Phrase structure and configurationality in Ossetic

Oleg Belyaev

(Interdisciplinary Scientific and Educational School of Lomonosov Moscow State University "Preservation of the World Cultural and Historical Heritage"; Institute of Linguistics RAS)

Abstract

In this paper, I analyze Ossetic clause-level word order in terms of LFG. Ossetic word order is not constrained by the grammatical functions of elements, but by their formal features and information-structure functions. Specifically, interrogatives, negative pronouns, most subordinators and foci have to be positioned in the preverbal position. I argue that this pattern cannot be captured by the standard model of $\overline{\text{X}}$ theory because the limit to two bar levels does not provide enough "slots" for all these elements, and the notions of "complement", "adjunct" and "specifier" themselves become devoid of meaning. Instead, I propose that Ossetic has to be analyzed as having more than two bar levels; the notion of projection thus becomes similar to the notion of "slot" in template-based inflectional morphology. In effect, this solution suggests a view of clause structure as a language-specific phenomenon that develops through processes akin to grammaticalization in the domain of morphology.

1 Introduction

Ossetic is an Iranian language spoken in the Caucasus.[1] Like all neighbouring languages and most modern Iranian languages, the dominant word order in Ossetic is SOV. All other word order variants are attested as well, and there is no evidence in favour of a configurational structure associated with specific grammatical functions, like in Englihs. However, Ossetic does severely constrain the linear position of a number of elements, namely preverbal negative pronouns, interrogatives and foci, and initial subordinators (C) and material that may precede these subordinators.

In this paper, I will argue that Ossetic word order is best treated in terms of discourse configurationality: preverbal material is organized in an $\overline{\text{X}}$-theoretic VP constituent, while the left periphery is positioned in a top-level CP constituent; a non-configurational S is sandwiched between these two strata. In terms of $\overline{\text{X}}$ theory, the number of "slots" that have distinct positioning rules is higher than that allowed by the traditional two-level $\overline{\text{X}}$ schema. Thus I argue that some languages, including Ossetic, may require more than two bar levels – something which is theoreticall allowed by modern versions of $\overline{\text{X}}$ theory, such as Lowe and Lovestrand (2020), but has not, to my knowledge, been proposed in the literature. In a wider perspective, such a solution seems to eliminate any theoretical significance

---

of the notions "specifier" and "complement", instead treating the $\overline{\mathrm{X}}$ theory as a kind of syntactic "Item-and-Arrangement" model.

The article is organized as follows. In section 2, I provide a description of word order in Ossetic clauses, characterizing the constraints on the distribution of the key elements which have fixed ordering: complex predicate components, negative pronouns, interrogatives and other preverbal foci, and subordinators. In 3, I interpret these data in terms of LFG, showing that the classical version of $\overline{\mathrm{X}}$ theory is inadequate for capturing the full complexity of Ossetic word order. I conclude that for Ossetic, the restriction to only two bar levels, and consequently, the distinction between complements and specifiers, should be abandoned in favour of a model of endocentricity where bar levels are more similar to linear slots in template-based models of morphology.

2 Word order in Ossetic

2.1 Overview

While Ossetic is described as an SOV language, all theoretically possible orders are grammatical, as shown in (1).

(1) a. *žawər alan-ə   fet:-a*
       Zaur   Alan-GEN see.PFV-PST.3SG
       'Zaur saw Alan.'

   b. *žawər fedta alanə*

   c. *alanə žawər fedta*

   d. *alanə fedta žawər*

But Ossetic is far from being a nonconfigurational language in the way, for example, some Australian languages are Austin and Bresnan (1996). It is more correctly described as discourse configurational with a grammaticalized preverbal area, with the closest typological counterparts being languages like Hungarian (É. Kiss 2004; Laczkó 2014).

In purely descriptive terms, the general organization of the Ossetic clause is summarized in (2). It can be divided into three fields: the verbal complex, i.e. the verb and the elements directly preceding it, which involves the most rigid linear constraints; the left periphery, i.e. the complementizer (Comp) and any preceding material; and the central area, i.e. elements between the left periphery and the verbal complex or to the right of the verbal complex – the least grammaticalized part of the clause, where constituents occur in relatively unconstrained order.

(2)   PreComp − Comp − PostComp − $\overbrace{\text{Foc} - \text{Wh} - \text{Adv} - \text{Neg} - \text{V}}^{\text{verbal complex}}$ −PostV
      $\underbrace{\phantom{\text{PostComp} - \text{Foc} - \text{Wh} - \text{Adv} - \text{Neg} - \text{V} - \text{PostV}}}_{\text{central area}}$

I will now demonstrate the ordering requiremenets one by one, starting from the bottom up.

2.2 The verbal complex

2.2.1 Complex predicates

The combination of a nonverbal component (NVC) and the light verb is the most tightly bound grouping in the Ossetic clause. Preverbs normally attach to the NVC, not to the lexical verb: *ləg kɜn-ən* (cut do-INF) → ***a-*ləg *kɜn-ən*** (PV-cut do-INF). These complexes can only be split by second-position clitics; wh-words, which are also preverbal, cannot occur between the NVC and the light verb:

(3) *žawər* ‹^OK ***kɜ-imɜ*** › *ba-nəχaš* ‹* ***kɜ-imɜ*** › *kot:-a?*
   Z.   who-COM PV-speech   do-PST.3SG
   '**With whom** did Zaur speak?'

(4) *a-ləg* *=ɜj* *kot:-on*
   PV-cut it.GEN do-PST.1SG
   'I cut it off.'

Complex predicates without a preverb can behave both as a single unit or as the combination of a finite verb and a noun, i.e. the interrogative can be found to the left of NVC or between NVC and the light verb:

(5) *žawər* ^OK***kɜ-imɜ*** *nəχaš* ^OK***kɜ-imɜ*** *kot:-a?*
   Zaur   who-COM speech   do-PST.3SG
   '**With whom** was Zaur speaking?'

2.2.2 Negative pronouns

The next elements linearly closest to the verb are negative pronouns (6) or negation markers (7), which are in complementary distribution: standard Ossetic has no double negation.

