Serbo-Croatian Word Order: A Logical Approach

Dissertation

Presented in Partial Fulfillment of the Requirements for the Degree
Doctor of Philosophy in the Graduate School of The Ohio State
University

By

Vedrana Mihaliček, B.A., M.A.

Graduate Program in Department of Linguistics

The Ohio State University

2012

Dissertation Committee:

Carl Pollard, Advisor

Brian Joseph

Michael White

Table of Contents

Page

List of Tables

Chapter 1: Introduction

Chapter 2: Framework

2.1 Background

In this chapter, we describe the framework that our theory of Serbo-Croatian grammar is expressed in. Our framework is closely related to Lambda Grammar (Muskens , 2003, 2007b) and ACG (de Groote , 2001). Variants of this framework, which we view as a descendant of Higher-Order Grammar (Pollard , 2004) and CVG (Pollard , 2011), have previously been called Linear Logic Based Grammar (Mihaliček , 2010b) and Pheno-Tecto Distinguished Categorial Grammar (originally in Smith (2010); also Mihaliček and Pollard (to appear)).

While obviously natural language expressions have both purely combinatorial syntactic properties (what sorts of arguments do they require? what sorts of expressions can they be arguments of?), and ordering properties (do they have to occur in some specific place in a clause or not? do they have to occur immediately to the left or right of some other expressions or not?), in many logical frameworks these two sets of properties are represented jointly, by a single component of the framework.

For example, in many mainstream versions of Categorial Grammar (Multi-modal Combinatory Categorial Grammar (MMCCG; Baldridge (2002)), Multi-modal Categorial Type Logics (MMCTL; Moortgat (1997); Bernardi (2002); Vermaat (2005)) and Multimodal Type Logical Grammar (MMTL; Morrill and Solias (1993); Morrill (1994); Morrill et al. (2007)), and non-multimodal versions of these frameworks), both combinatorial syntax and word order of expressions are represented by a single component of the framework which is formalized in Lambek Calculus. This results in inflexibility when it comes to dealing with freer word order. We see the development of multimodal versions of these frameworks as an attempt to compensate for the inflexibility that stems from representing combinatorial syntax and word order jointly, in the same part of the theory (see also (Muskens , 2003)).

Another approach to dealing with the same problem, exemplified by our and related frameworks, is simply to give up the assumption that word order and combinatorial syntax should be represented by a single grammatical component. We will ambiguously refer to these two sets of properties of linguistic expressions, but also the components of the theory that represent these sets of properties, as phenogrammar and tectogrammar respectively. This terminology originates to the best of our knowledge with Curry (1961).

Since Curry's original paper, many attempts have been made to dissociate syntactic combinatorics from linear order (e.g. Dowty (1996); Reape (1993, 1996);

Kathol (2000); Muskens (2003, 2007b); see also linear precedence rules of Pollard and Sag (1987) and GPSG (Gazdar et al. , 1985)), so this approach is not new. The separation of phenogrammar and tectogrammar has proven useful and important for elegantly analyzing phenomena such as quantifier scoping and medial extraction in English (Oehrle , 1994), and German word order (Reape , 1993; Kathol , 2000). Mihaliček and Pollard (to appear) argue that the separation of phenogrammar and tectogrammar considerably simplifies the analysis of interrogatives in English and Chinese, bringing out the underlying tectogrammatical similarities between the two languages and identifying phenogrammar as the locus of cross-linguistic variation with respect to different question forming strategies.

Essentially, the division of labor between phenogrammar and tectogrammar allows complex word order facts to be described somewhat independently in phenogrammar, without unnecessarily complicating the tectogrammar. This is important for describing a language such as Serbo-Croatian, which often allows semantically and syntactically insignificant reordering of expressions and discontinuities of various constituents.

So, our framework consists of three term/type calculi, which independently represent phenogrammar, tectogrammar and semantics. Each linguistic expression is then represented in our theory as a triple of term/type pairs, corresponding to the representation of that expression's word order and combinatorial syntactic properties, and its meaning.

In the remainder of this chapter, we describe each component separately and then sketch how the three components work together.

2.2 Phenogrammar

2.2.1 Terms and Types

To distinguish between different phenogrammatical (ordering) properties of expression, we assign them to different types. At the outset, we must distinguish between phonological words, and clitics. We assign all phonological words to type **p** and clitics to type **c**. For example, a phonological word such as the name *Marko*, and a clitic, such as the auxiliary *sam* 'am' are represented in the grammar as follows:

(1) a. $\vdash$ `marko` : **p**
read as: '`marko` is a term of type **p**',
i.e. the expression *Marko* is a phonological word

b. $\vdash$ `sam` : **c**
read as: '`sam` is a term of type **c**',
i.e. the expression *sam* is a clitic

Later we will define functions that represent cliticization, i.e. the construction of a larger phonological word out of a clitic and a phonological word.

However, simply distinguishing between clitics and phonological words is not sufficient to adequately describe the empirical domain. Phonological words and sequences of phonological words in Serbo-Croatian differ in terms of how freely they can order with respect to other constituents. So we introduce other phenogrammatical types which help us encode these distinctions.

First, every phonological word can be viewed as a length one string of phonological words. We introduce a type $\mathbf{s}$ for all strings of phonological words, i.e. $\mathbf{p}$-strings. So, the expression *Marko*, in addition to being represented as $\vdash \mathtt{marko} : \mathbf{p}$, also corresponds to the string $\vdash \mathtt{marko}_{\mathbf{s}} : \mathbf{s}$. While we can build strings out of phonological words, we cannot do so with the clitics.

Second, for any string of phonological words, we can define a set of strings whose member is exactly that string. Note that sets of strings are typically called *languages* in the formal language theory setting. So a set of $\mathbf{p}$-strings is called a $\mathbf{p}$-language. The type of sets of $\mathbf{p}$-strings is $\mathbf{s} \rightarrow \mathbf{t}$, abbreviated as $\mathbf{S}$. Here $\mathbf{t}$ is just the type of truth values. So starting with $\vdash \mathtt{marko}_{\mathbf{s}} : \mathbf{s}$, we can define:

(2) $\vdash \lambda_s.s = \mathtt{marko}_{\mathbf{s}} : \mathbf{S}$

This term denotes a set of strings which contains the string $\mathtt{marko}_{\mathbf{s}}$, i.e. $\{\mathtt{marko}_{\mathbf{s}}\}$. We will abbreviate such terms of type $\mathbf{S}$ as follows:

(3) $\vdash \mathrm{MARKO} : \mathbf{S}$
where $\mathrm{MARKO} =_{def} \lambda_s.s = \mathtt{marko}_{\mathbf{s}}$

Now, we don't stop here. In order to give an adequate description of the complexities of Serbo-Croatian word order, we build up even more complex types so we can make more fine grained distinctions in the grammar with respect to ordering possibilities.

First, just as we built up $\mathbf{p}$-strings out of terms of type $\mathbf{p}$ (i.e. phonological words), we construct strings of sets of $\mathbf{p}$-strings, i.e. strings of $\mathbf{p}$-languages. The type $\mathtt{z}$ is the type of strings of $\mathtt{p}$-languages. Since the type of $\mathtt{p}$-languages is $\mathtt{S}$, we will also refer to the type $\mathtt{z}$ as the type of $\mathtt{S}$-strings.

Seen as a length one string of languages, the expression *Marko* is represented as follows:

(4) $\vdash \mathtt{MARKO}_{\mathtt{z}} : \mathtt{z}$

Finally, just as we could define sets of $\mathtt{p}$-strings (i.e. $\mathtt{p}$-languages), we can define sets of $\mathtt{S}$-strings (i.e. $\mathtt{S}$-languages). The type of $\mathtt{S}$-languages is $\mathtt{S} \rightarrow \mathtt{t}$ abbreviated as $\mathtt{Z}$.

Seen as a set of $\mathtt{S}$-languages, the expression *Marko* is represented as follows:

(5) $\vdash \lambda_z . z = \mathtt{MARKO}_{\mathtt{z}} : \mathtt{Z}$

Table 2.1 summarizes various types of phenogrammatical constants that we have.

From these phenogrammatical types, we construct more complex types by means of the $\rightarrow$ type constructor. So, for example, $\mathtt{s}$ is the type of $\mathbf{p}$-strings, $\mathtt{s} \rightarrow \mathtt{s}$ is the type of functions from strings to strings, $\mathtt{p} \rightarrow \mathtt{s}$ is the type of functions from phonological words to $\mathbf{p}$-strings, $\mathtt{z} \rightarrow \mathtt{Z}$ is the type of functions from $\mathtt{S}$-strings to $\mathtt{S}$-languages, etc.

2.2.2 Functions

Many linguistic expressions will be represented phenogrammatically as functional terms. For example, an attributive adjective might be represented as a term of type $\mathtt{z} \rightarrow \mathtt{z}$, i.e. a function which takes an $\mathtt{S}$-string argument (the noun which

TYPE:	REFERRED TO AS:	EXAMPLE:
$\mathtt{c}$	clitics	$\vdash \mathtt{sam} : \mathtt{c}$
$\mathtt{p}$	phonological words	$\vdash \mathtt{marko} : \mathtt{p}$
$\mathtt{s}$	strings of phonological words; $\mathtt{p}$-strings	$\vdash \mathtt{marko}_{\mathtt{s}} : \mathtt{s}$
$\mathtt{S}$	sets of $\mathtt{p}$-strings; $\mathtt{p}$-languages	$\vdash \mathtt{MARKO} : \mathtt{S}$
$\mathtt{z}$	strings of $\mathtt{p}$-languages; $\mathtt{S}$-strings	$\vdash \mathtt{MARKO}_{\mathtt{z}} : \mathtt{z}$
$\mathtt{Z}$	sets of $\mathtt{S}$-strings; $\mathtt{S}$-languages	$\vdash \lambda_z.z = \mathtt{MARKO}_{\mathtt{z}} : \mathtt{Z}$

Table 2.1: Phenogrammatical Types and Non-Logical Constants.

it modifies), and outputs an $\mathtt{S}$-string (the concatenation of the adjective and the noun).

However, in this section, we focus on special phenogrammatical functions, which transform or combine different types of phenogrammatical terms, but do not themselves represent any linguistic expressions.

First, we have a pair of functions which combine clitics with their hosts. These functions combine a clitic with a phonological word, and output another phonological word.

(6) a. $\vdash \#_{pc} : \mathtt{c} \rightarrow \mathtt{p} \rightarrow \mathtt{p}$

b. $\vdash \#_{ec} : \mathtt{p} \rightarrow \mathtt{c} \rightarrow \mathtt{p}$

For example, if $\vdash \mathtt{grad} : \mathbf{p}$ (*grad* is a phonological word), and $\vdash \mathtt{u} : \mathbf{c}$ (*u* is a clitic), then $\vdash \mathtt{u}\#_{pc}\mathtt{grad} : \mathbf{p}$, i.e. *u grad* is a phonological word consisting of the clitic *u* procliticized onto the phonological word *grad*.

Similarly, if $\vdash \mathtt{grad} : \mathbf{p}$, and $\vdash \mathtt{je} : \mathbf{c}$, then $\vdash \mathtt{grad}\#_{ec}\mathtt{je} : \mathbf{p}$, i.e. *grad je* is a phonological word consisting of the clitic *je* encliticized onto the phonological word *grad*.

Second, we define an operation that combines $\mathbf{p}$-strings into larger $\mathbf{p}$-strings, i.e. concatenation.

(7) $\vdash \cdot : \mathbf{s} \rightarrow \mathbf{s} \rightarrow \mathbf{s}$

For example, if $\vdash \mathtt{marko}_{\mathbf{s}} : \mathbf{s}$ and $\vdash \mathtt{spava}_{\mathbf{s}} : \mathbf{s}$, i.e. they are both $\mathbf{p}$-strings, then we can concatenate them and construct a $\mathbf{p}$-string $\vdash \mathtt{marko}_{\mathbf{s}} \cdot \mathtt{spava}_{\mathbf{s}} : \mathbf{s}$.

Concatenation is associative, meaning that if s, t, u are strings, $(s \cdot t) \cdot u = s \cdot (t \cdot u)$, i.e. rebracketing is allowed. However, concatenation is not commutative, so a string $s \cdot t$ is not the same as the string $t \cdot s$, i.e. reordering of smaller strings inside a string is not allowed.

We have a special constant $\vdash \mathtt{e}_{\mathbf{p}} : \mathbf{s}$, called the empty string, which is the identity for concatenation. This means that for any string s, $s \cdot \mathtt{e}_{\mathbf{s}} = s = \mathtt{e}_{\mathbf{s}} \cdot s$.

Further, we have a function $\vdash \mathbf{toS} : \mathbf{p} \rightarrow \mathbf{s}$ which converts phonological words into $\mathbf{p}$-strings of length one. So, if $\vdash \mathtt{marko} : \mathbf{p}$, then $\vdash (\mathbf{toS}\ \mathtt{marko}) : \mathbf{s}$. We abbreviate terms such as $(\mathbf{toS}\ \mathtt{marko})$ as $\mathtt{marko}_{\mathbf{s}}$.

We also have functions that prefix ($\mathbf{cns_p}$) and suffix ($\mathbf{snc_p}$) a phonological word to a $\mathbf{p}$-string.

(8) a. $\vdash \mathbf{cns_p} : \mathbf{p} \to \mathbf{s} \to \mathbf{s}$
b. $\vdash \mathbf{snc_p} : \mathbf{s} \to \mathbf{p} \to \mathbf{s}$

For example, if $\vdash \texttt{marko} : \mathbf{p}$ and $\vdash \texttt{spava}_\mathbf{s} : \mathbf{s}$, then $\vdash (\mathbf{cns_p}\ \texttt{marko}\ \texttt{spava}_\mathbf{s}) : \mathbf{s}$, and $(\mathbf{cns_p}\ \texttt{marko}\ \texttt{spava}_\mathbf{s})$ is the same thing as the string $(\texttt{marko}_\mathbf{s} \cdot \texttt{spava}_\mathbf{s})$.

Similarly, $\vdash (\mathbf{snc_p}\ \texttt{spava}_\mathbf{s}\ \texttt{marko}) : \mathbf{s}$, and $(\mathbf{snc_p}\ \texttt{spava}_\mathbf{s}\ \texttt{marko})$ is the same thing as the string $(\texttt{spava}_\mathbf{s} \cdot \texttt{marko}_\mathbf{s})$.

Further, we have functions that take a $\mathbf{p}$-string and output the phonological word that is its prefix ($\mathbf{fst_p}$) or suffix ($\mathbf{lst_p}$).

(9) a. $\vdash \mathbf{fst_p} : \mathbf{s} \to \mathbf{p}$
b. $\vdash \mathbf{lst_p} : \mathbf{s} \to \mathbf{p}$

For example, $\mathbf{fst_p}(\texttt{marko}_\mathbf{s} \cdot \texttt{spava}_\mathbf{s}) = \texttt{marko}$, and $\mathbf{lst_p}(\texttt{marko}_\mathbf{s} \cdot \texttt{spava}_\mathbf{s}) = \texttt{spava}$.

Associated with the functions $\mathbf{cns_p}$ and $\mathbf{fst_p}$, and $\mathbf{snc_p}$ and $\mathbf{lst_p}$, we have the functions $\vdash \mathbf{rst_p} : \mathbf{s} \to \mathbf{s}$ and $\vdash \mathbf{tsr_p} : \mathbf{s} \to \mathbf{s}$ respectively, so that the following equalities hold, for any $\mathbf{p}$-string s:

(10) a. $\mathbf{cns_p}(\mathbf{fst_p}\ s)\ (\mathbf{rst_p}\ s) = s$
b. $\mathbf{snc_p}(\mathbf{lst_p}\ s)\ (\mathbf{tsr_p}\ s) = s$

These equalities just mean that if you take off the prefix (suffix) of some string s, and then reattach that prefix (suffix) to what's left of s, you just get the same $\mathbf{p}$-string s back.

Third, we define language concatenation (*fusion*) which takes some two sets of $\mathbf{p}$-strings and outputs another set of $\mathbf{p}$-strings by concatenating all the strings in the input sets of strings.

(11) a. $\vdash \bullet_{\mathbf{S}} : \mathbf{S} \to \mathbf{S} \to \mathbf{S}$

b. $\bullet_{\mathbf{S}} =_{def} \lambda_{STu}.\exists_{st}[(S\ s) \wedge (T\ t) \wedge u = s \cdot t]$

For example, if S denotes the set of strings $\{\mathtt{marko}_{\mathbf{s}}, \mathtt{ana}_{\mathbf{s}}\}$, and T denotes the set of strings $\{\mathtt{spava}_{\mathbf{s}}, \mathtt{radi}_{\mathbf{s}}\}$, then $S \bullet_{\mathbf{S}} T$ denotes the set of strings:

(12) $\{\mathtt{marko}_{\mathbf{s}} \cdot \mathtt{spava}_{\mathbf{s}}, \mathtt{marko}_{\mathbf{s}} \cdot \mathtt{radi}_{\mathbf{s}}, \mathtt{ana}_{\mathbf{s}} \cdot \mathtt{spava}_{\mathbf{s}}, \mathtt{ana}_{\mathbf{s}} \cdot \mathtt{radi}_{\mathbf{s}}\}$.

There are two special constants of type $\mathbf{S}$, $\vdash 0_{\mathbf{S}} : \mathbf{S}$, the null $\mathbf{p}$-language which doesn't contain any strings, and $\vdash 1_{\mathbf{S}} : \mathbf{S}$, the singleton $\mathbf{p}$-language which contains only the empty string.

We also have counterparts of functions that we just defined for phonological words ($\mathbf{p}$), $\mathbf{p}$-strings ($\mathbf{s}$) and $\mathbf{p}$-languages ($\mathbf{S}$), defined for $\mathbf{p}$-languages ($\mathbf{S}$), $\mathbf{S}$-strings ($\mathbf{z}$) and $\mathbf{S}$-languages ($\mathbf{Z}$).

The operation $\vdash \circ : \mathbf{z} \to \mathbf{z} \to \mathbf{z}$ is concatenation for $\mathbf{S}$-strings, which has the same properties as $\vdash \cdot : \mathbf{s} \to \mathbf{s} \to \mathbf{s}$, except that it operates on $\mathbf{S}$-strings, and not $\mathbf{p}$-strings. The counterpart of the logical constant $\vdash \mathrm{e}_{\mathbf{p}} : \mathbf{s}$ is $\vdash \mathrm{e}_{\mathbf{S}} : \mathbf{z}$.

The counterpart of $\mathbf{p}$-language fusion is $\vdash \bullet_{\mathbf{Z}} : \mathbf{Z} \to \mathbf{Z} \to \mathbf{Z}$, the $\mathbf{S}$-language fusion. The empty $\mathbf{S}$-language is $\vdash 0_{\mathbf{Z}} : \mathbf{Z}$ and the singleton $\mathbf{S}$-language is $\vdash 1_{\mathbf{Z}} : \mathbf{Z}$.

We also have the function $\vdash \mathtt{toZ} : \mathtt{S} \to \mathtt{z}$, which is the counterpart of $\vdash \mathtt{toS} : \mathtt{p} \to \mathtt{s}$, and turns $\mathtt{p}$-languages (of type $\mathtt{S}$) into length one strings of $\mathtt{p}$-languages (of type $\mathtt{z}$).

Finally, the functions $\mathtt{cns}_{\mathtt{p}}$, $\mathtt{fst}_{\mathtt{p}}$ and $\mathtt{rst}_{\mathtt{p}}$ on the one hand, and $\mathtt{snc}_{\mathtt{p}}$, $\mathtt{lst}_{\mathtt{p}}$ and $\mathtt{tsr}_{\mathtt{p}}$ on the other, have as their counterparts $\mathtt{cns}_{\mathtt{S}}$, $\mathtt{fst}_{\mathtt{S}}$ and $\mathtt{rst}_{\mathtt{S}}$, and $\mathtt{snc}_{\mathtt{S}}$, $\mathtt{lst}_{\mathtt{S}}$ and $\mathtt{tsr}_{\mathtt{S}}$ respectively.

We also need to define a few special functions which are going to be crucial for analyzing Serbo-Croatian word order because they allow us to control the degree of word order freedom of a given expression.

First, we define the function $\vdash \mathtt{PER} : \mathtt{z} \to \mathtt{Z}$, called permutation, which takes some $\mathtt{S}$-string (of type $\mathtt{z}$) and constructs an $\mathtt{S}$-languages (of type $\mathtt{Z}$) consisting of all possible reorderings of the smaller $\mathtt{S}$-strings in the original $\mathtt{S}$-string. For example, $\mathtt{PER}(x \circ y \circ z)$ denotes a set of six $\mathtt{S}$-strings, namely $\{x \circ y \circ z,\ x \circ z \circ y,\ y \circ x \circ z,\ y \circ z \circ x,\ z \circ x \circ y,\ z \circ y \circ x\}$. As the reader might already be suspecting, the output of this function will represent expressions in which smaller constituents are allowed to freely reorder, for example finite verbs and their noun phrase arguments.

Second, we define a more restrictive function $\vdash \odot : \mathtt{z} \to \mathtt{z} \to \mathtt{Z}$, called shuffle, which interleaves two $\mathtt{S}$-strings. For example, if $x = v \circ w$, and $x' = y \circ z$, then $x \odot x'$ denotes the following set of strings:

(13) $\{v \circ w \circ y \circ z,\ v \circ y \circ w \circ z,\ v \circ y \circ z \circ w,$
$y \circ v \circ w \circ z,\ y \circ v \circ z \circ w,\ y \circ z \circ v \circ w\}$

So, shuffling x into x' means constructing a set of $\mathbf{S}$-strings in which smaller strings in x and x' can reorder with respect to one another so long as the relative order of the smaller strings in x and x' is retained. Going back to the example above, $w \circ v \circ y \circ z$ is not in the set denoted by $x \odot x'$, because in x, v precedes w.

Third, we define a function $\vdash \cup_{\mathbf{S}} : \mathbf{S} \rightarrow \mathbf{S} \rightarrow \mathbf{S}$, called language union, which constructs a single $\mathbf{p}$-language out of two $\mathbf{p}$-languages.

(14) $\cup_{\mathbf{S}} =_{def} \lambda_{STs}.(S\ s) \vee (T\ s)$

If P contains the strings $s \cdot t$ and u, and T contains the strings $p \cdot u$ and t, then $P \cup T$ is the set of all strings that are either in P or T, i.e. it denotes the set $\{s \cdot t,\ u,\ p \cdot u,\ t\}$. We also define a counterpart of $\cup_{\mathbf{S}}$ for $\mathbf{S}$-languages, $\vdash \cup_{\mathbf{Z}} : \mathbf{Z} \rightarrow \mathbf{Z} \rightarrow \mathbf{Z}$.

Fourth, we define a function $\vdash \mathbf{L} : \mathbf{z} \rightarrow \mathbf{S}$, called linguification, which takes an $\mathbf{S}$-string and outputs a set of $\mathbf{p}$-strings.

(15) a. $(\mathbf{L}\ \mathrm{e}_{\mathbf{S}}) = 1_{\mathbf{S}}$
b. $\forall_{Pz}.(\mathbf{L}\ (\mathbf{cns_z}\ P\ z)) = P \bullet (\mathbf{L}\ z)$

For example, $(\mathbf{L}\ (\mathrm{MARKO}_{\mathbf{z}} \circ \mathrm{SPAVA}_{\mathbf{z}}))$ is $(\mathrm{MARKO} \bullet \mathrm{SPAVA})$ which is the set of $\mathbf{p}$-strings that contains exactly one string, $\mathtt{marko}_{\mathbf{s}} \cdot \mathtt{spava}_{\mathbf{s}}$.

Fifth, we define a function $\vdash \mathbf{k} : \mathbf{Z} \rightarrow \mathbf{S}$, called compaction, which takes a set of $\mathbf{S}$-strings and then unions the linguifications of all $\mathbf{S}$-strings in that set resulting in a $\mathbf{p}$-language.

(16) a. $\vdash (\mathbf{k}\ 0_{\mathbf{Z}}) = 0_{\mathbf{S}}$
b. $\vdash \forall_{Zv}.(\mathbf{k}(Z \cup_{\mathbf{Z}} \lambda_z.z = v)) = (\mathbf{k}\ Z) \cup_{\mathbf{S}} (\mathbf{L}\ v)$

CONSTANT	NOTES
$\vdash \mathtt{e}_{\mathbf{p}} : \mathbf{s}$	null $\mathbf{p}$-string; identity for $\mathbf{p}$-string concatenation $(\cdot)$
$\vdash \mathtt{e}_{\mathbf{S}} : \mathbf{z}$	null $\mathbf{S}$-string; identity for $\mathbf{S}$-string concatenation $(\circ)$
$\vdash 0_{\mathbf{S}} : \mathbf{S}$	the empty $\mathbf{p}$-language
$\vdash 0_{\mathbf{Z}} : \mathbf{Z}$	the empty $\mathbf{S}$-language
$\vdash 1_{\mathbf{S}} : \mathbf{S}$	the singleton $\mathbf{p}$-language; identity for $\mathbf{p}$-language fusion $(\bullet_{\mathbf{S}})$
$\vdash 1_{\mathbf{Z}} : \mathbf{Z}$	the singleton $\mathbf{S}$-language; identity for $\mathbf{S}$-language fusion $(\bullet_{\mathbf{Z}})$

Table 2.2: Phenogrammatical Logical Constants.

For example, $\mathbf{k}\ (\mathbf{PER}\ (\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{SPAVA}_{\mathbf{z}}))$ denotes the set of $\mathbf{p}$-strings $\{\mathtt{marko}_{\mathbf{s}} \cdot \mathtt{spava}_{\mathbf{s}}, \mathtt{spava}_{\mathbf{s}} \cdot \mathtt{marko}_{\mathbf{s}}\}$.

Table 2.2 lists all logical constants that we introduced, and Table 2.3 summarizes all the phenogrammatical functions.

2.3 Tectogrammar

2.3.1 Preliminaries

Recall that the tectogrammatical component of the grammar is not concerned with word order but primarily with argument/functor relations. Since expressions have different argument requirements, we distinguish between various tectogrammatical types of expressions, where each type roughly corresponds to a syntactic category.

We take the stance that since inflectional features, such as case, gender, number and person, influence the argument requirements of an expression, they need to

FUNCTION	NOTES
$\vdash \#_{pc} : \mathbf{c} \rightarrow \mathbf{p} \rightarrow \mathbf{p}$	procliticization to a phonological word
$\vdash \#_{ec} : \mathbf{p} \rightarrow \mathbf{c} \rightarrow \mathbf{p}$	encliticization to a phonological word
$\vdash \cdot : \mathbf{s} \rightarrow \mathbf{s} \rightarrow \mathbf{s}$	concatenation for **p**-strings
$\vdash \circ : \mathbf{z} \rightarrow \mathbf{z} \rightarrow \mathbf{z}$	concatenation for **S**-strings
$\vdash \bullet_{\mathbf{S}} : \mathbf{S} \rightarrow \mathbf{S} \rightarrow \mathbf{S}$	**p**-language fusion
$\vdash \bullet_{\mathbf{Z}} : \mathbf{Z} \rightarrow \mathbf{Z} \rightarrow \mathbf{Z}$	**S**-language fusion
$\vdash \cup_{\mathbf{S}} : \mathbf{S} \rightarrow \mathbf{S} \rightarrow \mathbf{S}$ $\vdash \cup_{\mathbf{Z}} : \mathbf{Z} \rightarrow \mathbf{Z} \rightarrow \mathbf{Z}$	**p**-language union **S**-language union
$\vdash \mathbf{cns_p} : \mathbf{p} \rightarrow \mathbf{s} \rightarrow \mathbf{s}$ $\vdash \mathbf{cns_S} : \mathbf{S} \rightarrow \mathbf{z} \rightarrow \mathbf{z}$	prefixes a **p**-string prefixes an **S**-string
$\vdash \mathbf{fst_p} : \mathbf{s} \rightarrow \mathbf{p}$ $\vdash \mathbf{fst_S} : \mathbf{z} \rightarrow \mathbf{S}$	outputs the prefix of a **p**-string outputs the prefix of an **S**-string
$\vdash \mathbf{rst_p} : \mathbf{s} \rightarrow \mathbf{s}$ $\vdash \mathbf{rst_S} : \mathbf{z} \rightarrow \mathbf{z}$	outputs a **p**-string without its prefix outputs an **S**-string without its prefix
$\vdash \mathbf{snc_p} : \mathbf{p} \rightarrow \mathbf{s} \rightarrow \mathbf{s}$ $\vdash \mathbf{cns_S} : \mathbf{S} \rightarrow \mathbf{z} \rightarrow \mathbf{z}$	suffixes a **p**-string suffixes an **S**-string
$\vdash \mathbf{lst_p} : \mathbf{s} \rightarrow \mathbf{p}$ $\vdash \mathbf{lst_S} : \mathbf{z} \rightarrow \mathbf{S}$	outputs the suffix of a **p**-string outputs the suffix of an **S**-string
$\vdash \mathbf{tsr_p} : \mathbf{s} \rightarrow \mathbf{s}$ $\vdash \mathbf{tsr_S} : \mathbf{z} \rightarrow \mathbf{z}$	outputs a **p**-string without its suffix outputs an **S**-string without its suffix
$\vdash \mathbf{L} : \mathbf{z} \rightarrow \mathbf{S}$	constructs a **p**-language out of a string of **p**-languages
$\vdash \mathbf{k} : \mathbf{Z} \rightarrow \mathbf{S}$	constructs a **p**-language out of an **S**-language
$\vdash \mathbf{PER} : \mathbf{z} \rightarrow \mathbf{Z}$	constructs the set of all permutations of some **S**-string
$\vdash \odot : \mathbf{z} \rightarrow \mathbf{z} \rightarrow \mathbf{Z}$	constructs an **S**-language by interleaving two **S**-strings

Table 2.3: Phenogrammatical Functions.

be represented in the tectogrammar. For example, the verb *spavam* 'sleep' needs a first person singular subject, but *spava* 'sleeps' needs a third person singular subject. In our grammar, they will be treated as distinct expressions assigned to different tectogrammatical types since they have different argument requirements.

