This chapter has two main goals. The first one is to introduce the reader to the grammatical properties of Avar (§2.2) whereas the other one is to establish a clause structure that would serve as the basis for the discussion in the chapters to come (§2.3). Before we see how the particular properties of Avar map onto a syntactic hierarchy, several basic notions must be introduced and a few general remarks on the architecture of the grammar given. This is done in the remainder of this section.

2.1 Introduction

2.1.1 The basics

Since the subject matter of this thesis, operator–variable dependencies, belongs in the domain of the syntax–semantics interface, I find it useful to introduce, in brief, the key notions of what I take to be the structure-building module (syntax) and the interpretational component (semantics). I do this in sections 2.1.1.1 and 2.1.1.2 respectively. Readers familiar with the fundamentals of generative syntax might find it useful to go directly to §2.2, which, in turn, can be skipped by those who are fairly comfortable with the morphosyntax of Northeast Caucasian languages.
2.1.1 Syntax

The present thesis adopts a broadly minimalist approach to syntax, syntax being understood in the narrow sense as a computational system that generates potentially arbitrarily complex expressions (Chomsky 1993, 1995, 2007). The core syntactic operation is merge, which can be defined as a function combining two elements from the domain of syntactic expressions and yielding a set containing those syntactic expressions. The output of an application of merge is therefore a symmetrical set and its definition is given in (1):

\[
\text{merge}(\alpha, \beta) = \{ \alpha, \beta \}
\]

I follow the conventions generally accepted in the generative literature and note the sets resulting from the application(s) of merge, such as \{\alpha, \beta\} above, as either labelled brackets (2a) or branching nodes mostly referred to as trees (2b, although I will use a revised version of the tree notation as explained later in this subsection). I use \gamma here for the resulting syntactic object.

\[
\begin{align*}
\text{(2) a. } \text{[ } \gamma \alpha \beta \text{ ]} & & \text{b. } \gamma \\
& & \alpha \beta
\end{align*}
\]

The set notation, in turn, will be reserved to represent numerations — collections of items drawn from the lexicon that syntax will operate on whilst building the structure. To exemplify, in order to create a complex object like this houndstooth jacket, the computational system will draw the following elements from the lexicon (the lexicon itself being, uncontroversially, a repository of elements that contains information on their various properties):

\[
\text{(3) } N = \{ \text{this, houndstooth, jacket} \}
\]

Syntactic operations like merge apply to the elements of \(N\) in any possible order, the resulting objects receiving an interpretation at the interfaces. Given

---

1. This view of merge as a binary operation is, although classic, not the only possible one. See Zwart 2009 for an alternative proposal whereby merge is understood as a unary operation that takes an object from the numeration and merges it to another object in the workspace, resulting in an ordered pair instead of an unordered set of Chomsky (1995) et seq.

2. The notion of labelling, or projection, has become problematic as the generative enterprise started to move away from preset phrase structure. In this thesis I do not discuss the nature of endocentricity and invite the interested reader to consult Chomsky (2013) and Adger (2013) for an overview of the problems and two very different solutions.

3. I will also use the set notation in Chapter 5 to represent sets of alternative propositions that focus-sensitive expressions evoke.
that N contains three elements, merge can apply to them in the following ways (since the outputs of merge are sets, the order of elements is irrelevant, a fact obscured by the tree notation):

\[(4)\]

\begin{align*}
\text{a. } & \text{merge}(\text{houndstooth, jacket}) = \{\text{houndstooth, jacket}\} \\
& \text{merge}(\{\text{houndstooth, jacket}\}, \text{this}) = \{\text{this, houndstooth, jacket}\} \\
\text{b. } & \text{merge}(\text{this, houndstooth}) = \{\text{this, houndstooth}\} \\
& \text{merge}(\{\text{this, houndstooth}\}, \text{jacket}) = \{\text{jacket, this, houndstooth}\} \\
\text{c. } & \text{merge}(\text{this, jacket}) = \{\text{this, jacket}\} \\
& \text{merge}(\{\text{this, jacket}\}, \text{houndstooth}) = \{\text{houndstooth, this, jacket}\}
\end{align*}

\[(5)\]

a. \\
\hspace{1cm} \text{this} \\
\hspace{2cm} \text{houndstooth} \\
\hspace{3cm} \text{jacket} \\

b. \\
\hspace{1cm} \text{jacket} \\
\hspace{2cm} \text{this} \\
\hspace{3cm} \text{houndstooth} \\

c. \\
\hspace{1cm} \text{houndstooth} \\
\hspace{2cm} \text{this} \\
\hspace{3cm} \text{jacket}

Once there are no elements left in N the complex syntactic object is ready to be sent to the interfaces to receive a semantic interpretation and a phonological realisation. Whilst there appears to be nothing wrong with either (5b) or (5c) from the point of view of pronunciation, neither of them can receive the right semantic interpretation, which is reflected in their unacceptability. This means, amongst other things, that syntax cannot process the information pertaining to selection.

The items drawn from the lexicon, or lexical items, I take to be collections of features.\(^4\)

Coming back to merge, the general consensus in the literature is that it exists in two kinds — external and internal (Chomsky 2004). External merge, which has been illustrated in (1) and (2) above, introduces new material from the numeration into the derivation. In contrast, internal merge (also called remerge, or movement) operates on elements already present in the structure by remerging a copy of that element with the already built structure. For the purposes of this thesis internal merge, remerge and movement are treated as synonyms and I will use them interchangeably.

---

\(^4\) I choose ‘collection’ over ‘bundle’ or ‘structure’ purely because I wish to remain agnostic as to the exact way of putting the features together to create a lexical item; see Adger & Svenonius (2011) for a detailed discussion of these issues as well as Adger (2010) for an explicit formalisation.
Using internal merge allows one to derive long-distance dependencies, which will be represented in this thesis with the trace notation, whereby the unpronounced copy is notated as t, oftentimes with a subscripted index, and the two positions related by movement are connected with an arrow. This is illustrated immediately below.

\[
\begin{align*}
(6) & & \text{a.} & \theta & \alpha & \eta & \zeta & \epsilon & \delta & \gamma & \beta & \alpha \\
& & \downarrow & \alpha & \eta & \zeta & \epsilon & \delta & \gamma & \beta & \alpha \\
& & \downarrow & \alpha_1 & \eta & \zeta & \epsilon & \delta & \gamma & \beta & t_1
\end{align*}
\]

In (6) the syntactic object \(\alpha\), originally in a sisterhood relationship with \(\beta\), has undergone syntactic movement to \(\theta\), or has become internally merged with, the node labelled \(\eta\). This is indicated by replacing the base position of \(\alpha\) with the trace symbol \(t\), and by connecting the two nodes with an arrow. The tree in (b) is identical to the one in (a) but for the appearance of a numerical index on the nodes related by a movement dependency, which is especially useful in trees where more than one element moves.

To paraphrase the foregoing discussion in Pesetsky's (2013) terms, internal merge has the following two properties: the \textit{c-command} property (in the sense of \(\alpha\) c-commanding \(t\) in (6)) and the \textit{multidominance} property (in the sense of \(\alpha\) being dominated by both \(\theta\) and \(\gamma\)). C-command can be viewed in both representational and derivational terms, a derivational definition being given in (7) below.

\[
(7) \quad \textit{C-command} \quad \alpha \textit{ c-commands } \beta \textit{ and all and only elements contained by } \beta \textit{ iff } \alpha \textit{ was merged with } \beta \textit{ in the course of the derivation.} \quad (\text{based on Epstein et al. (1998: 32)})
\]

Arrows of a different kind, dashed, notate the Agree relation between features on various lexical items. In (8) below, \(\alpha\), whilst still in its base position, carries a feature \([x:y]\), where \(x\) is the feature's attribute and \(y\) its value. A node higher than \(\alpha\) in the structure, say \(\zeta\), has a feature with a matching attribute but without a value.
The majority of the current generative literature presupposes that it is these feature-valuation operations, or probe–goal relations, which are responsible for triggering syntactic movement, either directly or with the help of dedicated movement-triggering features (edge features, or EPP-features in earlier work). In this thesis I depart from this tradition and follow Chomsky 2007, 2013 in viewing internal merge as an operation that is not driven by the featural needs of either the element undergoing it or those of the derived position.

Syntactic categories are organised in a functional sequence (cf. Starke 2001) roughly corresponding to Grimshaw’s (1997) notion of an extended projection that the outputs of merge must respect in order for the derivation to be interpretable.

An assumption widely shared in the generative literature is that syntactic derivations consist of subderivations that become opaque upon being completed, preventing any further syntactic operations such as internal merge from applying. These locality constraints are often taken to correspond to Ross’s (1967) “islands”, and most current approaches couched in the minimalist framework formalise these as phases of Chomsky (2001, 2008), although equating them has been shown to be problematic (Boeckx 2012). Since in this thesis I adopt a different view of locality, viz. that of derivation layering (Zwart 2009 et seq.), and given that an in-depth discussion of the concept of phases is beyond my immediate concerns, I limit myself to noting that if the approach of Zwart (2009) et seq. is on the right track, phases as a separate domain become redundant. The interested reader will find detailed introductions to the phase theory in Gallego (2010), Citko (2014), among others, and an explicit proposal of how the CED effects (=islands) follow from this notion in Müller (2010).
The notation: Spanning

In notating the dependencies created by merge as tree-like diagrams I adopt, after Bye & Svenonius (2012), Svenonius (2012), the use of the Telescope principle (Brody 2000, Brody & Szabolcsi 2003), whereby the intermediate projection level — or the bar level — is omitted from the tree entirely, and the node's right daughter is its complement whilst the left one is its specifier.

As the Minimalist Programme places particular emphasis on the interfaces, I follow Halle & Marantz (1993) in adopting an approach to morphology whereby morphological exponents are inserted in the representation post-syntactically. This insertion is notated, in line with Bye & Svenonius (2012), Svenonius (2012) and Ramchand & Svenonius (2014), via the squiggly lines. The choice of notation is, for the most part, aesthetic, motivated not least of all by the peculiar shape that trees for head-final languages tend to take, especially those where specifiers, like complements, linearise to the left of the heads. The notation used in this thesis is more economical of the vertical space and arguably reflects the properties of syntactic representations more accurately by focusing on the complement line (Bye & Svenonius 2012: 433). It also allows me to sidestep the problem of discussing the nature of head-final word orders, i.e. abstract away from whether sentences in languages like Avar, Turkish or Japanese are generated head final as a macroparametric option, or whether these orders result from multiple roll-up complement-to-specifier movements as proposed in Julien (2002).

To introduce the notation, let us consider a (simplified) syntactic representation of the following sentence containing a ditransitive predicate CM–ič– ‘sell’:

\[
\text{(9) } \text{di- ca du- e ſaka b–ič- ana.} \\
\text{1SG:OBL-ERG 2SG:OBL-DAT COW.ABS N–sell-PST} \\
\text{‘I sold you a/the cow.’}
\]

The traditional tree in (10a), even though necessarily incomplete, as it does not contain any information as to how the various functional heads in the clausal spine come to be pronounced as suffixes attached to the root, occupies twice as much vertical space as its Telescoped version in (10b).\footnote{The two structures also differ with respect to EPP: whereas the subject undergoes movement to Spec,T in (a), it remains inside the v-layer in the tree under (b). The distinction is immaterial at this point but I will suggest, in \$4.5.2.4, that this movement to Spec,T is not restricted to subjects alone.}
The projection, or complement, line in (10b) should be read bottom-up starting from V, with the elements dominating it corresponding to suffixes. The asterisk diacritic on V indicates that the complex morphological word is to be pronounced at V (as opposed to Appl, v or T) to ensure that it is linearised to the right of both the complement and the specifier.

What needs stressing at this point is that the tree in (10b) is a representation which the interfaces read the information off. In particular, it makes no claims as to the nature of merge, which I take to be binary, in accordance with most current work on the matter.

Having introduced the relevant syntactic notions, one more aspect of the notation that I find worth mentioning concerns the mapping from syntax to morphology. Most contemporary approaches to the syntax–morphology interface hold it that morpheme insertion, or lexical insertion, happens after the syntactic computation has run its course, i.e. after a numeration has been exhausted. The most famous approach of this kind has come to be known as Distributed Morphology (Halle & Marantz 1993, Embick & Noyer 2007) but it is not the only one possible, as evinced by the existence of frameworks like Nanosyntax (Caha 2009), where morpheme insertion is also performed after the syntactic computation has run its course but the need of having such operations as fusion or fission does not arise because morphemes are viewed as
being able to correspond to units bigger than just the terminal nodes. In this thesis, therefore, I subscribe to the view of morphology as operating after the narrow-syntactic computation has been done without going into the technical details of these postsyntactic processes.

2.1.1.2 Semantics

The preceding subsections will have hinted at the conclusion that, the organisation of the syntax being fairly simple (ideally reducible to just merge, Chomsky 2013), the interfaces must be set up in such a way as to interpret the “right” outputs only. As far as semantics is concerned, the main principle guiding semantic interpretation is that of compositionality, i.e. the conjecture that the meaning of a complex expression is calculated based on (i) the meanings of its parts and (ii) their mode of combination. Put differently, semantic interpretation is sensitive to certain syntactic relations (mostly dominance and c-command).