(6) a. *žawər-ə* ***ni-či*** (**n**ɜ) *žon-ə*
      Z.-GEN   NEG-who   NEG know-PRS.3SG
      'No one knows Zaur.'

   b. ***niči*** *žonə žawərə*

   c. * ***niči*** *žawərə žonə*

   d. * *žawərə žonə* ***niči***

(7) å*žawər-ə* ***nɜ*** *žon-ən*
    Z.-GEN   NEG know-PRS.1SG
    'I do **not** know Zaur.'

3

a. * nʒ žawərə žonən

b. * žawərə žonən nʒ

When negative pronouns occur clause-initially, second-position clitics intervene between them and the verb, just as with complex predicates (8). But there are also two clitics which are second-position in affirmative sentences, but become preverbal in negative sentences: *=wal* 'more' and *=ma* 'yet' (9).

(8) *ni-sə* **=jən** *rat:-on*
NEG-what he.DAT give.PFV-PST.1SG

'I gave **him** nothing.'

(9) *žawər =ən* *ni-sə* **=ma** / **=wal** *rat:-a*
Z. he.DAT NEG-what yet more give.PFV-PST.3SG

'Zaur didn't give him anything **yet** / **more**.'

2.2.3 Preverbal foci

Interrogatives in Ossetic must occur preverbally, and narrow foci usually do as well, as seen from the question-answer pair in (10).

(10) A. *zul* [**či**]_FOC *ba-lχʒt:-a?*
bread who PV-buy-PST.3SG

'**Who** bought bread?'

B. *zul* [**alan**]_FOC *ba-lχʒt:-a*
bread alan PV-buy-PST.3SG

'**Alan** bought bread.'

While Interrogatives are preverbal, they must precede negation and negative pronouns (and, by extension, NVCs of complex predicates), see (11).

(11) a. **či** **ni-sə** *žon-ə?*
who NEG-what know-PRS.3SG

'**Who** knows **nothing**?'

b. * **ni-sə či** *žon-ə?*

This also applies to preverbal subordinators (12), all of which are either synchronically or historically identical to interrogatives, see Erschler (2012).

(12) *kʷəd =zə š-fidar kʒn-ʒn iš iron ʒvžag,* **kʷə ni-či** *=jəl*
how it.IN PV-firm do-NMLZ is Ossetian language if NEG-who it.SUPER

nəχaš kʒn-a, *wʒd?*
speech do-SBJV.3SG then

'How is the Ossetic language to be strengthened **if no one** speaks it?'

(ONC)

4

As for the relative positioning of interrogative and non-interrogative narrow foci, if the clause contains both a focused NP and an interrogative, the former must precede the latter or otherwise be postverbal:

(13) a. **či** *fet:-a* **ɜrmɜšt žawər-ə**?
         who see.PFV-PST.3SG only Zaur-GEN
         '**Who** saw **only Zaur**?'

     b. **ɜrmɜšt žawərə či** *fet:a?*

     c. * **či ɜrmɜšt žawərə** *fedta?*

Finally, mysteriously enough, adverbs in the comparative degree can also (optionally) intrude after the interrogative:

(14) *žawər-ɜj* ⟨^OK**rɜvz-dɜr**⟩ *ni-či* ⟨*****rɜvz-dɜr**⟩ *žʁor-ə.*
     Zaur-ABL fast-COMPAR NEG-who run-PRS.3SG
     'No one runs **faster** than Zaur.'

(15) *žawər-ɜj* **či** **rɜvz-dɜr** *žʁor-ə?*
     Zaur-ABL who fast-COMPAR run-PRS.3SG
     '**Who** runs **faster** than Zaur?'

This applies not only to adverbs, but also to NPs containing comparative adjectives:

(16) *səmɜ* *nɜ= qɜw-ə* **či** **χʷəždɜr** *lɜg* *u?*
     I.wonder our village-IN who best man be.PRS.3SG
     'I wonder who is the **best** man in our village?'   (ONC)

The adverbs do not seem occupy the same position as NVC of complex predicates, because negative phrases and the verb cannot be broken up in this way:

(17) *****žawər-ɜj* *ni-či* **rɜvz-dɜr** *žʁor-ə*
     Z.-ABL NEG-who.NOM fast-ADD run-PRS.3SG
     ('No one runs faster than Zaur.')

2.2.4   Summary

The detailed structure of the verbal complex is summarized in (18).

(18)   FOC – WH – ADV – NEG – NVC – V

2.3   The central area

Outside the verbal complex, word order is much less constrained. The only constraint on material that precedes the verbal complex is that it cannot be focal (19b).

(19) a. [*žawər*]_{TOP} [**alan-ɜn**]_{FOC} *rat:-a* [**ɜχsa**]_{FOC}
        Zaur   Alan-DAT   give.PFV-PST.3SG   money
        (What did Zaur give to whom?) 'Zaur gave Alan **money**.'

    b. # [**alanɜn**]_{FOC} [*žawər*]_{TOP} *rat:a* [**ɜχsa**]_{FOC}

Postverbal material can be both topical and focal, in any order. Postverbal focus is illustrated in (13a). A topical postverbal NP is illustrated in (20).

(20) *ɜmɜ sə   fɜ-w-a,   sə   kɜn-a   [**lɜg**]_{TOP}?..*
     and what PV-be-SBJV.3SG what do-SBJV.3SG   man
     'And who should **the man** be, what shall (he) do?..'   (ONC)

2.4 The left periphery

Most subordinators in Ossetic are preverbal (PV) and occupy the same position as interrogatives. But there are five subordinators which tend to be initial (nPV): *sɜmɜj* 'in order that', *salənmɜ* 'while', *kɜd* 'if', *səma* 'as if', *iwgɜr* 'if, since'. These can technically be located anywhere before the verb (up to the interrogative preverbal position), but in Belyaev (2014) it has been argued that they actually have two positioning variants: in C^0 or in the standard interrogative preverbal position. For example, the grammaticality of (21) degrades the farther the subordinator is positioned after the second position, but improves to normal when it is preverbal.

