ʊm}/ \\ \text{PFRAME} & (\ldots(\ \cdot\ )_\sigma)_\omega \\ \text{DEP} & \text{LEFT} \\ \text{HOST} & [\text{IDENT}\ +] \end{array} \right]$

b. $\langle$ [Agr], $\left\{ \begin{array}{c} (\uparrow \text{SUBJ PERS}) = 1 \\ (\uparrow \text{SUBJ NUM}) = \text{SG} \\ \langle\langle @\text{PST} \rangle\rangle \end{array} \right\} \rangle \xrightarrow{\nu} \left[\begin{array}{ll} \text{PHON.REP} & /\text{ʊm}/ \\ \text{PFRAME} & (\ldots(\ \cdot\ )_\sigma)_\omega \\ \text{DEP} & \text{LEFT} \end{array} \right]$

(29) *šarī* *āz* *x̌ac* *biroz-um*
tomorrow I.NOM water drink.PRS-1SG
'Tomorrow I will drink water.'

Neither of these VIs realizes the PRED feature because they are compatible with both overt and pro drop subjects. However, there is one clitic marker, *=af*, which is incom-

patible with overt subjects (see example 15 above); I model this by having its VI also instantiate the PRED feature (30). The analysis of the fossilized combination *uf=af*, which seems to allow an overt pronoun to co-occur with *=af*, is unclear at the moment; further information is needed on the morphosyntactic status of this combination.

(30) $\langle$ [Agr], $\left\{ \begin{array}{c} (\uparrow \text{SUBJ PERS}) = 3 \\ (\uparrow \text{SUBJ NUM}) = \text{PL} \\ (\uparrow \text{SUBJ PRED}) = \text{`pro'} \\ \langle\langle @\text{PST} \rangle\rangle \\ \langle\langle (\uparrow \text{OBJ}) \rangle\rangle \end{array} \right\} \rangle \xrightarrow{\nu} \left[\begin{array}{ll} \text{PHON.REP} & /\text{af}/ \\ \text{PFRAME} & ((\ )(\ \cdot\ )_{\sigma})_{\omega} \\ \text{DEP} & \text{LEFT} \end{array} \right]$

4.5 Zero forms, irregular forms and spanning

There are two phenomena in the behaviour of Bartangi person-number markers that do not have a simple solution in the current version of LrFG: zero realizations of certain person-number combinations and the existence of irregular 3rd person singular present forms.

Recall that there are certain slots in the paradigms of the person-number markers (Tables 3 and 4) that have zero realizations: 3rd person singular in the past/perfect and 2nd person singular in the present. Normally, zero realization is modeled in LrFG through a mechanism called Pac-Man spanning: a VI can co-instantiate an adjacent c-structure node if the latter cannot be realized by any VI in the vocabulary. Pac-Man spanning has not been extensively discussed in the literature, but it seems to be implicitly assumed that it can only occur with nodes that map to the same f-structure; otherwise it would be possible for arbitrary nodes to Pac-Man span their neighbours, making the mechanism too powerful. Thus, in this case, we expect Agr to be Pac-Man spanned with one of the verbal projections; but this is not possible, because none of the verbal projections are adjacent to Agr. Hence, to model zero marking, we will have to resort to rules such as (31), which is an awkward way to model something that should follow from the structure of the paradigm.

(31) $\text{AgrP} \rightarrow \underset{(\uparrow \text{SUBJ})=\downarrow}{\text{XP}} \{ \underset{\uparrow=\downarrow}{\text{Agr}} | \left\{ \begin{array}{l|l} (\uparrow \text{PERS}) = 3 & \varepsilon \\ (\uparrow \text{NUM}) = \text{SG} & (\uparrow \text{PERS}) = 2 \\ @\text{PST} & (\uparrow \text{NUM}) = \text{SG} \\ & @\text{PRS} \end{array} \right\} \} \underset{\uparrow=\downarrow}{\text{TP}}$