2.3.2 Representing Inflectional Features

To represent inflectional features, we introduce special tectogrammatical types for each kind of feature. These types are special because no linguistic expressions are assigned to these types—they don't directly correspond to syntactic categories. However, they will help us encode inflectional information on other tectogrammatical types, that is, terms of the inflectional types will serve as parameters that define families of tectogrammatical types.[1]

Cse is the type of case features. Terms of this type represent specific case values: nom for nominative, gen for genitive, dat for dative, acc for accusative and inst for instrumental.

Gdr is the type of gender features. Terms of this type represent specific gender values: m for masculine, f for feminine and n for neuter.

Terms of type **Prs** are 1, 2 and 3 for first, second and third person.

Finally, terms of type **Num** are sg for singular and pl for plural.

[1]See de Groote and Maarek (2007) for a similar representation of inflectional features in a dependent type system.

2.3.3 N and NP type families

All nouns are assigned to a type in the **N** family. The types in this family are parametrized by terms of type **Cse**, **Gdr** and **Num**, since these features are relevant for determiner and adjective agreement with nouns. So, for example, $\mathbf{N}_{\text{nom, f, sg}}$ is the type of nominative singular feminine nouns, $\mathbf{N}_{\text{inst,m,pl}}$ is the type of instrumental plural masculine nouns, and so on.

All noun phrases, whether lexical or phrasal, are assigned to a type in the **NP** family. The types in this type family are parametrized by terms of type **Cse**, **Gdr**, **Num** and **Prs**, since these features are relevant for subject-verb agreement, and object selection. $\mathbf{NP}_{\text{dat,n,sg,3}}$ is, for example, the type of dative neuter third person singular noun phrases, and $\mathbf{NP}_{\text{acc,f,pl,3}}$ is the type of accusative feminine third person plural noun phrases.

For a more detailed exposition, as well as arguments for representing all these features on noun and noun phrase types, we direct the reader to Chapter 3.

2.3.4 S type family

All clauses are assigned to a type in the **S** family. This family of types is parametrized by terms of type **K** and **Nat**. Terms of type **K** encode different types of clauses: m(ain), e(mbedded), q(uestions) or inf(initival).

The type **Nat** is the type of natural numbers. As parameters, they will be used to enforce the placement and ordering of enclitics in a clause. This strategy will be explained in more detail in Chapter 5 (see also Morrill and Gavarró (1992) for a similar use of natural number parameters). For now we just mention that clauses with no enclitics placed inside of them are associated with the parameter 6, and as more and more enclitics are placed in the clause, the natural number parameter is lowered, so that a clause whose parameter is 0 cannot have any more clitics placed inside of it.

2.3.5 ⊸ types

We can construct more complex tectogrammatical types out of types in the **N**, **NP** and **S** family with the type constructor ⊸, the linear implication.[2] Such implicative types encode syntactic dependencies between expressions and their arguments.

For example, an intransitive verb which needs a subject noun phrase to construct a sentence would be associated with the type $\mathbf{NP}_{\mathrm{nom}} \multimap \mathbf{S}_{\mathrm{m}}$, a complementizer which converts a main clause into an embedded clause with the type $\mathbf{S}_{\mathrm{m}} \multimap \mathbf{S}_{\mathrm{e}}$, and a determiner which constructs noun phrases out of nouns with the type $\mathbf{N} \multimap \mathbf{NP}$.[3]

[2]Among categorial frameworks which also use linear implication as the type constructor in the tectogrammatical component are ACG (de Groote (2001)) and Lambda Grammar (Muskens (2003, 2007b)). Since linear implication is insensitive to the order of hypotheses, it can only be used as the main tectogrammatical type constructor in frameworks which distinguish between phenogrammar and tectogrammar, where phenogrammar is designated to keep track of linear order, and so the tectogrammatical component need not be order-sensitive.

[3]In these examples of functional types we abstracted away from many type parameters for illustrative purposes.

2.3.6 $\prod$ types

While it is necessary to keep track of inflectional features of expressions, many expressions are vague with respect to some subset of inflectional features. For example, present tense verbs in Serbo-Croatian require a nominative subject of a specific number and person, but they do not care what the gender of their subject noun phrase is. Past participles, on the other hand, require that their subject be nominative and of a specific gender and number, but do not care about the subject's person. Prepositions require a noun phrase argument of a specific case, but do not care about their argument's number, gender or person.

Continuing with the example of present tense verbs, one option would be to simply list each version of the verb. For example, *spava* 'sleeps' requires a third person singular nominative subject of any gender. Since there are three genders, we could list three versions of this verb, one for each gender:

(17) a. $\vdash \text{spava}_{\text{f}} : \mathbf{NP}_{\text{nom,f,sg,3}} \multimap \mathbf{S}_{\text{m,6}}$
b. $\vdash \text{spava}_{\text{m}} : \mathbf{NP}_{\text{nom,m,sg,3}} \multimap \mathbf{S}_{\text{m,6}}$
c. $\vdash \text{spava}_{\text{n}} : \mathbf{NP}_{\text{nom,n,sg,3}} \multimap \mathbf{S}_{\text{m,6}}$

However, listing each version of a vague expression in the lexicon is not just an inelegant solution that substantially increases the number of lexical entries in the grammar; if we were to do that we would in a sense also be missing a linguistic generalization.

We address this problem by introducing dependent product types, which, intuitively, help us abstract away from the value of some parameter, and allow us to obtain more specific versions of lexical entries by supplying that value.

So, a more expedient way to represent the same verb *spava* 'sleeps' tectogrammatically is as follows:

(18) a. $\vdash \lambda_{g:\mathbf{Gdr}}\mathrm{spava}_g : \prod_{g:\mathbf{Gdr}}[\mathbf{NP}_{\mathrm{nom},g,\mathrm{sg},3} \multimap \mathbf{S}_{\mathrm{m},6}]$

This term is not specified for gender. But, given the three terms of type **Gdr**, m, f, and n we can obtain three more specific versions of this term via the product elimination rule (which we return to in later in the chapter). So, instead of being non-logical axioms that have to be asserted, the three tectogrammatical versions of this verb listed in (17) are now theorems.

The product types should be understood as universally quantifying over all terms of a given type. So the tectogrammatical term in (18) above means that for each τ of type **Gdr**, there's a more specific version of the term that has precisely τ as its gender parameter.

While the parametrization of types essentially allowed us to simulate subtyping by allowing us to define type families, product types let us 'underspecify' tectogrammatical terms for certain features.

2.4 Semantics

2.4.1 Preliminaries

We assume a hyperintensional semantic theory along the lines of Pollard (2008) (see Thomason (1980) and Muskens (2005), Muskens (2007a) for versions of hyperintensional semantics with somewhat different technical assumptions). While we believe this choice to be well motivated (we direct the reader to Pollard (2008) for a detailed discussion of problems with traditional possible world semantics), our choice of semantic theory is inessential in the context of this project; our theory of Serbo-Croatian grammar is equally compatible with a more mainstream Montague-style possible world semantics.

2.4.2 Entailment

While in standard possible world semantics, the type of possible worlds is treated as basic, and propositions are analyzed as sets of possible worlds, on our approach propositions are taken as basic and the type of possible worlds is defined to be a certain sets of propositions.

This set up has (desirable) consequences regarding entailment. In standard possible world semantics, entailment is modeled as subset inclusion, an antisymmetric relation on sets of worlds. This prevents us from distinguishing mutually entailing propositions, since they are represented in the theory as the same set of possible worlds.

On our approach, however, propositions are modeled as members of a *pre*-boolean algebra *pre*-ordered by entailment. Entailment is axiomatized as a reflexive, transitive, but *not* antisymmetric relation on propositions. This way, it is possible for equivalent (mutually entailing) propositions to be distinct.

2.4.3 Types

The hyperintensional semantic theory that we are using is expressed in classical higher order logic. The underlying logic provides us with the basic types $\mathbf{t}$ (truth values) and $\mathbf{n}$ (natural numbers).

In addition to these types, we introduce as basic types $\mathbf{p}$ (propositions) and $\mathbf{e}$ (individual entities).[4] We call the types $\mathbf{p}$ and $\mathbf{e}$, and any implicative types constructed out of these by means of the type constructor $\rightarrow$, *hyperintensional types*. These types are used to model linguistic meanings.

We recursively define the function **Ext** mapping hyperintensional types to the corresponding extensional types. Here, $\mathcal{S}$ and $\mathcal{S}'$ are metavariables over hyperintensional types:

(19) a. $\mathbf{Ext}(\mathbf{e}) = \mathbf{e}$
b. $\mathbf{Ext}(\mathbf{p}) = \mathbf{t}$
c. $\mathbf{Ext}(\mathcal{S} \rightarrow \mathcal{S}') = \mathcal{S} \rightarrow (\mathbf{Ext}\mathcal{S})$

[4]For expository simplicity, we depart from Pollard (2008) in not distinguishing between the extensional type $\mathbf{e}$ and the corresponding hyperintensional type $\mathbf{i}$ (individual concepts). In particular, the meaning of a name is the same as its reference.

The type of possible worlds $\mathbf{w}$ is constructed out of the basic types in such a way that the interpretation of the type $\mathbf{w}$ is the set of ultrafilters of the preboolean prealgebra that interprets the type $\mathbf{p}$. Specifically. $\mathbf{w} =_{def} [\mathbf{p} \to \mathbf{t}]_{\mathsf{u}}$, where $\vdash \mathsf{u} : (\mathbf{p} \to \mathbf{t}) \to \mathbf{t}$ is a predicate on sets of propositions that picks out those sets of propositions that are ultrafilters (see Pollard (2008) for details of this construction).

Concomitantly, we introduce a family of constants $\mathbf{ext}_{\boldsymbol{S}} : \boldsymbol{S} \to \mathbf{w} \to (\mathbf{Ext}\boldsymbol{S})$ (where $\boldsymbol{S}$ is a variable ranges over the hyperintensional types) interpreted as a polymorphic function that maps a hyperintension and a world to the extension of that hyperintension at that world, as follows:

(20) a. $\vdash \forall_{x:\mathbf{e}} \forall_{w:\mathbf{w}} [(\mathbf{ext}_{\mathbf{e}}\ x\ w) = x]$
b. $\vdash \forall_{p:\mathbf{p}} \forall_{w:\mathbf{w}} [(\mathbf{ext}_{\mathbf{p}}\ p\ w) = p@w]$
c. $\vdash \forall_{f:\boldsymbol{S} \to \boldsymbol{S}'} \forall_{w:\mathbf{w}} [(\mathbf{ext}_{\boldsymbol{S} \to \boldsymbol{S}'}\ f\ w) = \lambda_{x:\boldsymbol{S}} (\mathbf{ext}_{\boldsymbol{S}'} (f\ x)\ w)]$

Here the notation '$p@w$' abbreviates $(\mu_{\mathsf{u}}\ w\ p)$, where μ_{u} denotes the embedding of the set of worlds into the set of sets of propositions.[5]

In general, however, we will only refer to extensional types in axioms which relate special hyperintensional constants (e.g. exists, and, etc.) to their extensional counterparts. Representations of linguistic meaning in the grammar are hyperintensional terms, and here we only mention how these can be extensionalized for completeness' sake.

[5]The kind of HOL we employ follows Lambek and Scott (1986) in having machinery for forming (separation-style) subtypes. Thus, if $\mathcal{S}$ is a type and σ an $\mathcal{S}$-predicate (term of type $\mathcal{S} \to \mathbf{t}$), then there is a type $\mathcal{S}_\sigma$ interpreted as the subset of the interpretation of $\mathcal{S}$ that has the interpretation of σ as its characteristic function; and there is a constant μ_σ that denotes the subset embedding.

2.5 Putting it all together

2.5.1 Signs

Lexical entries are triples of typed lambda terms, however, in practice, we often omit tectogrammatical terms altogether and write lexical entries in the following format:

(21) Γ[6] $\vdash$ `pheno term` : **`pheno type`**; **tecto type**; semantic term : **semantic type**

We call a *sign* any such triple of typed terms, including lexical entries (non-logical axioms) as well as larger expressions constructed out of the lexical entries via the inference rules (essentially, theorems of our grammar).

Sometimes, we will also omit the phenogrammatical and the semantic type and write signs in the following form:

(22) $\Gamma \vdash$ `pheno term`; **tecto type**; semantic term

Table 2.4 declares typesetting conventions for object language or metalanguage expressions for each of the three calculi, which we have thus far been using implicitly.

[6] Γ is a metavariable over contexts which are multisets of triples of typed variables, and while lexical entries, as non-logical axioms in our theory, typically do not depend on any assumptions, our analysis of the inherent reflexive in Chapter 5 will require that a lexical entry have a non-empty context, i.e. introduce a hypothesis.

	TECTOGRAMMAR	SEMANTICS	PHENOGRAMMAR
terms	acc, pl	ana, (love ana)	`marko`
types	**N**, **NP**, **S**, **Gdr**, **Num**, **Prs**, **Cs**	**e**, **p**, **t**	**`p`**, **`s`**, **`S`**, **`z`**, **`Z`**
term variables	*u,v,w,x,y,z*	*x,y,z;P,Q;p,q*	*`a, b, c, ...;`* *`p, q, r, s, t, u;`* *`P, Q, R, S, T, U;`* *`v, w, x, y, z;`* *`V, W, X, Y, Z`*
type variables	***T′, T″,...***	***S',S",...***	***`P′, P″, ...`***
term metavariables	τ', τ'',...	σ', σ'',...	ϕ', ϕ'', ...
variable metavariables	v', v', ...	***v'***, ***v'***, ...	***`v′`***, ***`v′`***, ...
type metavariables	$\mathcal{T}', \mathcal{T}''$, ...	$\mathcal{S}', \mathcal{S}''$, ...	$\mathcal{P}', \mathcal{P}''$, ...

Table 2.4: Typesetting conventions.

2.5.2 Rules

The rules operate on signs. The grammar has three logical rules governing the behavior of the implicative type constructors in the type system of each component. We will, however, present them omitting the phenogrammatical and semantic type, and the name we use for the rules references the tectogrammatical type constructor $\multimap$.

The [Ax] rule allows us to introduce hypotheses (i.e. triples of typed variables). Once a hypothesis is introduced, it is stored in the context. Informally, this rule allows us to introduce traces or gaps.

(23) $$\frac{}{\boldsymbol{v}; \mathcal{T}; \boldsymbol{V} \vdash \boldsymbol{v}; \mathcal{T}; \boldsymbol{V}}\,[\mathrm{Ax}]$$

Informally, the [$\multimap$E] rule allows two signs to combine into a larger sign, so long as the tectogrammatical type of one sign ($\mathcal{T}$) is the argument type of the other sign ($\mathcal{T} \multimap \mathcal{T}'$). It is our analogue of Merge in MGG.

More formally, this rule is just the implication elimination rule for each of the three calculi, and in the term calculi it is accompanied by function application.$\cup$ is supposed to denote multiset union.

(24) $$\frac{\Gamma \vdash \phi; \mathcal{T} \multimap \mathcal{T}'; \sigma \qquad \Delta \vdash \phi'; \mathcal{T}; \sigma'}{\Gamma \cup \Delta \vdash (\phi\ \phi'); \mathcal{T}'; (\sigma\ \sigma')}\,[\multimap\mathrm{E}]$$

The [$\multimap$ I] rule allows us to discharge any hypotheses, i.e. bind traces. This rule consists of implication introduction in each of the three calculi, and in the term calculi it is accompanied by λ abstraction.

(25) $$\frac{\Gamma \cup \boldsymbol{v}; \mathcal{T}; \boldsymbol{v} \vdash \phi; \mathcal{T}'; \sigma}{\Gamma \vdash \lambda_{\boldsymbol{v}}.\phi; \mathcal{T} \multimap \mathcal{T}'; \lambda_{\boldsymbol{v}}.\sigma} [\multimap\text{I}]$$

In addition to these three inference rules, we have to state a rule governing the behavior of the dependent product type constructor, $\prod$, which we have introduced in the tectogrammatical type system. Recall that this type constructor allowed us to abstract away from inflectional (or other) features when writing (parts of) lexical entries of expressions which are vague with respect to a subset of such features. It allows us to manage tectogrammatical information in a more economical way, but the elimination of this type constructor has no consequence for the phenogrammatical or the semantic calculus.

This type constructor will only even be introduced lexically in our grammar, i.e as part of a lexical entry (non-logical axiom), so we only give the $\prod$ elimination rule:

(26) $$\frac{\Gamma \vdash \phi; \lambda_{v}.\tau : \prod_{v:\mathcal{T}'} \mathcal{T}; \sigma}{\Gamma \vdash \phi; \tau[\tau'/v] : \mathcal{T}[\tau'/v]; \sigma} [\textstyle\prod\text{E}]$$

side condition: there is a term τ' of type $\mathcal{T}'$, i.e. $\vdash \tau' : \mathcal{T}'$

Chapter 3: Basic Word Order

3.1 Introduction

In this chapter, we analyze simple declarative Serbo-Croatian sentences consisting of an intransitive, transitive or a ditransitive verb, its object(s) and subject, and adverbial modifiers. We consider lexical noun phrases (names, pronouns and quantificational lexical noun phrases), as well as those consisting of a noun and possibly an attributive adjective, quantificational determiner or a postnominal modifier. Verbs which require complements other than noun phrases will not be considered here; instead, we return to those in later chapters.

Here, we also present a general theory of procliticization in Serbo-Croatian, since we will analyze prepositional adverbials and prepositional postnominal modifiers, and prepositions are proclitics in Serbo-Croatian. The enclitics, however, will not be considered in this chapter but in Chapter 5 which is entirely dedicated to encliticization in Serbo-Croatian.

Since relative clauses may contain enclitics, and their analysis will therefore interact with the analysis of enclitics, we will postpone their analysis and only consider other kinds of postnominal modifiers.

The purpose of this chapter, other than to analyze word order in simple clauses, is to introduce basic mechanisms of combination in the grammar which are essential to the theory. As we analyze more complex constituents in the later chapters, we will retain the fundamental assumptions laid out in this chapter concerning the combination of verbs with their subjects and objects, and the construction of declarative clauses in general.

3.2 Data

3.2.1 Lexical Noun Phrases

Agreement

Lexical noun phrases in Serbo-Croatian (names and pronouns) are marked for case, number, gender and person. There are five distinct cases[7], two numbers, three persons for each number, and three genders. There is some syncretism in the paradigm. The examples of lexical noun phrases below show some of the different combinations of these inflectional features. It is worth noting that only

[7] Traditionally, Serbo-Croatian is said to have seven distinct cases. However, we will not consider vocative case-marked noun phrases as they are always extraclausal, that is, they never occur as arguments of any other expressions. Further, while traditionally dative and locative are considered functionally distinct case, there is no difference in form between dative and locative noun phrases (or nouns). Therefore, *dative* in our grammar subsumes the traditional dative and locative cases.

	CASE	GENDER	NUMBER	PERSON
Ane	genitive	feminine	sg	3rd
njemu	dative	masculine	sg	3rd
ono	nominative	neuter	sg	3nd
one	nominative	feminine	pl	3rd
Marka	genitive/accusative	masculine	sg	3rd
vi	nominative	masculine/feminine	pl	2nd
mnom	instrumental	masculine/feminine	sg	1st
tobom	instrumental	masculine/feminine	sg	2nd
nama	dative/instrumental	masculine/feminine	pl	1st
njih	genitive/accusative	masculine/feminine/neuter	pl	3rd

Table 3.1: Examples of inflectional feature combinations on lexical noun phrases.

3rd person noun phrases can be neuter, and that only pronouns can be 1st or 2nd person; that is, all non-pronominal noun phrases are 3rd person. Here we are interested in the inflectional morphology in so far as it influences the distribution of noun phrases in the language, i.e. the syntactic properties of noun phrases.

Clearly, case influences the distributional properties of noun phrases since it determines whether they can be subjects, or objects of certain verbs or prepositions. For example, only nominative noun phrases can be subjects.[8] The verb

[8] In Serbo-Croatian, constructions with dative experiencers are pervasive. There are essentially two types of dative experiences constructions, (i) those which require a nominative argument as well, and (ii) those which do not. For example:

(i) Ani su trebali udžbenici.
Ana$_{DAT,f,sg,3}$ are$_{3,pl}$ need$_{ppl,m,pl}$ textbooks$_{NOM,m,pl,3}$
'Ana needed textbooks'

(ii) Ani je bilo hladno.
Ana$_{DAT,f,sg,3}$ is$_{sg,3}$ be$_{ppl,sg,n}$ cold$_{NOM,n,sg}$

voljeti 'to love' requires an accusative object, *bojati se* 'to be afraid' requires a genitive object, *zadiviti* 'to impress' requires a dative and an instrumental object. The prepositions are equally picky, with, for example, *za* 'for' requiring an accusative object, *iz* 'from' requiring a genitive object, *prema* 'towards' requiring a dative object, and *sa* 'with' requiring an instrumental object. Since differently case marked noun phrases are not interchangeable, case is syntactically significant.

The person, number and gender marking is relevant for subjects. Finite verbs[9] agree with their subjects in number and person, but not gender. However, Serbo-Croatian has periphrastic tenses composed from a finite auxiliary and a non-finite participle. In these constructions, the auxiliary agrees with the subject in person

'Ana was cold'

In constructions like (i), the verb agrees with the nominative argument in person, number and gender, while in constructions like (ii) with no nominative argument the verb is always neuter singular. Only nominative noun phrases induce verbal agreement.

Further, apart from inducing verbal agreement, nominative noun phrases are also special because only they control the interpretation of reflexives, including the pronominal reflexive *sebe* 'self' (see Chapter 4), and the subject-oriented possessive *svoj*. A dative experiencer cannot be coreferential with a reflexive.

So, we will consider subjects nominative noun phrases which induce verbal agreement and control the interpretation of reflexives. We will consider constructions like (ii) *subjectless*, while in constructions like (i) we will call the nominative argument that the verb agrees with the subject. All subjects are nominative; however, not all nominative noun phrases are subjects, cf. predicative structures, Chapter 4.

[9]Serbo-Croatian has three non-periphrastic tenses: present, aorist and imperfect. However, of these, only the present tense is in wide-spread use, aorist and imperfect having been replaced by a periphrastic past tense. We will therefore largely exclude aorist and imperfect from our analysis, with one notable exception which we will return to in Chapter 5, namely, the aorist of *biti* 'to be'. The latter are enclitic verbal forms used to form conditional mood. Keeping this one exception in mind, *finite verbs* will in general be used to refer to present tense verbs, since both the future tense and the most widely used past tense are periphrastic and use present tense auxiliaries.

and number, while the participle agrees with the subject in number and gender. The examples below illustrates this agreement pattern.

(27) finite verbs are gender-neutral:

a. Oni spavaju.
they$_{NOM,m,pl,3rd}$ sleep$_{pl,3rd}$
'They (masculine) sleep'

b. One spavaju.
they$_{NOM,f,pl,3rd}$ sleep$_{pl,3rd}$
'They (feminine) sleep'

c. Ona spavaju.
they$_{NOM,n,pl,3rd}$ sleep$_{pl,3rd}$
'They (neuter) sleep'

(28) finite verbs agree with the subject in number and person:

a. * Marko spavaju.
Marko$_{NOM,m,sg,3rd}$ sleep$_{pl,3rd}$
[intended: 'Marko sleeps']

b. * Mi spavaju.
we$_{NOM,m/f,pl,1st}$ sleep$_{pl,3rd}$
[intended: 'We sleep']

(29) participles in past tense agree with subjects in gender and number:

a. Marko je spavao.
Marko$_{NOM,m,sg,3rd}$ is$_{sg,3rd}$ sleep-PPL$_{m,sg}$
'Marko slept'

b. * Ana je spavao.
Ana$_{NOM,f,sg,3rd}$ is$_{sg,3rd}$ sleep-PPL$_{m,sg}$
[intended: 'Ana slept']

c. * Oni je spavao.
they$_{NOM,m,pl,3rd}$ is$_{sg,3rd}$ sleep-PPL$_{m,sg}$
[intended: 'They (masculine) slept']

At first glance, in simple sentences only case of object noun phrases seems to matter. In other words, If a verb or a preposition need an accusative object, any accusative noun phrase will do, regardless or gender, number or person of that noun phrase. However, there are instances where the gender and number, in addition to case, matter even for non-nominative noun phrases.

First, it is possible to relativize on noun phrases in any grammatical case. In such relative clauses, the relative pronoun must agree in gender and number with the modified noun phrase which requires the number and gender information to be recorded on the noun phrase. Below is an example of such a relative clause, where the relative pronoun *koju* agrees with the noun phrase *Ana* in gender and number:

(30)

Ana,	koju	znam	sa	fakulteta,	je
$\text{Ana}_{NOM,f,sg,3}$	$\text{which}_{ACC,f,sg}$	$\text{know}_{sg,1}$	from	$\text{college}_{GEN,m,sg,3}$	$\text{is}_{sg,3}$
moja	najbolja	prijateljica.			
$\text{my}_{NOM,f,sg,3}$	$\text{best}_{NOM,f,sg,3}$	$\text{friend}_{NOM,f,sg,3}$			

'Ana, who I know from college, is my best friend'

Second, in object-control constructions, the person of the object matters as well. Consider the following example:

(31)

a.

Marko	nagovara	Anu	da	vozi.
$\text{Marko}_{NOM,m,sg,3}$	$\text{convinces}_{sg,3}$	$\text{Ana}_{ACC,f,sg,3}$	COMP	$\text{drive}_{sg,3}$

'Marko convinces Ana to drive'

b.

Marko	nagovara	nas	da	vozimo.
$\text{Marko}_{NOM,m,sg,3}$	$\text{convinces}_{sg,3}$	$\text{we}_{ACC,f/m,pl,1}$	COMP	$\text{drive}_{pl,1}$

'Marko convinces us to drive'

With *nagovarati* 'convince' and other object-control verbs, the embedded verb has to agree in number and person with the matrix object. When an object noun phrase controls an adjective, the adjective has to agree with the object in gender and number:

(32) Marko smatra Anu pametnom.
Marko$_{NOM,m,sg,3}$ considers$_{sg,3}$ Ana$_{ACC,f,sg,3}$ smart$_{INST,f,sg}$
'Marko considers Ana smart'

Here, the accusative object *Anu* and the instrumental adjective *pametnom* must agree in gender and number. In order for this sentence to be composed, the accusative noun phrase *Anu* has to carry information about its number and gender, so that agreement with the adjective can be induced. Because of examples like these, we conclude that number, person and gender are syntactically relevant even for non-nominative noun phrases.

Word Order

It is uncontroversial that in Serbo-Croatian a verb, its subject and any objects can freely order with respect to one another.[10] A sentence consisting of an intransitive verb and its subject can be pronounced two different ways:

(33) a. Vesna spava.
Vesna$_{NOM}$ sleeps/is sleeping

[10]The utterances of the same sentence with different word orders are semantically (i.e. truth-conditionally) identical. It's not clear that one could even argue for significant pragmatic differences, at least in these simple cases. For example, each of the utterances above could be offered as an answer to both *Who is sleeping?* and *What's Vesna doing?*. See Progovac (2005) for a claim that different word orders are associated with pragmatic differences.

'Vesna sleeps/is sleeping'[11]

b. Spava Vesna.

A sentence consisting of a transitive verb, and its subject and object, can be pronounced six different ways (3!):

(34) a. Vesna voli Marka.
Vesna$_{NOM}$ loves Marko$_{ACC}$
'Vesna loves Marko'

b. Vesna Marka voli.

c. Voli Vesna Marka.

d. Voli Marka Vesna.

e. Marka Vesna voli.

f. Marka voli Vesna.

Finally, a sentence consisting of a ditransitive verb, and its subject and objects can be pronounced twenty-four different ways (4!):

(35) a. Vesna predstavlja Marka Ani.
Vesna$_{NOM}$ introduces Marko$_{ACC}$ Ana$_{DAT}$
'Vesna introduces Marko to Ana'

b. Vesna Marka predstavlja Ani.

c. Vesna Marka Ani predstavlja.

d. Predstavlja Marka Ani Vesna.

e. Marka Ani predstavlja Vesna.

f. Ani Vesna Marka predstavlja.

g. etc.

Therefore, the grammar must in general allow for free ordering of verbs and their noun phrase arguments.

[11]The imperfective present tense verb *spava* could be interpreted as denoting a habitual or an ongoing activity, hence the dual gloss. Present tense verbs in main clauses will in general be given in the imperfective form in the examples, and we will henceforth suppress the dual gloss of such verbs' meaning.