(21) ‹**sɜmɜj**› *žawər* ‹**sɜmɜj**› *rajšom*   ‹$^?$**sɜmɜj**› *alan-ɜn* ‹$^{??}$**sɜmɜj**› *ɜχsa*
     PURP   Z.   tomorrow   A.-DAT   money

     ‹**sɜmɜj**› *rat:-a*   ‹***sɜmɜj**›
            give-SBJV.3SG

     '**so that** Zaur gives money to Alan tomorrow'

Corpus frequencies also confirm this generalization: in the vast majority of examples (even correcting for the length of the clause), nPV subordinators are initial, second-position, or preverbal; in fact, there is only one true example of a "third position" nPV subordinator in the sample of Belyaev (2014).

If is is true that these subordinators should be placed in a fixed position in C, whatever is to the left should be treated as located in the specifier or adjunct of CP. What purpose this dislocation serves in Osseic is unclear: it does not seem to be straightforwardly associated with topic status, for example. This question is outside the scope of this paper; what is important here is to establish the syntactic structure of the Ossetic clause, not the semantic or information-structure properties of its constituents as such.

3 Analysis

3.1 Overview

From the discussion above, it is clear that linear order in Ossetic does not encode grammatical functions. Therefore, following standard LFG logic, a traditional English-style configurational structure cannot be assumed. We are left with the following possibilities:

1. Fully flat S, with the order of elements defined via regular expressions.

2. A hierarchical clause structure, of which there can be the following variants

   a **VP – S – CP**: a three-level clause where VP is the verbal complex, CP is the left periphery, while S contains the rest of the material "sandwiched" between them. A possible problem with this approach is that the standard $\overline{\text{X}}$ model leaves us with less than enough dedicated positions for the preverbal elements whose position is thus constrained.

   b **VP – IP – S – CP**: same as above, but with an additional IP layear. This leaves enough $\overline{\text{X}}$-theoretic positions but requires motivation for lexical material to be present in I.

   c Unlimited $\overline{\text{V}}$ adjunction in the style of Japanese and Korean in Sells (1994, 1995). In this case, linear order should be constrained in some other way.

   d A cartographic approach in the style of **rizzi1977** and subsequent work, with a host of separate projections (AspP, TenseP, AgrP, TopP, FocP, etc.). This is formally compatible with LFG architecture but is not typically applied, because projections are only stipulated when there is lexical material to fill them. L_RFG (Melchin, Asudeh, and Siddiqi 2020) introduces projections for inflectional features, but this is not completely equivalent to a full cartographic structure, and the specifier positions of these projections are not immediately identifiable with information-structure functions. Therefore, I will not consider cartography here, as such an analysis requires too many special assumptions to work in LFG.

In what follows I will discuss the relative merit of the first four alternatives in detail.

3.2 Linear schema vs. VP/S/CP

As stated above, the simplest analysis is to assume a completely flat clause structure:

(22)   S → $\text{XP}_{\text{"TOP"}}* \text{C XP}* \text{XP}_{\text{FOC}}* \text{XP}_{\text{COMPAR}}* \text{XP}_{\text{NEG}}*$ (Neg) XP_{PRED} V XP$*$

This is not an adequate analysis for at least two reasons. First, the verbal complex forms a prosodic word, or at least a prosodic phrase, with the verb (Abaev

1949). The exact acoustic characterization of the verbal complex has not been investigated, but the fact that immediately preverbal constituents are more tightly bound with the verb than other constituents is demonstrated by the fact that the initial ɜ- vowel of the verb is elided if the preceding element ends in a vowel (23). The same process occurs within noun phrases, which also form tight prosodic units (24), but does not occur across noun phrases (25).

(23)   *sə*   *(ɜ)r-səd-i?*
      what   PV-go-PST.3SG
      'What happened?'

(24)   *žawər-ə*   *(ɜ)fšəmɜr*
      Zaur-GEN   brother
      'Zaur's brother'

(25)   *žawər-ə* **(ɜ)fšad*   *ba-jjɜft-a*
      Z.-GEN   army
      'The army caught up with Zaur.'

Such prosodic facts cannot be taken as definite proof of constituency, because it is known that mismatches between prosody and syntax do occur. But, other things being equal, treating the verbal complex as a single constituent leads to a simpler mapping from c-structure to prosody, and should hence be preferred. However, Ossetic prosody is still in need of a separate, detailed study before definitive conclusions can be drawn.

But prosody is not the only motivation behind treating the left periphery and the VP as configurational. Another piece of evidence is coordination. One or more clauses can be coordinated under a single nVP subordinator; material to the left of the subordinator, if any, is shared across clauses (26).

(26)   [CP *sɜlək:* **sɜmɜj** [S *amond-žən*   *fɜndag-əl*   *ba-ft-a*]   *ɜmɜ* [S
      Ts.   PURP   happy-PROP   way-SUPER   PV-fall-SBJV.3SG   and

      iš-kʷə   *aχʷər-mɜ*   *fɜndag ššar-a*]],   **wəj**   **təχχɜj**
      INDEF-when   learning-ALL   way   find.PFV-SBJV.3SG   that[GEN]   for

      '**So that** Tsælykk falls upon a happy path and one day finds a road to learning.'   (ONC)

In contrast, all preverbal elements have to be repeated if more than one verb form is used (27). Scoping an interrogative over a coordinate phrase consisting of coordinated V + NP pairs is ungrammatical (28)–(29).

(27)   *...ɜmɜ ɜž kɜsɜj*   *žon-ən,*   *məšajnag* [VP *či*   *rat:-a*]
      and   I   from.where   know-PRS.1SG   donation   who   give.PFV-PST.3SG

      ɜmɜ [VP *či*   *nɜ*   *rat:-a*],   *wəj.*
      and   who   NEG   give.PFV-PST.3SG   that.DEM

      '...and how do I know who gave the donation and who didn't?'   (ONC)

(28) *__kɜm-ɜn__ rat:-a   žawər činəg ɜmɜ ba-lɜvar kot:-a   alan
    who-DAT give.PFV-PST.3SG Z.   book   and PV-gift   do-PST.3SG A.
    tetrɜd?
    workbook

    ('**To whom** did Zaur give a book, and Alan present a workbook?')

(29) *__kɜj__   fet:-a   alan ɜmɜ fe-qʷəšt-a   soslan?
    who.GEN see.PFV-PST.3SG A.   and PV-hear-PST.3SG S.