Where I discuss semantic interpretation, the formal framework is that of Heim & Kratzer (1998), one where syntactic expressions are interpreted by $\left[ \right]$ — the interpretation function — directly (i.e. without an intermediate translation step). The composition rules themselves are introduced where relevant.\[6\]

Presuppositions and partial functions

In addition to such standard interpretation rules as Function Application or Predicate Abstraction, one aspect of Heim & Kratzer’s (1998) system that features prominently in Chapter 5 is their notation for presupposition. In that chapter I follow Heim & Kratzer (1998) and separate the presupposed material from the rest of the information by putting it between a colon and a full stop:

$\lambda f : \alpha . f$  

Heim & Kratzer (1998) treat presuppositions as definedness conditions on semantic values: the function $f$ is defined if and only if $\alpha$ is satisfied, and undefined otherwise.

6. See Haug 2014 for a discussion of the advantages and disadvantages of Heim & Kratzer’s (1998) treatment of presupposition, the disadvantages that automatically carry over to most analyses that take Heim & Kratzer (1998) as their point of departure, as well as Coppock (2014) for an alternative, based on Beaver & Krahmer 2001, that is arguably immune to the problem.
2.2 Overview of Avar grammar

Avar (self-nomination: maʃarul maʃ ‘language of the mountains’) is the biggest language of the Avar-Andic branch within the Northeast Caucasian, or Nakh-Daghestanian, language family. It is spoken mainly in the western and southern parts of the Russian Caucasus republic of Daghestan, and the Balaken, Zaqatala regions of north-western Azerbaijan. According to the 2010 census, there are some 703,000 native speakers of Avar in the Russian Federation (compared to 744,000 eight years previously), and speakers of smaller Avar-Andic languages use it as a *lingua franca.* Although Avar has enjoyed a special status as a language of instruction with lots of published material, it is continuously being replaced by Russian as both the language of instruction and the *lingua franca* within the Republic of Daghestan, particularly so amongst the younger speakers.

The language has existed in written form (Arabic-based script) since the 17th century. In the 19th century a Russian Imperial Army major-general Baron von Uslar undertook an attempt at developing a Russian-based alphabet for

---

2.2. Overview of Avar grammar

<table>
<thead>
<tr>
<th></th>
<th>Lab</th>
<th>Den</th>
<th>Alveolar</th>
<th>Pal</th>
<th>Vel</th>
<th>Uvu</th>
<th>Epiglot</th>
<th>Glot</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nasal</td>
<td>m</td>
<td>n</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Flosive</td>
<td>voiced</td>
<td>b</td>
<td>d</td>
<td>g</td>
<td>?</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>voiceless</td>
<td>p</td>
<td>t</td>
<td>k</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ejective</td>
<td>t'</td>
<td></td>
<td>k'</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Affricate</td>
<td>voiced</td>
<td>c</td>
<td>č</td>
<td>l</td>
<td>q</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>ejective</td>
<td>c'</td>
<td>č'</td>
<td>l'</td>
<td>q'</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Fricative</td>
<td>voiced</td>
<td>s</td>
<td>š</td>
<td>l</td>
<td>x</td>
<td>χ</td>
<td>h</td>
<td></td>
</tr>
<tr>
<td></td>
<td>voiced</td>
<td>v</td>
<td>z</td>
<td>ž</td>
<td>y</td>
<td>ϔ</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Trill</td>
<td>r</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Approximant</td>
<td>l</td>
<td>j</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 2.1: Avar consonants

several Northeast Caucasian languages including Avar, of which he also published the very first grammar (von Uslar 1889). The resulting alphabet was phoneme-based and digraph-free, at the cost of having a number of additional characters. In the Soviet period, the script was first changed to Roman and then back to Cyrillic, and Avar became one of several major languages in Dagestan that had a special status: it was a language of instruction and learning in elementary and secondary schools, and a language of media and emerging literature.

A note on romanisation

As just mentioned, the current orthography for Avar is the Russian alphabet with one diacritic character, palochka ‘stick’, and given the abundance of consonants in the language’s phoneme inventory (illustrated in Table 2.1), those are often rendered as digraphs. For the purposes of this study I have decided to use a simplified mixture of transcription and transliteration, where gemination is only conveyed where it is realised orthographically in the original.8,9

8. The romanisation I adopt here differs very slightly from that in Yamada 2013. To avoid confusion, the Roman-to-Cyrillic correspondences are as follows, where the apostrophe represents ejection: a = a, b = b, v = w, г = g, gь = y, гь = h, гI = ġ, d = d, e = e, ж = ž, з = z, и = i, й = j, к = k, кь = q’, къ = k’, l = l, ль = l, м = m, н = n, о = o, п = p, р = r, с = s, t = t, тI = t’, у = u, ф = f, х = χ, хь = q, хь = x, хI = ĥ, ц = c, цI = c’, ч = č, чI = č’, ш = š, щ = š, Ϲ = e, ю = ju, я = ja, ъ = ʔ.
9. The romanisation adopted here is similar to the orthography of Standard Avar in not making a distinction between the lateral affricate, and the lateral fricative, which itself can be strong and weak, all of which are standardly written as ďs, or ř in the system adopted here.
As mentioned at the very beginning of this report, developing a theory of operator–variable dependencies for a free word order language like Avar implies having a credible syntax that could easily be mapped onto semantics, given the working principle of compositional interpretation. In the following few sections I attempt to do precisely this paying particular attention to subject–object asymmetries and identifying constraints on word order permutations, although I will remain largely agnostic as to the exact mechanism(s) effecting these. Before this can be done, however, I present an overview of Avar grammar that should be both detailed enough for the purposes of this thesis and concise enough so as not to distract from the main subject matter.

2.2.1 Sources and methodology

In line with the general methodology of theoretical and typological approaches to natural language syntax, I have chosen to use grammaticality and acceptability judgements of Avar native speakers as the primary source of linguistic data for the current project. I find it necessary, however, to complement this sort of approach to data collection by referring to (i) existing descriptions of Avar (von Uslar 1889; Bokarev 1949; Madieva 1980; Alekseev & Ataev 1997) and related languages (cf. Kibrik 1999, Lyutikova 2000 for Tsakhur, Kibrik 2001 for Bagwalal or Khalilova 2009 for Khwarshi), and (ii) written and published material in Avar collected from newspapers, magazines and other web-based resources.

Not being an Avar speaker myself, I employ translations and examples modelled on attested structures to elicit the speakers’ judgements using Russian as the mediating language (see Matthewson 2004, 2011 for strong arguments in favour of this approach over certain others, such as reliance on collections of texts). It is thus extremely important to have speakers clearly understand what they are being asked, particularly so because of the notorious difficulties in obtaining reliable judgements regarding scope-related phenomena (Szabolcsi 2010: §6), which in the case at hand is the scope of focus and question particles. Parts of the data on focus marking presented in Chapter 5 were elicited using the questionnaire on the semantics of focus-sensitive expressions in Renans, Zimmermann & Greif (2011).

In the subsections that follow I present an overview of Avar grammar so as to introduce the reader to the main properties of the language. To reiterate, this description is to be thought of as a brief sketch rather than a rigorous descriptive grammar. We begin with word order.
2.2.2 Basic word order

Although Avar, like the other Northeast Caucasian languages, displays a considerable freedom of word order (Testelec 1998a), it is still possible to identify a basic, ‘unmarked’, order:

(12) Subject > Indirect_Object > Direct_Object > Verb

The two sentences below illustrate this basic order, ignoring temporal and other adjuncts for the moment.

(13) was-as insu- e ruq’ b-a- l- e- b b-ugo
    son-ERG father.OBL-DAT house.ABS n–build–PRS–PTCP–N n–be.PRS
    ‘The son is building (his) father a house.’

(14) di- ca du- e ʕaka b–iĉ- ana
    1SG:OBL-ERG 2SG:OBL-DAT cow.ABS n–sell–PST
    ‘I sold you a/the cow.’

In both (13) and (14) there are three noun phrases to be thought of, descriptively, as agent (wasas ‘son’ in (13) and dica ‘I’ in (14)), indirect object (insue ‘father’ in (13), due ‘you.DAT’ in (14)), and direct object (ruq ‘house’ in (13), ʕaka ‘cow’ in (14)).

The unmarkedness of SOV is true not only of monoclausal constructions but also of cases of causativisation (15) and clausal complementation (16), which contains an infinitival clause:

(15) uĉitel- as rasuli-da kayat qwaz-a– b–una
    teacher-ERG Rasul-LOC letter.ABS write-CAUS–N-PST
    ‘The teacher made Rasul write a letter.’ [SOV]

Causative stems are formed by merging the base stem of a verb with that of the verb ha–CM-ize ‘do/make’, which contains a slot for a gender agreement marker. This whole cluster behaves like a single word from the prosodic and orthographic perspective, although this behaviour is subject to inter-speaker variation. I return to the morphosyntactic composition of the Avar verb in §2.2.5.

(16) untaras- e 恂urg w–ix- ize b–ol’- ana
    sick.man-DAT surgeon.ABS m–see–INF n–want–PST
    ‘The patient wanted to see the surgeon.’ [SOV]

10. I am using unmarked here as a purely descriptive label without attributing to it any theoretical significance.
However, alternative word orders are also very frequent, with the verb’s internal argument instantiated by a nominal phrase (17–19) or a clause (19–20):

(17) b–os- un b–ugo faka he– w či- jas
N–buy-CVB N–be.PRS cow.ABS that–M:ABS man-ERG
That man bought a cow.’

(18) avtomobil- al b–ač- ula niže-ca t’orš:el
automobile-ERG N–transport-PRS we- ERG grain.ABS
“We transport the grain with the car.’

(19) di- e b–ol’- un b–ugo wac- as institut łuʕ- ize
1SG-DAT N–want-CVB N–be.PRS brother-ERG institute.ABS finish-INF
‘I want (my) brother to graduate.’

(20) a– b–do– b heč’o= go razil- ana dol di- e
this–N that–N COP:NEG=EMPH agree-PST 3PL.ABS 1SG:OBL-DAT
jas l’e- ze
daughter.ABS give-INF
‘They agreed to give their daughter to me (as a wife) straight-out.’

As can be seen from the four sentences above, noun phrases can appear dislocated to the left or right periphery of the clause. For the purposes of this thesis I make the more or less standard assumption that such deviations from the SOV order are derived via overt syntactic movement, whilst also leaving open the possibility of these orders being base generated and involving no syntactic movement (see Brody & Szabolcsi 2003 and Adger, Harbour & Watkins 2009 for base-generation analyses of certain word-order patterns in Hungarian and Kiowa, respectively). We will see in the chapters to come that these disruptions of the basic word order result in particular information- and discourse-structural interpretations (which does not entail, I argue, that the movements underlying their derivation have to be triggered by dedicated information-structural features).

Returning to the issue of the default word order, it will become obvious from the exposition to follow that head finality in Avar is not restricted to verb phrases — the language can in principle be characterised as head-final.

### 2.2.3 Nouns and agreement

In Avar, like in other Northeast Caucasian languages, parts of speech can be easily distinguished morphologically. Not only is the inflectional morphology
2.2. Overview of Avar grammar

Avar nominal morphology is fairly rich: there is abundant case marking as well as gender/noun class agreement on verbs, adjectives and adpositions that overlaps with number agreement. There are four noun classes: I for masculine, II for feminine, III for neuter and inanimate and IV for plural objects; in this thesis, however, rather than following the convention of using Roman numerals to refer to the relevant noun class, I shall utilise mnemonic M, F, N, PL for the purposes of clarity. Plural agreement on nominal modifiers and verbs is expressed via distinct affixes, and I am using PL for both of them. The nouns themselves have distinct stems for singular and plural, as well as oblique stems.

Absolutive is treated as the ‘direct case’, and is derived from the singular and plural ‘direct stems’; all other cases are morphologically oblique, being derived from corresponding oblique stems (see Figure 2.2 for morphological details).

The most productively used plural affix in the absolutive case is –l, being followed by the somewhat less productive –abi. In order to form the ‘direct plural’ from wac ‘brother’ or ču ‘horse’, the corresponding ‘direct singulars’ take on –l, whilst certain other nouns like gali ‘step’ or mašina ‘car’ use –abi to the same end. This is illustrated in (21).

(21) a. wac ‘brother’ — wac-al ‘brothers’; ču ‘horse’ — ču-jal ‘horses’
   b. gali ‘step’ — gal-ab-abi ‘steps’; mašina ‘car’ — mašin-abi ‘cars’
   c. a number of other, less productive, suffixes

Before the ‘oblique plural’ can be formed, the final segment of the ‘direct plural’ stem is changed to -z- in the following ways:

(22) -l → -z:: wac-a-l ‘brother.dirl.pl → wac-a-z- ‘brother.obl.pl’
     -i → -a-z:: mašin-a-bi ‘car.dirl.pl → mašin-abi-a-z- ‘car.obl.pl’

In the singular, the direct-to-oblique transition is subject to a fair amount of variation determined for the most part by the declension class a particu-
lar noun belongs to, and the oblique stem corresponds in some cases to the erg-marked noun. Some of the more common patterns are given in (23).

(23)  -∅→s-:  wac-∅ 'brother-dir.sg' → wac-as- 'brother-obl.sg'
      -∅→l-:  jas-∅ 'girl-dir.sg' → jas-al- 'girl-obl.sg'
      -∅→i-:  t'oy-∅ 'roof-dir.sg' → t'oy-i- 'roof-obl.sg'

Given the fact that the oblique stem may fully correspond to the noun’s ergative form, as is the case for wac and jas in (23) above, there is potential for confusion with respect to the glossing conventions. In this thesis I use both OBL and ERG and rely on the ambiguity being resolved with the help of the surrounding context.