We also note that quantificational lexical noun phrases can freely order with respect to other clausal constituents. Further, if a sentence contains two quantificational pronouns, regardless of the word order, the sentence will be ambiguous. Consider the following example:

(36) a. Neko voli svakoga.
somebody$_{NOM,m,sg}$ loves$_{sg,3}$ everybody$_{ACC,m,sg}$
'Somebody loves everybody'
b. Neko svakoga voli.
c. Svakoga neko voli.
d. Svakoga voli neko.
e. Voli neko svakoga.
f. Voli svakoga neko.

Regardless of which of the six possible ways it's pronounced, the sentence above remains ambiguous between the two readings, namely 'there is some person who loves everybody' and 'for every person there is somebody who loves them'.

3.2.2 Phrasal Noun Phrases

Determiner-less Noun Phrases

While Serbo-Croatian has quantificational, demonstrative and possessive determiners, none of them are obligatory. Singular count nouns, bare or with modifiers, can occur as arguments of verbs or prepositions. For example:

(37) Djevojka spava.
girl$_{NOM,f,sg}$ sleeps$_{sg,3}$
'A/The girl sleeps'

The meaning of the bare noun *djevojka* 'girl' is ambiguous between an indefinite and a definite interpretation. The same is true in cases where modifiers occur with the noun. When nouns (with or without modifiers) occur as subjects, they always induce 3rd person agreement with the verb.

Attributive Adjectives

Nouns and adjectives in Serbo-Croatian are marked for case, number and gender and they have to agree in terms of these features. The example below shows the general agreement pattern.

(38) a. Dobri studenti uče.
good$_{NOM,m,pl}$ studenti$_{NOM,m,pl}$ study
'Good students study'

b. * Dobar studenti uče.
good$_{NOM,m,sg}$ studenti$_{NOM,m,pl}$ study
[intended: 'Good students study']

c. * Dobrih studenti uče.
good$_{GEN,m,pl}$ studenti$_{NOM,m,pl}$ study
[intended: 'Good students study']

d. * Dobre studenti uče.
good$_{NOM,f,sg}$ studenti$_{NOM,m,pl}$ study
[intended: 'Good students study']

There does not seem to be a consensus as to the empirical facts concerning the placement of attributive adjectives in Serbo-Croatian. Consider the sentence below:

(39) a. Ana kupuje novi auto.
Ana$_{NOM,f,sg,3}$ buys new$_{ACC,m,sg}$ car$_{ACC,m,sg}$
'Ana buys/is buying a new car'

In my judgment, the sentence above can be pronounced 24 different ways, i.e. all permutations of the verb, the subject NP, the noun and the attributive adjective are possible, with no change in meaning.

On the other hand, Leko (1999), officially discussing Bosnian, claims that attributive adjectives must occur immediately to the left of the noun they modify, while Zlatić (1997), officially discussing Serbian noun phrases, allows for the adjective and the noun to permute, but not for them to in general appear discontinuously in a clause. All three sets of judgments however include free permutation of a verb and its noun phrase arguments, which we discussed in Chapter 3.

Postnominal Modifiers

Postnominal modifiers in Serbo-Croatian include some predicative phrases[12] such as certain prepositional and adjectival phrases, possessive genitive noun phrases which are not predicative but used in circumstances when a possessive determiner cannot be formed, for morphological reasons, and relative clauses. The examples below show a variety of postnominal modifiers in Serbo-Croatian. The modifiers have been enclosed in square brackets.

(40) predicative phrases:

a. Djevojka [iz Beograda] voli Marka.
girl$_{NOM,f,sg}$ from Belgrade$_{GEN,m,sg}$ loves Marko$_{ACC,m,sg}$
'The girl from Belgrade loves Marko'

[12] By 'predicative phrases' we mean phrases which can be complements of the verb *biti* 'be'. Not all predicative phrases can occur as postnominal modifiers in Serbo-Croatian.

b. Ana daje poklone [vrijedne sto dolara].
Ana$_{NOM,f,sg}$ gives presents$_{ACC,m,pl}$ worth$_{ACC,m,pl}$ hundred dollars$_{GEN,m,pl}$
'Ana gives presents worth $100'

(41) possessive phrases:

a. Ana zna druga [moje sestre].
Ana$_{NOM,f,sg}$ knows friend$_{ACC,m,sg}$ my$_{GEN,f,sg}$ sisterGEN,f,sg
'Ana knows my sister's friend'

b. Plate [američkih glumaca] su ogromne.
salaries$_{NOM,f,pl}$ American$_{GEN,m,pl}$ actors$_{GEN,m,pl}$ are huge$_{NOM,f,pl}$
'Salaries of American actors are huge'

(42) relative clauses:

a. Djevojka [koju Ana zna sa fakuteta] dolazi.
girl$_{NOM,f,sg}$ which$_{ACC,f,sg}$ Ana$_{NOM,f,sg}$ knows from college$_{GEN,m,sg}$ arrives
'The girl who Ana knows from college is arriving'

b. Marko ima druga [kojeg Ana zna].
Marko$_{NOM,m,sg}$ has friend$_{ACC,m,sg}$ which$_{ACC,m,sg}$ Ana$_{NOM,f,sg}$ knows
'Marko has a friend who Ana knows'

All postnominal modifiers must occur immediately to the right of the noun they modify, which is why we refer to them as *postnominal*. For example, Serbo-Croatian in general doesn't allow extraposition of relative clauses (Browne (1974))[13]. In the following example, post nominal modifiers are enclosed in square brackets and the modified noun is underlined. They are all ungrammatical:

(43) a. * Djevojka voli [iz Beograda] Marka.

[13]If a relative clause modifies a pronoun, then extraposition if possible, but never if it modifies a noun. This phenomenon seems syntactic in nature and doesn't seem to correlate with whether the relative clause is restrictive or non-restrictive

b. * Ana zna [moje sestre] druga.
c. * Dolazi [koju Ana zna sa fakulteta] djevojka.
d. * Djevojka dolazi [koju Ana zna sa fakulteta]

There is no agreement between the modified noun and a prepositional phrase or a possessive genitive noun phrase; they can modify nouns of any case, number and gender. All postnominal modifiers must remain contiguous, so that no main clause material, including the noun they are modifying, may break up a multi-word postnominal modifier.

Even though in my judgment an attributive adjective and a noun that it modifies can occur discontinuously in a sentence, if there is also a postnominal modifier, the attributive adjective can no longer detach. However, so long as the noun and the adjective remain contiguous and immediately to the left of the postnominal modifier, they can still permute. The examples below illustrates this pattern:

(44) a. Marko ima dobrog druga iz Beograda.
Marko$_{NOM,m,sg}$ has good$_{ACC,m,sg}$ friend$_{ACC,m,sg}$ from Belgrade$_{GEN,m,sg}$
'Marko has a good friend from Belgrade'
b. Marko ima druga dobrog iz Beograda.
c. * Marko dobrog ima druga iz Beograda.
d. * Marko ima druga iz Beograda dobrog.
e. etc.

Note that both Leko (1999)'s and Zlatić (1997)'s pattern of judgments is far less complex. If they require the adjective and the noun to be contiguous anyways, presumably they must be contiguous in the presence of postnominal modifiers as well. We will sketch an analysis of all these judgments in the next section.

Quantificational Determiners

In my judgment, just like attributive adjectives, quantificational determiners can detach from their argument noun. Unlike attributive adjectives, they can do so even when the noun has postnominal modifiers. Consider the examples below:

(45) a. Neka djevojka koju Ana zna dolazi.
some$_{NOM,f,sg}$ girl$_{NOM,f,sg}$ which$_{ACC,f,sg}$ Ana$_{NOM,f,sg}$ knows arrives
'Some girl who Ana knows from college is arriving'
b. Neka dolazi djevojka koju Ana zna.
c. Djevojka koju Ana zna dolazi neka.
d. Djevojka koju Ana zna neka dolazi.
e. etc.

(46) a. Svi ljudi iz Beograda dolaze.
all$_{NOM,m,pl}$ people$_{NOM,m,pl}$ from Belgrade$_{GEN,m,sg}$ arrive
'All the people from Belgrade are arriving'
b. Svi dolaze ljudi iz Beograda.
c. Ljudi iz Beograda dolaze svi.
d. Ljudi iz Beograda svi dolaze
e. etc.

In the next section we will analyze this complicated pattern of judgments, where attributive adjectives, postnominal modifiers and quantificational determiners are all associated with different word order possibilities. Presumably, Zlatić (1997) and Leko (1999) require quantificational determiners to remain contiguous with the rest of the noun phrase, just like attributive adjectives, which we will also show how to represent in the grammar. Further, we will show how to represent Zlatić (1997)'s requirement that quantificational determiners always be left-most in the noun phrase, preceding any attributive adjectives.

3.2.3 Adverbial Modifiers

Non-Prepositional Adverbial Modifiers

Adverbial phrases can freely order with respect to the verb and its noun-phrase arguments:

(47) a. Vesna vozi brzo.
Vesna$_{NOM}$ drives fast
'Vesna drives fast'
b. Vesna brzo vozi.
c. Vozi Vesna brzo.
d. Vozi brzo Vesna.
e. Brzo Vesna vozi.
f. Brzo vozi Vesna.

If the adverbial expression consists of a degree and an adverb, the degree and the adverb must remain contiguous and the degree must precede the adverb. The sequence of the degree and the adverb can, however, freely order with respect to the other clausal constituents. The following examples illustrate the relevant pattern.

(48) Vesna vozi veoma brzo.
Vesna$_{NOM}$ drives very fast
'Vesna drives very fast'

(49) the degree and the adverb must remain contiguous:
a. * Veoma Vesna brzo vozi.
b. * Vozi veoma Vesna brzo.
c. * Brzo Vesna veoma vozi.
d. * Vozi brzo Vesna veoma.
e. etc.

(50) the degree must precede the adverb:

a. * Vesna vozi brzo veoma.
b. * Vozi brzo veoma Vesna
c. etc.

(51) the degree+adverb sequence can freely order with respect to other constituents:

a. Vesna veoma brzo vozi.
b. Veoma brzo Vesna vozi.
c. Veoma brzo vozi Vesna.
d. Vozi veoma brzo Vesna.
e. Vozi Vesna veoma brzo.

Given this data, the grammar must have a way of both (i) allowing free reordering of constituents, and (ii) ensuring that certain multi-word phrases remain contiguous and internally ordered, while freely reordering as a unit with respect to other phrases.

Prepositional Adverbial Modifiers

Prepositions in Serbo-Croatian are clitics, which means that they are not phonological words (i.e., they are not associated with a lexical pitch accent; see Godjevac (1999, 2000); also see Zec and Inkelas (1990) for a slightly different formulation to the same effect). They are *pro*clitics, which means that they must attach to a phonological word to their right.

A more conservative set of judgments about prepositional phrases in Serbo-Croatian would be that the preposition must occur immediately to the left of its argument noun phrase, procliticizing onto the first phonological word therein, and the entire prepositional phrase must remain contiguous. The permissible order within the noun phrase that is an object of a preposition is determined by one's judgments about the order within noun phrases in general, for example whether one believes that the adjective or the quantificational determiner must precede the noun or not. The order within the noun phrase in turn determines which phonological word the preposition procliticizes onto.

However, in my permissive judgment, whether the prepositional phrase must remain contiguous or not, depends on whether its object noun phrase must remain contiguous or not. For example, if the preposition's noun phrase object consists of an adjective and a noun, the preposition can procliticize onto either the noun or the adjective. The two parts of the prepositional phrase can then occur discontinuously in an utterance, provided that the part that includes the preposition precedes the part that does not. The example below illustrates this pattern.

(52) a. U velikom gradu Ana živi.
in big$_{DAT}$ city$_{DAT}$ Ana$_{NOM}$ lives
'Ana lives in a big city'
b. U gradu velikom Ana živi.
c. U velikom Ana živi gradu.
d. U gradu Ana živi velikom.
e. U velikom Ana gradu živi.
f. U gradu Ana velikom živi.
g. Ana u gradu živi velikom.
h. Ana u velikom živi gradu.
etc., but:
i. * Velikom Ana živi u gradu.
j. * Gradu Ana živi u velikom.

etc.

The same pattern is evident in my judgment if the noun phrase contains a quantificational determiner. However, since splitting a noun and its postnominal modifier is in general not possible, if such a noun phrase is the object of the preposition then the whole prepositional phrase must remain contiguous. We will sketch an analysis of both the less permissive (and easier to analyze) set of judgments whereby the entire prepositional phrase must remain contiguous, and the more permissive set of judgments whereby under certain conditions, depending on the structure of the object noun phrase, the prepositional phrase can be made discontinuous.

3.3 Analysis

3.3.1 Lexical Noun Phrases

Representation of Lexical Noun Phrases in the Grammar

To account for the different agreement properties of noun phrases, we assign them various tectogrammatical types which reflect their inflectional properties. We start with the basic tectogrammatical types **Cse**, **Gdr**, **Num**, and **Prs** which 'house' the different inflectional features, such that:

(53) a. **Cse** = {nom, gen, dat, acc, inst}
b. **Gdr** = {m, f, n}
c. **Num** = {sg, pl}
d. **Prs** = {1,2,3}

The tectogrammatical type of noun phrases is then dependent on terms of **Cse**, **Gdr**, **Num**, and **Prs**, so that the grammar can distinguish between noun phrases based on their case, gender, number and person respectively. Starting with the basic type $\mathbf{NP}_{u:\mathbf{Cse},v:\mathbf{Gdr},x:\mathbf{Num},y:\mathbf{Prs}}$ without any specified parameters, by combining it with different terms of type **Cse**, **Gdr**, **Num**, and **Prs**, we obtain more specific noun-phrase types such as:

(54) $\mathbf{NP}_{\text{gen,f,sg,3}}$ genitive feminine singular 3rd person noun phrases
$\mathbf{NP}_{\text{dat,m,sg,3}}$ dative masculine singular 3rd person noun phrases
$\mathbf{NP}_{\text{nom,n,sg,3}}$ nominative neuter singular 3rd person noun phrases

Semantically, lexical noun phrases such as the ones we considered in this chapter denote individuals, so they are represented in the grammar as constants of type $\mathbf{e}$. Phenogrammatically, we analyze them as denoting length one strings of languages, so they are represented as terms of type $\mathbf{z}$. Below are full lexical entries for three different case forms of two names:

(55)
a. $\vdash \mathtt{MARKO}_{\mathbf{z}} : \mathbf{z}; \mathbf{NP}_{\text{nom,m,sg,3}}; \mathsf{marko} : \mathbf{e}$
b. $\vdash \mathtt{MARKA}_{\mathbf{z}} : \mathbf{z}; \mathbf{NP}_{\text{acc,m,sg,3}}; \mathsf{marko} : \mathbf{e}$
c. $\vdash \mathtt{MARKU}_{\mathbf{z}} : \mathbf{z}; \mathbf{NP}_{\text{dat,m,sg,3}}; \mathsf{marko} : \mathbf{e}$
d. $\vdash \mathtt{VESNA}_{\mathbf{z}} : \mathbf{z}; \mathbf{NP}_{\text{nom,f,sg,3}}; \mathsf{vesna} : \mathbf{e}$
e. $\vdash \mathtt{VESNU}_{\mathbf{z}} : \mathbf{z}; \mathbf{NP}_{\text{acc,f,sg,3}}; \mathsf{vesna} : \mathbf{e}$
f. $\vdash \mathtt{VESNI}_{\mathbf{z}} : \mathbf{z}; \mathbf{NP}_{\text{dat,f,sg,3}}; \mathsf{vesna} : \mathbf{e}$

Combining Lexical Noun Phrases with Verbs

We start with an intransitive verb, the present tense forms of *spavati* ‘to sleep’. Recall that finite verbs require a nominative subject which agrees with them in terms of person and number, but it doesn’t care about the gender of its subject. Further, the verb has to combine itself with the subject in a way which will allow them to freely reorder with respect to one another. We associate *spava* 'sleeps' with the following lexical entry:

(56) $\vdash \lambda_{v}.\mathbf{PER}(v \circ \mathtt{SPAVA}_{\mathbf{z}}) : \mathbf{z} \to \mathbf{Z}; \prod_{x:\mathbf{Gdr}}[\mathbf{NP}_{\mathrm{nom},x,\mathrm{sg},3} \multimap \mathbf{S}_{\mathrm{m},6}]; \mathsf{sleep} : \mathbf{e} \to \mathbf{p}$

Semantically, the verb is a function from of individuals to propositions. Phenogrammatically, the verb is looking for an argument of type $\mathbf{z}$. Once it combines with such an argument, it creates a string of languages, and then constructs a set of all permutations of that string of languages. This allows the verb and the subject to freely order with respect to one another.

In very general and informal terms, the tectogrammatical type of this verb reflects the fact that, given an appropriate subject noun phrase (represented by the noun phrase parameters), it can construct a declarative sentence. Syntactically, *spava* 'sleeps' doesn't care about the gender of its subject which is why it's associated with a dependent product type. Focusing on the tectogrammatical calculus only, whose terms we will typically suppress, by using the $\prod$ elimination rule we can obtain three tectogrammatical versions of *spava*, depending on which term of type **Gdr** it combines with. Those three tectogrammatical versions of *spava* are listed below.

(57) $\vdash \mathrm{spava}_{\mathrm{m}} : \mathbf{NP}_{\mathrm{nom,m,sg},3} \multimap \mathbf{S}_{\mathrm{m},6}$

(58) $\vdash \mathrm{spava}_{\mathrm{f}} : \mathbf{NP}_{\mathrm{nom,f,sg},3} \multimap \mathbf{S}_{\mathrm{m},6}$

(59) $\vdash \mathrm{spava}_{\mathrm{n}} : \mathbf{NP}_{\mathrm{nom,n,sg},3} \multimap \mathbf{S}_{\mathrm{m},6}$

The tectogrammatical result type of *spava* is $\mathbf{S}_{\mathrm{m},6}$. Similar to noun phrase types, we also have a family of clause types, indexed by a term of type $\mathbf{K}$={e,m,q} and

Nat={0,1,2,3,...}. The first parameter, of type **K**, refers to the kind of clause in question: e(mbedded), m(ain) or q(uestion). The second parameter is a natural number which will help us enforce the order in the enclitic cluster.[14] We will return to **K** and **N** and their role in the grammar in later chapters. For now, suffice it to say that finite verbs typically build sentences associated with the tectogrammatical type $\mathbf{S}_{\mathrm{m},6}$.

Now we can already construct simple sentences. Below is a derivation of one such sentence, omitting phenogrammatical and semantic types.

(60)
$$\frac{\vdash \lambda_v.\mathbf{PER}(v \circ \mathtt{SPAVA}_{\mathbf{z}}); \mathbf{NP}_{\mathrm{nom,m,sg,3}} \multimap \mathbf{S}_{\mathrm{m},6}; \textsf{sleep} \qquad \vdash \mathtt{MARKO}_{\mathbf{z}}; \mathbf{NP}_{\mathrm{nom,m,sg,3}}; \textsf{marko}}{\vdash \mathbf{PER}(\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{SPAVA}_{\mathbf{z}}); \mathbf{S}_{\mathrm{m},6}; (\textsf{sleep marko})}\ [\multimap E]$$

Since the tectogrammatical type of the object noun phrase and the argument type of the verb are the same, we can use the [$\multimap$E] rule to combine them. [$\multimap$E] is accompanied by function application in the phenogrammatical and semantic term calculi. The resulting sign, the conclusion of the proof above, is the representation of the sentence *Marko spava* 'Marko sleeps' in the grammar. Tectogrammatically, this sign is a declarative sentence, and semantically it denotes the proposition that Marko sleeps.

The phenogrammatical part of the conclusion is the term $\vdash \mathbf{PER}(\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{SPAVA}_{\mathbf{z}}) : \mathbf{Z}$ which denotes a set that contains exactly two strings of languages,

[14]See Morrill and Gavarró (1992) for the original use of the natural number parameter to enforce the order of clitics in a clause. Also note that in practice, we only need numbers 0-6, and tectogrammatical types of signs will only make reference to those numbers.

$\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{SPAVA}_{\mathbf{z}}$ and $\mathtt{SPAVA}_{\mathbf{z}} \circ \mathtt{MARKO}_{\mathbf{z}}$, corresponding to the two ways this sentence could be pronounced, *Marko spava* and *Spava Marko* respectively.

Note that the grammar doesn't treat *Marko spava* and *Spava Marko* as distinct expressions at all, but as one and the same declarative sentence. The grammar generates a single sign whose phenoterm, because it denotes a set of strings of languages, specifies all the different ways that a single sentence could be pronounced.

Lexical entries of transitive verbs are specified similarly. Below is the lexical entry for *voli* 'loves'.[15]

(61) $\vdash \lambda_{vw}.\mathbf{PER}(w \circ \mathtt{VOLI}_{\mathbf{z}} \circ v) : \mathbf{z} \to \mathbf{z} \to \mathbf{Z};$
$\prod_{v,x:\mathbf{Gdr},w:\mathbf{Num},y:\mathbf{Prs}}[\mathbf{NP}_{\mathrm{acc},v,w,y} \multimap \mathbf{NP}_{\mathrm{nom},x,\mathrm{sg},3} \multimap \mathbf{S}_{\mathrm{m},6}];$
$\mathsf{love} : \mathbf{e} \to \mathbf{e} \to \mathbf{p}$

Phenogrammatically, this verb takes two arguments of type $\mathbf{z}$ (the object and the subject noun phrase), and outputs the set of all permutations of the string obtained by concatenating the subject, the verb and the object. Tectogrammatically, *voli* needs an accusative noun phrase of any gender, number and person as its object, and a 3rd person singular nominative noun phrase of any gender as its subject, to construct a declarative main clause.

[15]Note that the $\mathbf{S}$-string that the transitive verb builds and then permutes consists of the subject, followed by the verb, followed by the object. In practice, it makes no difference how we order these arguments since they all get permuted anyways. However, the reader can take our habit of writing phenogrammatical terms for verbs as permutations of the subject-verb-object sequence as a nod to the standard claim that Serbo-Croatian is underlyingly an SVO language; see *inter alia* Godjevac (2000); Progovac (2005).

From the lexical entries for the verb, the subject and the object, in two steps of [$\multimap$E] we can obtain the following sign, which represents the sentence *Marko voli Vesnu* 'Marko loves Vesna' in the grammar:

(62) $\vdash \mathbf{PER}(\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{VOLI}_{\mathbf{z}} \circ \mathtt{VESNU}_{\mathbf{z}}) : \mathbf{Z}; \mathbf{S}_{\mathrm{m},6}; (\mathsf{love\ marko\ vesna}) : \mathbf{p}$

Tectogrammatically, this sign is a declarative main clause. Semantically, it expresses the proposition that Marko loves Vesna. Phenogrammatically, it denotes a set of exactly six strings of languages, which correspond to the six different ways of pronouncing this sentence. Therefore, the grammar predicts that all of the following are possible pronunciations: *Marko voli Vesnu*, *Marko Vesnu voli*, *Voli Marko Vesnu*, *Voli Vesnu Marko*, *Vesnu voli Marko* and *Vesnu Marko voli*.

Lexical entries for ditransitive verbs such as *predstavlja* 'introduces' are given in a similar fashion. Below is a lexical entry for one tectogrammatical version of that verb, obtained by combining it with appropriate gender, number, and person parameters to construct the sentence *Marko predstavlja Vesnu Ani* 'Marko introduces Vesna to Ana', and the sign that the grammar generates for that sentence.

(63) $\vdash \lambda_{vxw}.\mathbf{PER}(w \circ \mathtt{PREDSTAVLJA}_{\mathbf{z}} \circ v \circ x) : \mathbf{z} \rightarrow \mathbf{z} \rightarrow \mathbf{z} \rightarrow \mathbf{Z};$
$\mathbf{NP}_{\mathrm{acc,f,sg,3}} \multimap \mathbf{NP}_{\mathrm{dat,f,sg,3}} \multimap \mathbf{NP}_{\mathrm{nom,m,sg,3}} \multimap \mathbf{S}_{\mathrm{m},6};$
$\mathsf{introduce} : \mathbf{e} \rightarrow \mathbf{e} \rightarrow \mathbf{e} \rightarrow \mathbf{p}$

(64) $\vdash \mathbf{PER}(\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{PREDSTAVLJA}_{\mathbf{z}} \circ \mathtt{VESNU}\mathbf{z} \circ \mathtt{ANI}_{\mathbf{z}}) : \mathbf{Z}; \mathbf{S}_{\mathrm{m},6};$
$(\mathsf{introduce\ vesna\ ana\ marko}) : \mathbf{p}$

Phenogrammatically, the sentence *Marko predstavlja Vesnu Ani* 'Marko introduces Vesna to Ana' is represented as the term $\vdash \mathbf{PER}(\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{PREDSTAVLJA}_{\mathbf{z}} \circ \mathtt{VESNU}\mathbf{z} \circ$

$\mathtt{ANI}_{\mathbf{z}})$: $\mathbf{Z}$. This term denotes a set of exactly twenty-four strings of languages which correspond to the twenty-four different ways of pronouncing this sentence.

Quantificational Lexical Noun Phrases

Recall that quantificational lexical noun phrases can freely order with respect to other clausal constituents, just like any other noun phrases, so the sentence below can be pronounced six different ways, without a change in meaning.

(65) Ana voli svakoga.
Ana$_{NOM,f,sg,3}$ love$_{pl,3}$ everybody$_{ACC,m,sg,3}$
'Ana loves everybody'

Our general strategy with respect to quantification will be to use quantifier lowering (Oehrle (1994)), whereby quantificational noun phrases combine with 'gappy' sentences, i.e. sentences missing a noun phrase, and scope over such constituents, which are semantically properties of individuals. Phenogrammatically, quantificational noun phrases lower themselves into and take the place of the 'gap', hence the name quantifier lowering. Below are the lexical entries required to construct the sentence above.

(66) a. $\vdash \mathtt{ANA}_{\mathbf{z}} : \mathbf{z}; \mathbf{NP}_{\text{nom,f,sg,3}}; \mathsf{ana} : \mathbf{e}$

b. $\vdash \lambda_{vw}.\mathbf{PER}(w \circ \mathtt{VOLI}_{\mathbf{z}} \circ v) : \mathbf{z} \to \mathbf{z} \to \mathbf{Z}; \mathbf{NP}_{\text{acc,m,sg,3}} \multimap \mathbf{NP}_{\text{nom,f, sg,3}} \multimap \mathbf{S}_{\text{m,6}}; \mathsf{love} : \mathbf{e} \to \mathbf{e} \to \mathbf{p}$

c. $\vdash \lambda_{F}.(F\ \mathtt{SVAKOGA}_{\mathbf{z}}) : (\mathbf{z} \to \mathbf{Z}) \to \mathbf{Z}; (\mathbf{NP}_{\text{acc,m,sg,3}} \multimap \mathbf{S}_{\text{m,6}}) \multimap \mathbf{S}_{\text{m,6}};$ $\mathsf{everyone} : (\mathbf{e} \to \mathbf{p}) \to \mathbf{p}$

In the lexical entry for *svakoga*, the semantic term $\vdash \mathsf{everyone} : (\mathbf{e} \to \mathbf{p}) \to \mathbf{p}$ is an abbreviation for $\vdash (\mathsf{every\ person}) : (\mathbf{e} \to \mathbf{p}) \to \mathbf{p}$. The hyperintensional generalized quantifier $\vdash \mathsf{every} : (\mathbf{e} \to \mathbf{p}) \to (\mathbf{e} \to \mathbf{p}) \to \mathbf{p}$ is related to its extensional counterpart via the following meaning postulates:

(67) $\forall_{PQw}[(\mathsf{every}\ P\ Q)@w = \forall_x[(P\ x)@w \to (Q\ x)@w]]$

Below we show the step-by-step derivation of the sentence above. We omit phenogrammatical and semantic types, as well as the non-case **NP** parameters for typesetting reasons.

(68)

$$\dfrac{\vdash \lambda_{vw}.\mathbf{PER}(w \circ \mathtt{VOLI}_{\mathbf{z}} \circ v); \mathbf{NP}_{\mathrm{acc}} \multimap \mathbf{NP}_{\mathrm{nom}} \multimap \mathbf{S}_{\mathrm{m,6}}; \mathsf{love} \qquad \dfrac{}{x; \mathbf{NP}_{\mathrm{acc}}; x \vdash x; \mathbf{NP}_{\mathrm{acc}}; x}\,[\mathrm{Ax}]}{x; \mathbf{NP}_{\mathrm{acc}}; x \vdash \lambda_w.\mathbf{PER}(w \circ \mathtt{VOLI}_{\mathbf{z}} \circ x); \mathbf{NP}_{\mathrm{nom}} \multimap \mathbf{S}_{\mathrm{m,6}}; (\mathsf{love}\ x)}\,[\multimap\mathrm{E}]$$

The first step is to introduce a hypothesis or a trace via [Ax]. The trace is of the same type as the verb's first argument–tectogrammatically, an accusative noun phrase; phenogrammatically, a string of languages; and, semantically, an individual. Then the verb and the hypothesis combine via [⊸E], but the hypothesis is still kept track of in the context, to the left of the turnstile. Now we can proceed and combine the verb phrase with its subject:

(69)

$$\dfrac{\dfrac{x; \mathbf{NP}_{\mathrm{acc}}; x \vdash \lambda_w.\mathbf{PER}(w \circ \mathtt{VOLI}_{\mathbf{z}} \circ x); \mathbf{NP}_{\mathrm{nom}} \multimap \mathbf{S}_{\mathrm{m,6}}; (\mathsf{love}\ x) \qquad \vdash \mathtt{ANA}_{\mathbf{z}}; \mathbf{NP}_{\mathrm{nom}}; \mathsf{ana}}{x; \mathbf{NP}_{\mathrm{acc}}; x \vdash \mathbf{PER}(\mathtt{ANA}_{\mathbf{z}} \circ \mathtt{VOLI}_{\mathbf{z}} \circ x); \mathbf{S}_{\mathrm{m,6}}; (\mathsf{love}\ x\ \mathsf{ana})}\,[\multimap\mathrm{E}]}{\vdash \lambda_x \mathbf{PER}(\mathtt{ANA}_{\mathbf{z}} \circ \mathtt{VOLI}_{\mathbf{z}} \circ x); \mathbf{NP}_{\mathrm{acc}} \multimap \mathbf{S}_{\mathrm{m,6}}; \lambda_x(\mathsf{love}\ x\ \mathsf{ana})}\,[\multimap\mathrm{I}]$$

After combining the verb phrase with the subject, we withdrew the hypothesis, i.e. bound the trace via [$\multimap$I]. This means that the hypothesis no longer appears in the context, and all occurrences of variables that originated with the hypothesis have been bound in the succedent. Now the quantificational noun phrase can combine with this 'gappy' sentence.