    ('**Who** did Alan see and Soslan hear?')

Were the clause a flat S, there would be no asymmetries between coordination patterns: either all clause subconstituents could be coordinated, or none. Therefore, we need at least a CP projection to explain why C can scope over coordinated constituents to its right, and a VP projection to capture the inseparability of the verbal complex. The material that is sandwiched between these projections can be assumed to be contained in a flat S, because the coordination of constituents containing such material is unconstrained: constituents to the left of the VP can be coordinated, sharing a right-adjacent constituent (30)–([ex:coord-lefts2]), or constituents to the right of the VP can be coordinated, sharing a left-adjacent constituent (32). I am aware of no evidence that could demonstrate that any of these constructions involve ellipsis in the form of Right Node Raising or a similar pattern. In a flat structure, this symmetricity of coordination can be accounted for by the finite-state mechanism of non-constituent coordination proposed in Maxwell and Manning (1996).

(30) *žawər-ə fɜnd-ə,*   [$_{CP}$ *sɜmɜj* [*alan a-nəχaš   kɜn-a*
    Z.-GEN   want-PRS.3SG   PURP   A.   PV-speech do-SBJV.3SG

    ažɜmɜt-imɜ$_i$] *ɜmɜ* [*šošlan š-sɜt:ɜ   kɜn-a   asɜmɜž-ɜn$_j$*
    A.-COMIT   and   S.   PV-ready   do-SBJV.3SG   A.-DAT

    televizor] *šɜ=$_{i+j}$ xɜzar-ə*].
    TV.set   their house-IN

    'Zaur wants **so that** [Alan speeks to Azæmæt$_i$] and [Soslan repairs for Atsæmæz$_j$ the TV] in their$_{i+j}$ house.'

(31) *žawər* [*alan-ɜn   ba-lɜvar   kot:-a*] *ɜmɜ* [*šošlan-ɜn   a-wɜj*
    Zaur   Alan-DAT PV-present do-PST.3SG   and   Soslan-DAT PV-sale

    kot:-a] *čingʷə-tɜ*
    do-PST.3SG book-PL

    'Alan presented the books to Alan and sold (them) to Soslan.'

(32) *mɜn*   *fɜnd-ə,*   [CP *sɜmɜj* *wə-sə*   *mašinɜ* [*š-sɜtːɜ*
    me.GEN   want-PRS.3SG   PURP   that-ATTR   car   PV-prepared

     kɜn-a   *žawər*] *ɜmɜ* [*a-wɜj* *kɜn-a*   *alan*]]
     do-SBJV.3SG Zaur   and   PV-sale do-SBJV.3SG Alan

    'I want [Zaur to repair] and [Alan to sell] this car.'

This leads us to adopt a c-structure for the Ossetic clause that roughly corresponds to (33).

(33)
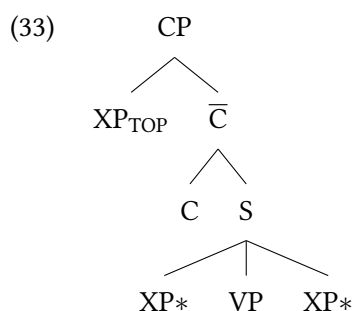
3.3 The structure of VP

The next question to consider is how to model the internal structure of the verbal complex, which we have established to be the VP. At least five distinct "slots" for preverbal material are required:

1. nonverbal components of complex predicates;

2. negative indefinites and negation;

3. comparative adverbs and NPs;

4. interrogative pronouns and complementizers;

5. focal non-wh NPs.

Once again, remaining within standard $\overline{\text{X}}$ theory, there are two options: either VP is assumed to be flat, or the elements should be distributed among the $\overline{\text{X}}$-theoretic positions in some way.
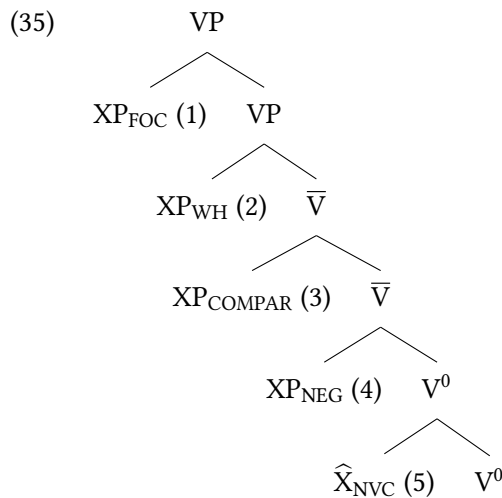
3.3.1 Flat structure

A flat VP structure is not an option either, because coordination facts suggest that each class of elements is attached at its own level of projection, and coordination cannot violate the "levels" of the verbal complex:

(34) a. *či* [*ni-kɜj* *warž-ə*] *ɜmɜ* [*ni-kɜ-wəl*
   who NEG-who.GEN love-PRS.3SG and NEG-who-SUPER

   ɜwwɜnd-ə]?
   trust-PRS.3SG

   'Who loves no one and trusts in no one?'

   b. *[*či* *ni-kɜj*] *ɜmɜ* [*kɜj* *ni-či*] *warž-ə?*
   who NEG-who.GEN and WHO.GEN NEG-who love-PRS.3SG

   ('Who loves no one and is loved by no one?')

If both negative indefinites and interrogatives were VP complements, both would be non-constituent coordination, and there should be no difference, just as we have seen with S-level constituents above.

3.3.2 $\overline{\text{X}}$ schema

The only way to fit all this positions into the standard $\overline{\text{X}}$ schema for VP is to use adjunction to the largest extent and assume that NVCs of complex predicates are non-projecting words, as shown in the tree in (35).

(35)

```
                VP
              /    \
      XP_FOC (1)    VP
                  /    \
           XP_WH (2)    V̄
                      /    \
              XP_COMPAR (3)  V̄
                            /    \
                    XP_NEG (4)    V⁰
                                /    \
                          X̂_NVC (5)   V⁰
```

This works for the data at hand, but turns $\overline{\text{X}}$ theory into nothing else but a system for labeling nodes in endocentric structures, and not a very convenient one at that. The distinction between complements, specifiers, and adjuncts, in particular, becomes completely blurred: most of the elements in question are iterable, and all except NVCs can be associated with any GF (argument or adjunct), but only $\text{XP}_{\text{COMPAR}}$ and NP_{FOC} are treated as adjuncts, for no other reason but to fit five elements into the $\overline{\text{X}}$ schema with three bar levels in total.