2.2.4 Nominal syntax

In noun phrases, the head noun is linearly always phrase-final with respect to its modifiers such as adjectives, demonstratives or relative clauses.\(^ {11,12} \)

2.2.4.1 Prenominal modifiers

Adjectives and demonstratives precede the head noun and agree with it in noun class (CM); the agreement marker is realised as a suffix, unlike the cases of agreement on the verb, on which see § 2.2.5. This is exemplified in (24) for adjectives and (25) for demonstratives, the agreement marker appearing in boldface.

(24)  a.  lik’a-w was ‘good–m boy’
      b.  lik’a-j jas ‘good–f girl’
      c.  lik’a-b žo ‘good–n thing’

\(^ {11} \) I prefer, for the time being, to refrain from joining the debate as to whether the Avar noun phrase projects a (possibly null) determiner, thus being structurally parallel to a clause (Abney 1987), or whether the D-layer is absent from the structure altogether, since it is difficult to evaluate the impact either view may have on the main subject matter of this thesis, perhaps with the exception of determiners probably not being syntactic heads, as that would make Avar head-initial in the nominal domain.

\(^ {12} \) I also leave the discussion of the ordering restrictions on the placement of the various elements internal to the noun phrase, as well as the mechanisms underlying these restrictions, to future work.
2.2. Overview of Avar grammar

(25) a. he-w was ‘that–m boy’
    b. he-j jas ‘that–f girl’
    c. he-b žo ‘that–n thing’

Adjectives and determiners, although possessing full declension paradigms in both singular and plural, always appear in their absolute form when modifying a noun phrase regardless of the case marking on the noun, their oblique forms being reserved for standalone uses.

(26) a. χwali- ca hiq’-ana χera- w či- jas- da
deanth.obl-erg ask- pst old- m:abs man-obl-loc
‘Death asked the old man.’
    b. "χwali- ca hiq’-ana χera- s- da či- jas- da
deanth.obl-erg ask- pst old- obl-loc man-obl-loc
    c. χwali- ca hiq’-ana χera- s- da
deanth.obl-erg ask- pst old- obl-loc
‘Death asked the old one.’

As can be observed from the three sentences above, the locative case marker on čijasda ‘man.loc’ must not be shared by the adjective (26a–b) unless the adjective is coerced into a nominal, or substantivised (26c). In this respect Avar is different from certain other languages with adjective declensions such as Russian or Estonian, where the modifying element shares the case of the noun.

It seems to me that this is the right moment to give a sketch of how this pattern can be given a very natural treatment in a Layered-Derivational framework of Zwart 2009 et seq., before proceeding with the grammatical description.

The pattern to be derived is the contrast between (26a) and (26b), where the adjective χéra-cm may only appear in the absolutive case when modifying a locative-marked noun. Suppose, following Zwart 2009, that the noun phrase old man is created, from a numeration of its own (any one of those given in (27), or something similar in spirit, depending on one’s favourite analysis of the noun phrase), in a derivational layer distinct from the one where the rest of the clause is being derived.

(27) a. {√old, √man, a, n}
    b. {old, man}
    c. {Adj, N}
Merge will then apply to the elements of the numeration — which I have chosen to look like (27d) — in any possible order, and create a number of hierarchically structured expressions that will, or fail to, receive an interpretation at the interfaces as outlined in §2.1 above. The crucial point here is that the numeration being exhausted, the output of merge can be put into the next derivational layer — the one corresponding to the vP, for concreteness — as a single opaque object. Its numeration is given in (28):\(^{13,14}\)

\[
\begin{align*}
\{ \text{App}, \nu, \sqrt{\text{ask}}, \text{pro, death, } & \text{xeraw, } \\
\text{či} & \\
\} 
\end{align*}
\]

Let us make a further assumption that the argument corresponding to the addressee of the question in (26) gets its LOC case in the same derivational layer; given the numeration in (28) case assignment will most plausibly be done by the applicative head App, but since the complex internal structure of the noun phrase is invisible to the case assigner, the case marker is affixed to the rightmost element of the complex noun phrase (presumably in the post-syntactic component, since for the purposes of the narrow syntax such notions as ‘prefix’, ‘word’ or ‘suffix’ are devoid of any meaning).

Having presented a sketch of how the Layered-Derivations framework might provide one with a handle on the absence of case marking on Avar adjectives, let us go back to the nominal domain and consider another type of nominal modifiers, relative clauses.\(^{15}\)

---

\(^{13}\) The sentence under consideration contains one complication in the form of pro-drop of the internal argument of hiq‘- ‘ask’. To see that it can be there, consider (i), where it appears, overtly, as lalareb żo ‘unknown thing’:

(i) la- la- r- e- b żo lal- e- s- da hiq‘e

know-PRES-NEG-PTCP-N thing.ABS know-PRES-PTCP-OBL-LOC ask-IMP

‘What you don’t know ask the person that does.’ (Gimbatov 2006)

\(^{14}\) That on its way from the first numeration to the second the adjective has transformed from xera–CM to xeraw by undergoing morphological agreement with the (masculine) noun. Since the current thesis sidesteps the issue of the place of agreement in the architecture of the grammar entirely, I prefer not to spell out the mechanism, suffice it to say that it is only natural for it to be confined to a particular derivational layer.

\(^{15}\) Another type of nominal construction — proper names — behave, from the point of view of case marking, in a fashion identical to regular modified nouns in that only the rightmost element inflects for case, all the other elements to its left in their default (i.e. absolutive) form. It should therefore be fairly straightforward to extend the Layered-Derivational analysis sketched...
Relativisation, which is the subject matter of Chapter 3, proceeds with a gap in the relativisation site, meaning that relative clauses feature neither a relative nor a resumptive pronoun. The verb takes on participial morphology, the participle patterning with the other modifiers in not realising, morphologically, the case of the head noun (the underscore in 29 corresponds to the gap left by the relativisation of the agent):

\[(29) \quad \text{narkotikal r- ič- ul- e- w či- jas} \]
\[
\text{drugs.ABS PL-sell-PRES-PTCP-M man-ERG}
\]

‘drugs dealer’ (lit.: ‘drugs-selling man’)

As can be seen from (29), the participle has two positions for agreement markers—a prefix and a suffix, which makes the participles’ “dual” (both verbal and adjectival) nature all the more salient: the agreement prefix, as well as the agreement trigger, correspond to the ones of the verb, whereas the agreement, or perhaps concord, suffix is that of the adjective. In the example at hand the verb cm-ić- ‘sell’ agrees with its plural-marked internal argument, narkotikal ‘drugs’, whereas the participle, by virtue of modifying a masculine head noun, appears with a masculine suffix.

out above to the declension of proper names, as well, which would have to be generated in a separate derivational layer before appearing, as an atomic item, in the position where they get morphological case. This is illustrated, exclusively for presentational reasons, in (i) for the ergative and the absolutive only.

(i) ABS: ramazan ʕabdulat’ipov
ERG: ramazant*-as ʕabdulat’ipov-as

As the typological evidence would suggest, however, the two structures do not have to behave on a par, since we do find languages where adjectival modifiers inflect together with the head noun and whose proper names behave as their Avar counterparts. This is illustrated in (ii) for Estonian valge maja ‘white house’ and the name of the incumbent President of Estonia.

(ii) Estonian common and proper nouns

\[
\begin{array}{ll}
\text{NOM} & \text{valge maja} \quad \text{Toomas Hendrik Ilves} \\
\text{GEN} & \text{valge maja} \quad \text{Toomas Hendrik Ilvese} \\
\text{PTV} & \text{valget maja} \quad \text{Toomas Hendrik Ilvest} \\
\text{ILL} & \text{valgesse maja} \quad \text{Toomas Hendrik Ilvesesse} \\
\text{INESS} & \text{valges majas} \quad \text{Toomas Hendrik Ilveses} \\
\text{EL} & \text{valgest majast} \quad \text{Toomas Hendrik Ilvesest} \\
\text{ALL} & \text{valgele majale} \quad \text{Toomas Hendrik Ilvesele}
\end{array}
\]

Moreover, the elements conventionally characterised as ‘modifiers’ do not have to behave in a uniform fashion either. To stay with Estonian, where adjectives do inflect for case together with the head noun, as shown above, it is clear that certain possessors in that language are obligatorily invariant. This heterogeneous behaviour can be taken to suggest that derivational layers need not be opaque to interface processes.
2.2.4.2 Case marking

Avar is a morphologically ergative language to the extent that all noun phrases, as well as pronouns, deverbal nouns and nominalisations are capable of assuming the ergative case marking. Unlike in many ergative languages, there is no tense/aspectual split in Avar ergative alignment — descriptively the direct object (and intransitive subject) always appears with nominative/absolutive case marking, except in the case of causativisation.\footnote{Readers familiar with Coon (2013a,b,c) might find this statement surprising, since these works cite Northeast Caucasian amongst languages featuring TAM-split ergativity. It is, to the best of my knowledge, still an open question whether the so-called biabsolutive construction present, to an extent, in all Northeast Caucasian languages is an instance of an aspect-based split (Forker 2012; Gagliardi et al. 2014).}

The core cases, to the exclusion of the absolutive, which is the citation form, and their affixes, are represented in Table 2.2 — in addition to these Avar, as is typical of Northeast Caucasian languages, has several locative cases with distinct locative series, depicted in Figure 2.3 (Creissels 2008, Daniel & Ganenkov 2009), which are analysed, from a nosynthetic perspective, in Pantcheva (2011). Unlike ABS, which corresponds to the direct stem, all the other case forms are based on either the oblique stem of the relevant nominal discussed on p. 20 above or, as in the case of most of the locatives, on the LOC/SUPESS stem, which is in turn derived from the oblique stem (see Appendix A for examples of declension). In what follows I provide a brief outline of the uses to which some of the cases are put.

The prototypical function of the ergative case is to mark the subject of a transitive clause (the A-argument, to use the typological term), whereas the absolutive is reserved for both P- and S-arguments (the patient/theme argument in a transitive clause and an intransitive verb’s only argument respectively, cf. Comrie 1978). The case marking in the transitive context is illustrated in (30a), whereas the absolutive marking on S is shown in (30b).

\footnotesize
(i) a. hel nuχ ha–b-ul e–l r–ugo  
3PL.ABS way.ABS make–PRS–PTCP–PL be.PRS  
‘They are in the state of building a road. They build a road.’

b. hez nuχ ha–b-ul e–b b–ugo  
3PL.ERG way.ABS make–PRS–PTCP–N be.PRS  
‘They are building the road.’  
\cite{Bokarev(1949:113), cit. ex Forker(2012:81)}

Given the absence of a clear-cut, precise definition of a split-ergative system I follow the tradition of Caucasian linguistics in treating Avar ergativity as unsplit. The same view is taken in a recent article on Avar relativisation (Polinsky, Gallo, et al. 2012).
2.2. Overview of Avar grammar

<table>
<thead>
<tr>
<th>Case</th>
<th>Exponent</th>
<th>Example</th>
</tr>
</thead>
<tbody>
<tr>
<td>Ergative</td>
<td>-∅//-ca</td>
<td>(30a)</td>
</tr>
<tr>
<td>Genitive</td>
<td>-ul/-l</td>
<td>(34)</td>
</tr>
<tr>
<td>Dative</td>
<td>-(j)e</td>
<td>(35)</td>
</tr>
</tbody>
</table>

Table 2.2: Core cases in Avar

(30) a. χan- as hes- uqe či w–it'– ana khan-ERG he.OBL-APL man.ABS m–send-PST
   ‘The khan sent a man to him.’ (Alekseev & Ataev 1997: 45)

b. son c’ad b–ana yesterday rain.ABS N–go.PST
   ‘It rained yesterday.’

The absolutive case is also assigned to the nominal part of the predicate in copular constructions, as is the case for both muhamad ‘Muhammad’ and učitel ‘teacher’ in (31), where the nominal part of the predicate is učitel:

(31) muhamad učitel w–ugo Muhammad.ABS teacher.ABS m–be.FRS
   ‘Muhammad is a teacher.’

As is typical of erg-marking (Dixon 1994: 57), Avar ergatives also function as instrumental arguments, or, put differently, Avar displays an ERG/INS syncretism (Palancar 2009).

(32) ebel- al han noso- ca q’ot’– ana mother-ERG meal.ABS knife-ERG chop-PST
   ‘Mother chopped the meat with a knife.’

Besides the instrumental, the ergative case also appears on certain adjuncts (reason in (33a) and time in (33b)):

(33) a. Šemer kwana-jaʃ čax- al= gi unt- ulaan much food- ERG tummy-PL=CNJ hurt-PST:ITER
   ‘And (their) tummies hurt because of the abundance of food.’

b. heb= go sordo-jaʃ wac– al roq’o– r– e ʃ:w–ana that=EMPH night-ERG brothers-ABS home.OBL–PL-Ill get-PST
   ‘Brothers got home that same night.’ (Alekseev & Ataev 1997: 45)

Turning now to the core cases that remain (i.e. the genitive and the dative), their functions are as follows. The genitive marks what could very loosely be
referred to as the possessor in a variety of constructions (e.g. possessive proper and part–whole relations, both illustrated in (34a) by a line from a poem whose authorship I have not been able to establish);\(^\text{17}\) in addition to that, it is also assigned to the complement of a limited number of adpositions (on which see §2.2.4.4 below), as illustrated in (34b) for haq‘ałul ‘about’.