(70)

$$\frac{\vdash \lambda_F.(F\ \mathtt{SVAKOGA}_{\mathbf{z}}); \mathbf{NP}_{\mathrm{acc}} \multimap \mathbf{S}_{\mathrm{m},6}; \textsf{everyone} \qquad \vdash \lambda_x \mathbf{PER}(\mathtt{ANA}_{\mathbf{z}} \circ \mathtt{VOLI}_{\mathbf{z}} \circ x); \mathbf{NP}_{\mathrm{acc}} \multimap \mathbf{S}_{\mathrm{m},6}; \lambda_x(\textsf{love}\ x\ \textsf{ana})}{\vdash \mathbf{PER}(\mathtt{ANA}_{\mathbf{z}} \circ \mathtt{VOLI}_{\mathbf{z}} \circ \mathtt{SVAKOGA}_{\mathbf{z}}); \mathbf{S}_{\mathrm{m},6}; \textsf{everyone}\ \lambda_x.(\textsf{love}\ x\ \textsf{ana})} [\multimap \mathrm{E}]$$

Once the quantificational noun phrase combines with the 'gappy' sentence, the term $\vdash \mathtt{SVAKOGA}_{\mathbf{z}} : \mathbf{z}$ takes the place of variable $\vdash x : \mathbf{z}$. The resulting phenoterm denotes a set of six strings of languages, corresponding to the six different pronunciations of this sentence. Semantically, the sentence is analyzed as expressing the expected universally quantified proposition, and tectogrammatically it is an ordinary declarative sentence.

3.3.2 Phrasal Noun Phrases

Representing Lexical Nouns in the Grammar

Just as with noun phrase types, we have a family of noun types indexed by gender, case and number parameters. No person parameters are necessary since nouns do not participate in verbal agreement for which the person parameter is required; only noun phrases agree with verbs. Below we give lexical entries for a few noun phrases, to illustrate the tectogrammatical noun types.

(71) a. $\vdash \mathtt{DJEVOJKA}_{\mathbf{z}} : \mathbf{z}; \mathbf{N}_{\mathrm{acc,f,sg}}; \textsf{girl} : \mathbf{e} \to \mathbf{p}$

b. $\vdash$ $\texttt{DJEVOJKE}_{\mathbf{z}} : \mathbf{z}; \mathbf{N}_{\text{gen,f,sg}}; \mathsf{girl} : \mathbf{e} \rightarrow \mathbf{p}$
c. $\vdash$ $\texttt{DJEVOJKOM}_{\mathbf{z}} : \mathbf{z}; \mathbf{N}_{\text{inst,f,sg}}; \mathsf{girl} : \mathbf{e} \rightarrow \mathbf{p}$
d. $\vdash$ $\texttt{STUDENT}_{\mathbf{z}} : \mathbf{z}; \mathbf{N}_{\text{nom,m,sg}}; \mathsf{student} : \mathbf{e} \rightarrow \mathbf{p}$
e. $\vdash$ $\texttt{STUDENTA}_{\mathbf{z}} : \mathbf{z}; \mathbf{N}_{\text{gen,m,sg}}; \mathsf{student} : \mathbf{e} \rightarrow \mathbf{p}$
f. $\vdash$ $\texttt{STUDENTOM}_{\mathbf{z}} : \mathbf{z}; \mathbf{N}_{\text{inst,m,sg}}; \mathsf{student} : \mathbf{e} \rightarrow \mathbf{p}$

Phenogrammatically, we treat nouns as strings of languages (type $\mathbf{z}$), while semantically we analyze them as expressing functions from individuals to propositions (type $\mathbf{e} \rightarrow \mathbf{p}$).

Quantifying Nouns in Absence of Determiners

Recall that Serbo-Croatian doesn't have obligatory determiners. Consider the example below.

(72) Djevojka spava.
girl$_{NOM,f,sg,3}$ sleeps$_{sg,3}$
'A/The girl sleeps'

Here, the bare noun *djevojka* 'girl' occurs as the subject of the 3rd person singular verb *spava* 'sleeps' and can mean either 'a girl' or 'the girl'. Therefore, we have to have a general schema that will convert nouns into quantificational noun phrases.

We convert nouns into quantificational noun phrases in two steps. First, we state a rule schema that converts nouns of a given case, gender and number into noun phrases of the same case, gender and number. This will be a purely tectogrammatical schema, in the sense that the phenogrammatical and semantic terms and types will not be affected by it at all. Next, we state a rule schema that converts signs which are tectogrammatically noun phrases, but semantically denote properties of individuals, into quantificational noun phrases.

The reason we are doing this in two steps has to do with enforcing ordering restrictions imposed on the attributive adjective in the presence of a postnominal modifier, as we will see later.

Below is a rule schema, [NC], that converts nouns into noun phrases. It allows any noun to be treated, tectogrammatically, as a 3rd person noun phrase of the same case, gender and number, since only pronouns can be of 2nd or 1st person. The phenogrammatical and the semantic portion of the sign remain unchanged.

(73)

$$\frac{\vdash \phi : \mathbf{z}; \mathbf{N}_{\tau':\mathbf{Cse},\tau'':\mathbf{Gdr},\tau''':\mathbf{Num}}; \sigma : \mathbf{e} \to \mathbf{p}}{\vdash \phi : \mathbf{z}; \mathbf{NP}_{\tau':\mathbf{Cse},\tau'':\mathbf{Gdr},\tau''':\mathbf{Num},3}; \sigma : \mathbf{e} \to \mathbf{p}} \text{[NC]}$$

Below is the rule that turns nouns whose tectogrammatical type is noun phrase, but which have the semantic type $\mathbf{e} \to \mathbf{p}$, to be converted into full blown quantificational noun phrases.

(74)

$$\frac{\vdash \phi : \mathbf{z}; \mathbf{N}P_{\tau':\mathbf{Cse},\tau'':\mathbf{Gdr},\tau''':\mathbf{Num},3}; \sigma : \mathbf{e} \to \mathbf{p}}{\vdash \lambda_F.(F\,\phi) : (\mathbf{z} \to \mathbf{Z}) \to \mathbf{Z}; (\mathbf{NP}_{\tau':\mathbf{Cse},\tau'':\mathbf{Gdr},\tau''':\mathbf{Num},3} \multimap \mathbf{S}_{\mathrm{m},6}) \multimap \mathbf{S}_{\mathrm{m},6}; \; (\mathsf{exists}\;\sigma) : (\mathbf{e} \to \mathbf{p}) \to \mathbf{p}} \text{[Quant]}$$

The hyperintensional constant $\vdash \mathsf{exists} : (\mathbf{e} \to \mathbf{p}) \to (\mathbf{e} \to \mathbf{p}) \to \mathbf{p}$ is related to its extensional counterpart via the following meaning postulate:

(75) $\forall_{PQw}[(\mathsf{exists}\; P\, Q)@w = \exists_x[(P\,x)@w \wedge (Q\,x)@w)]]$

As mentioned before, determiner-less noun phrases in Serbo-Croatian can also have definite meaning, in addition to the indefinite meaning which [Quant] introduces. Deriving such definite interpretations of noun phrases requires positing another rule similar to [Quant]. We will not pursue that here, instead focusing on the indefinite interpretation only. Below we show how to apply [NC] and [Quant] to turn the noun *djevojka* 'girl' into a quantificational noun phrase. We omit phenogrammatical and semantic types.

(76)

$$\dfrac{\dfrac{\vdash \mathtt{DJEVOJKA}_{\mathbf{z}}; \mathbf{N}_{\mathrm{nom,f,sg}}; \mathsf{girl}}{\vdash \mathtt{DJEVOJKA}_{\mathbf{z}}; \mathbf{NP}_{\mathrm{nom,f,sg,3}}; \mathsf{girl}}\,[\mathrm{NC}]}{\vdash \lambda_F.(F\ \mathtt{DJEVOJKA}_{\mathbf{z}}); (\mathbf{NP}_{\mathrm{nom,f,sg,3}} \multimap \mathbf{S}_{\mathrm{m,6}}) \multimap \mathbf{S}_{\mathrm{m,6}}; (\mathsf{exists\ girl})}\,[\mathrm{Quant}]$$

A sentence such as *Djevojka spava* 'A girl sleeps' would be represented in the grammar by the following sign:

(77) $\vdash \mathbf{PER}(\mathtt{DJEVOJKA}_{\mathbf{z}} \circ \mathtt{SPAVA}_{\mathbf{z}}) : \mathbf{Z}; \mathbf{S}_{\mathrm{m,6}}; ((\mathsf{exists\ girl})\ \lambda_x.(\mathsf{sleep}\ x)) : \mathbf{p}$

This sentence expresses an existentially quantified proposition that there exists a girl who sleeps, and it can be pronounced two ways, *Djevojka spava* or *Spava djevojka*.

Attributive Adjectives

In this section, we show how to represent in the framework different generalizations about nouns and attributive adjectives which were presented earlier in this chapter.

First, we analyze the most permissive empirical generalization concerning attributive adjectives, namely, that they can be detached from the noun they are modifying so that the noun and the adjective can individually freely reorder with respect to other clausal constituents.

Tectogrammatically, adjectives combine with nouns of a certain gender, case and number, and output a sign with the same tectogrammatical type. Semantically, their type is $(\mathbf{e} \rightarrow \mathbf{p}) \rightarrow (\mathbf{e} \rightarrow \mathbf{p})$.

Phenogrammatically, we assume that attributive adjectives are of type $z \rightarrow z$ and combine with nouns via ordinary string of languages concatenation. Therefore, the result of combining an attributive adjective with a noun is a length two string of languages. This will allow the adjective and the noun to individually freely reorder with respect to other constituents in the sentence.

Suppose we are trying to generate the sentence we mentioned earlier in the chapter, *Ana kupuje novi auto* 'Ana is buying a new car'. Below are the required lexical entries.

(78)
a. $\vdash \mathtt{ANA}_{\mathbf{z}} : \mathbf{z}; \mathbf{NP}_{\mathrm{nom,f,sg,3}}; \mathsf{ana} : \mathbf{e}$
b. $\vdash \lambda_{vw}.\mathbf{PER}(v \circ \mathtt{KUPUJE}_{\mathbf{z}} \circ w) : \mathbf{z} \rightarrow \mathbf{z} \rightarrow \mathbf{Z}; \mathbf{NP}_{\mathrm{acc,m,sg,3}} \multimap \mathbf{NP}_{\mathrm{nom,f,sg,3}} \multimap \mathbf{S}_{\mathrm{m,6}}; \mathsf{buy} : \mathbf{e} \rightarrow \mathbf{e} \rightarrow \mathbf{p}$
c. $\vdash \mathtt{AUTO}_{\mathbf{z}} : \mathbf{z}; \mathbf{N}_{\mathrm{acc,m,sg}}; \mathsf{car} : \mathbf{e} \rightarrow \mathbf{p}$
d. $\lambda_x.\mathtt{NOVI}_{\mathbf{z}} \circ x : \mathbf{z} \rightarrow \mathbf{z}; \mathbf{N}_{\mathrm{acc,m,sg}} \multimap \mathbf{N}_{\mathrm{acc,m,sg}}; \lambda_{Py}.(P\ y)\ \mathsf{and}\ (\mathsf{new}\ y) : (\mathbf{e} \rightarrow \mathbf{p}) \rightarrow (\mathbf{e} \rightarrow \mathbf{p})$

We can combine the noun and the adjective (53a) and then apply [NC] and [Quant] to the resulting sign to obtain (53b).

(79) a. $\vdash \texttt{NOVI}_{\mathbf{z}} \circ \texttt{AUTO}_{\mathbf{z}} : \mathbf{z}; \mathbf{N}_{\text{acc,m,sg}}; \lambda_{y}.(\mathsf{car}\ y)\ \mathsf{and}\ (\mathsf{new}\ y) : \mathbf{e} \rightarrow \mathbf{p}$

b. $\vdash \lambda_{F}.(F\ (\texttt{NOVI}_{\mathbf{z}} \circ \texttt{AUTO}_{\mathbf{z}})) : (\mathbf{z} \rightarrow \mathbf{Z}) \rightarrow \mathbf{Z}; (\mathbf{NP}_{\text{acc,m,sg,3}} \multimap \mathbf{S}_{\text{m,6}}) \multimap \mathbf{S}_{\text{m,6}};$
$\mathsf{exists}(\lambda_{y}.(\mathsf{car}\ y)\ \mathsf{and}\ (\mathsf{new}\ y)) : (\mathbf{e} \rightarrow \mathbf{p}) \rightarrow \mathbf{p}$

The verb combines with an introduced hypothesis that is the object noun phrase trace, then combines with the subject *Ana*. Finally, when that hypothesis is withdrawn via [$\multimap$I] (i.e., when the accusative trace is bound), the quantificational noun phrase *novi auto* 'a new car' can combine with it. The result is given below:

(80) $\vdash \mathbf{PER}(\texttt{ANA}_{\mathbf{z}} \circ \texttt{KUPUJE}_{\mathbf{z}} \circ \texttt{NOVI}_{\mathbf{z}} \circ \texttt{AUTO}_{\mathbf{z}}) : \mathbf{Z}; \mathbf{S}_{\text{m,6}};$
$\mathsf{exists}(\lambda_{y}.(\mathsf{car}\ y)\ \mathsf{and}\ (\mathsf{new}\ y))(\lambda_{x}.(\mathsf{buy}\ x\ \mathsf{ana})) : \mathbf{p}$

The phenoterm of this sign denotes a set of twenty-four strings of languages, corresponding to the twenty-four possible pronunciations of this sentence, according to the most permissive empirical generalization which we are currently considering.

For the intermediate case, where the adjective and the noun are allowed to freely reorder with respect to one another, but must stay contiguous in the sentence, we give the following lexical entry for *novi* 'new'.

(81) $\lambda_{x}.\mathbf{toZ}(\mathbf{k}(\mathbf{PER}(\texttt{NOVI}_{\mathbf{z}} \circ x))) : \mathbf{z} \rightarrow \mathbf{z}; \mathbf{N}_{\text{acc,m,sg}} \multimap \mathbf{N}_{\text{acc,m,sg}};$
$\lambda_{Py}.(P\ y)\ \mathsf{and}\ (\mathsf{new}\ y) : (\mathbf{e} \rightarrow \mathbf{p}) \rightarrow (\mathbf{e} \rightarrow \mathbf{p})$

Semantically and tectogrammatically, everything is the same. The difference is entirely phenogrammatical. While the phenogrammatical type of the adjective is the same, instead of merely concatenating itself with the noun and creating a length two string of languages, as in the most permissive grammar, here the adjective permutes itself with the noun via $\mathbf{PER}$, then compacts the resulting set of strings of languages into a set of strings via $\mathbf{k}$. Finally, that set of strings is turned into a length one string of languages via $\mathbf{toZ}$. This ensures that while the adjective and noun can reorder with respect to one another, they cannot be made discontinuous in a clause.

In this grammar, the sentence *Ana kupuje novi auto* 'Ana is buying a new car' is represented by the following sign:

(82) $\vdash \mathbf{PER}(\mathtt{ANA}_{\mathbf{z}} \circ \mathtt{KUPUJE}_{\mathbf{z}} \circ \mathbf{toZ}(\mathbf{k}(\mathbf{PER}(\mathtt{NOVI}_{\mathbf{z}} \circ \mathtt{AUTO}_{\mathbf{z}})))) : \mathbf{Z}; \mathrm{S}_{\mathrm{m},6};$ $\mathsf{exists}(\lambda_y.(\mathsf{car}\ y)\ \mathsf{and}\ (\mathsf{new}\ y))(\lambda_x.(\mathsf{buy}\ x\ \mathsf{ana})) : \mathbf{p}$

Looking at the phenoterm in more detail, $\mathbf{k}$ compacts $\mathbf{PER}(\mathtt{NOVI}_{\mathbf{z}} \circ \mathtt{AUTO}_{\mathbf{z}}))$ into a set of strings (type $\mathbf{S}$) which contains exactly two strings, $\mathtt{novi}_{\mathbf{s}} \circ \mathtt{auto}_{\mathbf{s}}$ and $\mathtt{auto}_{\mathbf{s}} \circ \mathtt{novi}_{\mathbf{s}}$. That set of strings is then turned into a length one string of languages (type $\mathbf{z}$). Therefore, while the length one string of languages $\mathbf{toZ}(\mathbf{k}(\mathbf{PER}(\mathtt{NOVI}_{\mathbf{z}} \circ \mathtt{AUTO}_{\mathbf{z}})))$ can freely permute with respect to $\mathtt{ANA}_{\mathbf{z}}$ and $\mathtt{KUPUJE}_{\mathbf{z}}$, the adjective and the noun cannot be made discontinuous.

For the most restrictive option, where the adjective and the noun must remain contiguous and the adjective must precede the noun, we give the following lexical entry for *novi* 'new'.

(83) $\lambda_x.\mathbf{toZ}(\mathbf{L}(\mathtt{NOVI}_{\mathbf{z}} \circ x)) : \mathbf{z} \to \mathbf{z}; \mathrm{N}_{\mathrm{acc,m,sg}} \multimap \mathrm{N}_{\mathrm{acc,m,sg}}; \lambda_{Py}.(P\ y)\ \mathsf{and}\ (\mathsf{new}\ y) :$ $(\mathbf{e} \to \mathbf{p}) \to (\mathbf{e} \to \mathbf{p})$

The adjective now concatenates itself with the noun, and then immediately 'linguifies' that string of languages into a set of strings via $\mathbf{L}$. $\mathbf{toZ}$ then turns that set of strings into a length one string of languages. In this grammar, the sentence *Ana kupuje novi auto* 'Ana is buying a new car' is represented by the following sign:

(84) $\vdash \mathbf{PER}(\mathrm{ANA}_{\mathbf{z}} \circ \mathrm{KUPUJE}_{\mathbf{z}} \circ \mathbf{toZ}(\mathbf{L}(\mathrm{NOVI}_{\mathbf{z}} \circ \mathrm{AUTO}_{\mathbf{z}}))) : \mathbf{Z}; \mathbf{S}_{\mathrm{m},6};$
$\mathsf{exists}(\lambda_y.(\mathsf{car}\ y)\ \mathsf{and}\ (\mathsf{new}\ y))(\lambda_x.(\mathsf{buy}\ x\ \mathsf{ana})) : \mathbf{p}$

In the phenoterm of this sign, $\mathbf{L}(\mathrm{NOVI}_{\mathbf{z}} \circ \mathrm{AUTO}_{\mathbf{z}})$ denotes a set of strings with exactly one member, $\mathtt{novi}_{\mathbf{s}} \circ \mathtt{auto}_{\mathbf{s}}$. So, the adjective must precede the noun, and while the length one string of languages $\mathbf{toZ}(\mathbf{L}(\mathrm{NOVI}_{\mathbf{z}} \circ \mathrm{AUTO}_{\mathbf{z}}))$ can freely reorder with respect to other length one strings of languages in the sentence, it cannot be made discontinuous.

Postnominal Modifiers

Recall that it is uncontroversial that postnominal modifiers must remain contiguous and occur immediately to the right of the noun they are modifying. In this section, we will analyze postnominal modifiers that are prepositional phrases, thereby simultaneously giving our theory of procliticization in Serbo-Croatian.

As in the previous section, we start with the most permissive generalization, building on the set of judgments according to which, under ordinary circumstances, an attributive adjective and a noun can detach and freely reorder with respect to other clausal constituents. In this permissive grammar, a sequence of a noun and an attributive adjective is phenogrammatically represented as a length two string of languages, which allows them to freely permute.

In the presence of postnominal modifiers, however, a noun and an attributive adjective can still reorder with respect to one another, but must remain contiguous and occur immediately to the left of the postnominal modifier. The example below illustrates the judgments patterns that we are currently considering.

(85) a. Marko ima dobrog druga iz Beograda.
Marko$_{NOM,m,sg,3}$ has$_{sg,3}$ good$_{ACC,m,sg}$ friend$_{ACC,m,sg}$ from Belgrade$_{GEN,m,sg}$
'Marko has a good friend from Belgrade'
b. Marko ima druga dobrog iz Beograda.
c. Marko dobrog druga iz Beograda ima.
d. Marko druga dobrog iz Beograda ima.
e. * Marko dobrog ima druga iz Beograda.
f. * Marko dobrog druga ima iz Beograda.
g. * Marko ima iz Beograda dobrog druga.
h. etc.

Note that last two examples, where the prepositional phrase and the adjective+noun sequence occur discontinuously are grammatical, but not on the relevant interpretation. That is, they are compatible with the interpretation on which Marko and his good friend met in Belgrade (the prepositional phrase being interpreted as adverbial). However, the target interpretation is the one on which the friend is from Belgrade (the prepositional phrase being interpreted as modifying the noun).

The first challenge we have to address is the fact that a postnominal modifier restricts the word order possibilities for the noun and the adjective. Informally, it's as if the adjective has to 'know' whether there is also a postnominal modifier or not, and then 'behave' accordingly. We account for this by demanding that any attributive adjectives combine with the noun *before* any postnominal modifiers do so. Then, when the postnominal modifier appears, it 'freezes' the entire phrase, preventing adjectives from escaping.

More formally, we accomplish this via tectogrammatical typing. Whereas attributive adjectives require an argument whose type is in the **N** family, we analyze postnominal modifiers as requiring an argument whose type is in the **NP** family. In other words, we require that the target of postnominal modification be nouns that have undergone the [NC] rule, which changes their tectogrammatical type from a noun type to a corresponding noun phrase type, while leaving its phenogrammar and semantics unchanged, so that they still denote strings of languages, and their semantic type is still $\mathbf{e} \rightarrow \mathbf{p}$.

For example, the type of *druga* 'friend' is $\mathbf{N}_{acc,m,sg}$, as is the type of *dobrog druga* 'good friend'. That phrasal noun can then undergo the rule [NC], which changes its tectogrammatical type to $\mathbf{NP}_{acc,m,sg,3}$. At this stage, a postnominal modifier can apply and output something of the same noun phrase type, so that *dobrog druga iz Beograda* 'good friend from Belgrade' would be of type $\mathbf{NP}_{acc,m,sg,3}$. Since attributive adjectives require arguments of a certain noun, not noun phrase, type, no more attributive adjectives could apply because of a type mismatch. In this way, the grammar forces all attributive adjectives to combine with the noun before any postnominal modifiers do.

Since postnominal modifiers do not agree with the nouns they modify, they will in general be associated with the following tectogrammatical type:

(86) $\prod_{x:\mathbf{Cse},y:\mathbf{Gdr},z:\mathbf{Num},3}[\mathbf{NP}_{x:\mathbf{Cse},y:\mathbf{Gdr},z:\mathbf{Num},3} \multimap \mathbf{NP}_{x:\mathbf{Cse},y:\mathbf{Gdr},z:\mathbf{Num},3}]$

Intuitively, this type just captures the fact that whatever the case, gender or number of its argument, the postnominal modifier will combine with it and then output something with the same agreement features.

Let's consider the preposition *iz* 'from' in the context of postnominal modifiers. It has to combine with its argument noun phrase and procliticize onto some word in its argument. Then, it combines with a noun that it modifies, and must ensure that all chunks of that noun occur immediately to the left of the prepositional phrase.

Below is the version of the lexical entry for *iz* 'from' which occurs in the sentence *Marko ima dobrog druga iz Beograda* 'Marko has a good friend from Belgrade' which we examined above.

(87)
$$\begin{array}{l} \vdash \lambda_{vw}.\mathtt{toZ}[\mathtt{k}(\mathtt{PER}\ w) \bullet (\lambda_s.\exists_t[(\mathtt{k}\ (\mathtt{PER}\ v)\ t) \wedge \\ s = (\mathtt{iz\#}(\mathtt{fst_p}\ t))_{\mathtt{s}} \cdot (\mathtt{rst_s}\ t)])] : \mathbf{z} \rightarrow \mathbf{z} \rightarrow \mathbf{z}; \\ \mathbf{NP}_{\text{gen,m,sg,3}} \multimap \mathbf{NP}_{\text{acc,m,sg,3}} \multimap \mathbf{NP}_{\text{acc,m,sg,3}}; \\ \lambda_{xPy}.(\mathsf{from}\ x\ y)\ \mathsf{and}\ (P\ y) : \mathbf{e} \rightarrow (\mathbf{e} \rightarrow \mathbf{p}) \rightarrow (\mathbf{e} \rightarrow \mathbf{p}) \end{array}$$

Semantically, the preposition needs an argument of type $\mathbf{e}$ (its argument noun phrase), then an argument of type $\mathbf{e} \rightarrow \mathbf{p}$ (the noun that the prepositional phrase will modify). The constant $\vdash \mathsf{from} : \mathbf{e} \rightarrow \mathbf{e} \rightarrow \mathbf{p}$ expresses a binary relation on individuals.

Now we examine the phenogrammatical term in more detail. Intuitively, the variable v stands for the noun phrase argument of the preposition, while w stands for the noun that the prepositional phrase modifies.

The preposition permutes the noun it modifies, then compacts it into a string of languages which is expressed by the subterm $\mathbf{k}\ (\mathbf{PER}\ w)$. The reason it permutes the noun is that, given the set of judgments we are considering now, the adjective and the noun can freely order with respect to each other before the postnominal modifier. So, if the noun is *dobrog druga* 'good friend', it predicts that both *dobrog druga* and *druga dobrog* are possible.

As for its argument noun phrase, the preposition permutes it as well, for the same reasons as in the case of the modified noun. Then it compacts the resulting set of strings of languages into a set of strings. t is one string in that set. This is all expressed in the subterm $(\mathbf{k}\ (\mathbf{PER}\ v)\ t)$.

The constant $\vdash \# : \mathbf{c} \to \mathbf{p} \to \mathbf{p}$ takes a clitic and and phonological word and procliticizes the clitic onto that phonological word resulting in another phonological word. The preposition *iz* in this way procliticizes onto the first word of its argument noun phrase, which is expressed in the subterm $(\mathtt{iz}\#(\mathbf{fst_p}\ t\ ))$. This new phonological word is then turned into a length one string and concatenated with the rest of the noun phrase, expressed in the subterm $(\mathtt{iz}\#(\mathbf{fst_p}\ t))_\mathbf{s} \cdot (\mathbf{rst_s}\ t)$.

We then fuse the set of all strings constructed in this way, by procliticizing the preposition onto the first phonological word of the noun phrase, and then putting it together with the rest of the noun phrase, with the set of strings we obtained by compacting the permutations of the modified noun. The result is a set of strings in which the postnominal modifier, with the appropriately placed clitic, occurs to the right of all modified noun material.

This language is then converted into a length one string of languages via **toZ**, which ensures that the entire phrasal noun, with the postnominal modifier and perhaps some attributive adjectives as well, remains contiguous and that it is impervious to any permutations the verb may require of its arguments.

More concretely, we show how to construct the noun phrase *dobrog druga iz Beograda* 'a good friend from Belgrade'. First we combine the preposition and its argument noun phrase, which results in the following sign:

(88) $\vdash \lambda_w.\mathtt{\mathbf{toZ}}[\mathbf{k}(\mathbf{PER}\,w) \bullet (\lambda_s.\exists_t[(\mathbf{k}\,(\mathbf{PER}\,\mathrm{BEOGRADA}_{\mathbf{z}})\,t) \wedge s = (\mathtt{iz\#}(\mathbf{fst_p}\,t))_{\mathbf{s}} \cdot (\mathbf{rst_s}\,t)])]; \mathbf{NP}_{\mathrm{acc,m,sg,3}} \multimap \mathbf{NP}_{\mathrm{acc,m,sg,3}}; \lambda_{Py}.(\mathsf{from\ belgrade}\ y)\ \mathsf{and}\ (P\,y)$

In this case the object of the preposition is a one-word noun phrase, so *iz* procliticizes onto `beograda`, which ultimately results in a length one string $(\mathtt{iz\#beograda})_{\mathbf{s}}$. That string is then concatenated with $\mathtt{e}_{\mathbf{p}}$, since $\mathtt{e}_{\mathbf{p}}$ is the identity for string concatenation. So, phenogrammatically, the second argument of $\bullet$ in the sign above simply denotes the set of strings $\{(\mathtt{iz\#beograda})_{\mathbf{s}}\}$.