Irregular person-number forms in Bartangi are only attested in the present tense (which is fitting, considering that agreement markers are affixal in the present) and, as discussed above, belong to three types: irregular 3rd person singular (such as *kiǩt* 'does' from *kin-* 'do'), contracted forms (*sām* 'I go' for *sāw-um*, *sān* 'we/they go' for *sāw-an*) and final consonant dropping in the imperative (*sā* 'go!' but *sāw* 'thou goest'). The standard way of modelling such inflectional irregularities in LrFG would be through vocabulary spans, but, again, this mechanism is not available because the verbal root, v and Agr are not adjacent.

I suggest that the impossiblity of modelling these forms as irregular is due to a contradiction in the architecture of LrFG. In modelling affixes through host identification by c-command, it abandons the idea that c-structure adjacency is equivalent to linear adjacency of morphemes. Yet the notion of spans remains dependent on linear adjacency,

observing the same restrictions as lexical sharing (Wescoat 2002, Lowe 2016). My proposal is to disconnect spanning from c-structure adjacency and model it on the same principles as host identification, namely, c-command and f-structure mapping. This can be done by adopting the following two principles:

(32) **Spanning**
C-structure nodes that map to the same v-structure (= are spanned) must be in a successive c-command relation and map to the same f-structure. The linear position of the phonological form of a span corresponds to the node that does not asymmetrically c-command any other node in the span (i.e. its **lowest** node).

(33) **Pac-Man spanning**
If a c-structure node n_1 has no suitable realization, then, if there is a node n_2 such that $\phi(n_1) = \phi(n_2)$ and **closest**(n_1,n_2), $\nu(n_1) = \nu(n_2)$.

In other words, spans can be non-adjacent as long as the nodes involved map to the same f-structure and are in a c-command relationship; Pac-Man spanning proceeds downwards, to the closest (Asudeh, Bögel & Siddiqi 2023) node it c-commands that maps to the same f-structure.

These are not meant to be fully formalized definitions; the notion of "linear position of a span" is especially unclear. I propose the following informal generalization to capture the ordering of VIs (which applies to the mapping of v-structures to p-structure):

(34) a. Phonological forms of v-structures lacking the feature DEP (independent VIs) are ordered according to the c-structure order of the lowest nodes instantiated by each v-structure.

b. Phonological forms of v-structures having the feature DEP (affixes and clitics) are ordered relative to the v-structure occupying their HOST feature according to the value of DEP (LEFT, RIGHT).

The intuition behind these preliminary definitions seems clear enough to capture the cases at hand. Zero realization of 2nd person singular present and 3rd person singular past is now possible: since there is no VI in the vocabulary that is compatible with these features, Pac-Man spanning ensures that Agr is mapped to the same v-structure as the T node, as in Figure 6.

Irregular forms can now be defined as vocabulary spans, as in the following VI for the short 1st person singular present form of the verb 'go' (35), which corresponds to the tree in Figure 7. The form is spanned across several nodes, but positioned according to the linear location of the verbal root and v.

(35) $\langle$ [Agr,$\sqrt{}$,v,T], $\left\{\begin{array}{c}(\uparrow \textsc{pred}) = \text{'go'} \\ (\uparrow \textsc{subj pers}) = 1 \\ (\uparrow \textsc{subj num}) = \textsc{sg} \\ @\textsc{prs}\end{array}\right\}\rangle \xrightarrow{\nu}$ *sām*

The extended approach to spanning also allows us to capture the existence of a "zero" realization of the verb 'do'[16] in the present tense, as seen in the following examples:

[16]For Shughni, it has been observed in Parker (2023) with complex predicates only, but Maksim Melenchenko (p. c.) has demonstrated that it is also possible with the lexical verb 'do'.