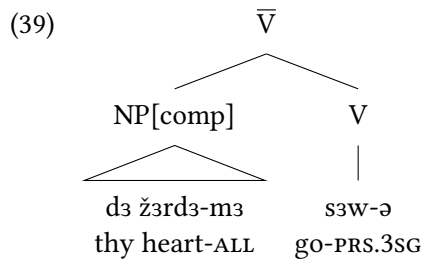Futhermore, the tree in (35) assumes that NVCs of complex predicates are treated as non-projecting words, as in (36).

(36)

```
              V⁰
            /    \
          N̂       V⁰
          |        |
        a-ləg     kot:-a
        PV-cut   do-PST.3SG
```
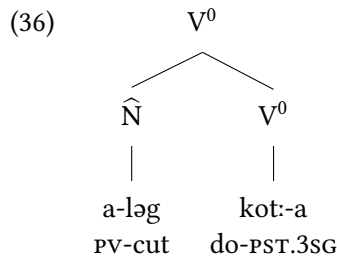
This is indeed plausible for most NVCs, which cannot be phrasal:

(37) a. ^{OK}*zul-ə*   *a-ləg*
          bread-GEN PV-cutting
          'the cutting of bread'

  b. *zul-ə   *a-ləg*   *kot:-on*
        bread-GEN PV-cutting do-PST.1SG
        ('I cut the bread.', lit. 'I did the cutting of bread.')
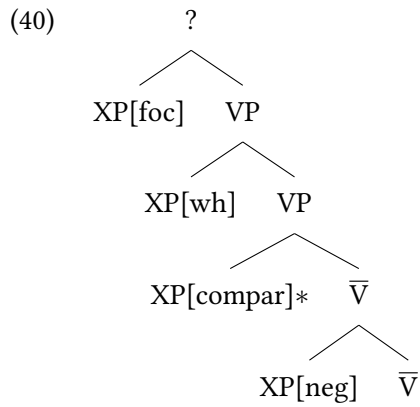
However, some complex verbs, in certain contexts, can have a phrasal NVC, and it is located exactly where expected in the schema:

(38) **sə**   [*dʒ= žɜrdʒ-mɜ*] *sɜw-ə?*
      what   thy heart-ALL   go-PRS.3SG
      'What do you like?'

Therefore, we need a full-fledged VP complement position to accommodate such NVCs, which should be treated as in (39).

(39)

```
                 V̄
               /    \
          NP[comp]    V
           /  \        |
        dʒ  žɜrdʒ-mɜ   sɜw-ə
       thy heart-ALL  go-PRS.3SG
```
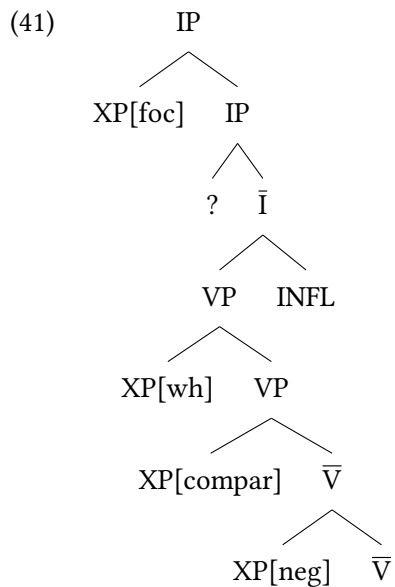
Therefore, the tree in (35) cannot be correct. Since bar level 1 is occupied by a complex predicate slot, we now have four **iterable** classes of elements (negative pronouns, comparative NPs, wh-words, other foci) that have to be distributed over **three** remaining positions: V̄ adjunct, VP adjunct, and Spec-VP. Even if multiple specifiers are allowed (disregarding that SpecVP is a strange position for comparative adverbs), this does not fit into the schema: there is no available label for the uppermost node in (40), unless one resorts to parameterized rules for different types of adjuncts.

(40)

                         ?
              ┌──────────┴──────────┐
          XP[foc]   VP
                          ┌─────────┴─────────┐
                      XP[wh]   VP
                                    ┌──────────┴──────────┐
                               XP[compar]* $\overline{\text{V}}$
                                                ┌────────┴────────┐
                                            XP[neg] $\overline{\text{V}}$

3.3.3 IP/VP split?

Another possibility is to introduced an IP layer into the clause structure, which gives us more than enough positions for all elements (41).

(41)

                    IP
         ┌──────────┴──────────┐
     XP[foc]   IP
                      ┌────────┴────────┐
                    ? $\overline{\text{I}}$
                           ┌──────────┴──────────┐
                         VP   INFL
                ┌──────────┴──────────┐
            XP[wh]   VP
                          ┌──────────┴──────────┐
                     XP[compar] $\overline{\text{V}}$
                                      ┌────────┴────────┐
                                  XP[neg] $\overline{\text{V}}$

But in LFG, constituents like IP cannot be taken for granted: it should first be demonstrated that the head node I can be filled by any lexical material. In fact, there is no evidence for IP in Ossetic. There are no periphrastic TAM paradigms. All periphrastic constructions that do exist, e.g. the periphrastic passive with *sɜw-* 'go' (42) and the periphrastic resultative with *wə-* 'be' (43) do not obey any fixed ordering.

(42) *šk'ola-studi* «*bonvɜrnon* **arɜšt** **ɜr-səd** 2001 *až-ə*
school-studio morning_star make.PTCP PV-go[PST.3SG] year-IN

cχinval-ə ...
Tskhinval-IN

'The school-studio "Morning Star" was created in 2001 in Tskhinval ...'

(ONC)

(43) 31 *dekabr-ɜj* =*nɜm* *abon-ə* *ong* *iw* *fədrakɜnd* =*dɜr* *ne*
december-ABL we.ALL today-GEN until one crime ADD NEG

r-səd **arɜšt**
PV-go[PST.3SG] make.PTCP

'Not one crime **was committed** here (lit. "at us") from December 31st until today.'