Then, we combine the adjective with the noun, and then apply the rule [NC] which results in the following sign:

(89) $\vdash \mathrm{DOBROG}_{\mathbf{z}} \circ \mathrm{DRUGA}_{\mathbf{z}} : \mathbf{z}; \mathbf{NP}_{\mathrm{acc,m,sg,3}}; (\mathsf{good\ friend}) : \mathbf{e} \rightarrow \mathbf{p}$

Now *iz Beograda* can combine with *dobrog druga*, resulting in the following sign:

(90) $\vdash \mathtt{\mathbf{toZ}}[\mathtt{\mathbf{k}}(\mathtt{\mathbf{PER}}\,(\mathtt{DOBROG}_{\mathbf{z}} \circ \mathtt{DRUGA}_{\mathbf{z}})) \bullet (\lambda_s.\exists_t[(\mathtt{\mathbf{k}}\,(\mathtt{\mathbf{PER}}\,\mathtt{BEOGRADA}_{\mathbf{z}})\ t) \wedge$
$s = (\mathtt{iz\#}(\mathtt{\mathbf{fst}}_{\mathbf{p}}\ t))_{\mathbf{s}} \cdot (\mathtt{\mathbf{rst}}_{\mathbf{s}}\ t)])]; \mathbf{NP}_{\mathrm{acc,m,sg,3}};$
$\lambda_y.(\text{from belgrade } y) \text{ and } (\text{good friend } y) : \mathbf{e} \to \mathbf{p}$

Looking at the phenoterm, it denotes a length one string of languages constructed out of the set that contains exactly two strings, $\mathtt{dobrog}_{\mathbf{s}} \cdot \mathtt{druga}_{\mathbf{s}} \cdot (\mathtt{iz\#beograda})_{\mathbf{s}}$, and $\mathtt{druga}_{\mathbf{s}} \cdot \mathtt{dobrog}_{\mathbf{s}} \cdot (\mathtt{iz\#beograda})_{\mathbf{s}}$. Now we can apply the [Quant] rule, and turn this sign into a quantificational noun phrase which can then combine with a sentence with an accusative 'gap'. Even though verbs allow free reordering of themselves and their noun phrase arguments, this noun phrase has to stay intact because it is a length one string of languages and the verb can't take it apart. The entire noun phrase can however permute with respect to the subject and the verb.

For the less permissive sets of judgments with respect to the ordering of attributive adjectives and nouns, we can simplify the lexical entry given above for the preposition *iz* 'from' and just give the following one:

(91) $\vdash \lambda_{vw}.\mathtt{\mathbf{toZ}}[(\mathtt{\mathbf{L}}\ w) \bullet (\lambda_s.\exists_t[(\mathtt{\mathbf{L}}\ v\ t) \wedge s = (\mathtt{iz\#}(\mathtt{\mathbf{fst}}_{\mathbf{p}}\ t))_{\mathbf{s}} \cdot (\mathtt{\mathbf{rst}}_{\mathbf{s}}\ t)])] :$
$\mathbf{z} \to \mathbf{z} \to \mathbf{z}; \mathbf{NP}_{\mathrm{gen,m,sg,3}} \multimap \mathbf{NP}_{\mathrm{acc,m,sg,3}} \multimap \mathbf{NP}_{\mathrm{acc,m,sg,3}};$
$\lambda_{xPy}.(\text{from } x\ y) \text{ and } (P\ y) : \mathbf{e} \to (\mathbf{e} \to \mathbf{p}) \to (\mathbf{e} \to \mathbf{p})$

This lexical entry works both for the grammar where the adjective and noun can reorder with respect to one another but must remain contiguous, and for the grammar where the adjective must precede the noun. Here, the modified noun and the argument noun phrase of the preposition are not required to permute. Everything else is the same as in the lexical entry for the most permissive grammar.

Quantificational Determiners

Recall that in the most permissive grammar quantificational determiners can in general be detached from their argument noun, even if that noun contains a postnominal modifier, in contrast to attributive adjectives which may not do so. In the less permissive case, the quantificational determiners have to immediately precede the argument noun (Zlatić (1997)).

For the more permissive grammar, we add the following lexical entry for the quantificational determiner *svaka* 'every'.

(92) $\vdash \lambda_{vF}.(F\,(\mathtt{SVAKA}_{\mathbf{z}} \circ v)) : \mathbf{z} \rightarrow (\mathbf{z} \rightarrow \mathbf{Z}) \rightarrow \mathbf{Z}; \mathbf{NP}_{\mathrm{nom,f,sg,3}} \multimap (\mathbf{NP}_{\mathrm{nom,f,sg,3}} \multimap \mathbf{S}_{\mathrm{m,6}}) \multimap \mathbf{S}_{\mathrm{m,6}}; \lambda_{xP}.(\mathsf{every}\ x)P : (\mathbf{e} \rightarrow \mathbf{p}) \rightarrow (\mathbf{e} \rightarrow \mathbf{p}) \rightarrow \mathbf{p}$

This determiner has to pick up an argument of an appropriate noun phrase type first, but otherwise it works just like the quantificational pronouns we considered earlier in the chapter. Phenogrammatically, it concatenates itself with its first argument, and the resulting string of languages is simply lowered into the 'gap' of its second argument, the finite verb phrase.

Since the verbs in general permute themselves with their arguments and in this case the noun phrase which contains *svaka* 'every' is of length greater than one, the quantificational determiner can be detached from the rest of the noun phrase and freely reorder with respect to other constituents. This is the case no matter what the internal structure of its first argument, i.e. whether its first argument contains postnominal modifiers, or attributive adjectives, or not. So we predict that, in the permissive grammar, a sentence such as *Svaka djevojka spava* 'Every girl sleeps' can be pronounced six different way, and a sentence such as *Svaka djevojka iz Beograda spava* ‘Every girl from Belgrade sleeps’ can be pronounced also six different ways, since *svaka* can detach from the rest of the noun phrase material but *djevojka iz Beograda*, because of how we analyzed postnominal modification, must remain contiguous.

For the less permissive grammar which requires that the quantificational determiner occur immediately to the left of its argument, we give the following lexical entry:

(93) $\vdash \lambda_{vF}.(F\ (\mathbf{toZ}(\mathbf{L}(\mathrm{SVAKA}_{\mathbf{z}} \circ v)))) : \mathbf{z} \to (\mathbf{z} \to \mathbf{Z}) \to \mathbf{Z}; \mathbf{NP}_{\mathrm{nom,f,sg,3}} \multimap (\mathbf{NP}_{\mathrm{nom,f,sg,3}} \multimap \mathbf{S}_{\mathrm{m,6}}) \multimap \mathbf{S}_{\mathrm{m,6}}; \lambda_{xP}.(\text{every } x)P : (\mathbf{e} \to \mathbf{p}) \to (\mathbf{e} \to \mathbf{p}) \to \mathbf{p}$

The only difference is in the phenoterm. In this case, the quantificational determiner concatenates itself with its first argument. The resulting string is turned into a length one string of languages. This ensures that once the quantificational noun phrase combines with the verb phrase the noun phrase cannot be discontinuous and the determiner must occur on its left periphery. In this grammar, a sentence such as *Svaka djevojka spava* ‘Every girl sleeps’ and a sentence such as *Svaka djevojka iz Beograda spava* ‘Every girl from Belgrade sleeps’ are predicted to be pronounceable two different ways, since all the noun phrase material must remain contiguous.

3.3.3 Adverbial Modifiers

Single-Word Adverbs

Tectogrammatically, we analyze adverbial expressions as verb phrase modifiers. This means that adverbs need a verb phrase argument, and output something of the same type. However, in our grammar there is strictly speaking no verb phrase type, since various finite verb phrases are tectogrammatically distinguished in terms of the kind of subject required to form a sentence. In other words, they differ in terms of number, gender and person parameters of their subject noun phrase.[16] Adverbs need to be able to combine with a verb phrase requiring a subject of such-and-such gender, number and person, and output a modified verb phrase that retains those same subject requirements, so that the verb/subject agreement is retained.

We define **Adv** as an abbreviation for the following dependent product type:

(94) $\mathbf{Adv} =_{def} \prod_{w:\mathbf{Gdr},x:\mathbf{N},y:\mathbf{Prs}}[(\mathbf{NP}_{\mathrm{nom},w,x,y} \multimap \mathbf{S}_{\mathrm{m},6}) \multimap \mathbf{NP}_{\mathrm{nom},w,x,y} \multimap \mathbf{S}_{\mathrm{m},6}]$

The dependent product type above, which we associate with adverbs, ensures that the gender, number and person parameters of the subject required by the non-modified argument verb phrase are also required by the resulting, modified verb phrase.

[16]Since adverbs can occur not just in main declarative clauses, but also in interrogative and embedded clauses, and can modify participial and infinitival verb phrases as well, we have to generalize their tectogrammatical types not just with respect to the subject noun phrase parameters but also with respect to the resulting clause type parameters. In this chapter, we will abstract away from that by only dealing with adverbs in main declarative clauses, and we will return to this issue in later chapters when discussing other clause types.

Now we can give the following lexical entry for the adverb *brzo* 'fast':

(95) $\vdash \lambda_{Fvw}.\exists_x[(Fvx) \wedge (\mathtt{BRZO}_{\mathbf{z}} \odot x \;\; w)] : (\mathbf{z} \rightarrow \mathbf{Z}) \rightarrow \mathbf{z} \rightarrow \mathbf{Z}$;
$\mathbf{Adv}$; fast : $(\mathbf{e} \rightarrow \mathbf{p}) \rightarrow \mathbf{e} \rightarrow \mathbf{p}$

Examining the phenogrammatical term of this sign in more detail, we see that the adverb first combines with the argument of type $\mathbf{z} \rightarrow \mathbf{Z}$ (F, the verb phrase). The variable v is a placeholder for the subject. Recall that verbs build a set of all permutations of the string of languages consisting of itself, the subject and any objects, so the subterm Fv above denotes the set of strings obtained by combining the verb phrase with its subject. x is one string of languages in that set. A sentence that contains the adverb denotes a set of strings of languages w, where w is any string obtained by shuffling the adverb into x.

To give a concrete example, we will construct the sentence *Marko vozi brzo* 'Marko drives fast'. Below are the lexical entries for the tectogrammatically appropriate versions of *vozi* and the adverb.

(96) $\vdash \lambda_v.\mathbf{PER}(v \circ \mathtt{VOZI}_{\mathbf{z}}) : \mathbf{z} \rightarrow \mathbf{Z}; \mathbf{NP}_{\mathrm{nom},m,\mathrm{sg},3} \multimap \mathbf{S}_{\mathrm{m},6}$; drive : $\mathbf{e} \rightarrow \mathbf{p}$

(97) $\vdash \lambda_{Fvw}.\exists_x[(Fvx) \wedge (\mathtt{BRZO}_{\mathbf{z}} \odot x \;\; w)] : (\mathbf{z} \rightarrow \mathbf{Z}) \rightarrow \mathbf{z} \rightarrow \mathbf{Z}$;
$\mathbf{Adv}_{\mathrm{m,sg,3}}$; fast : $(\mathbf{e} \rightarrow \mathbf{p}) \rightarrow \mathbf{e} \rightarrow \mathbf{p}$

From these two signs and the lexical entry for *Marko* we can construct the following sign:

(98) $\vdash \lambda_w.\exists_x[(\mathbf{PER}(\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{VOZI}_{\mathbf{z}})x) \wedge (\mathtt{BRZO}_{\mathbf{z}} \odot x \;\; w)]; \mathbf{S}_{\mathrm{m},6}$; (fast drive marko) :
$\mathbf{p}$

The subterm $\mathbf{PER}(\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{VOZI}_{\mathbf{z}})$ denotes a set which contains exactly two strings of languages, $\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{VOZI}_{\mathbf{z}}$ and $\mathtt{VOZI}_{\mathbf{z}} \circ \mathtt{MARKO}_{\mathbf{z}}$. By shuffling $\mathtt{BRZO}_{\mathbf{z}}$ into each of these two strings, we obtain the following set of strings of languages, which is precisely what the whole phenoterm of the sign above denotes:

(99)
$$
\begin{array}{l}
\{\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{VOZI}_{\mathbf{z}} \circ \mathtt{BRZO}_{\mathbf{z}}, \\
\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{BRZO}_{\mathbf{z}} \circ \mathtt{VOZI}_{\mathbf{z}}, \\
\mathtt{VOZI}_{\mathbf{z}} \circ \mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{BRZO}_{\mathbf{z}}, \\
\mathtt{VOZI}_{\mathbf{z}} \circ \mathtt{BRZO}_{\mathbf{z}} \circ \mathtt{MARKO}_{\mathbf{z}}, \\
\mathtt{BRZO}_{\mathbf{z}} \circ \mathtt{VOZI}_{\mathbf{z}} \circ \mathtt{MARKO}_{\mathbf{z}}, \\
\mathtt{BRZO}_{\mathbf{z}} \circ \mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{VOZI}_{\mathbf{z}}\}
\end{array}
$$

The six strings of languages in this set correspond exactly to the six possible pronunciations of the sentence *Marko vozi brzo* 'Marko drives fast'.

At this point the reader may be wondering why we are shuffling in the adverb by using $\odot$, instead of letting it permute with the constituents in the sentence by using **PER**. It doesn't matter which of those functions we choose for the simple sentences we are considering here in which all verbal arguments are lexical noun phrases. However, when we extend the grammar to deal with more complex constituents such as phrasal noun phrases and sentential complements, if the adverb were to introduce **PER** in its phenogrammatical term, it could wreck any pre-existing structures and islands of inflexible order already established in the verb phrase. Shuffling it in allows enough flexibility to get the possible word orders, without interfering too much with the existing word order in the verb phrase. We will see examples of this in later chapters.

Adverbial Degrees

Next we turn to cases where an adverbial expression occurs with a degree. Recall that the degree must immediately precede the adverb it modifies, but the entire degree+adverb sequence can be freely ordered with respect to other clausal constituents. Semantically, we must analyze adverbial degrees as adverbial modifiers, of type $((\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}) \to (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$. To preserve verb/subject agreement, we define **Deg** to be the following tectogrammatical type:

(100) $\mathbf{Deg} =$

$$_{def}\prod_{\mathrm{nom},w:\mathbf{Gdr},x:\mathbf{N},y:\mathbf{Prs}}[((\mathbf{NP}_{\mathrm{nom},w,x,y} \multimap \mathbf{S}_{\mathrm{m},6}) \multimap \mathbf{NP}_{\mathrm{nom},w,x,y} \multimap \mathbf{S}_{\mathrm{m},6}) \multimap (\mathbf{NP}_{\mathrm{nom},w,x,y} \multimap \mathbf{S}_{\mathrm{m},6}) \multimap \mathbf{NP}_{\mathrm{nom},w,x,y} \multimap \mathbf{S}_{\mathrm{m},6}]$$

Below we focus on the phenogrammatical part of the lexical entry for the adverbial degree *veoma* 'very'. Note that in the phenoterm below, $\vdash F : \mathtt{z} \to \mathtt{Z}$ and $\vdash G : (\mathtt{z} \to \mathtt{Z}) \to \mathtt{z} \to \mathtt{Z}$.

(101) $$\lambda_{GFvw}.\exists_{xy}[(F\ \ v\ \ x) \wedge (G\ (\lambda_{z}.1_{\mathtt{S}})\ \mathrm{e}_{\mathtt{S}}\ y) \wedge w = \mathtt{toZ}(\mathtt{L}(\mathrm{VEOMA}_{\mathtt{z}} \circ y)) \odot x] : ((\mathtt{z} \to \mathtt{Z}) \to \mathtt{z} \to \mathtt{Z}) \to (\mathtt{z} \to \mathtt{Z}) \to \mathtt{z} \to \mathtt{Z}$$

So, the degree first combines with the adverb (G). The resulting sign then combines with the verb phrase (F), and finally the subject (v). The subterm $(F\ \ v)$ stands for the verb phrase combined with the subject. It denotes a set of that contains all permutations of the string consisting of the subject and the verb phrase. The subterm $(F\ \ v\ \ x)$ means that x is a string in $(F\ \ v)$.

As for the adverb, G, the degree essentially 'destroys' all the argument slots in G, by feeding it the empty $\mathtt{S}$-language, $1_{\mathtt{S}}$, and then the empty $\mathtt{S}$-string, $\mathrm{e}_{\mathtt{S}}$. In the case of the adverb *brzo* 'fast', $(G\,(\lambda_z.1_{\mathbf{S}})\;\mathtt{e}_{\mathbf{S}})$ would amount to the set of strings of languages that contains exactly one string of languages, namely, $\mathtt{BRZO}_{\mathbf{z}}$. Call that string y.

The degree then concatenates itself with y, the adverb, and then 'linguifies' the resulting string via $\mathbf{\mathtt{L}}$ thereby creating a set of strings. Finally, that set of strings is turned into a length-one string of languages via $\mathbf{\mathtt{toZ}}$, which is then shuffled into x, a string in the set of all permutations of the verb phrase and the subject.

More concretely, the sign representing the sentence *Marko vozi veoma brzo* 'Marko drives very fast' has the following phenoterm:

(102) $$\lambda_w.\exists_{xy}[(\mathbf{\mathtt{PER}}(\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{VOZI}_{\mathbf{z}})x) \wedge \exists_{x'}[(1_{\mathbf{S}}x') \wedge (\mathtt{BRZO}_{\mathbf{z}} \odot x' \;\; y)] \wedge w = \mathbf{\mathtt{toZ}}(\mathbf{\mathtt{L}}(\mathtt{VEOMA}_{\mathbf{z}} \circ y)) \odot x] : \mathbf{\mathtt{Z}}$$

The fact that the degree+adverb string of languages (of length two) is 'linguified' into the set $\{\mathtt{veoma}_{\mathbf{S}} \cdot \mathtt{brzo}_{\mathbf{S}}\}$ which is then turned into length-one string of languages, ensures that the degree and the adverb remain contiguous and exactly in that order. Since $\odot$ (and also $\mathbf{\mathtt{PER}}$) is a function of strings of languages, it cannot pull apart the degree+adverb unit, since it has been turned into an atomic string of languages. $\odot$ cannot 'see' the internal structure of the range of that length one string of languages, i.e. it doesn't have access to the set of strings $\{\mathtt{veoma}_{\mathbf{S}} \cdot \mathtt{brzo}_{\mathbf{S}}\}$.

Prepositional Adverbial Modifiers

Recall that in the less permissive grammar, prepositional phrases must remain contiguous. The preposition must occur immediately to the left of its argument noun phrase, and no discontinuities in the noun phrase are allowed either.

In the more permissive grammar, a discontinuity is allowed only if the chunk of the noun phrase that the preposition procliticized onto precedes in the sentence the chunk of the noun phrase that the preposition did not procliticize onto.

We start with the less permissive set of judgments first, because the phenogrammatical part of the lexical entry for the preposition is very similar to the one we gave in the case of prepositional postnominal modifiers.

Suppose we are trying to generate the sentence *Ana živi u velikom gradu* 'Ana lives in a big city'. We give the following lexical entry for the preposition that builds a verb phrase modifier:

(103) $$\lambda_{yFvw}.\exists_{xzt}[(F\ v\ x) \wedge (\mathbf{L}\ v\ t) \wedge z = \mathbf{toZ}(\lambda_s.s = (\mathrm{u}\#(\mathbf{fst_p}\ t))_{\mathbf{s}} \cdot (\mathbf{rst_s}\ t)) \wedge ((z \odot x)\ w)] : \mathbf{z} \to (\mathbf{z} \to \mathbf{Z}) \to \mathbf{z} \to \mathbf{Z}; \mathbf{NP}_{\text{dat, m,sg,3}} \multimap \mathbf{Adv}; \lambda_x \mathsf{in} : \lambda_{yPx}.(((\mathsf{in}\ y)\ P)\ x)$$

Phenogrammatically, this sign is very similar to the sign for the preposition which builds a postnominal modifier. The preposition procliticizes onto the first word in its object noun phrase. The grammar treats the whole prepositional phrase as a length one string of languages. Therefore, when such an adverbial prepositional phrase is shuffled into the sentence, it cannot be made discontinuous.

For the more permissive grammar, we give the following lexical entry for the preposition that builds a verb phrase modifier:

(104) $\lambda_{yFvw}.\exists_{xzt}[(F\ v\ x) \wedge (\mathbf{k}\ (\mathbf{PER}\ y)t) \wedge z = \mathbf{toZ}(\lambda_s.s = (\mathrm{u}\#(\mathbf{fst_p}\ t))_{\mathbf{s}} \circ \mathbf{toZ}(\mathbf{rst_s}\ t)) \wedge ((z \odot x)\ w)] : \mathbf{z} \to (\mathbf{z} \to \mathbf{Z}) \to \mathbf{z} \to \mathbf{Z};$
$\mathbf{NP}_{\text{dat, m,sg,3}} \multimap \mathbf{Adv}; \lambda_{yPx}.(((\mathsf{in}\ y)\ P)\ x)$

Phenogrammatically, the difference between this lexical entry and the analogous entry in the less permissive grammar, is that the prepositional phrase is treated as a string of languages of length two, not one. The first string is the preposition and its host, and the second string is the remainder of the preposition's object noun phrase. Therefore, when the adverbial prepositional phrase is shuffled into the sentence, the two strings can appear discontinuously, but the first string, consisting of the proclitic and its host, must always precede the second string.

3.4 Conclusion

In this chapter we have analyzed some simple Serbo-Croatian sentences, but more importantly, we have illustrated how the grammar works, and seen the basic effects of some essential phenogrammatical functions such as **PER**, $\odot$, **L**, **k** and **toZ**, as well as the cliticization function #, which will continue to play an important role in our theory of Serbo-Croatian word order. We have also made several generalizations about the representations of Serbo-Croatian expressions in the grammar which are summarized below.

	TECTOGRAMMATICAL TYPE	SEMANTIC TYPE	PHENOGRAMMATICAL TYPE
1	**N** family	$\mathbf{e} \rightarrow \mathbf{p}$	$\mathbf{z}$
2	**NP** family	$\mathbf{e} \rightarrow \mathbf{p}$	$\mathbf{z}$
3	**NP** family	$\mathbf{e}$	$\mathbf{z}$
4	$(\mathbf{NP} \multimap \mathbf{S}) \multimap \mathbf{S}$	$(\mathbf{e} \rightarrow \mathbf{p}) \rightarrow \mathbf{p}$	$(\mathbf{z} \rightarrow \mathbf{z}) \rightarrow \mathbf{z}$

Table 3.2: Summary of noun and noun phrase types.

All agreement features of nouns and noun phrases are built into the tectogrammatical types in the **N** and **NP** family. In order the account for the different word order possibilities, nouns and noun phrases are assigned to different combinations of tectogrammatical, phenogrammatical and semantic types, summarized in the table below.

The first type in the table above corresponds to lexical nouns. Signs of that type are arguments and results of adjectival modification. Finally, signs of that type can undergo the [NC] rule which changes their tectogrammatical type from some **N** type to the corresponding **NP** type.

The signs that tectogrammatically have **NP** types, but otherwise behave just like nouns, are the result of the application of the [NC] rule. They are also arguments and results of postnominal modification. Finally, they can be arguments of determiners as well as undergo the [Quant] rule.

The signs that tectogrammatically have **NP** types but semantically denote individuals are non-quantificational lexical noun phrases such as proper names.

Finally, the fourth type in the table above consists of quantificational noun phrases. This includes lexical quantificational noun phrases, the signs that result from the application of the [Quant] rule, as well as noun phrases that contain quantificational determiners.

Intransitive, transitive and ditransitive verbs combine with either lexical noun phrases which denote individuals, or are eventually picked up by quantificational noun phrases as arguments. Phenogrammatically, such verbs combine with arguments of type **`z`** and via **`PER`** construct sets of all permutations of themselves and their arguments (type **`Z`**), to account for the free ordering of verbs, and their objects and subjects. Quantificational noun phrases lower themselves into the 'gap' site, and therefore also participate in the free ordering of verbs and their arguments.

Adverbial phrases combine with verb phrases and output modified verb phrases with the same subject requirements. Phenogrammatically, they shuffle themselves into various permutations of the verb and its arguments via $\odot$. Adverbial degrees construct length-one strings of languages from the degree+adverb unit via **`L`** and **`toZ`**, which ensures that the degree always immediately precedes the adverb as they are shuffled into the sentence.

Finally, prepositions, whether in the case of noun or verb phrase modification, procliticize onto the first word of their object noun phrase. Depending on the prepositional phrase in question, and how permissive one's grammar is, the prepositional phrase is either turned into a length one string which requires it to remain intact, or it is turned into a length two string, which allows it to occur discontinuously in the sentence.

Chapter 4: Embedding, Predicative and Control

4.1 Introduction

Whereas in the previous chapter we considered only verbs whose arguments are noun phrases, in this chapter we turn our attention to verbs with more complex predicates. In particular, we analyze embedded declarative clauses, subject and object control structures and predicative complements.

We think it's necessary to first provide our general theory of embedding, control and predication, before we analyze enclitics, because (i) embedded clauses and controlled finite verb phrases are domains for enclitic placement, and (ii) the set of enclitics in Serbo-Croatian includes predicative and control verbs. In this chapter, we focus only on non-clitic predicative and control verbs, but we build our analysis of clitic verbs on the basic assumptions laid out here.

4.2 Embedded Declarative Clauses

4.2.1 Data

Embedded declarative clauses in Serbo-Croatian must obligatorily occur with a complementizer. There are different complementizers and the two most common declarative complementizers are *da* and *što*. Clause embedding expressions require an embedded clause with a certain complementizer; i.e. not all types of embedded clauses are compatible with all embedding expressions. Although here we mainly focus on clauses headed by *da*, the obligatoriness of a particular kind of complementizer is illustrated in the examples below.

(105) a. Ana misli da Marko spava.
Ana$_{NOM,f,sg,3}$ think$_{sg,3}$ DA Marko$_{NOM,m,sg,3}$ sleep$_{sg,3}$
'Ana thinks that Marko is sleeping'

b. * Ana misli što Marko spava.

c. * Ana misli Marko spava.

(106) a. Ani smeta što Marko stalno spava.
Ana$_{DAT,f,sg,3}$ bother$_{sg,3}$ ŠTO Marko$_{NOM,m,sg,3}$ always sleep$_{sg,3}$
'It bothers Ana that Marko is always sleeping'

b. * Ani smeta da Marko stalno spava.

c. * Ani smeta Marko stalno spava.

The kinds of constituents which can freely order inside a main declarative clause can also freely reorder inside an embedded declarative clause. The complementizer, however, must be leftmost in the embedded clause.

(107) a. Ana misli da Marko voli Vesnu.
Ana$_{NOM,f,sg,3}$ think$_{sg,3}$ DA Marko$_{NOM,m,sg,3}$ love$_{sg,3}$ Vesna$_{ACC,f,sg,3}$
'Ana thinks that Marko loves Vesna'

b. Ana misli da Marko Vesnu voli.
c. Ana misli da voli Vesnu Marko.
d. Ana misli da Vesnu Marko voli.
e. * Ana misli voli da Marko Vesnu.
f. * Ana misli voli Marko da Vesnu.

Embedded clauses cannot be made discontinuous. That is, main clause material cannot occur inside the embedded clause:

(108) a. Ana misli da Marko spava.
Ana$_{NOM,f,sg,3}$ think$_{sg,3}$ DA Marko$_{NOM,m,sg,3}$ sleep$_{sg,3}$
'Ana thinks that Marko is sleeping'
b. * Ana da Marko misli spava.
c. * Misli da Ana Marko stalno spava.
d. etc.

Embedded clauses also have to occur on the right edge of the matrix clause and cannot freely reorder with respect to the verb and the subject:

(109) a. Ana misli da Marko spava.
Ana$_{NOM,f,sg,3}$ think$_{sg,3}$ DA Marko$_{NOM,m,sg,3}$ sleep$_{sg,3}$
'Ana thinks that Marko is sleeping'
b. * Ana da Marko spava misli.
c. * Misli da Marko stalno spava Ana.
d. etc.

The only exception to this are matrix clause adverbial expressions, which may occur immediately to the right of the embedded clause, though they still can't occur inside of the embedded clause. Below we consider a sentence in which the matrix verb is compatible with an adverbial prepositional phrase with a dative complement (denoting a location), but the embedded verb is not.

(110) a. Ana je rekla da će Marko doći na
Ana$_{NOM,f,sg,3}$ is$_{sg,3}$ say$_{ppl,f,sg}$ DA will$_{3,sg}$ Marko$_{NOM,m,sg,3}$ come$_{inf}$ at
sastanku, ali meni je kasnije rekla da neće
meeting$_{DAT,m,sg}$ but I$_{DAT,sg,3}$ is$_{sg,3}$ later say$_{ppl,f,sg}$ DA not-will$_{sg.3}$
doći.

'Ana said at the meeting that Marko would/will come, but later she told me he wouldn't/won't
#'Ana said that Marko would/will come to the meeting, but later she told me he wouldn't/won't'

b. Ana je rekla na sastanku da će Marko doći.

c. Ana je na sastanku rekla da će Marko doći.

d. Na sastanku je Ana rekla da će Marko doći.

e. * Ana je rekla da će na sastanku Marko doći.