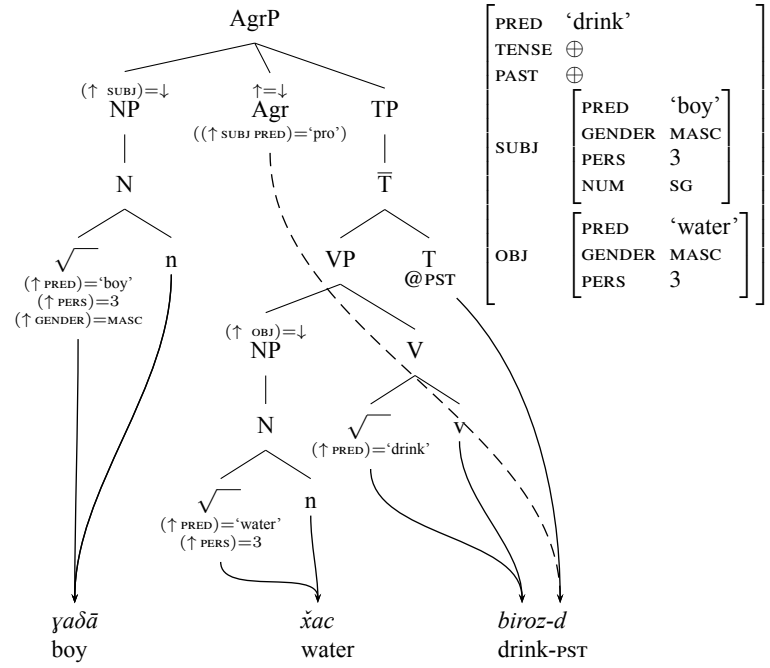

Figure 6: Long-distance Pac-Man spanning of zero Agr

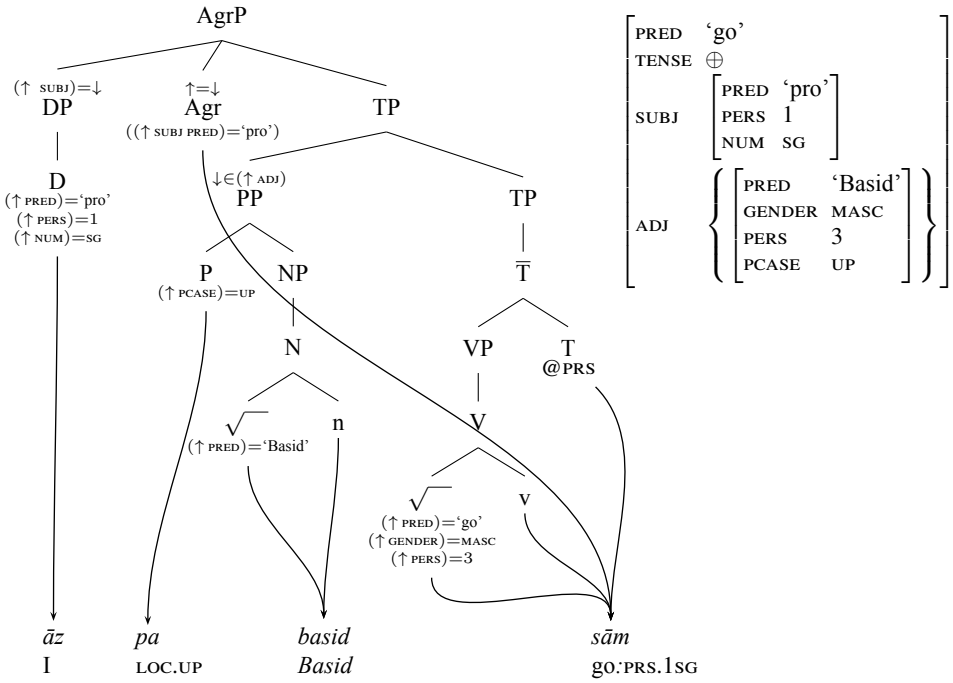

Figure 7: Long-distance vocabulary spanning ('I go to Basid')

(36) *āz* *kor-um* / *kor* *kin-um*
I.NOM work-1SG work do.PRS-1SG
'I am working.'