(ONC)

Therefore, adding I as a fixed position for the auxiliary verb, or assuming that the finite verb is sometimes found in I, does not lead to any helpful generalizations. Stipulation of IP in Ossetic is redundant.
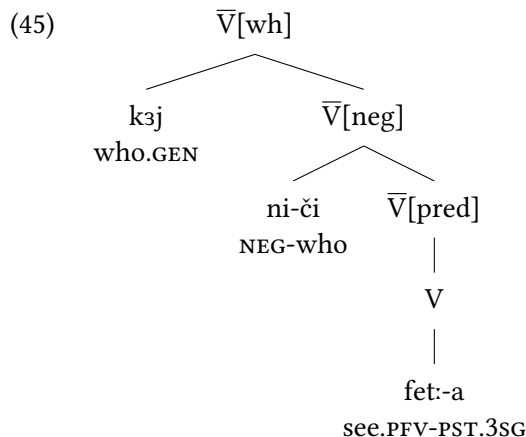
3.3.4 Sells-style adjunction?

Another option is to adopt Sells' analysis of Japanese and Korean (Sells 1994, 1995). Sells assumes that $\overline{V}$ is the maximal projection. All verbal dependents are adjoined to $\overline{V}$; the Ossetic version of Sells' analysis would look like in (44).

(44)
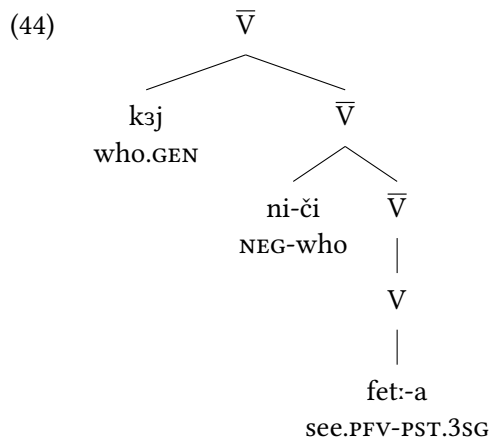
Unlike the flat structure analysis, this correctly captures the coordination facts because binary branching is used. But the relative order of elements has to be captured by another mechanism, which could be f-precedence or rule parameterization. The former introduces additional complexity and duplicates the role of c-structure rules, which are supposed to simultaneously model linear order and dominance relations. Parameterized rules will work (45), but their introduction further blurs the role of $\overline{X}$ theory: if they are freely used, all $\overline{X}$-theoretic positions can be reduced to adjunction, with parameters doing the work of bar levels.

Basically, this smuggles the cartographic approach into LFG, but via a different route.

(45)

$\overline{\text{V}}$[wh]
- k₃j
  who.GEN
- $\overline{\text{V}}$[neg]
  - ni-či
    NEG-who
  - $\overline{\text{V}}$[pred]
    - V
      - fet:-a
        see.PFV-PST.3SG

3.4 Reinterpretation of $\overline{\text{X}}$ theory

What all these solutions (except Sells-adjunction, which is empirically problematic for Ossetic) have in common is that they use $\overline{\text{X}}$ theory in ways it was not originally meant to be. Perhaps, then, one should face the facts and reinterpret the role of $\overline{\text{X}}$-theoretic positions? The classical version of $\overline{\text{X}}$ theory uses three projection levels (unlike Jackendoff 1977, who used four) and, consequently, awards a central role to the distinction between complements and specifiers. But we can see that actually the maximum bar level is a language-specific parameter: some languagres, like Japanese and Korean, can apparently get away with just $\overline{\text{X}}$. Similarly, we can assume that some languages require more than two levels if more than two endocentric positions are involved. In fact, this is already envisaged in the theory of Lowe and Lovestrand (2020), where the maximum level of projection is not viewed as universal.
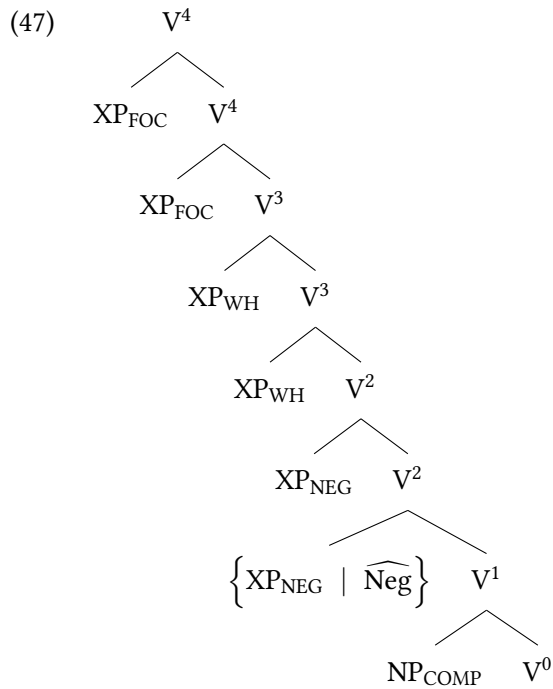
Under this interpretation, we need to distinguish between only two types of $\overline{\text{X}}$-theoretic rules: adjunction and complementation; adjunction does not increase the bar level, while complementation does. Both can apply at any bar level. In effect, this means that the role of $\overline{\text{X}}$ theory is the same as that of template morphology in languages with rich systematic affix ordering patterns. Consider the analysis of the Ashti Dargwa verb form in (46).