The adverbial in the (a) sentence cannot be modifying the embedded clause, that is, it cannot mean 'to the meeting', so it must be a matrix adverbial. That adverbial can, just like adverbials in general, freely reorder with respect to other clausal constituents (b-d), but cannot occur inside of the embedded clause (e).

If we pick an adverbial that is compatible with both the matrix and the embedded verb, and place that adverbial on the right edge of the sentence, ambiguity will arise:

(111) a. Ana nam je rekla da je Marko
Ana$_{NOM,f,sg,3}$ we$_{DAT,pl,1}$ is$_{sg,3}$ say$_{ppl,f,sg}$ DA is$_{sg,3}$ Marko$_{NOM,m,sg,3}$
došao tek juče.
come$_{ppl,m,sg}$ only yesterday

'Ana told us that Marko came only yesterday'
'Ana told us only yesterday that Marko came'

b. Ana nam je rekla da je Marko došao tek juče, i rekla je da je tu već nekoliko dana.

'Ana told us only yesterday that Marko came and she said he'd been here a few days already'

c. Ana nam je rekla da je Marko došao tek juče, iako je trebao doći prije nekoliko dana.

'Ana told us that Marko came only yesterday, even though he was supposed to come a few days ago'

The sentence (a) is ambiguous with respect to the adverbial interpretation. In sentences (b) and (c) we provide disambiguating context to draw out each possible interpretation of the adverbial.

4.2.2 Analysis

So far we've only been concerned with clauses whose **K** parameter is m, i.e. main declarative clauses, since the ultimate result type of finite verbs is $\mathbf{S}_{\mathrm{m},6}$. We analyze complementizers such as *da* as expressions that turn main declarative clauses into embedded declarative clauses, whose type is $\mathbf{S}_{\mathrm{e},6}$.

Suppose we are trying to generate the sentence *Ana misli da Marko spava* 'Ana thinks that Marko is sleeping'. Below are the lexical entries for the complementizer *da* and the sentence embedding verb *misli*:

(112) $\vdash \lambda_{Xw}.\exists_{v}[(X\ v) \wedge w = (\mathtt{DA}_{\mathbf{z}} \circ v)] : \mathbf{Z} \to \mathbf{Z}; \mathbf{S}_{\mathrm{m},6} \multimap \mathbf{S}_{\mathrm{e},6}; \lambda_{q}.q : \mathbf{p} \to \mathbf{p}$

(113) $\vdash \lambda_{Xvw}.\exists_{y}[(\mathbf{PER}(v \circ \mathtt{MISLI}_{\mathbf{z}})\ y) \wedge w = y \circ \mathbf{toZ}(\mathbf{k}\ X)] : \mathbf{Z} \to \mathbf{z} \to \mathbf{Z};$
$\mathbf{S}_{\mathrm{e},6} \multimap \mathbf{NP}_{\mathrm{nom,\ f,sg,3}} \multimap \mathbf{S}_{\mathrm{m},\ 6}; \mathsf{think} : \mathbf{p} \to \mathbf{e} \to \mathbf{p}$

Tectogrammatically, the complementizer picks up a main declarative clause and outputs an embedded declarative clause. Semantically, it's an identity function on propositions, meaning that it does not affect the meaning of its complement clause. Phenogrammatically, its lexical entry ensures that the complementizer occurs on the left edge of the embedded clause, with the rest of the embedded clause material reordering insofar as the embedded verb allows such reordering of itself and its arguments.

The sentence embedding verb *misli* 'thinks' needs an embedded clause argument and a subject arguments to make a main declarative sentence. Semantically, it expresses a relation between an individual and a proposition. Phenogrammaticaly, such a verb allows free reordering of itself with its subject. However, it turns its sentential complement into a length one string of languages which ensures that no embedded clause material can escape into the matrix clause and that no matrix clause material may occur inside the embedded clause. It then requires the embedded clause to occur on the right edge of the main clause, after some permutation of the itself and its subject.

Below we show how to derive the sentence *Ana misli da Marko spava* 'Ana thinks that Marko is sleeping'. First, the complementizer combines with the declarative main clause *Marko spava* 'Marko sleeps' and turns it into an embedded clause with the same meaning.

(114)

$$\frac{\vdash \lambda_{Xw}.\exists_{v}[(X\ v) \wedge w = (\mathtt{DA}_{\mathbf{z}} \circ v)]; \mathbf{S}_{\mathrm{m},6} \multimap \mathbf{S}_{\mathrm{e},6}; \lambda_{q}.q \qquad \vdash \mathbf{PER}(\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{SPAVA}_{\mathbf{z}}); \mathbf{S}_{\mathrm{m},6}; (\mathsf{sleep\ marko})}{\vdash \lambda_{w}.\exists_{v}[(\mathbf{PER}(\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{SPAVA}_{\mathbf{z}})\, v) \wedge w = (\mathtt{DA}_{\mathbf{z}} \circ v)]; \mathbf{S}_{\mathrm{e},6}; (\mathsf{sleep\ marko})}\ [\multimap E]$$

Next, the sentence embedding verb combines with its sentential complement, resulting in the following sign:

(115) $\lambda_{xy}.\exists_{z}[(\mathbf{PER}(x \circ \texttt{MISLI}_{\mathbf{z}})\ z) \wedge y = z \circ \mathtt{toZ}(\mathbf{k}\ (\lambda_{w}.\exists_{v}[(\mathbf{PER}(\texttt{MARKO}_{\mathbf{z}} \circ \texttt{SPAVA}_{\mathbf{z}})\ v) \wedge w = (\texttt{DA}_{\mathbf{z}} \circ v)])))] : \mathtt{z} \to \mathtt{Z}; \mathbf{NP}_{\text{nom, f,sg,3}} \multimap \mathbf{S}_{\text{m, 6}}; \mathsf{think}(\mathsf{sleep\ marko}) : \mathsf{e} \to \mathsf{p}$

Finally, this verb phrase can combine with the subject, resulting in the following sign:

(116) $\lambda_{y}.\exists_{z}[(\mathbf{PER}(\texttt{ANA}_{\mathbf{z}} \circ \texttt{MISLI}_{\mathbf{z}})\ z) \wedge y = z \circ \mathtt{toZ}(\mathbf{k}\ (\lambda_{w}.\exists_{v}[(\mathbf{PER}(\texttt{MARKO}_{\mathbf{z}} \circ \texttt{SPAVA}_{\mathbf{z}})\ v) \wedge w = (\texttt{DA}_{\mathbf{z}} \circ v)])))] : \mathtt{Z}; \mathbf{S}_{\text{m, 6}}; (\mathsf{think}(\mathsf{sleep\ marko})\mathsf{ana}) : \mathsf{p}$

Tectogrammatically, this is a main declarative clause. Semantically it expresses the proposition that Ana thinks that Marko sleeps. Phenogrammatically, it denotes a set of string of languages, each of which consists of some permutation of the verb and the subject, followed by the length one string of languages constructed out of the embedded clause in which the complementizer is always the first string.

Since we analyze adverbial expressions as shuffling into the sentence via $\odot$, we predict that matrix adverbials will be able to occur anywhere in the main clause, including at its right edge after the embedded clause. Since the verb turns the embedded clause into a length one string of languages, the matrix adverb will not, however, be able to occur inside the embedded clause.

	SINGULAR		PLURAL	
	IMPF	PF	IMPF	PF
	full/clitic	full	full/clitic	full
1	jesam/sam	budem	jesmo/smo	budemo
2	jesi/si	budeš	jeste/ste	budete
3	jeste/je	bude	jesu/su	budu

Table 4.1: The verb *biti* 'be' paradigm.

4.3 Predicatives

4.3.1 Data

In this section, we are concerned with complements of the verb *biti* 'be'. It has a perfective and an imperfective present tense paradigm, the latter consisting of full and enclitic forms. Both paradigms are presented in the table below.

This verb can take a variety of complements, including predicative adjectives, predicative prepositional phrases, noun phrases, passive participles and past participle, the latter being used in the periphrastic past tense construction.

(117) a. Igor je pametan.
Igor$_{NOM,m,sg,3}$ is$_{sg,3}$ smart$_{NOM,m,sg}$
'Igor is smart'

b. Vi ste iz Beograda.
you$_{NOM,pl,2}$ su$_{pl,2}$ from Belgrade$_{GEN,m,sg}$
'They are from Belgrade'

c. Ona je studentkinja.
she$_{NOM,f,sg,3}$ is$_{sg,3}$ student$_{NOM,f,sg,3}$
'She is a student'

d. Knjiga je pročitana.
book$_{NOM,f,sg,3}$ is$_{sg,3}$ read$_{pass,nom,f,sg}$
'The book is/has been read'

e. Mi smo pročitali knjigu.
we$_{NOM,pl,1}$ are$_{pl,1}$ read$_{ppl,m,pl}$ book$_{ACC,f,sg,3}$
'We read the book'

Predicative prepositional phrases do not agree with the matrix subject at all. Adjectives and passive participles, which have the same morphology as adjectives, must occur in the nominative case in predicative contexts. Predicative adjectives, passive participles and past participles must agree with the subject in number and gender.

Predicative noun phrases also must occur in the nominative case, but do not necessarily have to agree with the subject in gender and number. For example, in addition to (c) above, where the subject and the predicative noun phrase agree in gender and number, the following are also possible:

(118) a. Ona je student.
she$_{NOM,f,sg,3}$ is$_{sg,3}$ student$_{NOM,m,sg,3}$
'She is a student'

b. Huligani su veliki problem u našem društvu.
hooligan$_{NOM,m,pl,3}$ are$_{pl,3}$ big$_{NOM,m,sg}$ problem$_{NOM,m,sg}$ in our$_{DAT,n,sg}$ society$_{DAT,n,sg}$
'Hooligans are a big problem in our society'

In the next chapter, we will analyze the clitic forms of the verb *biti*. In this chapter we will abstract away from that complicating factor and only consider the full forms *biti*. Here we only mention, but do not analyze, the conditional mood construction which consists of the aorist of *biti* and a past participle, since there are no non-clitic forms of the aorist of *biti*. Below, *biste* is glossed as 'would' but it is really an aorist clitic of *biti*.

(119) Vi biste kupili to.
you$_{NOM,pl,2}$ would$_{pl,2}$ buy$_{ppl,m,pl}$ that$_{ACC,n,sg,3}$
'You would buy that'

Considering only predicative structures that contain non-clitic forms of the copula, the word order is largely unrestricted. For example, a predicative adjective, or a predicative noun phrase can freely order with respect to the verb and the subject:

(120) a. Igor jeste pametan.
Igor$_{NOM,m,sg,3}$ is$_{sg,3}$ smart$_{NOM,m,sg}$
'Igor is smart'
b. Pametan jeste Igor.
c. Pametan Igor jeste.
d. Jeste Igor pametan.
e. Jeste Pametan Igor.
f. Igor pametan jeste.

(121) a. Mi jesmo studenti.
we$_{NOM,pl,1}$ are$_{pl,1}$ students$_{NOM,m,pl,3}$
'We are students'
b. Mi studenti jesmo.
c. Studenti mi jesmo.
d. Studenti jesmo mi.
e. Jesmo studenti mi.
f. Jesmo mi studenti.

Verbal predicative complements, that is passive and past participles, can also freely order with respect to other clausal constituents. If these participles have arguments of their own, those arguments can also be freely ordered in the sentence.

(122) a. Ana jeste dobila pismo.
Ana$_{NOM,f,sg,3}$ is$_{sg,3}$ received$_{ppl,f,sg}$ letter$_{ACC,n,sg,3}$
'Ana received a letter'
b. Ana pismo jeste dobila.
c. Jeste dobila Ana pismo.
d. Pismo Ana jeste dobila.
e. etc.

(123) a. Pismo jeste poslano Ani.
letter$_{NOM,n,sg,3}$ is$_{sg,3}$ sent$_{pass,NOM,n,sg}$ Ana$_{DAT,f,sg,3}$
'A letter was sent to Ana'
b. Jeste Ani poslano pismo.
c. Pismo jeste Ani poslano.
d. Ani pismo jeste poslano.
e. etc.

As for prepositional phrases, as in the case of adverbial prepositional phrases, we will entertain two sets of judgments. A less permissive set of judgments requires that the entire predicative prepositional phrase remain contiguous, but freely order with respect to other clausal constituents. A more permissive set of judgments accepts discontinuities in the prepositional phrase so long as the chunk which contains the preposition precedes the chunk of the prepositional phrase which does not.

Below we present an analysis of these predicative structures, and also try establish connections with the remainder of the grammar by exploring the relationships between predicative complements and the counterparts of those expressions that occur as verbal or nominal modifiers.

4.3.2 Analysis

We introduce a family of tectogrammatical types $\mathbf{Prd}_x$ which will be the result type of predicative phrases. We introduce a tectogrammatical type **D**, such that x in $\mathbf{Prd}_x$ is of type **D**. The terms of type **D** are ps,pl,n,a and pp for passive participles, past participles, noun phrases, adjectives and prepositional phrases respectively.

Past and Passive Participles

Recall that past and passive participles agree with the subject in gender and number. In addition, passive participles must occur in nominative case predicatively.

Suppose we are trying to construct a representation of the sentence *Marko jeste spavao* 'Marko slept'. Below we give the lexical entry for the past participle *spavao*.[17]

(124) $\vdash \texttt{SPAVAO}_{\mathbf{z}} : \mathbf{z}; \prod_{p:\mathbf{Prs}}[\mathbf{NP}_{\text{nom,m,sg},p} \multimap \mathbf{Prd}_{\text{pl}}]; \lambda_x.\mathsf{PST}(\mathsf{sleep}\ x) : \mathbf{e} \rightarrow \mathbf{p}$

[17]We abstract away from a tense analysis and simply assume that there is a propositional operator $\vdash \mathsf{PST} : \mathbf{p} \rightarrow \mathbf{p}$ which contributes the correct temporal interpretation.

spavao requires of its subject that it be a masculine singular nominative noun phrase, but it does not care about the person of the subject, i.e. *Ja jesam spavao* ‘I slept’ and *Ti jesi spavao* ‘You slept’ are both possible, in addition to many similar sentences with a 3rd person subject. For this reason, its tectogrammatical type is a dependent product. The version of the participle needed for the sentence *Marko jeste spavao* ‘Marko slept’ is given below:

(125) $\vdash \texttt{SPAVAO}_{\mathbf{z}} : \mathbf{z}; \mathbf{NP}_{\text{nom,m,sg,3}} \multimap \mathbf{Prd}_{\text{pl}}; \lambda_x.\mathsf{PST}(\mathsf{sleep}\ x) : \mathbf{e} \rightarrow \mathbf{p}$

The non-clitic 3rd person singular form *jeste* ‘is’ is represented by the following sign:

(126) $\vdash \lambda_{vw}.\mathbf{PER}(w \circ \texttt{JESTE}_{\mathbf{z}} \circ v) : \mathbf{z} \rightarrow \mathbf{z} \rightarrow \mathbf{Z};$
$\prod_{g:\mathbf{Gdr},d:\mathbf{D}}[(\mathbf{NP}_{\text{nom},g,\text{sg},3} \multimap \mathbf{Prd}_d) \multimap \mathbf{NP}_{\text{nom},g,\text{sg},3} \multimap \mathbf{S}_{\text{m},6}];$
$\lambda_{Px}.(Px) : (\mathbf{e} \rightarrow \mathbf{p}) \rightarrow \mathbf{e} \rightarrow \mathbf{p}$

The auxiliary *jeste* requires of its subject that it be a singular 3rd person singular nominative noun phrase but it doesn’t care about its gender. Also, it will take any predicative phrase as its complement. This is why its tectogrammatical type is a dependent product. Below is the tectogrammatical version of this verb needed for the sentence *Marko jeste spavao* ‘Marko slept’, looking for a masculine subject.

(127) $\vdash \lambda_{vw}.\mathbf{PER}(w \circ \texttt{JESTE}_{\mathbf{z}} \circ v) : \mathbf{z} \rightarrow \mathbf{z} \rightarrow \mathbf{Z};$
$(\mathbf{NP}_{\text{nom,m,sg,3}} \multimap \mathbf{Prd}_{\text{pl}}) \multimap \mathbf{NP}_{\text{nom,m,sg,3}} \multimap \mathbf{S}_{\text{m},6};$
$\lambda_{Px}.(Px) : (\mathbf{e} \rightarrow \mathbf{p}) \rightarrow \mathbf{e} \rightarrow \mathbf{p}$

Semantically, *jeste* predicates its complement’s meaning of its subject’s meaning. Phenogrammatically, it permutes itself with its arguments, resulting in a set of strings of languages. When we combine the copula with the participle we get the following sign:

(128) $\vdash \lambda_w.\mathbf{PER}(w \circ \mathtt{JESTE}_{\mathbf{z}} \circ \mathtt{SPAVAO}_{\mathbf{z}}) : \mathbf{z} \rightarrow \mathbf{Z};$
$\mathbf{NP}_{\text{nom,m,sg,3}} \multimap \mathbf{S}_{\text{m,6}}; \lambda_x.\mathsf{PST}(\mathsf{sleep}\ x) : \mathbf{e} \rightarrow \mathbf{p}$

Because of the tectogrammatical typing of the copula and the past participle it is impossible to introduce a subject with an inappropriate case, or gender, person or number features.

As another example, below we give a lexical entry for a past participle of a transitive verb *kupiti* 'buy' and the sign that represents the sentence *Ana jeste kupila knjigu* 'Ana bought a book'.

(129) $\vdash \lambda_v.\mathtt{KUPILA}_{\mathbf{z}} \circ v : \mathbf{z} \rightarrow \mathbf{z};$
$\prod_{g:\mathbf{Gdr},n:\mathbf{Num},p,p':\mathbf{Prs}}[\mathbf{NP}_{\text{acc},g,n,p'} \multimap \mathbf{NP}_{\text{nom,f,sg},p} \multimap \mathbf{Prd}_{\text{pl}}];$
$\lambda_{yx}.\mathsf{PST}(\mathsf{buy}\ y\ x) : \mathbf{e} \rightarrow \mathbf{e} \rightarrow \mathbf{p}$

(130) $\vdash \mathtt{ANA}_{\mathbf{z}} \circ \mathtt{JESTE}_{\mathbf{z}} \circ \mathtt{KUPILA}_{\mathbf{z}} \circ \mathtt{KNJIGU}_{\mathbf{z}} : \mathbf{Z}; \mathbf{S}_{\text{m,6}};$
$\mathsf{exists}\ (\mathsf{book})\ (\lambda_x.\mathsf{PST}(\mathsf{buy}\ x\ \mathsf{ana})) : \mathbf{p}$

Recall that for passive participles we introduced another term of type **D**, namely ps. Below we give a lexical entry for the passive participle *pročitana* 'read'.

(131) $\vdash \mathtt{PRO\check{C}ITANA}_{\mathbf{z}} : \mathbf{z}; \prod_{p:\mathbf{Prs}}[\mathbf{NP}_{\text{nom,f,sg},p} \multimap \mathbf{Prd}_{\text{ps}}];$
$\lambda_x.\mathsf{exists}(\mathsf{person})(\lambda_y.\mathsf{read}\ x\ y) : \mathbf{e} \rightarrow \mathbf{p}$

This passive participle requires of the subject that it be a nominative feminine singular noun phrase, of whatever person. Starting with the same lexical entry for *jeste* 'is' given earlier, we obtain the following tectogrammatical version of it which can combine with passive participles.

(132) $\vdash \lambda_{vw}.\mathbf{PER}(w \circ \mathtt{JESTE}_{\mathbf{z}} \circ v) : \mathbf{z} \rightarrow \mathbf{z} \rightarrow \mathbf{Z};$
$\prod_{g:\mathbf{Gdr}}[(\mathbf{NP}_{\mathrm{nom},g,\mathrm{sg},3} \multimap \mathbf{Prd}_{\mathrm{ps}}) \multimap \mathbf{NP}_{\mathrm{nom},g,\mathrm{sg},3} \multimap \mathbf{S}_{\mathrm{m},6}];$
$\lambda_{Px}.(Px) : (\mathbf{e} \rightarrow \mathbf{p}) \rightarrow \mathbf{e} \rightarrow \mathbf{p}$

The only difference between this lexical entry for *jeste* and the one given earlier is in the tectogrammatical type. Namely, the **D** parameter is instantiated as ps instead of pl.

Now we can represent the sentence *Knjiga je pročitana* 'A book is read' as follows:

(133) $\mathbf{PER}(\mathtt{KNJIGA}_{\mathbf{z}} \circ \mathtt{JESTE}_{\mathbf{z}} \circ \mathtt{PRO\check{C}ITANA}_{\mathbf{z}}) : \mathbf{Z}; \mathbf{S}_{\mathrm{m},6};$
$\mathsf{exists}(\mathsf{book})(\lambda_x.\mathsf{exists}(\mathsf{person})(\lambda_y.\mathsf{read}\ x\ y)) : \mathbf{p}$

As another example, below we give a lexical entry for a neuter singular passive participle of a ditransitive verb, and the representation of the sentence *Pismo jeste poslano Ani* 'A letter is/has been sent to Ana'.

(134) $\vdash \lambda_v.\mathtt{POSLANO}_{\mathbf{z}} \circ v : \mathbf{z} \rightarrow \mathbf{z};$
$\prod_{g:\mathbf{Gdr},n:\mathbf{Num}\,p,p':\mathbf{Prs}}[\mathbf{NP}_{\mathrm{dat},g,n,p'} \multimap (\mathbf{NP}_{\mathrm{nom,n,sg},p} \multimap \mathbf{Prd}_{\mathrm{ps}})];$
$\lambda_{zx}.\mathsf{exists}(\mathsf{person})(\lambda_y.\mathsf{send}\ x\ z\ y) : \mathbf{e} \rightarrow \mathbf{e} \rightarrow \mathbf{p}$

(135) $\mathbf{PER}(\mathtt{PISMO}_{\mathbf{z}} \circ \mathtt{JESTE}_{\mathbf{z}} \circ \mathtt{POSLANO}_{\mathbf{z}} \circ \mathtt{ANI}_{\mathbf{z}}) : \mathbf{Z}; \mathbf{S}_{\mathrm{m},6};$
$\mathsf{exists}(\mathsf{letter})(\lambda_x.\mathsf{exists}(\mathsf{person})(\lambda_y.\mathsf{send}\ x\ \mathsf{ana}\ y)) : \mathbf{p}$

Predicative Noun Phrases and Adjectives

Recall that predicative noun phrases have to be nominative but in general don't have to agree with subjects in gender or number. Consider the sentence *Ana jeste student* 'Ana is a student'.

We give the following lexical entry for the predicative version of *student*.

(136) $\vdash \mathtt{STUDENT}_{\mathbf{z}}; \prod_{g:\mathbf{Gdr},n:\mathbf{Num},p:\mathbf{Prs}}[\mathbf{NP}_{\mathrm{nom},g,n,p} \multimap \mathbf{Prd}_{\mathrm{n}}]; \mathsf{student} : \mathbf{e} \rightarrow \mathbf{p}$

Of course, if we wanted to enforce gender or number agreement between a predicative noun phrase and the subject, we could appropriately instantiate the relevant parameters. Also note that we are assuming that the semantic type of a predicative noun phrase is $\mathbf{e} \rightarrow \mathbf{p}$. We will return to this later in the chapter.

Below is the tectogrammatical version of the copula looking for a predicative noun phrase complement.

(137) $\vdash \lambda_{vw}.\mathbf{PER}(w \circ \mathtt{JESTE}_{\mathbf{z}} \circ v) : \mathbf{z} \rightarrow \mathbf{z} \rightarrow \mathbf{Z};$
$\prod_{g:\mathbf{Gdr}}[(\mathbf{NP}_{\text{nom},g,\text{sg},3} \multimap \mathbf{Prd}_{\text{n}}) \multimap \mathbf{NP}_{\text{nom},g,\text{sg},3} \multimap \mathbf{S}_{\text{m},6}];$
$\lambda_{Px}.(Px) : (\mathbf{e} \rightarrow \mathbf{p}) \rightarrow \mathbf{e} \rightarrow \mathbf{p}$

The sentence *Ana jeste student* 'Ana is a student' is represented in the grammar by the following sign:

(138) $\vdash \mathbf{PER}(\mathtt{ANA}_{\mathbf{z}} \circ \mathtt{JESTE}_{\mathbf{z}} \circ \mathtt{STUDENT}_{\mathbf{z}}) : \mathbf{Z}; \mathbf{S}_{\text{m},6}; (\text{student ana}) : \mathbf{p}$

For predicative adjectives, we must enforce gender and number agreement between the adjective and the subject. Consider the sentence *Marko jeste pametan* 'Marko is smart'. We give the following lexical entry for the predicative adjective.

(139) $\vdash \mathtt{PAMETAN}_{\mathbf{z}}; \prod_{p:\mathbf{Prs}}[\mathbf{NP}_{\text{nom},\text{m},\text{sg},p} \multimap \mathbf{Prd}_{\text{a}}]; \text{smart} : \mathbf{e} \rightarrow \mathbf{p}$

The sentence *Marko jeste pametan* 'Marko is smart' is represented in the grammar by the following sign.

(140) $\vdash \mathbf{PER}(\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{JESTE}_{\mathbf{z}} \circ \mathtt{PAMETAN}_{\mathbf{z}}) : \mathbf{Z}; \mathbf{S}_{\text{m},6}; (\text{smart marko}) : \mathbf{p}$

Predicative Prepositional Phrases

A predicative prepositional phrase does not care about the subjects agreement features, except that it be nominative. Recall that prepositions in Serbo-Croatian are proclitics. Phenogrammatically, we analyze prepositions in predicative prepositional phrases similar to postnominal prepositional phrases. That is, the preposition permutes its complements noun phrase, then procliticizes onto the first word of its complement. The entire prepositional phrase is then turned into a length one string of languages, ensuring that it remains contiguous.

Consider the sentence *Marko jeste iz Beograda* 'Marko is from Belgrade'. We give the following lexical entry for the preposition *iz* 'from' which takes a 3rd person masculine singular genitive noun phrase complement and builds a predicative prepositional phrase.

(141) $$\begin{array}{l} \vdash \lambda_{v}.\mathbf{toZ}(\lambda_{s}.\exists_{t}[(\mathbf{k}\ (\mathbf{PER}\ v)\ t) \wedge s = (\mathtt{iz}\#(\mathbf{fst_s}\ t))_{\mathbf{s}} \cdot (\mathbf{rst_s}\ t)]) : \mathbf{z} \rightarrow \mathbf{z}; \\ \mathbf{NP}_{\mathrm{gen,m,sg,3}} \multimap \prod_{g:\mathbf{Gdr},n:\mathbf{Num},p:\mathbf{Prs}}[\mathbf{NP}_{\mathrm{nom},g,n,p} \multimap \mathbf{Prd}_{\mathrm{pp}}]; \\ \lambda_{xy}.(\mathsf{from}\ x\ y) : \mathbf{e} \rightarrow \mathbf{e} \rightarrow \mathbf{p} \end{array}$$

This lexical entry accounts for the less permissive set of judgments because it does not allow any discontinuities in the prepositional phrase. According to the more permissive set of judgments, the predicative prepositional phrase, as in the case of adverbial prepositional phrases, may be split into two chunks, one consisting of the preposition procliticized onto some word of its complement noun phrase, and the other of the remainder of the complement noun phrase, with the condition that the chunk containing the preposition precede the other chunk in the sentence.

To account for this set of judgments, we analyze a predicative prepositional phrase as an expression which takes as an argument a finite sentence missing a predicative prepositional phrase. This allows the prepositional phrase to split itself into two chunks and then shuffle into the sentence. The phenogrammatical term of this sign is more like the permissive lexical entry for an adverbial prepositional phrase, and not like the one for a postnominal modifier prepositional phrase.

(142) $\vdash \lambda_{yFw}.\exists_{xzt}[(F\ \mathtt{e}_{\mathbf{s}}\ x) \wedge (\mathbf{k}\ (\mathbf{PER}\ y)t) \wedge z = \mathbf{toZ}(\lambda_s.s = (\mathtt{iz\#}(\mathbf{fst_s}\ t))_{\mathbf{s}}$
$\circ\mathbf{toZ}(\mathbf{rst_s}\ t)) \wedge ((z \odot x)\ w)] : \mathbf{z} \to (\mathbf{z} \to \mathbf{Z}) \to \mathbf{Z};$
$\mathbf{NP}_{\text{gen,m,sg,3}} \multimap \prod_{g:\mathbf{Gdr},n:\mathbf{Num},p:\mathbf{Prs}}[((\mathbf{NP}_{\text{nom},g,n,p} \multimap \mathbf{Prd}_{\text{pp}}) \multimap \mathbf{S}_{\text{m,6}}) \multimap \mathbf{S}_{\text{m,6}}];$
$\lambda_{yP}.(P\ (\text{from}\ y)) : \mathbf{e} \to ((\mathbf{e} \to \mathbf{p}) \to \mathbf{p}) \to \mathbf{p}$

Predication and Nominal Modification

In this section, we explore connections between predicative phrases and noun modifiers, in order to streamline the grammar as much as possible.