(34) a. dir k‘odo-\(^\text{1}\) ca ah- a- ra- b koč’-\(^\text{2}\) ol raʕ-\(^\text{1}\) abi
\(^{\text{1}}\text{1SG:GEN grandmother-\text{\textit{erg}}} \text{\textit{shout}}\text{-\textit{PST-PTCP-N}} \text{\textit{song}}\text{\textit{OBL-GEN word-PL}}\)
kkő-\(^\text{2}\) la
hold\text{-\textit{PRS}}
‘(I) remember the lyrics of the song sung by my grandmother.’

b. dayıstanal-\(^\text{1}\) ul kulturajal-\(^\text{2}\) ul wa etnografijał-\(^\text{1}\) ul
\(^{\text{1}}\text{Daghestan}\text{\textit{OBL-GEN culture}}\text{\textit{OBL-GEN}}\)
haq‘ałul dokumentali-\(^\text{2}\) jal film-al
\(^{\text{2}}\text{documentaries about the culture and ethnography of Daghestan’}
\text{\textit{about}}} \text{\textit{documentary-PL film-PL}}\)
\text{(http://hakikat.info/6-test.html)}

The (a) sentence above contains two \textit{GEN}-marked noun phrases, both corresponding to possessors of sorts: \textsl{dir} ‘my’ in \textsl{dir k‘odo} ‘my grandmother’ is a true possessor, whereas \textsl{koč‘ol} ‘song.\textit{GEN}’ stands for the ‘whole’ in a part–whole relation. The possession relation (between Daghestan and its culture and ethnography) is expressed in (34b) as well, but there we also see that the coordinated noun phrase in the genitive case is itself the dependent of a postposition.

Finally, the Avar \textbf{dative} marks either the goal/recipient argument (35a) or an experiencer subject of certain verbs such as \textit{cm–ół} ‘love’ (35b).

2.2. Overview of Avar grammar

(35) a. di- ca kin du- e hab b-ec' il- e- b
    1SG:OBL-ERG how 2SG:OBL-DAT this.ABS N–repay-FUT-PTCP–N
    ‘How will I give this back to you?’ (Alekseev & Ataev 1997: 84)

b. di- e mun w–ol' ula.
    1SG:OBL-DAT 2SG:ABS M–love–PRS
    ‘I love you.’

We will see in §2.3.1 that DAT-marked subjects of experiencer verbs behave like their ergative counterparts in the agentive environments with respect to such structure-sensitive phenomena as variable binding, causativisation, control and raising.

The core cases having been introduced, a few words should be said about the locative cases, some of which are frequently used in non-locative environments (see below). Similarly to the dative, the Avar locative (which is, strictly speaking, supess but which I from now on gloss as LOC) can be used to mark external arguments of certain non-agentive verbs, as shown in (36a) for cm–ix–‘see’ (some others belonging to this category are la–‘know’, boz–‘believe’, kwe–‘think’ etc.). In addition to the properly locative, its other functions are marking the applicative argument of certain verbs such as hig–‘ask’, see (36b), repeated from before, as well as expressing a temporal meaning (36c).18

(36) a. di- da heb b–ix– ana
    1SG:OBL-LOC that.ABS N–see–PST
    ‘I saw it.’

b. χwali- ca hig’–ana χera–w či– jas– da
    death.OBL-ERG ask–PST old– M man-OBL-LOC
    ‘Death asked the old man.’

    ‘That film was released in 1992.’

Just like LOC, the other locative cases are often used outside of purely locative

18. Given the information on the semantics of the ergative case marker provided at the beginning of this subsection, one might wonder if the temporal semantics can be expressed by marking son ‘year’ with the ergative case. This is indeed correct:

(i) heb kino b–aq– un b–uk’–ana 1992 son– al–
    ‘That film was released in 1992.’

The extent to which the distribution of temporal expressions with erg- and loc-case marking overlaps is to be further investigated.
contexts. In comparative constructions, for instance, the object of comparison carries an elative case marker to express the comparative degree (there being no comparative marking on the adjective):

\[(37) \quad \text{wac- as- da- sa jac lik’a- j j–igo} \]
\[
\text{brother-OBL-LOC-EL sister.ABS good-F F-be.PRS}
\]
\[
\text{‘Sister is better than brother.’} \quad \text{(von Uslar 1889: 91)}
\]

Such locative cases as subelative or apudessive mark the oblique arguments of hinq’- ‘fear’ (38a) and cm–alah- ‘look’ (38b) respectively:

\[(38) \]
\[
\begin{align*}
\text{a. } & \quad \text{sundu- l’-a mun hinq’-a- ra- w} \\
& \quad \text{what.OBL-SUB-EL 2SG:A BS fear- PST-PTCP-M}
\end{align*}
\]
\[
\text{‘What are you afraid of?’}
\]
\[
\begin{align*}
\text{b. } & \quad \text{kinalgo passažir- al hesu-q r- alah-ana} \\
& \quad \text{all.ABS passenger-PL he- APRESS M-look-PST}
\end{align*}
\]
\[
\text{‘All passengers looked at him.’}
\]
\[
\text{(http://hakikat.etnosmi.ru/one_stat.php?id=9380)}
\]

This concludes the necessarily brief introduction to the syntax of Avar noun phrases before we return to locativity in §2.2.4.4.

### 2.2.4.3 Pronouns

I begin the discussion of Avar pronouns with reflexives — a topic that I have addressed previously (Rudnev 2010, to appear, 2011).

#### Reflexive pronouns

Avar reflexives come in three varieties: a reduplicated reflexive, źi–w=go źinca=go, that is strictly local (39)\(^19\), a compound reflexive źi–w=go that can be both local and long-distance ((40) illustrates long-distance uses only); a simplex reflexive źi–w which is strictly non-local and (arguably) logophoric (41). The reduplicated and go-reflexives are both derivatives of the simplex reflexive formed by reduplication in the case of źi–w=go źinca=go and by adding an emphatic particle, =go, to form the compound reflexive źi–w=go.\(^20\) This distributional

\(^{19}\) When such a reflexive is used, which means in (almost) all local contexts, its reduplicated components can come in either order.

\(^{20}\) As in many languages of Dagestan, Tsakhur or Bagwalal, for example, one component of the reduplicated reflexive bears the case marking of the antecedent with the other element absolutive-marked.
2.2. Overview of Avar grammar

pattern of various reflexive-like elements holds of the majority of other North-east Caucasian languages as well (see Testelec & Toldova 1998 for an overview; as for descriptions and analyses of reflexivity in separate Daghestanian languages, see Toldova 1999 on Tsakhur, Lyutikova 2001 on Bagwalal, Khalilova 2009: §3.5 on Khwarshi).

(39) ʕali-ça žinca= go żi- w=go č’w-ana
Ali-erg self.erg=emph self.abs=m=emph kill-pst
‘Ali killed himself.’

As can be seen from (39), Avar reflexive pronouns inflect for case, and when in abs also carry a noun class marker. Since in all the other cases agreement marking is absent, and because the non-reduplicated complex reflexive can have non-local antecedents, sentences like (40) are referentially ambiguous.

(40) ebelal- da b–ix- ana maliki- ca žindi-e= go ruq’
mother.obl-loc n–see-pst Malik.obl-erg self= dat=emph house.abs
b–a l e– b
n–build-prs-ptcp–n
‘Mother₁ saw Malik₁ build her₁/himself₂ a house.’

(41) pat’imati-ca ab–una šadžati- da žindi-e čaj t’e= jilan
Patimat-erg say-pst Khadizhat-loc self= dat tea.abs pour.imp=compl
‘Patimat₁ told Khadizhat₂ to pour her₁/₂ some tea.’

In my previous work I have argued that this three-way distinction is illusory and that Avar only has one type of reflexive — žiwgo, which in both local and long-distance configurations is obligatorily interpreted as a bound variable. I have also argued that the simplex žiw is in fact a bona fide logophoric pronoun akin to those in African languages (Hagège 1974).

Foreshadowing the discussion of the configurational structure of the Avar clause in §2.3, the behaviour of reflexive pronouns in the contexts of variable binding will provide important evidence against the competing, non-configurational, approaches.

Reciprocal pronouns

Yamada (2013) presents a very detailed description of the syntax of reciprocal binding. Even though the analysis of the behaviour of reciprocal pronouns is not of direct concern to the present thesis, I choose to dedicate a couple of remarks to them, mainly because they appear to contradict, very directly, the
well-established typological generalisation that reciprocal pronouns must be c-commanded by their antecedent.

The main reciprocalisation strategy in Avar is to employ a reciprocal pronoun coco ‘one another’, which is a reduplicated form of co — the numeral meaning “one”. The reciprocal pronoun inflects for case in accordance with the declension rules in §2.2.3 and 2.2.4 but not for number, in the sense that it only inflects in the plural.21

(42) a. [šamil= gi ʕumar= gi ] coco- da r- ix- ana
   Shamil.ABS=CNJ ʕumar.ABS=CNJ RECP-LOC PL-SEE-PST

b. *[šamili- da= gi ʕumari-da= gi ] coco- l r- ix- ana
   Shamil-LOC=CNJ ʕumar- LOC=CNJ RECP-ABS PL-SEE-PST

‘Shamil and Omar saw each other.’

The coordinated noun phrase is the reciprocal pronoun’s antecedent, yet in (42a) it is marked with the absolutive case, the locative case marker, which is typically assigned to subjects of certain experiencer verbs, being carried by the reciprocal pronoun itself. Similarly, when an agentive transitive verb is being reciprocalised, it is the reciprocal pronoun that carries the ergative case marking, and the absolutive case is reserved for the antecedent. Although typologically rare, this behaviour of Avar reciprocal pronouns is unlikely to shed new light on the proper treatment of ergativity since, as Yamada (2013) convincingly argues, ergative case marking on the reciprocal pronoun is distinct from the regular ergative marking. The construction itself, however, remains by and large unanalysed, although I do not see any immediate problems for semantic accounts of reciprocity such as Dotlačil (2013).

**Personal and demonstrative pronouns**

Avar only has personal pronouns for the 1st and the 2nd person in both sg and pl, all of which have full inflectional paradigms with respect to case marking, just as their reflexive counterparts. Instead of 3rd person pronominals of the he type in English, demonstrative pronouns are employed. Crucially, and unlike he in English, they cannot be bound by a c-commanding quantifier phrase.22

---

21. Avar has two other reciprocalisation strategies — reduplication of the pronoun coja ‘one of the members’ and the use of dand=, a preverb expressing symmetrical reciprocity, — both of which are much less frequent (as well as less exotic-looking) than the one illustrated in the main text. I refer the interested reader to Yamada’s (2013) article for examples and discussion.

22. Given this duality of function of Avar demonstratives, they appear, in the examples in this thesis, glossed as either pronouns (he, she, it, etc.) or as demonstratives, depending on the
2.2. Overview of Avar grammar

<table>
<thead>
<tr>
<th></th>
<th>M</th>
<th>F</th>
<th>N</th>
<th>PL</th>
</tr>
</thead>
<tbody>
<tr>
<td>Proximal</td>
<td>haw</td>
<td>haj</td>
<td>hab</td>
<td>hal</td>
</tr>
<tr>
<td>Medial</td>
<td>hew</td>
<td>hej</td>
<td>heb</td>
<td>hel</td>
</tr>
<tr>
<td>Distal</td>
<td>dow</td>
<td>doj</td>
<td>dob</td>
<td>dol</td>
</tr>
</tbody>
</table>

Table 2.4: Avar demonstratives

Demonstrative pronouns vary along a variety of dimensions, the principal one being the relative-distance-from-\( x \), \( x \) either the speaker or the addressee. From this perspective Avar demonstratives can be split into proximal, distal and medial, similar to the this–that–yon opposition in some dialects of English. Demonstratives, like reflexives, display full case-marking paradigms; unlike reflexives, however, their case forms in oblique cases distinguish noun class as well. What is crucial for the three demonstratives in this group is that the object they modify appears on roughly the same horizontal level as either the speaker or the addressee. Their absolutive forms are listed in Table 2.4.

Another dimension of variation, in addition to what can be called proximity, involves the level at which the noun phrase modified by the demonstrative is located with respect to the speaker. If the object’s location is higher than that of the speaker, \( yo\text{-}CM \) is used; if the relation is the opposite one, \( lo\text{-}CM \) is to be employed. Just as was the case with the proximal, distal and medial demonstratives, these, too, have full inflectional paradigms as far as noun class, number and case are concerned.

Avar demonstratives having been introduced, let us turn now to the remaining personal pronouns — 1st and 2nd person pronouns, or indexicals. Their properties are as follows.

Firstly, they can be used referentially, just like first-person pronouns in other languages, such as \( I \) in English, by picking up the contextually relevant referent that includes or corresponds to one of the speech act participants.

(43) a. pavel b–ugo dir c’ar
Pavel N–be.PRS 1SG:GEN name
‘My name is Pavel.’

b. niže- ca ruq’ b–a- l- e- b b–ugo
IPL.OBL-ERG house.ABS N–build-PRS-PTCP–N N–be.PRS
‘We are building a house.’

syntactic environment where they appear.
Secondly, and probably because the language lacks 1st and 2nd person reflexives, Avar indexicals can be used locally in contexts of reflexivisation.

(44) a. du-da mun w-ix- un w-uge-w
   'Have you seen yourself?'

b. di-e dun w-ol'– ula
   1SG:OBL-DAT 1SG:ABS M–like–PRS
   'I like myself.'

Thirdly, these same indexicals can be interpreted as bound variables, which can be seen from the availability of ‘sloppy’ readings that obtain under ellipsis (such indexicals are also known as ‘fake indexicals’, possibly because the person features on them are not semantically interpreted).