(37) *āz* *salāt-um*
I.NOM salad-1SG
'I am making a salad.'

(38) *āz* *tar* *xāmb-um* *az* *xu*
I.NOM LOC.EQ shelter-1SG OBJ REFL
'I am hiding myself.'

This construction most commonly occurs with complex predicates where 'do' is a light verb, as in (36), but it can equally well be used when 'do' is used in the lexical sense, as in (37), where 'make a salad' can hardly be considered a complex predicate. Example (38) shows that this "zero" verb is not sensitive to the category of the preceding constituent, here attaching to a noun that forms part of a PP.

Under the approach I propose, this exceptional verb can be accounted for by treating these forms as spans of Agr, $\sqrt{\ }$, v and T that have clitic status, as in (39).

(39) $\langle$ [Agr,$\sqrt{\ }$,v,T], $\left\{\begin{array}{c}(\uparrow \textsc{pred}) = \text{'do'} \\ (\uparrow \textsc{subj pers}) = 1 \\ (\uparrow \textsc{subj num}) = \textsc{sg} \\ @\textsc{pst}\end{array}\right\}\rangle \xrightarrow{\nu} \left[\begin{array}{ll}\textsc{phon.rep} & /\text{ʊm}/ \\ \textsc{pframe} & (\ldots(\ \cdot\ )_{\sigma})_{\omega} \\ \textsc{dep} & \textsc{left}\end{array}\right]$

This correctly predicts that this "zero" version of the verb cannot be used when the corresponding full verb has a zero ending, as in the 2nd person singular (40). If we were to analyze (36)–(38) as a genuine zero realization of the present stem, we would have been able to Pac-Man span this "zero VI" with the Agr lacking an exponent.

(40) *tū* *kor* **(kin)?*
thou.NOM work do.PRS[2SG]
'Are you working?'

The fact that these forms are unavailable for the 3rd person singular (41) also suggests an analysis in terms of spans stored directly in the vocabulary, where a 3rd person clitic form of the verb 'do' is simply unavailable (presumably due to the complex morphonological combinations of *-t* with its host, which would require storing individual host + affix combinations in the vocabulary, as done for irregular 3rd person singular present forms).

(41) *asad* **kor-t* / $^{\text{OK}}$*kor* *kiǐt*
A. work-3SG work do.PRS:3SG
'Asad is working'

5 Conclusions

In this paper, I have provided an LrFG analysis of clause structure in Bartangi. Under this analysis, the language has a three-layer clause structure consisting of a VP that in-

cludes the direct object, TP that includes the subject, and AgrP that includes the clause-initial topic. In the area of clausal agreement, Bartangi demonstrates a mismatch between distribution and linear exponence: the clitic and affix person-number markers are in complementary distribution depending on tense, but have distinct linear positions. In LrFG, their distributional equivalence can be captured by consistently introducing the person-number features in the same c-structure node Agr; the difference in linear order is captured by competition between a clitic that leans to the topical constituent to its left in the past and perfect tenses, and an affix that attaches to the verb stem in the present tense, through LrFG's mechanism of host identification (Asudeh, Bögel & Siddiqi 2023). I have further proposed a modification to LrFG's notion of spanning to account for spans that seemingly instantiate non-adjacent nodes, allowing the modelling of irregular present-tense forms, including apparent "zero" realizations of the present stem of the verb 'do'. Whether this modification proves to be too powerful is to be decided in future work.

There are further features of Bartangi clause structure that require an explanation in future work. Bartangi allows scrambling to non-topical positions and postverbal arguments, as seen in (4), but the status of such elements is not accounted for under the current analysis. Whether arguments beyond subjects and direct objects have default structural positions for realization is also an open question. Bartangi also has a system of subordinating and coordinating conjunctions with a large number of distinct linear positions (preverbal, clause-initial, clause-final, second; see Belyaev 2025); these have to be integrated into the clause structure proposed in this paper.