As for past participles, they cannot be used to modify nouns at all, and in fact only appear in the periphrastic past tense and the conditional mood construction, the discussion of the latter having been deferred until the next chapter.[18]

All passive phrases, by which we mean a passive participle and any of its arguments excluding the subject, can be used as postnominal modifiers. However, recall that in predicative uses, passive participles must occur in nominative case. As postnominal modifiers, they must agree with the noun they are modifying in case (and number and gender, just like adjectives). For example:

[18]Past participles also occur without the copula in some kind of not very productive optative like construction, for example *Živjeli*$_{ppl}$ 'May we live on!' (used as 'Cheers!'), or, in an old Chernobyl joke *Tražila*$_{ppl}$ *te majka gajgerovim brojačem!* 'May your mother look for you with a Geiger counter!'. In this use, the past participles also occur in many profane expressions.

(143) a. Ana voli tursku kafu skuhanu sa puno šećera.
$\text{Ana}_{NOM,f,sg,3}$ $\text{love}_{sg,3}$ $\text{Turkish}_{ACC,f,sg}$ $\text{coffee}_{ACC,f,sg}$ $\text{cooked}_{pass,ACC,f,sg}$ with lots $\text{sugar}_{GEN,m,sg}$
'Ana likes Turkish coffee cooked with lots of sugar'

b. Marko ne voli hranu pripremljenu sa puno začina.
$\text{Marko}_{NOM,m,sg,3}$ $\text{not-love}_{sg,3}$ $\text{food}_{ACC,f,sg}$ $\text{prepared}_{pass,ACC,f,sg}$ with lots $\text{spices}_{GEN,m,pl}$
'Marko doesn't like food prepared with lots of spices'

While we can write a non-logical rule that maps predicative passive phrases to postnominal modifiers, since all predicative passive participles are nominative, we would only generate a small subset of passive postnominal modifiers, namely, only those that modify nominative nouns. For the (many) other cases, we would have to directly add lexical entries for such passive phrases.

For example, here is a non-logical rule, call it [psNP] that takes a predicative passive phrase and outputs a postnominal modifier of nominative nouns:

(144)

$$\frac{\vdash \phi : \mathbf{z}; \mathbf{NP}_{\text{nom},\tau,\tau',3} \multimap \mathbf{Prd}_{\text{ps}}; \sigma : \mathsf{e} \to \mathsf{p}}{\begin{array}{c}\vdash \lambda_v.\mathtt{toZ}(\mathtt{k}\ (\mathtt{PER}\ v) \bullet (\mathtt{k}\ (\mathtt{PER}\ \phi))) : \mathbf{z} \to \mathbf{z}; \mathbf{NP}_{\text{nom},\tau,\tau',3} \multimap \mathbf{NP}_{\text{nom},\tau,\tau',3}; \\ \lambda_{Py}.(P\ y)\mathsf{and}(\sigma\ y) : (\mathsf{e} \to \mathsf{p}) \to \mathsf{e} \to \mathsf{p}\end{array}}\ [\text{psNP}]$$

Recall that postnominal modifiers combine with nouns whose tectogrammatical types have been converted to the corresponding noun phrase types. Since all such phrases are 3rd person, the passive phrase must have its person parameter instantiated as 3, before it can undergo this rule. Here is an example of an output of this rule:

(145) $\vdash \lambda_v.\mathtt{toZ}(\mathbf{k}(\mathbf{PER}\ v) \bullet \mathbf{k}(\mathbf{PER}(\mathrm{POSLANO}_{\mathbf{z}} \circ \mathrm{ANI}_{\mathbf{z}}))) : \mathbf{z} \to \mathbf{z};$
$\mathbf{NP}_{\mathrm{nom,n,sg,3}} \multimap \mathbf{NP}_{\mathrm{nom,n,sg,3}};$
$\lambda_{Pz}.\mathsf{exists}(\mathsf{person})(\lambda_y.\mathsf{send}\ z\ \mathsf{ana}\ y)\ \mathsf{and}\ (P\ z) : (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$

If we want an appropriately case marked version of this passive phrase that can modify, say, instrumental nouns, we must assert a new lexical entry for the passive participle itself, like so:

(146) $\vdash \lambda_{wv}.\mathtt{toZ}(\mathbf{k}(\mathbf{PER}\ v) \bullet \mathbf{k}(\mathbf{PER}(\mathrm{POSLANIM}_{\mathbf{z}} \circ w))) : \mathbf{z} \to \mathbf{z} \to \mathbf{z};$
$\prod_{g:\mathbf{Gdr},n:\mathbf{Num}p:\mathbf{Prs}}[\mathbf{NP}_{\mathrm{dat},g,n,p} \multimap \mathbf{NP}_{\mathrm{inst,n,sg,3}} \multimap \mathbf{NP}_{\mathrm{inst,n,sg,3}}];$
$\lambda_{xPz}.\mathsf{exists}(\mathsf{person})(\lambda_y.\mathsf{send}\ z\ x\ y)\ \mathsf{and}\ (P\ z) : \mathbf{e} \to (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$

Further, all passive participles (just participles, not passive phrases) can be used as attributive adjectives. However, again the issue is that in attributive uses passive participles agree in nouns they modify in case (and number and gender). We will just assume that we have to directly assert lexical entries for passive participles which behave like attributive adjectives.

As for predicative adjectives, not all attributive adjectives can be used predicatively, for example *navodni* 'alleged'. But I can't think of and haven't found any examples of predicative adjectives which cannot be used attributively. So assuming we have lexical entries for predicative adjectives, we can give the following non-logical rule, call it [aN], which maps predicative adjectives to their attributive counterparts.

(147)

$$\frac{\vdash \phi : \mathbf{z}; \mathbf{NP}_{\mathrm{nom},\tau,\tau',3} \multimap \mathbf{Prd}_{\mathrm{a}}; \sigma : \mathbf{e} \to \mathbf{p}}{\vdash \lambda_v.\phi \circ v : \mathbf{z} \to \mathbf{z}; \mathbf{N}_{\mathrm{nom},\tau,\tau'} \multimap \mathbf{N}_{\mathrm{nom},\tau,\tau'};} \ [\mathrm{aN}]$$
$$\lambda_{Py}.(P\ y)\mathsf{and}(\sigma\ y) : (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$$

This rule converts a predicative adjective into an attributive adjective which behaves permissively, i.e. it's detachable from the remainder of the noun. We leave it to the reader to formulate an appropriate version of the rule that will output attributive adjectives which behave in accordance with more restrictive judgments concerning the ordering of noun phrase material discussed in the previous chapter.

As in the case of passive phrases, the conversion via this rule only works for nominative adjectives, since all predicative adjectives are nominative. We have to independently introduce lexical entries for differently case marked versions of such adjectives. We also have to introduce lexical entries of attributive adjectives, nominative and otherwise, which do not have predicative counterparts.

We analyze predicative noun phrases as derived from signs which are targets of postnominal modification, that is, possibly phrasal nouns which are semantically of type $\mathbf{e} \rightarrow \mathbf{p}$ but tectogrammatically have a noun phrase type. We independently motivated this tecto/semantic type 'mismatch' in the previous chapter to account for certain word order peculiarities concerning the word order within phrasal noun phrases which contain both attributive and postnominal modifiers. Now, we exploit the fact that we already have such signs in the grammar to generate the set of predicative noun phrases.

The following rule, call it [NPn], maps signs which are tectogrammatically nominative case marked noun phrases but semantically of type $\mathbf{e} \rightarrow \mathbf{p}$ into predicative phrases.

(148)

$$\frac{\vdash \phi : \mathbf{z}; \mathbf{NP}_{\mathrm{nom},\tau,\tau',3}; \sigma : \mathbf{e} \rightarrow \mathbf{p}}{\vdash \phi : \mathbf{z}; \prod_{g:\mathbf{Gdr},n:\mathbf{Num},p:\mathbf{Prs}}[\mathbf{NP}_{\mathrm{nom},g,n,p} \multimap \mathbf{Prd}_{\mathrm{n}}]; \sigma : \mathbf{e} \rightarrow \mathbf{p}} \text{[NPn]}$$

This rule has an additional advantage in that it allows modification of nouns to proceed as usual. Once any attributive adjectives and postnominal modifiers have combined with a nominative case marked noun, it can undergo this rule and become a predicative phrase. This version of the rule produces predicative phrases which do not agree with the subject in number or gender, in accordance with the empirical generalization presented earlier in this chapter. We leave it to the reader to formulate less permissive versions of this rule which would impose more stringent agreement requirements on the predicative phrase which is its output.

Finally, all predicative prepositional phrases can be used as postnominal modifiers. The non-logical rule below, call it [ppN], maps predicative prepositional phrases into postnominal modifiers. Recall that we entertained two very different lexical entries for predicative prepositional phrases, one which allows discontinuities in the prepositional phrase and the other one which doesn't. However, all the grammars converge on not allowing discontinuities in the noun+postnominal modifier sequence. We will accordingly give two formulations of the rule though both rules have to output the same kind of thing, a postnominal modifier which must remain contiguous and occur immediately to the right of the noun it modifies. We start with the version of the rule that would be added to the less permissive grammar where discontinuities in predicative prepositional phrases are not allowed.

(149) version 1 - less permissive grammar

$$\frac{\begin{array}{c}\vdash \phi : \mathtt{z}; \prod_{g:\mathbf{Gdr},n:\mathbf{Num},p:\mathbf{Prs}}[\mathbf{NP}_{\text{nom},g,n,p} \multimap \mathbf{Prd}_{\text{pp}}];\\ \sigma : \mathsf{e} \to \mathsf{p}\end{array}}{\begin{array}{c}\vdash \lambda_v.\mathtt{toZ}((\mathtt{L}\ v) \bullet (\mathtt{L}\ \phi)) : \mathtt{z} \to \mathtt{z}; \prod_{c:\mathbf{Cse},g:\mathbf{Gdr},n:\mathbf{Num}}[\mathbf{NP}_{c,g,n,3} \multimap \mathbf{NP}_{c,g,n,3}];\\ \lambda_{Px}.(P\ x)\mathsf{and}(\sigma\ x) : (\mathsf{e} \to \mathsf{p}) \to \mathsf{e} \to \mathsf{p}\end{array}}\ [\text{ppN}]$$

The output of this rule is a postnominal modifier which can combine with nouns (tectogrammatically associated with a noun phrase type) of any case, gender and number. The phenogrammatical term ensures that the prepositional phrase occur immediately to the right of the noun it modifies. The prepositional phrase and the noun form a length one string of languages which ensures that they remain continuous and in exactly that order.

Recall that in order for a predicative prepositional phrase to occur discontinuously in a sentence, we had to analyze it as a functor over sentences missing a predicative prepositional phrase. Now we give a rule that can be added to the more permissive grammar in which predicative prepositional phrases are allowed to occur discontinuously.

(150) version 2 - more permissive grammar

$$\frac{\vdash \phi : (\mathtt{z} \to \mathtt{Z}) \to \mathtt{Z}; \quad \prod_{g:\mathbf{Gdr},n:\mathbf{Num},p:\mathbf{Prs}}[((\mathbf{NP}_{\mathrm{nom},g,n,p} \multimap \mathbf{Prd}_{\mathrm{pp}}) \multimap \mathbf{S}_{\mathrm{m},6}) \multimap \mathbf{S}_{\mathrm{m},6}]; \sigma : ((\mathbf{e} \to \mathbf{p}) \to \mathbf{p}) \to \mathbf{p}}{\vdash \lambda_v.\mathtt{toZ}((\mathtt{L}\ v) \bullet (\mathtt{k}\ (\phi(\lambda_{xy}.y = \mathtt{e}_{\mathtt{S}})))) : \mathtt{z} \to \mathtt{z}; \quad \prod_{c:\mathbf{Cse},g:\mathbf{Gdr},n:\mathbf{Num}}[\mathbf{NP}_{c,g,n,3} \multimap \mathbf{NP}_{c,g,n,3}]; \lambda_{Qx}.\sigma(\lambda_P.P\,x) \text{ and } (Q\,x) : (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}} \text{[ppN]}$$

This complicated rule essentially outputs the same kind of signs as its counterpart in the less permissive grammar, by outputting prepositional phrases which are postnominal modifiers, and requiring them to occur contiguously and immediately to the right of the noun they modify.

4.4 Subject and Object Control

4.4.1 Data

Here, we consider verbs which are object or subject control verbs. By *subject control* verb we mean a verb which has a finite verb phrase complement (and possibly a noun phrase object as well), and whose subject, which occurs in the matrix clause, is also interpreted as the embedded subject. By *object control* verb we mean a verb which has a finite verb phrase complement and a noun phrase object, and whose object is interpreted as the embedded subject. We will perhaps abuse terminology and talk about the matrix subject (object) controlling not the non-existent embedded subject but controlling the finite verb phrase. We will also consider verbs whose objects control a noun phrase or an adjective.

	sg (full/clitic)	pl (full/clitic)
1	hoću/ću	hoćemo/ćemo
2	hoćeš/ćeš	hoćete/ćete
3	hoće/će	hoće/će

Table 4.2: The verb *htjeti* 'want, will' paradigm.

As we will see below, it is more instructive to think of these controlled embedded verb phrases as embedded sentences with a subject (nominative) gap, since they must occur with the complementizer, just like full embedded clauses.

Subject Control

One important subject control verb in Serbo-Croatian is *htjeti* 'want, will', which can take a variety of complements and also participates in the subject controlled future tense construction. There are enclitic and non-clitic present tense forms of *htjeti*. We show the full paradigm below, but we will return to its clitics forms in the next chapter and here only consider the full forms.

With the meaning 'want', *htjeti* can take a noun phrase complement (a), a full embedded sentence (b) or an embedded sentence with a subject gap (c), the latter being an instance of subject control.

(151) a. Ana hoće pivo.
Ana$_{NOM,f,sg,3}$ want$_{sg,3}$ beer$_{ACC,n,sg,3}$
'Ana wants a beer'

b. Ana hoće da Marko ode.
Ana$_{NOM,f,sg,3}$ want$_{sg,3}$ DA Marko$_{NOM,m,sg,3}$ leave$_{sg,3}$

'Ana wants Marko to leave'

c. Ana hoće da ode.
Ana$_{NOM,f,sg,3}$ want$_{sg,3}$ DA leave$_{sg,3}$
'Ana wants to leave'

Future tense has the same structure as the subject control sentence (c) above. While typically, only the clitic forms participate in future tense formation, in pragmatically marked contexts the full forms can have the future interpretation as well. The result is that the subject control versions of *htjeti* are ambiguous between the future meaning and the ordinary 'want' meaning. In the right context, the (c) sentence above can mean 'Ana will leave' and not 'Ana wants to leave'. Below we show an example where the full form of *htjeti* under contrastive focus expresses future tense.

(152) a. A: Nikad nećeš da dobiješ taj posao!
never not-will$_{sg,2}$ DA get$_{sg,2}$ that$_{ACC,m,sg}$ job$_{ACC,m,sg}$
'You'll never get that job'

b. B: Sigurno hoću!
definitely will$_{sg,1}$
'I definitely will'

(153) a. Možda ti nećeš, ali ja sigurno hoću da dobijem taj posao.
maybe you$_{NOM,sg,2}$ not-will$_{sg,2}$, but I$_{NOM,sg,1}$ definitely will$_{sg,1}$ DA get$_{sg,1}$ that$_{ACC,m,sg}$ job$_{ACC,m,sg}$
'Maybe you won't, but I definitely will get that job'

Modal verbs such as *smjeti* 'may', *trebati* 'need, should', *morati* 'must' and *moći* 'be able to' are also subject control verbs, as well as *znati* 'know' with the meaning 'know how/be able to do something'. Below are some examples.

(154) a. Ana mora da ode.
Ana$_{NOM,f,sg,3}$ must$_{sg,3}$ DA leave$_{sg,3}$
'Ana must leave'

b. Mi smijemo da odemo.
we$_{NOM,pl,1}$ may$_{pl,1}$ DA leave$_{pl,1}$
'We may leave'

c. Oni trebaju da odu.
they$_{NOM,pl,3}$ should$_{pl,3}$ DA leave$_{pl,3}$
'They should/need to leave'

d. Ona zna da vozi.
she$_{NOM,f,sg,3}$ know$_{sg,3}$ DA drive$_{sg,3}$
'She can/knows how to drive'

Some other subject control verbs in Serbo-Croatian are *pokušati* 'try', *obećati* 'promise' and *voljeti* 'love', the latter in the sense of 'like to do something'. Below are some examples.

(155) a. Marko je obećao da će da vozi.
Marko$_{NOM,m,sg,3}$ is$_{sg,3}$ promise$_{ppl,m,sg}$ DA will$_{sg,3}$ DA drive$_{sg,3}$
'Marko promised that he will/would drive'

b. Marko voli da vozi.
Marko$_{NOM,m,sg,3}$ love$_{sg,3}$ DA drive$_{sg,3}$
'Marko likes to drive'

Note that in subject control constructions the matrix subject has to agree in number and person with both the matrix and the embedded verb, as shown below:

(156) a. Marko mora da ode.
Marko$_{NOM,m,sg,3}$ must$_{sg,3}$ DA leave$_{sg,3}$
'Marko must leave'

b. * Marko mora da odemo$_{pl,1}$

c. * Marko mora da odete$_{pl,2}$

d. * Marko mora da odem$_{sg,1}$

Further, the gender information of the subject controller must be accessible as well. While gender marked verbal forms cannot occur in controlled verb phrases (i.e. past tense or conditional forms), the controlled verb phrase may be predicative and then the gender of the subject controller does matter. Consider the following examples:[19]

(157) a. Marko mora da bude oprezan.
Marko$_{NOM,m,sg,3}$ must$_{sg,3}$ DA is$_{sg,3}$ cautious$_{NOM,m,sg}$
'Marko must be cautious'

b. * Marko mora da bude opreznа$_{f,sg}$

c. Ana mora da bude oprezna.
Ana$_{NOM,f,sg,3}$ must$_{sg,3}$ DA is$_{sg,3}$ cautious$_{NOM,f,sg}$
'Ana must be cautious'

d. * Ana mora da bude oprezan$_{m,sg}$

These examples show that gender agreement between the matrix subject and a predicative adjective in a controlled verb phrase must be maintained. Therefore, not only the number and person but also the gender of the subject controller matters.

In addition to finite embedded clauses with a subject gap, all subject control verbs can also take an infinitival verb phrase complement. In this case, there is no person or number agreement between the infinitive and the matrix subject, as the infinitive does not carry any agreement information. The examples below show

[19]The form *bude* glossed as 'is' is a perfective present form of *biti* 'be', whereas *je* which we've also been glossing as 'is' is an imperfective present form of the same verb. There are restrictions on the tense and aspect of the verb in certain embedded environments, including controlled verb phrases, the details of which are unfortunately beyond the scope of this thesis.

some combinations of various subject controllers with controlled infinitival verb phrases.

(158) a. Mi moramo otići.
we$_{NOM,pl,1}$ must$_{pl,1}$ leave$_{inf}$
'We must leave'

b. Marko hoće dobiti posao.
Marko$_{NOM,m,sg,3}$ will$_{sg,3}$ get$_{inf}$ job$_{ACC,m,sg,3}$
'Marko wants to/will get a job'

c. Ja znam voziti bicikl.
I$_{NOM,sg,1}$ know$_{sg,1}$ drive$_{inf}$ bicycle$_{ACC,m,sg,3}$
'I can/know how to ride a bicycle'

However, if the controlled infinitival verb phrase contains a predicative adjective, gender of the subject controller does matter:

(159) a. Marko mora biti oprezan.
Marko$_{NOM,m,sg,3}$ must$_{sg,3}$ be$_{inf}$ cautious$_{NOM,m,sg}$
'Marko must be cautious'

b. * Marko mora biti oprezna$_{NOM,f,sg}$

c. Ana mora biti oprezna.
Ana$_{NOM,f,sg,3}$ must$_{sg,3}$ be$_{inf,3}$ cautious$_{NOM,f,sg}$
'Ana must be cautious'

d. * Ana mora da bude oprezan$_{NOM,m,sg}$

Further, the controlled infinitive may itself be a subject control verb, and if it embeds a finite controlled verb phrase, the person and number of the matrix subject controlled again matter:

(160) a. Ana mora pokušati da dođe.
Ana$_{NOM,sg,f,3}$ must$_{sg,3}$ try$_{inf}$ DA come$_{sg,3}$
'Ana must try to come'

b. * Ana mora pokušati da dođete$_{pl,2}$

In sum, regardless of whether the controlled verb phrase is finite of infinitival, the grammar must keep track of number, gender and person requirements of that verb phrase so the agreement with the matrix subject controller can be established.

As for word order in subject control sentences, the generalizations are largely the same as with embedded clauses which we discussed earlier in the chapter. The complementizer must occur leftmost in the embedded clause, and the rest of the embedded clause material may reorder freely, depending on the embedded verb. In the matrix clause, the embedded clause must occur rightmost, with the exception of adverbials, while the rest of the matrix clause material may freely reorder.

Inside infinitival verb phrase complements, constituents can undoubtedly freely reorder, for example:

(161) a. Oni moraju kupiti Ani poklon.
they$_{NOM,pl,3}$ must$_{pl,3}$ buy$_{inf}$ Ana$_{DAT,f,sg,3}$ present$_{ACC,m,sg,3}$
'They have to buy Ana a present'
b. Oni moraju Ani kupiti poklon.
c. Oni moraju poklon kupiti Ani.
d. Oni moraju Ani poklon kupiti.
e. etc.

In fact, in our judgment, the infinitive and its complements may freely reorder with respect to the matrix clause material, in contradistinction to embedded finite clauses:

(162) a. Oni moraju kupiti Ani poklon.
they$_{NOM,pl,3}$ must$_{pl,3}$ buy$_{inf}$ Ana$_{DAT,f,sg,3}$ present$_{ACC,m,sg,3}$

'They have to buy Ana a present'

b. Oni Ani moraju kupiti poklon.

c. Poklon oni moraju Ani kupiti.

d. Kupiti poklon oni Ani moraju.

e. etc.

In the analysis section we will show both how to extend the grammar to allow the infinitive and its complements to intermingle with the other matrix clause material, and how to extend the grammar so as to keep the infinitival verb phrases contiguous while allowing free reordering inside of them.

Object Control

In contrast to subject control, an infinitival verb phrase is never a possible complement of an object control verb. Instead, such verbs have to combine with a finite embedded clause with a subject gap.

Object control verbs in Serbo-Croatian include *zamoliti* 'ask, request', *natjerati* 'force, make', *nagovoriti* 'persuade' and *ponuditi* 'offer'. Below we show some examples of object control structures:

(163) a. Mi smo zamolili Anu da donese pivo.
we$_{NOM,pl,1}$ are$_{pl,1}$ ask$_{ppl,m,pl}$ Ana$_{ACC,f,sg,3}$ DA bring$_{sg,3}$ beer$_{ACC,n,sg,3}$
'We asked Ana to bring beer'

b. Marko je ponudio Ani da dođe u Ameriku.
Marko$_{NOM,m,sg,3}$ is$_{sg,3}$ offer$_{ppl,m,sg}$ Ana$_{DAT,f,sg,3}$ DA come$_{sg,3}$ in America$_{ACC,f,sg,3}$
'Marko offered to Ana to come to America'

Note that the noun phrase object of an object control verb could be in accusative or in dative, depending on the verb. Further, note that the embedded verb has to agree in person and number with the matrix object, which is one of the reasons we are keeping track of number on non-nominative noun phrases.

In addition to person and number, the gender of the object controller matters as well. First, as in the case of subject control, the controlled verb phrase may be predicative:

(164) a. Marko je zamolio Anu da bude pristojna.
Marko$_{NOM,m,sg,3}$ is$_{sg,3}$ ask$_{m,sg}$ Ana$_{ACC,f,sg,3}$ DA is$_{sg,3}$ polite$_{NOM,sg,f}$
'Marko asked Ana to be polite'

b. * Marko je zamolio Anu da bude pristojan$_{NOM,m,sg}$

Second, objects of some verbs control predicative complements and not verb phrases, as shown in the examples below.

(165) a. Marko je smatrao Anu pristojnom.
Marko$_{NOM,m,sg,3}$ is$_{sg,3}$ consider$_{ppl,m,sg}$ Ana$_{ACC,f,sg,s}$ polite$_{INST,f,sg}$
'Marko considered Ana polite'

b. * Marko je smatrao Anu pristojnim$_{INST,m,sg}$

c. Ana je nazvala Marka kretenom.
Ana$_{NOM,f,sg,3}$ is$_{sg,3}$ call$_{ppl,f,sg}$ Marko$_{ACC,m,sg}$ idiot$_{INST,m,sg}$
'Ana called Marko an idiot'

Therefore, as in the case of subject control, the object controlled constituent has to agree with the controller in number, person and gender.

As for word order in object control structures, the generalizations are the same as for subject control, whereby the complementizer must come first in the embedded clause and the embedded clause must remain contiguous and occur rightmost in the matrix clause. Other matrix clause material, including the object controller, may freely reorder with respect to the matrix verb and subject. The generalizations concerning the placement of adverbs are the same as well.

4.4.2 Analysis

Subject Control

We first analyze subject controlled finite verb phrases, then the infinitival ones. We analyze controlled finite verb phrases as embedded declarative clauses with a bound subject trace. For example, a controlled verb phrase *da ode* in the sentence *Ana hoće da ode* 'Ana wants to leave' is built up as follows. First, an appropriate subject trace is introduced and a declarative sentence is constructed:

(166)

$$\dfrac{\vdash \lambda_v.\mathbf{PER}(v \circ \mathtt{ODE}_{\mathbf{z}}); \mathbf{NP}_{\text{nom,f,sg,3}} \multimap \mathbf{S}_{\text{m,6}}; \text{leave} \qquad \dfrac{}{x; \mathbf{NP}_{\text{nom,f,sg,3}}; x \vdash x; \mathbf{NP}_{\text{nom,f,sg,3}}; x}[\text{Ax}]}{x; \mathbf{NP}_{\text{nom,f,sg,3}}; x \vdash \mathbf{PER}(x \circ \mathtt{ODE}_{\mathbf{z}}); \mathbf{S}_{\text{m,6}}; (\text{leave } x)}[\multimap\text{E}]$$

Second, the complementizer *da* combines with the sentence, turning it into an embedded sentence, and then the subject trace is bound, i.e. the hypothesis is withdrawn:

(167)

$$\dfrac{\dfrac{\lambda_{Xw}.\exists_v[(X\ v) \wedge w = (\mathtt{DA}_{\mathbf{z}} \circ v)]; \mathbf{S}_{\text{m,6}} \multimap \mathbf{S}_{\text{e,6}}; \lambda_q.q \qquad x; \mathbf{NP}_{\text{nom,f,sg,3}}; x \vdash \mathbf{PER}(x \circ \mathtt{ODE}_{\mathbf{z}}); \mathbf{S}_{\text{m,6}}; (\text{leave } x)}{x; \mathbf{NP}_{\text{nom,f,sg,3}}; x \vdash \lambda_w.\exists_v[(\mathbf{PER}(x \circ \mathtt{ODE}_{\mathbf{z}})\ v) \wedge w = (\mathtt{DA}_{\mathbf{z}} \circ v)]; \mathbf{S}_{\text{e,6}}; (\text{leave } x)}}{\lambda_{xw}.\exists_v[(\mathbf{PER}(x \circ \mathtt{ODE}_{\mathbf{z}})\ v) \wedge w = (\mathtt{DA}_{\mathbf{z}} \circ v)]; \mathbf{NP}_{\text{nom,f,sg,3}} \multimap \mathbf{S}_{\text{e,6}}; \lambda_x.(\text{leave } x)}[\multimap\text{I}]$$

The conclusion of the proof above is the kind of sign that can be a complement of a subject control verb–an embedded sentence with a bound subject trace. Note that the relevant agreement features of the missing embedded subject are recorded in the tectogrammatical type. This will ensure that the matrix subject and the embedded predicate agree in number, person, and gender. We give the following lexical entry for the control verb *hoće* 'wants'.

(168) $\lambda_{Fvw}.\exists_{xy}[(F \Subset_{\mathtt{S}} x) \wedge (\mathbf{PER}(v \circ \mathtt{HO\acute{C}E}_{\mathtt{z}})\ y) \wedge w = y \circ \mathtt{toZ}(\mathtt{L}\ x)] : (\mathtt{z} \to \mathtt{Z}) \to$
$\mathtt{z} \to \mathtt{Z}; \prod_{g:\mathbf{Gdr}}[(\mathbf{NP}_{\text{nom},g,\text{sg},3} \multimap \mathbf{S}_{\text{e},6}) \multimap \mathbf{NP}_{\text{nom},g,\text{sg},3} \multimap \mathbf{S}_{\text{m},6}];$
$\lambda_{Fx}.(\text{want}\ (F\ x)\ x) : (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$

This finite verb requires a third person singular subject, so it is only compatible with verb phrase complements which also require such subjects. It is not, however, inflected for gender, which is why it has a dependent product type. Whatever the gender feature of its complement verb phrase, *hoće* will require a subject of that gender. To construct the sentence *Ana hoće da ode* 'Ana wants to leave' we need a feminine version of the verb. Combining that verb with the embedded clause with a subject gap, and then with the matrix subject *Ana* we get the following sign:

(169) $\lambda_{w}.\exists_{xyv}(\mathbf{PER}(\mathtt{ODE}_{\mathtt{z}})\ v) \wedge x = (\mathtt{DA}_{\mathtt{z}} \circ v) \wedge (\mathbf{PER}(\mathtt{ANA}_{\mathtt{z}} \circ \mathtt{HO\acute{C}E}_{\mathtt{z}})\ y) \wedge$
$w = y \circ \mathtt{toZ}(\mathtt{L}\ x)] : \mathtt{Z}; \mathbf{S}_{\text{m},6}; (\text{want (leave ana) ana}) : \mathbf{p}$

Semantically, we predict that this sentence expresses the proposition that Ana wants that she (Ana) leave. Syntactically, it is a main declarative clause. Phenogrammatically, it denotes a set of strings of languages each of which consists of some permutation of the matrix verb and the subject, followed by a length one string of languages constructed out of the embedded sentence with a subject gap, just as in the case of ordinary embedded sentences. So we predict that this sentence can be pronounced exactly two ways, *Ana hoće da ode* and *Hoće Ana da ode*.