(45) di-ca dir tušman Ĕ-w-ana, hedingo muradi- ca= gi
    1SG:OBL-ERG 1SG:GEN enemy.ABS kill–PST too Murad.OBL-ERG=CNJ
    'I killed my enemy, and Murad did too.'
    = Murad, killed his, enemy [binding]
    = Murad killed my enemy [coreference]

The sentence in (45) features dir—a GEN-marked indexical pronoun—as the possessor in a possession relation, and the pronoun in the ellipsis site should, at least if we take the ellipsis parallelism more or less seriously, be identical, in one way or another, to the pronoun in the antecedent clause (Fiengo & May 1994). The availability of the bound-variable reading (x killed x’s enemy) in addition to the purely indexical reading whereby dir picks out the speaker is at least as puzzling for the purely indexical theories as the presence of the same reading in the English translation.

Finally, when embedded under attitudinal predicates such as ab–‘say’ in (46), Avar indexicals can undergo a reference shift, which means that they may optionally be interpreted relative to the embedded context rather than only being sensitive to the circumstances of evaluation.

(46) [marija dir c’ar= ] ilan ab- una he- l.
    Maria.ABS 1SG:GEN name= COMP say–PST she–ERG
    'She said [my/her] name was Maria'

The embedded clause in (46) contains a first-person possessive pronoun dir ‘my’, which can be interpreted as referring to either the person uttering (46), or the speaker of the saying event, i.e. hel ‘she.ERG’.

Both (45) and (46) present a problem for mainstream accounts of index-
2.2. Overview of Avar grammar

icality whereby indexicals are evaluated relative to the utterance context (D. Kaplan 1989) simply because it is true of neither case. In (45) the ϕ-features on the pronoun do not seem to be interpreted at all, whereas in (46), on the non-canonical reading, the possessive indexical dir ‘my’ embedded under an attitude predicate is interpreted relative to the embedded context rather than the utterance context.

2.2.4.4 Adpositions

We have seen in §2.2.4 that a variety of relations, mostly locational with respect to space, find their expression in Avar in the case morphology. In addition to the locative case series (illustrated in Table 2.3 on p. 27), however, the language also has a different sort of adpositions — postpositions — expressing locative, temporal, causal relations as well as some others. Besides these, Avar has various adverbs, including those with locative and temporal semantics.

The postpositions themselves are mostly homonymous with the adverbs of the same semantics, and have three important properties.\(^{23}\)

First, postpositions assign case to their complement, which is illustrated in (47) below. The assigned case is predominantly LOC, although we have already seen a genitive-assigning postposition — ḥaq’alul ‘about’ — during our discussion of the genitive case on p. 27 above.

(47) ɣor’ll ‘under’
  zobal- da ɣor’ll b–ugo ral’
  sky.OBL-LOC under N-be.PRS earth.ABS
  ‘The earth is under the sky.’ (von Uslar 1889: 231)

Next, postpositions undergo agreement with the ABS-marked argument in the clause, just like the verbs do, rather than an ABS-marked dependent within their c-command domain (provided they have a corresponding slot, see (48) for an illustration). Incidentally, this agreement pattern extends to some of the locative cases, such as INESS and ILL, as shown in the table.

---

\(^{23}\) The rule of thumb in distinguishing adverbs from postpositions is to establish whether the element under consideration is used in a particular environment with a dependent or on its own. If there is a dependent, we are dealing with a postposition, the standalone uses being characteristic of adverbs. This makes these adpositions similar to their counterparts in the Germanic languages, where superficially one and the same lexical item can be characterised as a preposition in one set of context (i.e. when having a dependent) and as a particle in another.
Both the verb (cm-uk’-) ’be’ and the postposition carry a noun class marker by virtue of having undergone agreement with the absolutive subject (masculine in (a) and plural in (b)). Observe that the plural marker — a suffix — on the postposition, –r, is not the same one as we saw earlier with adjectives, demonstratives and participles.

Finally, Avar postpositions can themselves inflect for case, or at least change the locative series, as can be glimpsed from (49).

The meaning contributed by the postpositions in (49) has changed from the purely locative one to a directional one, the relevant directional relations being “towards-the-object” in (49a), where we also see a change in case-marking on the postposition’s complement (from LOC to APx), and “from-the-object” in (49b).

This completes the presentation of the nominal domain. We can now turn to the Avar verb.

2.2.5 Verb

The present subsection is a brief description of the morphosyntax of the Avar verb. We begin by considering the categories constituting the extended projection, to use Grimshaw’s (1997) term, of the verb (§2.2.5.1), proceeding next, in §2.2.5.2, to review those derivational affixes which underlie the composition of deverbal forms, such as nominalisations (masdars), participles and conversbs, of relevance for the subject matter of this thesis. Appendix B contains derivational schemata for most of the synthetic forms as well as examples of those
forms for several common verbs featuring in this thesis.

To return to agreement, which has already been touched upon in §2.2.3 from the perspective of the noun phrase, Avar verbs are not uniform with respect to its morphological realisation. This non-uniformity is reflected in both the presence of an (overt) agreement marker and its position inside the verbal form in question. It should be added that, as the foregoing exposition has illustrated all of the patterns described below, I am not adding further examples.

As regards the presence of an agreement marker, a descriptive generalisation has it that, in the overwhelming majority of cases, verbs with consonant-initial stems (e.g. la- ‘know’, čwa- ‘kill’, qwa- ‘write’, l’e- ‘give’ etc.) systematically fail to display agreement, the exception to this generalisation being ha–CM- ‘do’, where the agreement marker does appear overtly. That the stem should begin with a vowel, on the other hand, is not sufficient for the verb to carry an agreement marker, as the case of, for instance, ab- ‘say’ or ah- ‘shout’, makes clear. But more often than not, verbs with vowel-initial stems, such as the already familiar cm–ix- ‘see’, cm–oł’- ‘like’, cm–ič- ‘sell’ or cm–uk’- ‘be’, tend to come equipped with an agreement marker.

The agreement marker itself, if present, is most of the time realised as a prefix — we have seen numerous examples of this already. It is much rarer that the marker of agreement on the verb takes the shape of a suffix, an example of such a pattern being ha–CM- ‘do’ as well as all of the causatives by virtue of their being built with its help.24,25

2.2.5.1 Verbal categories

Various verbal categories can be expressed in Avar as either synthetic or analytic forms of the verb, the distinction being somewhat blurred in certain cases. One of these unclear cases is causativisation, which, as already mentioned, was formed by combining a lexical verb with the “light” verb ha–CM- ‘do’. Recall from §2.2.2 the existence of a certain degree of inter-speaker variation with regards to the status of such constructions: one set of speakers treat the “light” verb as having incorporated into the more complex, and synthetic, causative

24. It is unlikely that the reasons precluding the agreement marker from appearing as a suffix lie in the particular shape of the marker itself, since the marker’s default place on, for instance, adpositions is the suffixal position, as we have seen in §2.2.4.4.
25. Other verbs with suffixal noun class markers include t’i–CM–lt’- ‘spread’ and t’o–CM–lt’- ‘see off’, both of which, in fact, consist of a preverb (t’i and t’o respectively) and CM–lt’- ‘send’ (Testelec 2008).
verb form, with others viewing the “light” verb as a full-fledged auxiliary, at least from the point of view of prosody. Let us review some of the categories starting with those that appear as synthetic forms.

**Synthetic verb forms**

Since we have already discussed causativisation, let us stay in what can, perhaps with a stretch, be called the v-domain of the clause, which is constituted, together with causativising morphemes, of markers of certain flavours of aspect.

The inceptive aspect is marked by adding the suffix -l- to the stem, the stem typically being a noun. The meaning conveyed is one of starting a new action.

(50) dos- ul pikru rit’uq-l- ana
he.OBL-GEN opinion.ABS true- INC-PST
‘His opinion was confirmed.’

The inceptive voice marker in (50) attaches to rit’uq ‘true’, verbalising it.

Another aspect realised synthetically is the iterative (51), which becomes antipassive when a transitive verb is involved (51b). The markers signalling such a transformation are many, the choice depending partly on the phonological make-up of the stem of the verb in question.

(51) a. muradi-ca t’ex c’al- ul- e- b b–ugo
Murad- ERG book.ABS read-PRS-PTCP-N N–be.PRS
‘Murad is reading a book.’

b. murad (*t’ex) c’al- d- ul- e– w w–ugo
Murad.ABS book.ABS read-ITER-PRS-PTCP-M M–be.PRS
‘Murad is studying.’

The addition of an iterative morpheme to c’al- ‘read' detransitivises it, which is reflected in the case marking on the argument.

We return to the remaining aspectual distinctions (progressive, perfect etc.) shortly when discussing the analytic verbal forms on p. 40.

26. I am not including the category of voice here because such voice, or diathesis, alternations, as, for instance, passivisation do not exist in Avar.
27. It is not entirely clear to me why the inceptive aspect should belong to the grammatical categories of the verb, since one of its most obvious properties is changing the word class from nominal/adjectival to verbal.
Moving on, and onto the inflectional domain, Avar tense markers can be split into present (traditionally referred to as general in the tradition of Caucasian linguistics), future and past (also frequently alluded to as aorist). Exponents for these vary depending, once again, on the conjugation class (see §B.2 for some examples of this variation); besides, certain verbs like cm–uk’- ‘be’ have a stem alteration in some of the tense forms. The infinitive, being distinct from the root, is also marked morphologically, the exact shape of the marker also being dependent on the conjugation class (see sentences (59) for illustrations). 28

(52) a. dun mac’al- ul ʕalimči w–ugo
    1SG.ABS language.OBL-GEN scholar.ABS M–be.PRS
b. dun mac’al- ul ʕalimči w–uk’-ana
    1SG.ABS language.OBL-GEN scholar.ABS M–be- PST
c. dun mac’al- ul ʕalimči w–uk’-ina
    1SG.ABS language.OBL-GEN scholar.ABS M–be- FUT
‘I am/was/will be a linguist.’

What (52) shows is that the present tense of cm–uk’- ‘be’ has a stem, ending in -g, that is different from the one used to host the past and future inflections. The tensed forms can themselves serve as stems for a number of verbal forms, a summary of which can be found in §B.1.

In addition to tense, the category of mood manifests itself in Avar, being represented, besides the indicative, by the imperative (53a), prohibitive (53b), three different optatives (54), and irrealis (55).

Forming the imperative depends on both the verb’s conjugation class and its transitivity; as for the prohibitive, its marker, -ge, is invariant and attaches to the verb’s present tense stem:

(53) a. a! c’al- el
    GO.IMP read-IMP
    ‘Go! Read!’
b. un- ge! c’alu- ge!
    GO.PRS-PROH read.PRS-PROH
    ‘Don’t go! Don’t read!’

28. Forker (2014) describes the general tense as expressing ‘characteristic properties or habitual situations’ without having a specific temporal reference. Since the tense marker that traditional grammars categorise as general in Avar nevertheless does have a specific reference to the present, I gloss it as PRS. The difference between it and the ‘regular’ present tense are most noticeable with the verb cm–uk’- ‘be’, which the reader is invited to consult in §B.2 of Appendix B.
Avar possesses three distinct forms of the optative, i.e. the forms intended at expressing the wish of the speaker, the differences between them remaining largely unexplored (Testelec 2008). The three forms are exemplified, in (54), with cm–aq‘ ‘fall’.

(54)  a. l’olo- l’a b–aq- ad
      saddle.OBL-SUBJ.LCL N-fall-OPT
      ‘May you die!’ (lit.: ‘May you be taken from under the saddle!’)

      b. k’jab=go b–aq-ja- w
         both= EMPH eye.ABS N-fall-OPT-M
         ‘May your eyes both fall out!’

      c. talaw b–aq-gi
         ulcer N-fall-OPT
         ‘May you be ridden with ulcers!’

A significant subset of contexts featuring an OPT-marked verb can be classified as maledictions.

We round off the discussion of the category of mood with irrealis frequently used in conditionals, as (55) demonstrates.

(55)  b–os- a- ra- b-ani bibliotekal-da- sa heb t’ex
      N-take-PST-PTCP-N-COND library.OBL-SUP-EL that.ABS book.ABS
      lik’ b–uk’-ina- an
      well N-be- FUT-IRR
      ‘If (you) took that book from the library, that would be lovely.’

      (Alekseev & Ataev 1997: 59)

There remains one category that we have not addressed yet, and that is polarity. I defer the discussion of negation, chiefly for presentational reasons, until §2.2.5.3, and proceed directly to an overview of the analytic forms.

**Analytic verb forms**

The two main groups in which the analytic forms can be divided are aspect and modality. I discuss them both in turn.

The perfect is formed by combining the converbial form of the lexical verb with cm–uk‘ ‘be’. If the auxiliary is in the present tense, the meaning that results is very similar to the resultative (56a), and if past, as in (56b), we are dealing with the pluperfect.
2.2. Overview of Avar grammar

The progressive form, which can both be present (57a) and past (57b), consists of a present participle of the lexical verb and cm–uk'-'be' in the present or past tense respectively.

(57) a. jacał t'ex c'al- ul- e- b b–ugo
girl.erg book.abs read-prs-ptcp–n n–be.prs
b. jacał t'ex c'al- ul- e– b b–uk'–ana
girl.erg book.abs read-prs-ptcp–n n–be-pst

‘The girl is/was reading a book.’