(46)

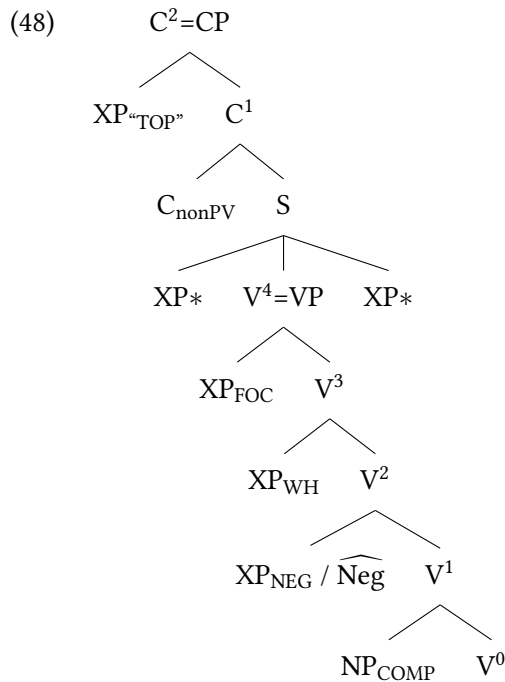| -4 nominal stem	-3 preverb(s)	-2 negation	-1 gender	0 base stem	+1 causative	+2 ... inflectional endings
aq-	ha-	ta-	r-	ic:	-ah	-a-d-i
high	up	NEG	F	stand.PFV	CAUS	PRET-1-SG
'I did not make her stand up.'						

The role of the numbers in (46) is exactly the same as that of the projection levels in the "extended $\overline{\text{X}}$-theoretic" anlysis of the Ossetic VP in (47), viz. to capture the relative ordering and hierarchical embedding of dependents in endocentric structures (in syntax or in morphology).

(47)

$$V^4$$

The tree structure (47) branches as follows:

- V^4
  - XP_{FOC} , V^4
    - XP_{FOC} , V^3
      - XP_{WH} , V^3
        - XP_{WH} , V^2
          - XP_{NEG} , V^2
            - $\{\, XP_{NEG} \mid \widehat{Neg} \,\}$, V^1
              - NP_{COMP} , V^0

In this structure, each bar level corresponds to one type of preverbal elements. This analysis is trivially compatible with Lowe and Lovestrand (2020). One implementational detail is that all preverbal constituents can iterate, but the last focal XP, wh-word and negative pronoun are treated separately (as complements, not adjuncts) because at least final wh-words have special status (Belyaev 2014). If this is not modeled in c-structure, all can be viewed as adjuncts. Finally, I do not position comparative adverbs in this tree: they can adjoin to V^2 or have their own bar level, in which the total number of bar levels in the Ossetic VP will increase to 5.

The new picture of the Ossetic clause as a whole that emerges from this approach is shown in (48).[2]

---

2. Technically, the node S is endocentric, because in Ossetic it is always headed by V. It is therefore possible to treat S as an additional (flat) projection of V as well, as was proposed in Jackendoff (1977), who argued that S is actually V''' in English.

(48)

C^2=CP
- XP$_{\text{“TOP”}}$ C^1
  - C_{nonPV} S
    - XP∗ V^4=VP XP∗
      - XP$_{FOC}$ V^3
        - XP$_{WH}$ V^2
          - XP$_{NEG}$ / $\widehat{Neg}$ V^1
            - NP$_{COMP}$ V^0

3.5 Rules

The formalization of this analysis is straightforward, although I cannot at present provide a rule for the lowest level (complex predicates), because their f-structure is unclear. The analysis should probably differ for the more "incorporated" and "idiomatic" types of predicates.

The position of interrogative and negative pronouns can be modeled using constraining equations. In (49), the V^2 rule introduces a defining equation for negation, which is "checked" by the negative element itself. This guarantees that if the rule is used, this position must contain a negative elemnet, and that negative elements themselves cannot appear outside of their dedicated positions. The machinery for interrogatives in (50) is the same.

(49) a. *ni-sə* D $(\uparrow \text{SPEC}) =_c \text{NEG}$

...

b. $V^2 \rightarrow$ XP V^1
$(\uparrow \text{GF})=\downarrow$ $\uparrow=\downarrow$
$(\downarrow \text{SPEC})=\text{NEG}$
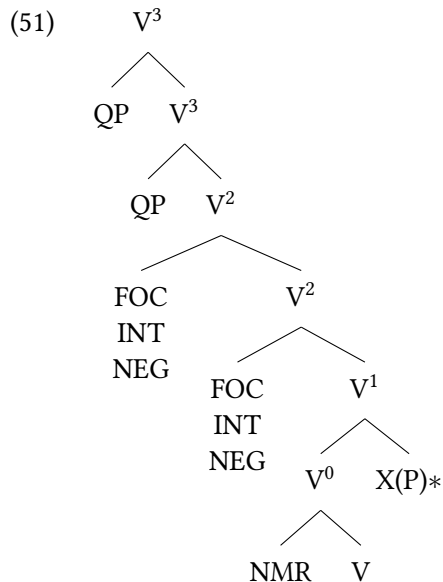
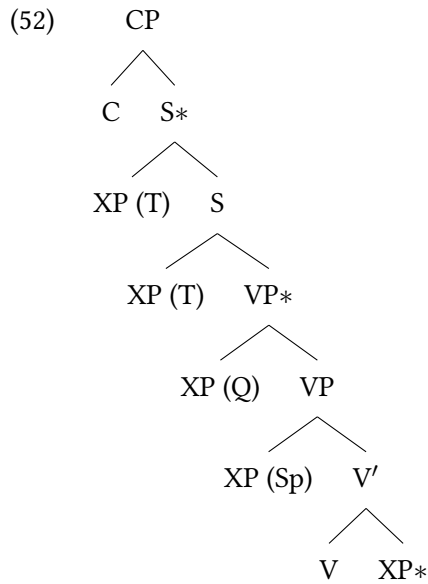(50) a. *sə* D $(\uparrow \text{SPEC}) =_c \text{WH}$

...

b. $V^3 \rightarrow$ XP V^2
$(\uparrow \text{GF})=\downarrow$ $\uparrow=\downarrow$
$\downarrow \in (\uparrow \text{DIS})$
$(\downarrow \text{SPEC})=\text{WH}$

3.6 Similar proposals

Ossetic is not the only language that has a strictly organized preverbal area. Hungarian is another such language which has been extensively analyzed in LFG. Like Ossetic, Hungarian combines a verbal complex built on the basis of information structure and structural NP types rather than grammatical functions with a relatively free word order outside this verbal complex. And in fact, a very similar proposal for the c-structure of Hungarian predicates has been made in Börjars, Payne, and Chisarik (1999) and Payne and Chisarik (2000). (Payne and Chisarik 2000, 204) propose the c-structure in () for Hungarian. The similarities with my analysis of Ossetic are in the use of more than two bar levels and in the special treatment of final foci, interrogatives and negative pronouns. The key difference is that Chisarik and Payne use Optimality Theory to model the relative order of some elements in the clause, while I believe that Ossetic word order can be described through phrase structure rules alone.