List of glossing abbreviations

CONJ: conjunction; D3: distal; F: feminine; INS: instrumental; LOC.DOWN: preposition denoting location below speaker; LOC.EQ: preposition denoting location at equal level with speaker; LOC.UP: preposition denoting location above speaker; M: masculine; NOM: nominative; OBJ: object marker; OBL: oblique; PST: past; REFL: reflexive; SG: singular; TR: transitive.

References

Asudeh, Ash, Bronwyn Bjorkman, Frances Dowle, Neil Myler, Daniel Siddiqi & Lisa Sullivan. 2024. Fusional morphology, metasyncretism, and secondary exponence: A morphemic, realizational approach to Latin declension. In Miriam Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *Proceedings of the LFG'24 conference*, 45–68. Konstanz: PubliKon. https://lfg-proceedings.org/lfg/index.php/main/article/view/49 (24 December, 2025).

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. 2022. Realizational morphosemantics in LrFG. In Miriam Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *Proceedings of the LFG'22 conference*, 21–40. Konstanz: University of Konstanz. https://lfg-proceedings.org/lfg/index.php/main/article/view/20 (2 October, 2025).

Austin, Peter & Joan Bresnan. 1996. Non-configurationality in Australian Aboriginal languages. *Natural Language and Linguistic Theory* 14(2). 215–268. https://doi.org/10.1007/BF00133684.

Belyaev, Oleg. 2022. Phrase structure and configurationality in Ossetic. In Miriam Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *Proceedings of the LFG'22 conference*. Konstanz: University of Konstanz. https://lfg-proceedings.org/lfg/index.php/main/article/view/26 (2 October, 2025).

Belyaev, Oleg I. 2025. Pozicionnye klassy sojuzov v bartangskom jazyke [Position classes of conjunctions in Bartangi]. *Lomonosov Philology Journal* 2. 70–87. https://doi.org/10.55959/MSU0130-0075-9-2025-48-02-6.

Bresnan, Joan. 1998. Morphology competes with syntax: Explaining typological variation in weak crossover effects. In Pilar Barbosa, Danny Fox, Paul Hagstrom, Martha McGinnis & David Pesetsky (eds.), *Is the best good enough? Optimality and competition in syntax*, 59–92. Cambridge, MA: MIT Press & MIT Working Papers in Linguistics.

Chistyakova, Daria. 2025. Split intransitivity in Shughni and Bartangi: A study in lexical microvariation. Talk at the Tenth International Conference on Iranian Linguistics (ICIL10), Viterbo.

Edelman, Joy I. 1990. *Sravnitel'naja grammatika vostočnoiranskix jazykov. morfologija. èlementy sintaksisa* [Comparative grammar of Eastern Iranian languages: Morphology, elements of syntax]. Moscow: Nauka.

Edelman, Joy I. & Šodixon P. Jusufbekov. 1999. Šugnanskij jazyk [Shughni]. In Vera S. Rastorgueva, Joy I. Edelman & Vjačeslav V. Moškalo (eds.), *Jazyki mira: Iranskie jazyki* [Languages of the World: Iranian languages]. Vol. III: *Vostočnoiranskie jazyki* [Eastern Iranian languages], 225–242. Moscow: Indrik.

Karamšoev, Dodxudo. 1978. *Kategorija roda v pamirskix jazykax (šugnano-rušanskaja gruppa)* [The category of gender in the Pamir languages (Shughni-Roshani branch)]. Dushanbe: Doniš.

Karamxudoev, Nodir. 1973. *Bartangskij jazyk (fonetika i morfologija)* [Bartangi language (phonetics and morphology)]. Dushanbe: Doniš.

Laczkó, Tibor. 2014. Essentials of an LFG analysis of Hungarian finite sentences. In Miriam Butt & Tracy Holloway King (eds.), *Proceedings of the LFG'14 conference*. Stanford, CA: CSLI Publications. https://typo.uni-konstanz.de/lfg-proceedings/LFGprocCSLI/LFG2014/papers/lfg14laczko2.pdf (13 September, 2020).