As regards the order of the elements of the analytical forms, the auxiliary need not always follow the lexical verb, as in (57), and can appear in a variety of positions in the clause:

(58) a. jacał t'ex b–ugo c'al- ul- e– b
girl.erg book.abs n–be.prs read-prs-ptcp–n
b. jacał b–ugo t'ex c'al- ul- e– b
girl.erg n–be.prs book.abs read-prs-ptcp–n

The meaning of the Avar progressive is similar to the English progressive. The much bigger issue of Northeast Caucasian verbal aspect has received very little attention in the literature.

Finally, prospective future is formed by combining the infinitive of the lexical verb with the present auxiliary to give the meaning similar to that of the futur immédiat in French, i.e. to describe an action or event that is about to happen. Alekseev & Ataev (1997) and Testelec (2008) report this form as contributing a deontic flavour, similarly to the is-to construction in English.

(59) q'isas b–os– ul– e– b q'o b–ač'– ine b–ugo
reckoning.abs n–take-prs-ptcp–n day.abs n–come-inf n–be.prs

‘The day of reckoning is going to come.’
In (59) the relevant form is *bačine bugo*.

Turning to modality, it is customary for the modal predicates such as *cm–oł’- ‘want’, kke- ‘have to’ etc. to combine with the infinitive of the lexical verb, as shown in (60a) for the epistemic reading of *beh- ‘may’, and in (60b) for its deontic interpretation.


b. hedina– b ak’al- da halt’–ul e w či č’uhara- w such – n place.obl-loc work-prs-ptcp–m man.abs arrogant–m w–uk’–ine beh- ula ro m–be- inf may-prs–neg ‘Someone working at such a place must not be arrogant.’

Having introduced some of the forms of the Avar verb, let us take a brief look at some of the language’s derivational morphology, and more specifically to various deverbal forms. We will, however, touch on both the modals and the infinitives again when discussing agreement patterns in the context of clausal complementation.

### 2.2.5.2 Derivational affixes

It is a well-established fact with an impressive amount of typological evidence that verbs may serve as a basis to derive other word classes such as (deverbal) nouns, participles, which in traditional grammar are frequently described as deverbal adjectives (62), and converbs, perhaps better known as adverbial participles or verbal adverbs. An example of a nominalisation can be seen in (61), from Dutch.

(61) a. een artikel schrijv-en det article write-inf ‘(to) write a paper’

b. het schrijven van een artikel det write.nmlz of an article ‘the writing of a paper’

A progressive participle is illustrated in (62), from English:

(62) a running boy
Finally, an example of a convert is presented in (63), this time from Russian.

(63) na- pisa- v pis’mo on uexal domoj [Russian]
ppfv-write-cvb letter.acc he.nom went home
‘Having written the letter, he went home.’

Avar allows all of these kinds of derivation. Deverbal nouns, which I refer to as masdars, following the tradition of Caucasian linguistics, correspond, morphologically, to the verb’s root followed by the thematic vowel (further followed, for some verbs and declension classes, by a glide or a nasal sonorant).

(64) aħmadi-ca gordu ganč’i-ca b–ek– i lik’a– b ẓo
Ahmed-pl erg window.abs stone-erg n-break-mds good-n thing.abs
b–ugo
n–be.prs
‘That Ahmed broke the window with a stone is a good thing.’

The masdar in (64) is b–eki, which is derived from cm–ek– ‘break’ by adding the thematic vowel -i. The Avar masdars’ verbal characteristics are reflected in their ability to undergo agreement with the absolutive noun phrase in their c-command domain (as made evident by the neuter agreement prefix b– on the masdar above). The masdars are fully nominal, however, in their external syntax, as it were, i.e. the properties and categories they may display — such as case marking — depending on the surrounding syntactic context. Since this chapter is intended to be but an introduction, I choose to move on to Avar participles now and discuss both internal and external syntax of masdars in §3.2 of the next chapter.

Avar participles, too, display heterogeneous properties in being verbal on the one hand (e.g. by undergoing agreement as if they were regular verbs) and either adjectival or nominal on the other.

(65) narkotik-āl r– ič– ul– e– b bak’
drug-pl abs-pl-sell-prs-ptcp-n place.abs
‘the place where the drugs are (being) sold’

The participial clause in (65) occupies the prenominal modifier position. Morphologically Avar distinguishes between present, past and future tense participles, and the participle ričuleb ‘selling’ in (65) contains the present tense morpheme -ul- that was introduced during the discussion of the category of tense on p. 38. The syntactic properties of participial clauses are addressed in detail in chapter 3.

Verbal adverbs, or converses, in Avar are derived by attaching dedicated
markers—either affixes or particles—to a selection of verbal or deverbal stems. These forms are typically interpreted as heading adjunct clauses of all manners and kinds. Let us consider (66) as illustrating temporal adjuncts (66a), as well as those of condition (66b) and reason (66c). I have bracketed the clausal adjuncts to enhance readability.

(66) a. [roq’o– w–uk’–ago ] c’al– ila dica t’ex
home–M m–be– cvb read–fut 1sg:erg book:abs
‘I shall read the book when I am home.’ (von Uslar 1889: 139)

b. [mun hani–w w–ugo– ni ] dun= gi w–uk’–ina
2sg:abs here–M m–be:prs-cvb:cond 1sg:abs=cnj m–be– fut
‘If you are here, so am I.’ (von Uslar 1889: 143)

c. [hani–w w–uge– hul ] ha–b–ula dica heb
here–M m–be:prs-cvb do–n:prs 1sg:erg that:abs
‘I do this because I am here.’ (von Uslar 1889: 140)

The convers in (66) differ from one another both with respect to the particular affixes (-go, -ni and -hul respectively, -hul apparently being the genitive-marked form of the present participle) and the stem to which those markers attach. The reader is invited to consult §B.1 of appendix B to see which verb derives from which stem.

2.2.5.3 Negation

The discussion of the category of polarity, and of negation more specifically, has had to wait until the present subsection because, morphologically speaking, it relies heavily on the familiarity with deverbal forms described in §2.2.5.2.

Sentential negation

Sentential negation in Avar can be realised by a number of distinct markers, depending on the tense of the event being negated. Suffixal -ro, for example, is added to present or future tensed forms to negate present (67) or future events (68).

that:abs pharmacy.obl-loc-trans n–sell:prs-neg
‘They do not distribute it through the pharmacy.’
(http://hakikat.info/276-samozdrav-gunar-tiokiab-apparat.html)
2.2. Overview of Avar grammar

(68)  
dica  hab mina  kidanigi  b–ič- ila- ro
1SG:ERG this house.ABS ever  N–sell-FUT-NEG
'I shall never sell this house.'

(www.youblisher.com/files/publications/64/378082/pdf.pdf)

It is perfectly transparent, then, that the form, or stem, hosting the negation marker is the one of the tensed verb.

The same negation marker may not be used, however, if one were to negate a past event (69a). Moreover, not only must the negation marker be a different one, -č'o instead of -ro, the stem that it attaches to is, unexpectedly, that of the masdar, and hence unmarked for tense (69c), whereas the combination of -č'o with the past tense stem is considered unacceptable (69b).

(69) a.  
*dica  hab mina  b–ič- ana-ro
1SG:ERG this house.ABS N–sell-PST-NEG

b.  
*dica  hab mina  b–ič- ana-č'o
1SG:ERG this house.ABS N–sell-PST-NEG

c.  
dica  hab mina  b–ič- i- č'o
1SG:ERG this house.ABS N–sell-MSD-NEG
'I did not sell this house.'

The two negation markers bear a certain degree of resemblance to the two negative copulas in the present tense, guro and heč'o, to which I return shortly. These seem to be described in the literature as suppletive stems of cm–uk’- ‘be’ (Alekseev & Ataev 1997), the present-tense copula being, for some reason, incompatible with the non-past negation marker -ro:

(70) a.  
dir  c’ar  pavel  b–ugo
1SG:GEN name.ABS Pavel.ABS N–be.PRS
'My name is Pavel.'

b.  
*dir  c’ar  muhammad  b–ugo – ro
1SG:GEN name.ABS Mohammed.ABS N–be.PRS-NEG

c.  
dir  c’ar  muhammad  guro
1SG:GEN name.ABS Mohammed.ABS COP:NEG:PRS
'My name is not Mohammed.'

29. See Rudnev (in prep.) for an analysis of the -č'o marker as projecting a locative-like structure (cf. Salanova 2007, 2011 for a similar approach to the morphosyntax of negation in Mebengokre), in which -č'o is a variant of the locative negative copula heč'o that takes two dependents: a vP-level nominalisation (i.e., a masdar clause) and an abstract location. The second negation marker, -ro, is analysed as a regular sentential negation marker.
The distribution of the two negative copulas requires further investigation. Testelec (2008) cites a hypothesis due to Kalinina (1993), and based on the data from the Andalal dialect, whereby predicational, identificational and characterisational copular clauses allow both negative copulas, as opposed to the possessive and locative ones, where only heč'о is acceptable. Alekseev & Ataev (1997), on the other hand, name contrastivity as the main factor underlying the use of guro. Their examples do not, however, contradict Kalinina’s (1993) generalisation. To illustrate the pattern, I give two predicational sentences in (71), guro and heč'о being allowed in both cases.

(71) a. muḥamad učitel guro / heč'о
Muhammad.abs teacher.abs cop:NEG:PRS cop:NEG:PRS
‘Muhammad is not a teacher.’

b. he– w či lik’a– w guro / heč’о
that–m man.abs good–m cop:NEG:PRS cop:NEG:PRS
‘That man is not good.’

An illustration of the second part of Kalinina (1993) observation is provided in (72a), which involves a negated locative statement, and (72b) a negative possessive.

(72) a. rasul šahar-al- da heč’о / *guro
‘Rasul is not in town.’

b. ʕali-l ładi heč’о / *guro
‘Ali hasn’t got a wife.’

One other set of environments in which heč'о is systematically chosen over guro are negated analytic tense forms whose affirmative counterpart would have contained the auxiliary in the present tense (CM–ugo) (Alekseev & Ataev 1997: 77). This is shown in (73).

(73) amma nile-ca żaq’a hal- ul b–ic– ine heč’о / *guro
but 1PL-ERG today this.OBL-GEN N-speak-INF cop:NEG:PRS
‘But we are not going to discuss this today.’

(http://maarulal.ru/2009/12/26/) The analytic form of the verb in (73) corresponds to the one for prospective future, which normally consists of an infinitive and an auxiliary, as discussed

30. The judgements as presented here differ from the ones in Testelec (2008), where heč’о is marked as unacceptable in sentences (72a–b).
2.2. Overview of Avar grammar

in §2.2.5.1 on p. 40. When negated, the auxiliary must appear as heč’o and may not be realised as with guro.

Besides being used as the negative copula in the present tense, guro serves as the marker of constituent negation, to which we now turn.

Constituent negation

There are two things to be remembered about constituent negation in Avar. The first one is that it is marked by inserting guro to the right of the focused constituent. Secondly, and most importantly, the verb may no longer appear in its finite form and must become a participle instead. Examples of this are given in (74), where I have also included sentential negation to make the paradigm complete.

(74)  a. jac j– ač’– ana  
     sister.abs F–arrive–pst 
     ‘(My) sister has arrived.’

b. jac j– ač’– in– č’o  
     sister.abs F–arrive–msd–neg 
     ‘(My) sister has not arrived.’

c. jac guro j– ač’– a– ra– j / *j– ač’– ana  
     ‘It was not my sister that arrived.’

The relevant contrast is between (74a), where the verb carries a past tense marker (i.e. is finite), and (74c), where the same verbal form is disallowed. The only other element distinguishing the two sentences is the presence of the constituent negation marker guro in (74c) and its absence in (74a). We shall keep coming back to this issue in all of the subsequent chapters, ultimately reaching the conclusion that sentences involving constituent negation, as well as a number of other focus-sensitive expressions, project a bicausal cleft-like structure around a relative clause, participialisation being, in Avar, the hallmark of relativisation. This proposal is put forward in chapter 5.

Negating an infinitive

An additional negation strategy is involved when one attempts to use a negative infinitive, as deciding to group it with sentential or constituent negation is a non-trivial matter. The verbal form is, morphologically speaking, a temporal convert such as the ones we have seen in §2.2.5.2, and whose meaning can be
loosely paraphrased as ‘whilst in the process of P’, P being a placeholder for whatever verb happens to be negated.\textsuperscript{31} The more or less literal translation of (75), then, is something along the lines of Father allowed me to be in the process of not going to school, with (76) meaning Murad allowed his wife to put the khinkal in the state of not being made.

\begin{verbatim}(75) insuca w–ič- ana dun škola– de ø-in- č'o= go father.erg m–let-pst isg.abs school.obl-lat m–go.msd-NEG=prt w–uk'–ine m–be- inf ‘Father allowed me not to go to school.’
\end{verbatim}

\end{verbatim}

Variants of (75) and (76) containing any one of the negation markers attached directly to the infinitive that is being negated are unacceptable. The three sentences below are intended as expressing the same meaning as (75), and speakers have no doubts rejecting them all.

\begin{verbatim}(77) a. " insuca w–ič- ana dun škola– de ø–ine guro father.erg m–let-pst isg.abs school.obl-lat m–go.inf neg
c. " insuca w–ič- ana dun škola– de ø–ine– č'o father.erg m–let-pst isg.abs school.obl-lat m–go.inf-neg
('Father allowed me not to go to school.'
\end{verbatim}

All of these sentences are identical but for the actual negation marker: (77a) attempts to treat the negation as constituent negation by combining the infinitival clause with guro, whereas (77b) and (77c) append a sentential negation marker — non-past and past, respectively — to the verb carrying the infinitival morphology. The distribution of various negation markers in Avar is, though fairly peculiar, orthogonal to the main issue of this thesis.