(51)

```
            V³
          /    \
        QP      V³
              /    \
            QP      V²
                  /    \
               FOC      V²
               INT    /    \
               NEG  FOC      V¹
                    INT    /    \
                    NEG   V⁰     X(P)∗
                        /   \
                      NMR    V
```

Laczkó (2014) argues against the OT-LFG approach to Hungarian and proposes a different model for the VP, but his view of overall Hungarian clause structure (52) is basically the same as my proposal for Ossetic: an S node sandiched between CP and VP. Laczkò assumes a binary-branching S, but this is not essential: the same analysis could be used for Ossetic, given that coordination at this level is available in any order.

(52)

```
            CP
           /  \
          C    S*
              /  \
         XP (T)   S
                 /  \
            XP (T)   VP*
                    /  \
               XP (Q)   VP
                       /  \
                  XP (Sp)  V′
                          /  \
                         V    XP*
```

It is remarkable that two languages that are both geographically and genetically very distant from each other demonstrate similar behaviour that independently leads to very similar solutions. This suggests that the clause structure pattern of Ossetic is by no means exotic, and may in fact be the norm for a certain type of non-configurational language.

4 Conclusions

In this paper, I have attempted to model Ossetic clause structure using standard LFG assumptions about c-structure. The problem with Ossetic word order is that the number of dedicated positions for different elements is so high that it requires either "going cartographic" or extending $\overline{\text{X}}$ theory to more than three projection levels, essentially eliminating "complement" and "specifier" as meaningful notions. Instead, what $\overline{\text{X}}$ theory models is which syntactic positions have grammaticalized in a given language. Just like lexicalist models of morphology do not assume a fixed schema but accept that languages may grammaticalize different types of affixes in different orders, this view of phrase structure suggests that elements like VP and IP are not innate building blocks of grammar, but syntactic structures that grammaticalize differently in different languages. Ossetic, for instance, has developed discourse configurational VP and CP but has not developed any intermediate projections. Hungarian arrived at a similar structure, but with different elements included in the VP and in different order. Classically configurational languages like English have taken a different route, grammaticalizing a VP and IP based on grammatical functions rather than discourse. This does not mean that these structures are completely random: again like in morphology, general laws of diachronic change mean that not all structures are equally likely

to grammaticalize. This way of approaching configurationality seems to offer a viable alternative to approaches that assume a uniform hierarchy of projections, of which cartography is perhaps the most consistent example.

Many questions remain unanswered, both for Ossetic and in a more general sense. The status of S itself is unclear: if it is always headed by V, should it rather be treated as a projection of V? Or should S be used to represent a "flat" clause level regardless of its endocentricity? A full analysis of Ossetic clause structure also requires a separate, detailed account of complex predicates and negation, as well as the establishment of the information-structure status of elements preceding nPV subordinators. Periphrastic constructions should also be looked at in more detail for evidence of a hierarchical c-structure. Finally, what constituency tests except prosody and coordination can be used to determine clause structure in Ossetic? A weak point of the LFG approach to c-structure seems to be that very often, the number of analytical possibilities (flat vs. branching structure, adjunct vs. specifier) exceeds the number of criteria that can reliably distinguish between them.

References

Abaev, V. I. 1949. *Osetinskij jazyk i fol'klor* [Ossetian language and folklore]. Vol. 1. Moscow and Leningrad: Izdatel'stvo Akademii nauk SSSR.

Austin, P., and J. Bresnan. 1996. "Non-configurationality in Australian aboriginal languages." *Natural Language and Linguistic Theory* 14 (2): 215–268.

Belyaev, O. I. 2014. "Korreljativnaja konstrukcija v osetinskom jazyke v tipologičeskom osveščenii" [Ossetic correlatives in typological perspective]. Cand. thesis, Lomonosov Moscow State University.

Börjars, K., J. Payne, and E. Chisarik. 1999. "On the justification for functional categories in LFG." In *Proceedings of the LFG '99 Conference,* ed. by M. Butt and T. H. King. S: CSLI Publications.

É. Kiss, K. 2004. *The syntax of Hungarian.* Cambridge: Cambridge University Press.

Erschler, D. 2012. "From preverbal focus to preverbal "left periphery": The Ossetic clause architecture in areal and diachronic perspective." *Lingua* 122 (6): 673–699.

Jackendoff, R. 1977. *X-bar syntax: a study of phrase structure.* Linguistic Inquiry Monographs 2. Cambridge, MA: MIT Press.

Laczkó, T. 2014. "Essentials of an LFG analysis of Hungarian finite sentences." In *Proceedings of the LFG '14 Conference,* ed. by M. Butt and T. H. King. Stanford, CA: CSLI Publications.

Lowe, J. J., and J. Lovestrand. 2020. "Minimal phrase structure: a new formalized theory of phrase structure." *Journal of Language Modeling* 8 (1): 1–52.

Maxwell, J. T., III, and C. D. Manning. 1996. "A theory of non-constituent coordination based on finite-state rules." In *On-line Proceedings of the First LFG Conference,* ed. by M. Butt and T. H. King, Rank Xerox, Grenoble, August 26–28, 1996.

Melchin, P. B., A. Asudeh, and D. Siddiqi. 2020. "Ojibwe agreement in Lexical-Realizational Functional Grammar." In *Proceedings of the LFG'20 Conference,* ed. by M. Butt and I. Toivonen, 268–288. S: CSLI Publications.

Payne, J., and E. Chisarik. 2000. "Negation and focus in Hungarian: An Optimality Theory account." *Transactions of the Philological Society* 98:185–230.

Sells, P. 1994. "Sub-phrasal syntax in Korean." *Language Research* 30 (2): 351–386.

———. 1995. "Korean and Japanese morphology from a lexical perspective." *Linguistic Inquiry* 26, no. 2 (Spring): 277–325.