Lowe, John J. 2016. English possessive 's: clitic *and* affix. *Natural Language & Linguistic Theory* 34. 157–195. https://doi.org/10.1007/s11049-015-9300-1.

Makarov, Yury. 2024. Shughni morphophonology from a synchronic perspective. Talk at the Research Seminar on Iranian languages, Moscow, 22 May. https://pamiri.online/static/uploads/makarov2024.pdf (2 October, 2025).

Makarov, Yury Yu. & Vladimir A. Plungian. 2023. K opisaniju šugnanskoj morfonologii: predvaritel'nye zamečanija [Towards a description of Shughni morphophonology: Preliminary remarks]. In Irina M. Kobozeva, Xenia P. Semyonova,

Natal'ja D. Svetozarova, Leonid M. Zaxarov & Alexander E. Kostyuk (eds.), *«…Vperëd i vverx po lestnice zvučaščej». Sbornik statej k 80-letiju Ol'gi Fëdorovny Krivnovoj* [A collection of papers for Olga Krivnova's 80th anniversary], 178–193. Moscow: Buki-Vedi. https://doi.org/10.5281/zenodo.16364199.

Melenchenko, Maxim G. 2023. Èvidencial'nyj perfekt v šugnanskom jazyke. [The evidential perfect in Shughni]. *Voprosy jazykoznanija* 6. 120–137. https://doi.org/10.31857/0373-658X.2023.6.120-137.

Muravyeva, Irina A. 1975. Nepozicionnye čeredovanija fonem v šugnanskom jazyke [Non-positional alternations of phonemes in Shughni]. In Vladimir A. Zvegincev (ed.), *Issledovanija po strukturnoj i prikladnoj lingvistike* [Studies in structural and applied linguistics], 124–135. Moscow: Izdatel'stvo MGU.

Parker, Clinton. *A grammar of the Shughni language*. Ph.D. dissertation, McGill University. https://escholarship.mcgill.ca/concern/theses/0v8385834 (22 December, 2025).

Payne, John & Erika Chisarik. 2000. Negation and focus in Hungarian: An Optimality Theory account. *Transactions of the Philological Society* 98. 185–230. https://doi.org/10.1111/1467-968X.00062.

Serdobolskaya, Natalia & Oleg Belyaev. 2022. Dative external possessors in Ossetic. In Miriam Butt, Jamie Y. Findlay & Ida Toivonen (eds.), *Proceedings of the LFG'22 conference*, 325–344. Konstanz: University of Konstanz. https://lfg-proceedings.org/lfg/index.php/main/article/view/25 (2 October, 2025).

Sergienko, Alexander. 2025. Alignment shift in Bartangi. Talk at the Tenth International Conference on Iranian Linguistics (ICIL10), Viterbo.

Snijders, Liselotte. *The nature of configurationality in LFG*. Ph.D. dissertation, University of Oxford. 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.

Stewart, Tom & Gregory Stump. 2007. Paradigm Function Morphology and the morphology/syntax interface. In Gillian Ramchand & Charles Reiss (eds.), *The Oxford handbook of linguistic interfaces*, 383–421. Oxford: Oxford University Press.

Stump, Gregory & Andrew Hippisley. 2011. Valence sensitivity in Pamirian past-tense inflection: A realizational analysis. In Agnes Korn, Geoffrey Haig, Simin Karimi & Pollet Samvelian (eds.), *Topics in Iranian linguistics*, 103–115. Wiesbaden: Dr. Ludwig Reichert Verlag. http://titus.uni-frankfurt.de/personal/agnes/paris09/stump_hippisley2011.pdf (2 October, 2025).

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).

Wescoat, Michael T. *On lexical sharing*. Ph.D. dissertation, Stanford University. https://searchworks.stanford.edu/view/4814040 (22 December, 2025).