\textsuperscript{31} I am grateful to Charlotte Lindenbergh (p.c.) for posing this question.
2.2. Overview of Avar grammar

Wh-questions

From the point of view of overt morphosyntax, Avar questions have at least one aspect in common with constituent negation, that aspect being the requirement that the verb may not remain in the same finite form as in the question’s declarative counterpart but should instead appear as a participle whenever a wh-phrase is present (cf. the unacceptable (78a) with (78b)).

(78) a. *kida jac j-ač’- ana
    when sister.ABS F-arrive-PST

b. kida jac j-ač’-a ra– j
    when sister.ABS F-arrive-PST-PTCP-F

‘When has (your/my/…) sister arrived?’

As for the position of the wh-phrases, these often appear at the left edge of the clause, as just illustrated. To anticipate the discussion in Chapter 4, this position does not correspond to the top link of an A-movement chain, instead being the predicate position of a pseudo-cleft.

The left-edge position, however, is not the only one available to Avar interrogative expressions, which can also remain in situ:

(79) a. mun šːaj ʔod-ul- e– j
    2SG.ABS why cry-PRS-PTCP-F

‘Why are you crying?’

b. du- ca niž bertal- e šːaj ah- i- č’- e– l
    2SG-ERG 1PL.ABS wedding-DAT why call-PST-NEG-PTCP-PL

‘Why didn’t you invite us to the wedding?’

c. di- ca kin du- e ha– b b– ec’- il- e– b
    1SG-ERG how 2SG-DAT this.ABS-N N-repay-FUT-PTCP-N

‘How will I give this back to you?’ (Alekseev & Ataev 1997: 84)

Interrogative phrases, moreover, can appear at the right edge of the clause, as evidenced by the acceptability of (80):

(80) du- e luh- a- ra– b žo šːi– b
    2SG-DAT happen-PST-PTCP-N thing.ABS what-N

‘What happened to you?’ (ibid.)

Multiple wh-questions are allowed, although reportedly dispreferred, and because of the preverbal surface positions of wh-elements it is often hard to see if any movement has taken place — a problem also characteristic of questions containing but one interrogative expression. I return to multiple wh-questions
Evidently, (81) above can in principle be understood as having several possible structures, i.e. with either one, both or neither one of the wh-phrases undergoing A-movement. As just mentioned, I argue in Chapter 4 that Avar wh-questions do not involve A-extraction of wh-phrases, regardless of the number of such elements in one given question.

Coming back to participialisation, it is not, strictly speaking, mandatory that every verb in every question should appear as a participle. Instead, the participial morphology on the verb seems to mark the scope of a question, as shown immediately below:

(82) jac- al- da la- ra [šːiw  w–ač’- a- ra– w]
    sister-obl-loc know-pres-NEG who.abs m-arrive-pst-ptcp-m
    ‘Sister does not know who arrived.’

The matrix verb in (82) appears in the present tense, giving rise to the embedded question reading; the verb in that question, on the contrary, must be a participle. This sentence containing an embedded clause, it seems an appropriate place to provide an overview of the strategies of clausal embedding available to an Avar speaker.

2.2.5.4 Clausal complementation

Avar allows for clausal complements to be embedded under the matrix verb via two distinct strategies—finite (with a complementiser) and non-finite. Unlike in languages like English, the use of the finite strategy of clausal embedding is severely restricted: it is only allowed for reported speech. All other embedding is non-finite, with the verbs in embedded clauses appearing as infinitive, masdars, participles, and, in the case of clausal adjuncts, as conversbs. In the examples that follow the clausal complements are bracketed to facilitate their identification.

32. It should be emphasised that I have been unable to verify the marginal acceptability of multiple questions reported by Alekseev & Ataev (1997), as speakers readily accept, and produce them during elicitation, so long as the necessary context is provided.
Non-finite embedded clauses can be further classified into infinitival (83), nominalised (84) and participial (82) varieties. Their subject position can be filled by either a null pronoun (ignoring for the moment the null pronoun’s status as a PRO or a pro) or an overt noun phrase.33

(83) a. untaras- e [χıırurg w–ix- ize ] b–ol’– ana
   sick.man-dat surgeon.abs m–see-inf n–want– pst
   ‘The patient wanted to see the surgeon.’

   b. dow [ęcord– eze ] ø–a– na
      he.abs swim–inf m–go–pst
      ‘He’s gone swimming.’

Embedded clauses behave very much like regular arguments, which can be seen in (83a), where they trigger gender agreement on the matrix verb. More concretely, the embedded verb in (83a), cm–ixize ‘see’, agrees with the absolutive object NP, χıırurg ‘surgeon’ thus taking the masculine agreement prefix. The matrix verb, too, has a slot for gender agreement, which is filled by the neuter prefix.

Certain matrix predicates are comparatively flexible with respect to the kind of clause they can take as their dependent. The verb la- ‘know’, for instance, is compatible with questions, as we have just seen in the preceding subsection, as well as masdars, as shown in (84a). It goes without saying that, just like know in English, la- ‘know’ can combine with non-clausal noun phrases.

(84) a. dos–da laraoan heb kayat heresija–b b–uk’–in
    he–loc know.pst.iter.neg that letter.abs fake– n n–be– msd
    ‘He didn’t know that the letter was fake.’

      1sg–loc pl–see–pst she–gen hands= cjn shake–prs–ptcp–pl
      r– uk’–in
      pl–be– msd
      ‘And I saw that her hands were trembling.’ (Alekseev & Ataev 1997: 105)

33. It is possible, in most cases, to identify which clause a particular noun phrase belongs to, as Avar displays the phenomenon similar to backward raising or backward control familiar from Tsez (Polinsky & Potsdam 2002, Potsdam & Polinsky 2012), by relying on the class of the verb. In (83a) the subject is in the dative case presumably assigned by the matrix verb cm–ol’- ‘want’. Had it been located inside the embedded clause, it would have to have been marked with the locative case, as was explained in §2.2.4.2. Deciding on the exact location of dow ‘he’ in (83b) is more difficult as in that case both verbs, by being intransitive, require an absolutive-marked subject. The constraints on these backwards-oriented phenomena in Avar are very poorly understood.
The matrix verb can, in addition to undergoing agreement with the masdar (which option was illustrated in (84a) above), agree with an absolutive argument contained within the masdar. It is this long-distance agreement that obtains in (84b) and accounts for the particular agreement prefix on CM–ix–‘see’, which would have to be neuter, had the verb only been able to agree with the masdar, but is in fact plural.

**Finite embedded clauses**

In reported discourse an embedding particle, which for the present purposes I take to serve as a complementiser, is adjoined to the immediate right of the embedded clause. These elements, =((j)ilan, =((j)in, =((j)an, are occasionally alluded to as quotative particles in the literature on indirect discourse and evidentiality.

The following are examples of attitude predicates that subcategorise for the finite clause with =((j)ilan: ab- ‘say/tell’, bic- ‘speak’, k’al- ‘say/talk’, har- ‘ask’, ahd- ‘yell’, šur- ‘whisper’, t’ad žub- ‘add’, lazab- ‘announce’ (technically a causative version of la- ‘know’), žawab l‘- ‘answer’ (lit. ‘answer give’). In (85), for instance, we are witnessing ab- ‘say/tell’ with a complement clause that is headed by =ilan, which cliticises to the embedded verb in the future tense.

(85) žin- ca ču b–ič- il= ilan ab- una wac- as self-ERG horse.ABS N–sell=FUT=COMP say-PST brother-ERG

‘Brother said that he would sell the horse.’

Interestingly, it is not just the finite forms of embedded verbs which =((j)ilan combines with, as it can also be adjoined directly to imperatives (86) and nominal predicates (87):

(86) nuž lik’ halt‘- e= jilan ab- una ebel- al 2PL.ABS well work-IMP=COMP say-PST mother-ERG

‘Mother told us to work hard.’

(87) asijat b–ugo dir c‘ar= ilan t‘ade=žub-ana he- l Asiyat.ABS N–be.PRS my name.ABS=COMP to= add-PST she-ERG

‘She added that [her/my] name was Asiyat.’

Observe that both (86) and (87) represent the so-called ‘shifted’ readings of indexical pronouns briefly mentioned in §2.2.4.3, whereby indexical elements get their interpretation with respect to the embedded context instead of the context of evaluation. As for the fact that =ilan has so flexible a distribution, it can, in principle, be taken as evidence of the embedded clause being represen-
2.3. Towards a configurational structure of an Avar clause

In discussing the theoretical significance of certain aspects of Avar morphosyntax, it is essential that two issues should be kept separate: (i) the existence of structural asymmetries between a verb’s arguments, and (ii) the cooccurrence, as well as ordering, restrictions on the position of affixes within an Avar morphological word. I address these three questions in §§2.3.1, 2.3.2 and 2.3.3 respectively. On a related note, a third point could be added concerning the ease of accounting for some of the aforementioned properties once a layered-derivational framework is assumed. This is discussed in §2.3.3.

2.3.1 Subject–object asymmetries

In this subsection I attempt to show that free word order notwithstanding, the Avar clause can still be treated as configurational to the extent that it dis-

34. There remains the possibility of these clauses being instances of mixed quotation, which have recently been argued to allow such operations as wh-movement to cross the quote’s boundary: (i) Who did Mary say that she would “never misunderstand ever again”? (Maier 2014: 7)

The question is whether a mixed-quotatiion account such as the one sketched by Maier (2014) can be extended to Avar =ilan clauses, and it is even less clear what the status of wh-extractions analogous to (i) is. However interesting, I defer the discussion of these questions to future work.

2.2.6 Summary

To sum up this section, I have sketched here an overview of the Avar grammar, concentrating mainly on the morphosyntactic peculiarities of the language in the nominal and verbal domain. We have also got acquainted with the default word order in the clause as well as deviations from that default order. We have, in addition, taken a cursory look at Å-dependencies, thus paving the path for the chapters to come, as well as introduced non-matrix clauses of various kinds that will also play an important part in establishing locality constraints on the derivation of operator–variable dependencies in Avar. Before such a derivation can be elaborated, however, a closer look at how these properties reflect the clause structure of the language is in order.

Toward a configurational structure of an Avar clause

In §3.2.4.5 we shall see preliminary evidence against such a conclusion, and establish that the =ilan clauses, unlike pieces of direct quotation, are transparent for such syntactic operations as relativisation.34
plays fairly robust asymmetries between arguments of transitive verbs. More specifically, I will show that ERG-marked NPs are hierarchically superior to ABS-marked ones. Before doing so I would like to summarise the best-known arguments against a configurational syntax of Northeast Caucasian languages from the literature.

Ergativity and agreement

The first such argument is rooted in morphological ergativity and agreement marking: indeed, the fact that subjects of transitive and intransitive predicates carry distinct case marking makes the application of constituency tests based on substitution difficult, as illustrated by the contrast between (88) from English and (89) from Avar.

(88) a. John died.
     b. John sells drugs.

In English an intransitive VP can be replaced by a transitive VP without triggering a change in case marking on the subject. This, however, is not the case in the ergative languages. 35

(89) a. ahmad yw-ana
     Ahmed.ABS die-PST
     ‘Ahmed died.’
     b. *ahmad narkotikal r- ič- ula
     Ahmed.ABS drugs.ABS PL–sell-PRS
     c. ahmadī-ca narkotikal r- ič- ula
     Ahmed-ERG drugs.ABS PL–sell-PRS
     ‘Ahmed sells drugs.’

The agreement problem is demonstrated by (89b), where the verb agrees with the P-argument in opposition to the English (88b). Besides these obvious cases there is also reason to believe that Avar might be a non-configuration language, to which we now turn. 36

35. It can be argued, as has been done by Kazenin & Testelec (1999) for Tsakhur, that distributional tests do in fact yield positive evidence for the VP being a constituent to the exclusion of the subject, with that evidence coming from the biabsolutive construction introduced on p. 25.
36. This is not to say that non-configurationality should be treated as a theoretical primitive or that it cannot be given a configurational treatment. On the contrary, there have been studies proposing very configurational analyses for languages like Passamaquoddy (Bruening 2001), Warlpiri (Legate 2002) and Kiowa (Adger, Harbour & Watkins 2009), to name just a few.
2.3. Towards a configurational structure of an Avar clause

Hallmarks of non-configurationality

Descriptively speaking, non-configurational languages are languages displaying the following characteristics (Hale 1983):

1. freedom of word order
2. freedom of argument omission
3. possibility of having discontinuous constituents

From what is known about Avar, it seems to fit this description of non-configurationality rather snugly: as regards the first hallmark, we have already witnessed the freedom of argument ordering in §2.2.2. The availability of pro-drop, illustrated in (91), can be taken to be analogous to the freedom of argument omission.

(91)

a. \( \text{pr} \) mašina b–ič- un, muradi-ca1 mina b–ana
car.abs n–sell-cvb Murad–erg house.abs n–build.pst

‘Having sold his car, Murad build a house.’

b. was-as1 mašina b–ič- un, insu- ca \( \text{pr} \)1 w–ux– ana
son–erg car.abs n–sell-cvb father–erg m–beat–pst

‘The son sold the car, and the father beat him up’ (Samedov 2003)

Furthermore, Kazenin (2009) reports, citing his own field notes but not adducing any data, Avar as allowing discontinuous constituents, albeit only noun phrases in the absolutive case. Avar thus displays all three hallmarks of non-configurationality, and given the absence of certain grammatical processes normally viewed as targeting exclusively the “subject” such as passivisation, the base clausal structure of most of the Northeast Caucasian languages can be argued to be as depicted in (92a), where all of the verb’s dependents are in a sisterhood relation, as opposed to the more asymmetric structure represented in (92b). Indeed, something like this “flat” structure is currently the consensus, at least in the typological literature, as to the general make-up of a Northeast Caucasian clause.37

37. This is indeed the set-up that has explicitly been argued for, for instance, Tsakhur in Kibrik (1999) and Bagwalal in Kibrik (2001).
The issue is exacerbated by the fact that such subjecthood diagnostics as control environments, which typically only target syntactic subjects, as exemplified in (93) for English, do not exclude absolutive arguments:

(93)   a. Mother sent her daughter to school
       b. Mother agreed [ PRO to send her daughter to school ]
       c. * The daughter agreed [ mother to send PRO to school ]

If, in Avar, the ergative argument is the syntactic subject, the prediction is for the Avar counterpart of (93c) to be as unacceptable. This prediction is not confirmed.

(94)   a. ebel- al jas [ _ c’al- ize ] j- it’- ana
        mother-erg daughter.abs PRO.abs study-inf f-send-pst
        ‘Mother sent her daughter to school.’

       b. ebel razil- ana [ _ jas c’al- ize j- it’- ize ]
          mother.abs agree-pst daughter.abs study-inf f-send-inf
        ‘Mother agreed to send her daughter to school’

       c. jas razil- ana [ ebel- al [ _ c’al- ize j- it’- ize ]
          girl.abs agree-pst mother-erg study-inf f-send-inf
        ‘The girl agreed for her mother to send her to study.’

The acceptability of (94c) contradicts the conclusion, formulated in Polinsky, Gallo, et al. (2012), that ‘[t]he ergative, and not the absolutive, participates in control structures’ (Polinsky, Gallo, et al. 2012: 270). Let us note, however, before taking this acceptability as proof of the absence of structural asymmetries between the transitive verb’s ergative and absolutive arguments, that it can have its roots in the availability of an alternative structure — the one involving object pro-drop, as sketched below:

(95)   jas razil- ana [ ebelal pro [ _ c’al- ize j- it’- ize ]
       girl.abs agree-pst mother.erg study-inf f-send-inf
Towards a configurational structure of an Avar clause

There are, however, several other subjecthood tests, with respect to which Avar ergative noun phrases unambiguously behave as c-commanding the absolutive noun phrases. First, another infinitival construction, raising (Keenan 1976), differs from control in only targeting subjects, which in Avar are hypothesised to correspond to ergative arguments of transitive verbs.

(96) a. šamili-ça tušman č’w-ana
   Shamil-ERG enemy.ABS kill-PST
   ‘Shamil has killed the enemy.’

   b. šamil luh-ana [tušman č’w-aze]
   Shamil.ABS begin-PST kill-INF
   ‘Shamil has started to kill the enemy.’

   c. ‘tušman luh-ana [šamili-ça č’w-aze]
   enemy.ABS begin-PST Shamil-ERG kill-INF
   ‘The enemy has started to be killed by Shamil.’

Second, the ergative and not the absolutive argument of a transitive verb is the addressee of an imperative — just as we would expect if the ergative argument were the canonical subject in nominative–accusative languages (although see Dixon 1994: §5.3 for arguments against this being a robust subjecthood test).

(97) pro t’ex c’al- e!
    pro.ERG book.ABS read-IMP
    ‘Read the book!’

In addition, it is the ERG-marked argument that changes its case (to LOC) when the verb is causativised, example (98b) repeated from p. 18 above. 38

(98) a. rasuli-ca kayat qw-ana
    Rasul-ERG letter.ABS write-PST
    ‘Rasul wrote a letter.’

38. This statement does not generalise across all possible ERG-marked noun phrases in the Avar language. One systematic exception involves ERG-marked reciprocal pronouns under causativisation, whose case-marking does not alter to LOC but remains ERG (Yamada 2013), as shown below.

(i) śisa-ca he- l coca-z / coca-ca č’wa-ze ha- r- una
    Isa-ERG that-PL.ABS RECP-ERG kill-INF CAUS-PL-PST
    ‘Isa made them kill each other.’  (Yamada 2013: 164)

This can hardly be considered to be countervidence to the claim that ERG-marked nominals c-command ABS-marked ones in transitive contexts, especially if Yamada (2013) is correct in treating predicates with reciprocal pronouns as intransitive.
The unacceptability of (98c) is unexpected if there is no asymmetry between the A- and P-arguments of qw-‘write’.

Finally, ergative arguments can bind absolutive arguments but the converse is not true (I provide examples involving quantificational antecedents to rule out the possibility of coreference which is frequently argued to be significantly less structure-sensitive than variable binding):39,40

(99) a. li- ca- nigī ži- w=go  č’w-ač’o
    who-erg-pol self=m=emph kill-pst.neg
    ‘Nobody has killed himself.’

b. *ži- ca= go  šiw- nigī č’w-ač’o
    self-erg=emph who.m.abs-pol kill-pst.neg
    (‘Nobody has killed himself.’)

The binding dependency between a quantificational antecedent and a reflexive pronoun can be established in (99a) but not in (99b), which is expected if the ergative argument c-commands the absolutive. The impossibility of a bound-variable interpretation appearing in (99b) is arguably due to the structural restrictions (=c-command) on variable binding, accidental coreference being ruled out by the non-referential nature of the antecedent. It is worth noting, too, that the unacceptability of (99b) cannot be the result of the reflexive preceding its antecedent, since the reversal of word order does not alter the judgement:

(100) *šiw- nigī ži- ca= go  č’w-ač’o
    who.m.abs-pol self-erg=emph kill-pst.neg
    (‘Nobody has killed himself.’)

39. It should be noted that both Toldova (1999) and Lyutikova (2001) provide hardly any evidence from variable binding; instead they use referential antecedents, which are independently known to be more liberal in the way they establish a coreferential relation with a reflexive, to reach the conclusion that the structure of the verb phrase in Tsakhur and Bagwalal respectively is “flat”.
40. We shall return to Principle C and its interaction with movement when discussing diagnostics of A-movement in §§3.2.3, 4.3.1.4.
Besides ERG-marked agents, this pattern is also very robust for dative- and locative-marked subjects of psych verbs such as cm–ol’- ‘love’ or cm–ix- ‘see’, in contrast to the same argument types in Tsez, another Northeast Caucasian language (Polinsky & Comrie 2003).

Taking into account the aforementioned facts, I believe there is enough reason to conclude that at least in the transitive constructions, the ergative argument asymmetrically c-commands the absolutive argument. The same is true for the dative and locative ‘subjects’ of experiencer verbs although there is evidence to suggest that these two cases are licensed in distinct positions in the clausal spine. Put differently, these facts are incompatible, in the absence of additional mechanisms or stipulations, with the structure in (92a).

2.3.2 **Affix ordering: from morphemes to the functional hierarchy**

Just as important as the relationship between the verb’s arguments is the question, also concerning both structure and order, of the mechanisms underlying the composition of Avar morphological words. One way of looking at things is to adopt what has come to be known a “cartographic” perspective of clause structure, whereby such categories as V, T and C from the GB era can be further decomposed into dozens of categories. In the particular case of the Avar verb, for instance, clauses would be built, besides the lexical root, on the basis of such functional heads, presumably coming from a universal repository of functional elements, as v (in some frameworks the head responsible for introducing the external argument), Asp to encode aspectual information, T and F expressing tense and finiteness, and eventually one derivational affix or another, with an occasional applicative head in the case of, for instance, ditransitive verbs.

The cartographic effort has given rise to a number of empirical generalisations regarding the (linear) ordering of various elements in the clause, the noun phrase and the adpositional phrases, from which the hierarchical structure can be gleaned. In an ideal world the attested ordering of the affixes encoding various types of information should be derived from the semantic scope that each of those elements takes in the interpretation of a particular syntactic object, with possible restrictions being reducible to some fundamental cognitive principle. In relation to the languages of the Caucasus, this has been proposed, by Korotkova & Lander (2010), to be what is responsible for the ordering of suffixes in Adyghe, a polysynthetic West Caucasian language, whose suffixes functioning as semantic operators are argued to modify the part of
the stem that (immediately) precedes it.\footnote{An important argument made in Korotkova & Lander 2010 regarding the flexibility of ordering displayed by a number of suffixes in Adyghe cannot, regrettably, be extended to Avar, where I have been unable to find any such affixes. It can, however, be the case that the semantic operators in question are realised in Avar as auxiliaries, such as modal verbs, which do seem to allow for some flexibility of ordering.}

Given the inventory of categories and morphemes realising them that we have come to know in §2.2, the functional hierarchy in the clausal domain, in the indicative mood for the sake of concreteness, would have to be a version of (101).

\begin{align*}
(101) & \quad \text{Force} > \text{Fin} > \text{T} > \text{Asp} > \text{Caus} > \text{Voice} > v > \text{Appl} > V/V\neg
\end{align*}

As is evident from (101), the list of categories lexicalised in Avar by dedicated morphemes is not especially vast, and it is inevitable that future work will bring about further refinements. Suffice it to say, the hierarchy as sketched above reflects two crucial properties of the Avar clause: first, it connects the order of morphemes within the morphological word to the scope of semantic operators encoded by those morphemes. It also ensures that the verb’s arguments display certain asymmetries by virtue of being introduced, via dedicated elements such as Appl, v or Caus, at different points in the derivation.

To take a concrete example, consider (14), repeated here as (102), which is built around the ditransitive predicate \(cm-\text{ič}-\) ‘sell’ and where all the arguments are expressed overtly.

\begin{align*}
(102) & \quad \text{di-} \quad \text{ca} \quad \text{du-} \quad \text{ē} \text{aka} \quad \text{b-} \quad \text{ič-} \quad \text{ana} \\
& \quad \text{1SG:OBL-ERG} \quad \text{2SG:OBL-DAT} \quad \text{COW.ABS} \quad \text{N-} \text{sell-PST} \\
& \quad \text{‘I sold you a/the cow.’}
\end{align*}

Its syntactic representation will, then, be as sketched in (103):
2.3. Towards a configurational structure of an Avar clause

The tree above should be read, in a bottom-up fashion, as the complex consisting of V–Appl–v–T–Fin–Force being spelled out as one morphological word, –ič-∅-∅-a-na, which becomes b–ič-∅-∅-a-na by hosting a neuter agreement prefix.

As far as the semantic interpretation is concerned of various portions of the clause, I am inclined to follow Ramchand & Svenonius (2014) and split it into three partitions: the event zone with various functors operating on events \( e_v \), the situation zone where events become constitutive parts of situations \( s_s \), and the proposition zone where propositions are created out of situations. The event zone corresponds to \( vP \), the situation zone is the TP, and everything above TP involves propositions.

### 2.3.3 Avar syntax in derivational layers

I have already given an illustration, in §2.2.4.2, of the workings of the layered-derivational approach (as put forth in Zwart 2009 et seq.) in the nominal domain. The clausal domain should not be any different: given the binary nature of merge and the presence of numerations, the interfaces will assign interpretations to whatever objects delivered to them once the numeration has been exhausted. Staying with the derivation of our example from the previous subsection, repeated here as (104), the numerations and derivational steps leading to this representation will be as follows.
A direct consequence of Zwart’s (2009) approach is that specifiers, such as the three arguments of argument-introducing elements notated as \( D \) in the tree above, are necessarily opaque by virtue of being outputs of separate, and distinct, derivational layers.\(^{42}\)

The crucial question at this point is whether the representation in (104) is derived by exhausting just one numeration (not counting those underlying the derivation of complex arguments) with all the relevant elements already in it, or whether a number of smaller numerations are involved in building the resulting structure (as would be the case in the Phase Theory, where the opacity of a domain is tied to that domain’s syntactic category). Derivation layering, on the other hand, makes no mandatory connection between a particular syntactic category and the object with that label being the output of a distinct derivation. Put differently, objects of different syntactic categories can be outputs of a distinct derivation (not just v and C, as in the Phase Theory), and not every vP and CP is opaque. The answer to the question about the number of numerations implicated in the derivation of (104) above should, ideally, be found in the notion of the numeration, and the selectional procedure responsible for its creation.\(^{43}\)

---

42. It seems to me that there is no reason to think of *dica, due* and *šaka* as not being syntactically complex, especially in light of recent analyses of pronouns as definite descriptions (Elbourne 2005, 2008).

43. Trotzke 2012 contains a preliminary discussion of opacity as a biproduct of derivation layering.
2.4 Chapter summary

The aims of this chapter were declared to be (i) introducing the framework, and (ii) providing a description of Avar grammar, both intended as paving the way for the discussion to follow. As regards (i), I have introduced, in § 2.1 the basic notions of linguistic minimalism as an interface-centric research programme. The next, and longest, section of the chapter introduced the reader to such aspects of Avar grammar as word order in both the verbal and nominal domains, case marking, pronominalisation and reflexivisation, verbal categories, negation, matrix and embedded clause types. Finally, in §2.3 I have argued that, despite appearances and contrary to much work on Northeast Caucasian languages, Avar could and should be given a configurational treatment in such domains as affix ordering and the structural asymmetries between arguments. Reiterating, only a very small selection of properties listed in the present chapter will be elucidated as the discussion unfolds, with all others remaining as promising avenues for future work.

This being said, we turn our attention to the syntax of Avar prenominal relative clauses.