Recall that infinitives are not inflected for gender, number and person. However, the infinitive could itself be a control verb, with a finite verb phrase complement, or a predicative verb, in which case the gender, number and person features matter to establish agreement with the matrix subject controller.

A subject control verb taking an infinitival complement cannot know in advance whether its complement is itself a control or a predicative verb (agreement matters) or is not a control or predicative verb (agreement does not matter). For example, in *Ana mora otići* 'Ana must leave' the agreement features do not matter. But in *Ana mora biti pristojna* 'Ana must be polite', the adjective *pristojna* and the matrix subject must agree in number and gender, while in *Ana mora pokušati da dođe* 'Ana must try to come', the embedded finite verb *dođe* and the matrix subject must agree in number and person. So a subject control verb like *mora* 'must', has to be prepared to deal with all these different complements and, if needed, make sure that the matrix subject is an appropriate controller.

For this reason, the agreement information must be recorded on the infinitival verb phrases in general, regardless of whether it's a control verb or not. In other words, a control verb taking an infinitival complement has to err on the side of too much information, just in case its infinitival complement is a control or predicative verb.

Suppose we're trying to construct a representation of the sentence *Ana hoće voziti bicikl* 'Ana wants to ride a bike' where the infinitive is not a control verb. We give the following lexical entry for the infinitive:

(170) $\vdash \lambda_{v}.\mathtt{VOZITI}_{\mathbf{z}} \circ v : \mathbf{z} \to \mathbf{z}; \prod_{g:\mathbf{Gdr},n:\mathbf{Num},p:\mathbf{Prs}}[\mathbf{NP}_{\mathrm{acc,m,sg,3}} \multimap \mathbf{NP}_{\mathrm{nom},g,n,p} \multimap \mathbf{S}_{\mathrm{inf,6}}];$
$\lambda_{xy}.(\mathsf{ride}\ x\ y) : \mathbf{e} \to \mathbf{e} \to \mathbf{p}$

Here, inf is a term of type **K**. It is introduced especially to distinguish infinitival verb phrases from finite ones. Note that there is a 'mismatch' between the phenogrammatical typing on the one hand, and the semantic and tectogrammatical typing on the other. While semantically, *voziti* needs two individual arguments, and tectogrammatically two noun phrase arguments (the object and the subject), phenogrammatically, it only needs one argument of type **z**. That is, phenogrammatically, *voziti* is expecting only one argument (the object noun phrase). This is to ensure that no non-sentences consisting of a nominative noun phrase and an infinitival verb phrase are ever generated by the grammar. The tectogrammatical and semantic typing ensures that the appropriate agreement information is available and that the correct interpretation is generated. The phenogrammatical typing ensures that there are no signs of type $\mathbf{S}_{\mathrm{inf,6}}$ in the grammar, as they do not correspond to any actual sentences of Serbo-Croatian.

Whether the constituents of the infinitival verb phrase are allowed to reorder with respect to the matrix clause constituents, or merely to reorder among themselves but remain contiguous in the matrix clause does not depend on the infinitive's but the subject control verb's lexical entry. Below we give two versions of *hoće* 'wants', one corresponding to the more permissive set of judgments, the other to the more conservative.

(171) $\vdash \lambda_{xy}.\mathbf{PER}(y \circ \mathtt{HO\acute{C}E}_{\mathbf{z}} \circ x) : \mathbf{z} \to \mathbf{z} \to \mathbf{Z}$;
$\prod_{g:\mathbf{Gdr}}[(\mathbf{NP}_{\mathrm{nom},g,\mathrm{sg},3} \multimap \mathbf{S}_{\mathrm{inf},6}) \multimap \mathbf{NP}_{\mathrm{nom},g,\mathrm{sg},3} \multimap \mathbf{S}_{\mathrm{m},6}]$;
$\lambda_{Fx}.(\mathsf{want}\ (F\,x)\ x) : (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$

(172) $\vdash \lambda_{xyz}.\exists_{vw}[(\mathbf{PER}(y \circ \mathtt{HO\acute{C}E}_{\mathbf{z}})\, v) \wedge (\mathbf{PER}\ x\ w) \wedge z = v \circ w] : \mathbf{z} \to \mathbf{z} \to \mathbf{Z}$;
$\prod_{g:\mathbf{Gdr}}[(\mathbf{NP}_{\mathrm{nom},g,\mathrm{sg},3} \multimap \mathbf{S}_{\mathrm{inf},6}) \multimap \mathbf{NP}_{\mathrm{nom},g,\mathrm{sg},3} \multimap \mathbf{S}_{\mathrm{m},6}]$;
$\lambda_{Fx}.(\mathsf{want}\ (F\,x)\ x) : (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$

These two lexical entries differ only in terms of their phenogrammatical term. The first lexical entry allows free reordering of all constituents in the infinitival verb phrase with matrix constituents, so that the sentence *Ana hoće voziti bicikl* 'Ana wants to ride a bike' is predicted to be pronounceable 24 different ways.

The second lexical entry allows permutation of the constituents in the infinitival verb phrase, and the permutation of the control verb and its subject, but demands that the subject and the verb occur to the left of the infinitival verb phrase, and that the infinitival verb phrase remain contiguous, so that the sentence *Ana hoće voziti bicikl* 'Ana wants to ride a bike' is predicted to be pronounceable 4 different ways.

Below are the signs representing the sentence *Ana hoće voziti bicikl* 'Ana wants to ride a bike' in the permissive and the more conservative grammar:

(173) $\vdash \lambda_{xy}.\mathbf{PER}(\texttt{ANA}_{\mathbf{z}} \circ \texttt{HOĆE}_{\mathbf{z}} \circ \texttt{VOZITI}_{\mathbf{z}} \circ \texttt{BICIKL}_{\mathbf{z}}) : \mathbf{Z}; \mathbf{S}_{\mathrm{m},6};$
$\mathsf{exist(bicycle)}\lambda_{y}.\mathsf{(want\ (ride}\ y\ \mathsf{ana)\ ana)} : (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$

(174) $\vdash \lambda_{z}.\exists_{vw}[(\mathbf{PER}(\texttt{ANA}_{\mathbf{z}} \circ \texttt{HOĆE}_{\mathbf{z}})v) \wedge (\mathbf{PER}\ (\texttt{VOZITI}_{\mathbf{z}} \circ \texttt{BICIKL}_{\mathbf{z}})\ w) \wedge z =$
$v \circ w] : \mathbf{Z}; \mathbf{S}_{\mathrm{m},6}; \mathsf{exist(bicycle)}\lambda_{y}.\mathsf{(want\ (ride}\ y\ \mathsf{ana)\ ana} : (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$

We can also imagine an intermediate option, where the infinitival verb phrase would the required to remain contiguous, and its constituents could reorder with respect to one another, but the matrix verb and subject could freely reorder with respect to the infinitival complement. In other words, this is just like the conservative option above, except that the infinitival verb phrase is not required to occur on the right edge of the matrix clause. The lexical entry below models this set of judgments:

(175) $\vdash \lambda_{xy}.\mathbf{PER}(y \circ \texttt{HOĆE}_{\mathbf{z}} \circ \mathbf{toZ}(\mathbf{k}\ (\mathbf{PER}\ x))) : \mathbf{z} \to \mathbf{z} \to \mathbf{Z};$
$\prod_{g:\mathbf{Gdr}}[(\mathbf{NP}_{\mathrm{nom},g,\mathrm{sg},3} \multimap \mathbf{S}_{\mathrm{inf},6}) \multimap \mathbf{NP}_{\mathrm{nom},g,\mathrm{sg},3} \multimap \mathbf{S}_{\mathrm{m},6}];$
$\lambda_{Fx}.\mathsf{(can}\ (F\,x)\ x) : (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$

With this lexical entry for the subject control verb, the grammar predicts that the sentence *Ana hoće voziti bicikl* 'Ana wants to ride a bike' is pronounceable 12 different ways. Below is the sign representing that sentence in this somewhat permissive grammar:

(176) $\vdash \mathbf{PER}(\texttt{ANA}_{\mathbf{z}} \circ \texttt{HOĆE}_{\mathbf{z}} \circ \mathbf{toZ}(\mathbf{k}\ (\mathbf{PER}\ (\texttt{VOZITI}_{\mathbf{z}} \circ \texttt{BICIKL}_{\mathbf{z}})))) : \mathbf{Z}; \mathbf{S}_{\mathrm{m},6};$
$\mathsf{exist(bicycle)}\lambda_{y}.\mathsf{(want\ (ride}\ y\ \mathsf{ana)\ ana} : (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$

We've simply assumed that we have two different lexical entries for each subject control verb, one looking for a finite verb phrase, the other for an infinitival verb phrase complements. Below we state a lexical, non-logical rule [SF], which maps each subject control verb looking for an infinitival complement to its finite verb phrase seeking counterpart. In the rule, ϕ is a metavariable over phenogrammatical terms of type $\mathtt{z} \rightarrow \mathtt{z} \rightarrow \mathtt{Z}$, and F is a variable of type $\mathtt{z} \rightarrow \mathtt{Z}$. For typographical reasons, we suppress the semantic component, as it remains unchanged.

(177)

$$\frac{\vdash \phi; \prod_{g:\mathbf{Gdr}}[(\mathbf{NP}_{\mathrm{nom},g,\tau,\tau'} \multimap \mathbf{S}_{\mathrm{inf},6}) \multimap \mathbf{NP}_{\mathrm{nom},g,\tau,\tau'} \multimap \mathbf{S}_{\mathrm{m},6}]}{\begin{array}{c}\vdash \lambda_{Fxy}.\exists_{vw}[(\phi \Subset_{\mathtt{S}} x\ w) \wedge (F \Subset_{\mathtt{S}} v) \wedge y = w \circ \mathtt{toZ}(\mathtt{L}\ v)];\\ \prod_{g:\mathbf{Gdr}}[(\mathbf{NP}_{\mathrm{nom},g,\tau,\tau'} \multimap \mathbf{S}_{\mathrm{e},6}) \multimap \mathbf{NP}_{\mathrm{nom},g,\tau,\tau'} \multimap \mathbf{S}_{\mathrm{m},6}]\end{array}}\ \text{[SF]}$$

This main impact of this rule is in the phenogrammatical term transformation, since instead of seeking a term of type $\mathtt{z}$ (infinitival verb phrase), the subject control verb now has to combine with a term of type $\mathtt{z} \rightarrow \mathtt{Z}$ (embedded clause with a subject gap). Tectogrammatically, the subject control verb's argument type has changed from $\mathbf{NP}_{\mathrm{nom},g,\tau,\tau'} \multimap \mathbf{S}_{\mathrm{inf},6}$ to $\mathbf{NP}_{\mathrm{nom},g,\tau,\tau'} \multimap \mathbf{S}_{\mathrm{e},6}$.

Object Control

Object control verbs cannot combine with infinitival verb phrases but instead only with embedded clauses with a subject gap. Suppose we are trying to construct a representation of the sentence *Marko nagovara Vesnu da dođe* 'Marko is persuading Ana to come'. The embedded sentence with the subject gap is constructed in exactly the same way as in the case of subject control. We give the following lexical entry for *nagovara* 'persuades':

(178) $\lambda_{zFvw}.\exists_{xy}[(F \in_{\mathbf{S}} x) \wedge (\mathbf{PER}(v \circ \mathtt{NAGOVARA}_{\mathbf{z}} \circ z)\ y) \wedge w = y \circ \mathbf{toZ}(\mathbf{L}\ x)] :$
$\mathbf{z} \to (\mathbf{z} \to \mathbf{Z}) \to \mathbf{z} \to \mathbf{Z};$
$\prod_{g,g':\mathbf{Gdr},n:\mathbf{Num},p:\mathbf{Prs}}[\mathbf{NP}_{\mathrm{acc},g,n,p} \multimap (\mathbf{NP}_{\mathrm{nom},g,n,p} \multimap \mathbf{S}_{\mathrm{e},6}) \multimap \mathbf{NP}_{\mathrm{nom},g',\mathrm{sg},3} \multimap$
$\mathbf{S}_{\mathrm{m},6}]; \lambda_{yFx}.(\text{persuade } y\ (F\ y)\ x) : \mathbf{e} \to (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$

Phenogrammatically, this object control verb works much like a subject control verb: it allows free reordering of itself, its subject and its accusative object, and forces the embedded clause with a subject gap, transformed into a length one string of languages, to occur on the right edge of the matrix sentence. Semantically, we analyze the verb as expressing a relation between an individual, a proposition and another individual.

Tectogrammatically, *nagovara* 'persuades' requires a nominative 3rd person subject of any gender. The g' variable stands for that 'missing' **Gdr** parameter of its subject. In addition, its object must be an accusative noun phrase, but its **Gdr**, **Num** and **Prs** parameters are not specified. Whatever those parameters turn out to be, however, *nagovara* will then require the controlled verb phrase to be missing a nominative noun phrase with the same **Gdr**, **Num** and **Prs** parameters as its object's. For example, the sentence *Marko nagovara Vesnu da dođe* 'Marko is persuading Ana to come' requires the following tectogrammatical version of *nagovara*, where the object and the 'missing' subject must be 3rd person singular feminine noun phrases, and the matrix subject must be masculine:

(179) $\lambda_{zFvw}.\exists_{xy}[(F \in_{\mathbf{S}} x) \wedge (\mathbf{PER}(v \circ \mathtt{NAGOVARA}_{\mathbf{z}} \circ z)\ y) \wedge w = y \circ \mathbf{toZ}(\mathbf{L}\ x)] :$
$\mathbf{z} \to (\mathbf{z} \to \mathbf{Z}) \to \mathbf{z} \to \mathbf{Z};$
$\mathbf{NP}_{\mathrm{acc,\ f,sg,3}} \multimap (\mathbf{NP}_{\mathrm{nom,f,sg,3}} \multimap \mathbf{S}_{\mathrm{e},6}) \multimap \mathbf{NP}_{\mathrm{nom,m,sg,3}} \multimap \mathbf{S}_{\mathrm{m},6};$
$\lambda_{yFx}.(\text{persuade } y\ (F\ y)\ x) : \mathbf{e} \to (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$

The whole sentence is represented in the grammar as follows:

(180) $\lambda_w.\exists_{xyv}[(\mathbf{PER}(\mathtt{DOĐE}_{\mathbf{z}})\ v) \wedge x = (\mathtt{DA}_{\mathbf{z}} \circ v) \wedge (\mathbf{PER}(\mathtt{MARKO}_{\mathbf{z}} \circ \mathtt{NAGOVARA}_{\mathbf{z}} \circ \mathtt{ANU}_{\mathbf{z}})\ y) \wedge w = y \circ \mathbf{toZ}(\mathbf{L}\ x)] : \mathbf{Z}; \mathbf{S}_{\mathrm{m},6};$
(persuade ana (come ana) marko) : **p**

Tectogrammatically, this is a declarative main clause. Semantically, it expresses the expected proposition. Phenogrammatically, the sign denotes a set of strings of languages consisting of some permutation of the verb, the subject and the object, followed by the embedded clause with the nominative gap which has been transformed into a length one string of languages. The grammar predicts that this sentence can be pronounced 6 different ways, which is correct.

Finally, we analyze verbs whose objects control a predicative adjective or a predicative noun phrase. We focus on controlled predicative adjectives, as in *Ana smatra Marka pristojnim* 'Ana considers Marko polite'. The adjective must occur in instrumental case, but it agrees with the accusative object in number and gender.

The verb *smatrati* 'consider' can also combine with a finite clause, as in *Ana smatra da je Marko pristojan*, which expresses the same meaning as its object control counterpart. Therefore, we analyze the meaning of this verb as a relation between an individual and a proposition.

Recall that adjectives which are complements of the copula must be nominative. It was to these kinds of adjectives that we assigned the tectogrammatical result type $\mathbf{Prd}_{\mathrm{a}}$. So that we don't get a case mismatch between a predicative adjective and the subject in such predicative structures with the copula, we will analyze this instrumental adjective that occurs as an argument of *smatrati* 'consider' as an ordinary attributive instrumental adjective. So, in the sentence *Ana smatra Marka pristojnim* 'Ana considers Marko polite' the adjective is associated with the following lexical entry:[20]

(181) $\lambda_v.\mathtt{PRISTOJNIM}_{\mathbf{z}} \circ v : \mathbf{z} \to \mathbf{z}; \mathbf{N}_{\text{inst,m,sg}} \multimap \mathbf{N}_{\text{inst,m,sg}}; \mathsf{polite} : \mathbf{e} \to \mathbf{p}$

Below is the lexical entry for the verb *smatra*.

(182) $\vdash \lambda_{vFw}.\mathbf{PER}(w \circ \mathtt{SMATRA}_{\mathbf{z}} \circ v \circ (F\, e_{\mathbf{S}})) : \mathbf{z} \to (\mathbf{z} \to \mathbf{z}) \to \mathbf{z} \to \mathbf{Z};$
$\prod_{g,g':\mathbf{Gdr},n:\mathbf{Num},p:\mathbf{Prs}}[\mathbf{NP}_{\text{acc},g,n,p} \multimap (\mathbf{N}_{\text{inst},g,n} \multimap \mathbf{N}_{\text{inst},g,n}) \multimap \mathbf{NP}_{\text{nom},g',\text{sg},3} \multimap$
$\mathbf{S}_{\text{m},6}]; \lambda_{xPy}.(\mathsf{consider}(P\, x)\; y) : \mathbf{e} \to (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$

This verb requires an accusative object, an instrumental adjective and a nominative singular 3rd person subject. Its tectogrammatical type ensures that its object and the adjective that its object controls have the same gender and number. The phenogrammatical term of this verb allows free reordering of all the constituents. The sentence *Ana smatra Marka pristojnim* 'Ana considers Marko polite' is represented in the grammar as follows.

(183) $\vdash \mathbf{PER}(\mathtt{ANA}_{\mathbf{z}} \circ \mathtt{SMATRA}_{\mathbf{z}} \circ \mathtt{MARKA}_{\mathbf{z}} \circ \mathtt{PRISTOJNIM}_{\mathbf{z}}) : \mathbf{Z}; \mathbf{S}_{\text{m},6};$
$(\mathsf{consider}(\mathsf{polite\ marko})\ \mathsf{ana}) : \mathbf{p}$

4.5 Conclusion

In this chapter, we analyzed embedded declarative clauses, and predicative and control structures. We analyzed controlled finite verb phrases, in both subject

[20]Here we consider only the most permissive version of the attributive adjective, of all the versions entertained in Chapter 3.

and object control structures, as embedded declarative clauses with a gap, without introducing any new tectogrammatical types.

We introduced a new **K** parameter for infinitival phrases. Following the observation that all and only subject control verbs can combine with infinitival verb phrases in addition to finite verb phrases, we gave a rule [FS] which takes a subject control verb subcategorized for an infinitival complement and outputs its counterpart which takes a finite complement.

Finally, we incorporated predicative complements into the grammar, by introducing a family of predicative types. Doing so enabled us to give a single lexical entry for a given form of the copula, which can take as its complement whichever kind of predicative phrase. We also explored the connection between predicative complements and postnominal modifiers and, where appropriate, devised non-logical rules which establish the connection between the two incarnations of certain classes of expressions.

In order to give a theory of encliticization in Serbo-Croatian we first had to analyze embedding, control and predicative structures both because embedded clauses are domains for clitic placement, and because all clitic verbs in Serbo-Croatian are either subject control or predicative verbs. Building on the analysis of these structures laid out here, in the next chapter we present our analysis of the Serbo-Croatian enclitic cluster.

Chapter 5: Enclitics

5.1 Introduction

Serbo-Croatian enclitics have attracted a lot of attention in the literature over the decades (see Browne (1974); Halpern (1995); Schütze (1994); Progovac (1996, 2005); Radanović-Kocić (1996); Penn (1999a); Bošković (2001, 2004) inter alia). In contrast to proclitics (such as prepositions which we examined in the previous chapter), enclitics attach to a prosodic word to their left, and together with their host form a new phonological word.

The enclitics, which include pronouns, reflexives and auxiliaries, all cluster together in a certain not entirely predictable order. The resulting enclitic cluster is extremely limited in terms of possible placement within the sentence. The ordering within the clitic cluster and the placement of the cluster in the sentence are the biggest challenges when it comes to constructing an analysis. In this chapter we first examine the empirical facts concerning the order and the placement of the clitic cluster, and then present our analysis.

5.2 Data

5.2.1 Order

Enclitics include dative, genitive and accusative pronominal clitics, the accusative reflexive pronoun and the inherent reflexive *se*, the interrogative complementizer *li* which occurs in polar interrogatives, and several different kinds of auxiliaries. The auxiliaries include the the clitic forms of *htjeti* 'want, will' used in future tense formation, and two different sets of clitics of the verb *biti* 'be'. The imperfective present tense clitics of *biti* 'be' are used as copula and for past tense formation, while its aorist clitics are used to construct conditional forms of verbs. Of all these enclitics, only *li* and the aorist of *biti* do not have corresponding non-clitic forms.

The enclitics cluster together and are strictly ordered as follows:

li<auxiliaries – {*je*}<dative<accusative<genitive<*se*<*je*

Here, *je* 'is', is the 3rd person singular present tense clitic of *biti* and occurs in a different slot in the clitic cluster than all other auxiliary clitics. Any violations of this order in the clitic cluster, or any attempt to make the cluster discontinuous, result in sharp ungrammaticality (Browne , 1974; Progovac , 1996, 2005; Radanović-Kocić , 1996; Franks and King , 2000).

While at first glance it may seem that the clitics are ordered by their syntactic categories, this cannot be maintained without much stipulation. First, the reflexive *se* is accusative but occurs in a different slot than other accusative clitics. In addition to *se* being a possible argument of a verb needing an accusative object, it is sometimes simply a part of a lexical verb:

(184) a. Ana vidi Maju u ogledalu.
Ana$_{NOM}$ sees Maja$_{ACC}$ in mirror
'Ana sees Maja in the mirror'

b. Ana **se** vidi u ogledalu.
Ana$_{NOM}$ REFL$_{ACC}$ in mirror
'Ana sees herself in the mirror'

(185) a. Ana **se** bavi lingvistikom.
Ana$_{NOM}$ does (professionally) linguistics$_{INST}$
'Ana does linguistics (professionally)'

b. Ana **se** boji Maje.
Ana$_{NOM}$ is-afraid-of Maja$_{GEN}$
'Ana is afraid of Maja'

While in (184b) *se* simply occurs as an argument of *vidi* 'sees', in (185a) and (185b) it is just a part of the verb and cannot be left out of the sentence or replaced by some other accusative expression, as if the verb and *se* constituted a phrasal idiom. I will call the *se* that occurs in (185) inherent reflexive. Both kinds of *se*, however, occur in the same slot in the clitic cluster, although they are tectogrammatically different.

Also note that *je*, 3rd person singular present tense *biti* 'be' (which functions as a copula and as an auxiliary in periphrastic past tense) is idiosyncratic in that it occurs at the end of the clitic cluster and not where all the other auxiliaries occur.

In sum, we will maintain that the order of enclitics in the cluster must be stipulated and cannot be derived from any general syntactic properties of the clitics, contrary to popular MGG views.

5.2.2 Placement

The even trickier issue is the placement of the enclitic cluster within the clause. Below we argue that its placement cannot be accounted for either purely syntactically or purely prosodically. This fact alone undermines the MGG analyses that I'm familiar with, since they all essentially assume either purely syntactically or purely prosodically conditioned placement.

Radanović-Kocić (1996) attempts to give a prosodic account of enclitic placement and states that the clitics need to come right after the first phonological phrase in their intonational phrase (Bošković (2001, 2004) accepts this generalization and makes use of it in his analysis). She however offers no definition of what constitutes a phonological phrase in Serbo-Croatian other than "one or more prosodic words" (p. 441). She suggests that degemination may be prohibited across a phonological phrase boundary, but has no phonetic evidence to substantiate this claim.

A phonological phrase is supposed to be a prosodic constituent larger than a prosodic word but smaller than an intonational phrase. It seems that Radanović-Kocić (1996) wants syntactic constituents in general to correspond to phonological phrases, and that a focused (presumably contrastively focused, given her examples) word within a constituent may constitute its own phonological phrase.

Note that both of the following are possible:

(186)	a.	Moja	**je**	sestra	došla.
		my_{NOM}	is	sister_{NOM}	arrived

'My sister arrived'

b. Moja sestra **je** došla.

Example (186a) shows the so-called 1W clitic cluster placement (after the first word), and example (186b) 1C clitic cluster placement (after the first constituent).

According to Radanović-Kocić (1996), *moja sestra* in (186b) forms a phonological phrase and so the example obeys her generalization. (186a), she claims, is only possible if the possessive is focused and forms its own phonological phrase. This allows her to maintain that in (186a) the clitic does come after the first phonological phrase in its intonational phrase, since the possessive is stipulated to constitute its own phonological phrase. However, the production study in Yu (2009) directly contradicts Radanović-Kocić (1996) in that no evidence for a prosodic break after *moja* in sentences like (186a) was found. It also doesn't seem to be the case that the clitic host in 1W examples has to be contrastively or otherwise focused.

Further, Godjevac (2000) in her extensive study of Serbo-Croatian intonation found no phonetic evidence for an intermediate prosodic constituent, even when the utterance is very long. While in very long utterances there is a periodic pitch reset (an effect she calls pleating), the points at which the pitch is reset do not seem to reliably correspond to syntactic constituent boundaries, and she analyzes such utterances as consisting of single intonational phrases.

So the idea that there are phonological phrases in Serbo-Croatian doesn't appear to be empirically grounded. Moreover, even if phonological phrases turn out to be in the Serbo-Croatian prosodic inventory, there doesn't seem to be a prosodic boundary precisely where Radanović-Kocić (1996) needs one for her generalization concerning the clitic cluster placement to go through.

Clearly, a prosodic generalization on clitic cluster placement cannot be stated in terms of phonological words alone since the cluster can attach to the right edge of the first or the second phonological word in an IP (as we saw above), or be delayed by many phonological words, e.g.:

(187) a. Roditelji uspešnih studenata **su se** razišli.
parents$_{NOM}$ successful$_{GEN}$ students$_{GEN}$ are dispersed
'The parents of successful students dispersed'
[adapted from Progovac (1996)]

b. Ona moja sestra koja je u Sarajevu **vas**
that$_{NOM}$ my$_{NOM}$ sister$_{NOM}$ who$_{NOM}$ is in Sarajevo you$_{GEN}$
se sjeća.
remembers
'That sister of mine who is in Sarajevo remembers you'
[adapted from Radanović-Kocić (1996)]

So, the generalization concerning the placement of the clitic cluster cannot be stated over prosodic constituents. However, their placement is clearly sensitive to prosodic factors since they cannot occur immediately to the right of proclitics, or be the first elements in an intonational phrase, because they need a host that is a phonological word.

Progovac (1996, 2005) attempts to give a syntactic generalization concerning the placement of the clitic cluster. She maintains that only expressions that are independently 'moveable' around the clause can host clitics (so does Bošković

(2004), see p.12). So, for example, Progovac would claim that an attributive adjective can host the clitics because it can in general be detached from the rest of the noun phrase, but since postnominal modifiers cannot be detached from the noun they are modifying, clitics cannot come between a noun and its postnominal modifier.

We believe this generalization to be largely correct. However, in some cases enclitics can in fact split constituents that otherwise must remain contiguous. While the examples may be somewhat degraded, it is possible to place the clitic cluster between a noun and a post-nominal modifier (Browne (1974); Halpern (1995), Aaron Halpern p.c. to Progovac (2005)).

(188) ? Sestra **će** moje prijateljice doći.
sister$_{NOM}$ will my$_{GEN}$ friend$_{GEN}$ arrive$_{INF}$
'My sister's friend will arrive'

The clitics can also occur right after the first conjunct, although co-ordinate structures in general cannot appear discontinuously in a clause in Serbo-Croatian, as shown below:

(189) a. ? Knjige **ću** i teke sutra kupiti.
books will and notebooks tomorrow buy$_{INF}$
'I'll buy books and notebooks tomorrow'
b. Knjige i teke **ću** sutra kupiti.
c. * Knjige sutra i teke **ću** kupiti.

In our analysis, we will first account for Progovac's generalization, namely that the set of potential clitic hosts coincides with the set of 'moveable' expressions; i.e.

expressions that can freely reorder with other clausal constituents. These expressions may be one prosodic word long (1W placement) or many prosodic words long (1C placement). In our grammar, they will turn out to correspond exactly to phenogrammatical terms that are length one strings of languages.

Then, we will add a rule that allows clitics to be placed in such a way that they break up otherwise unbreakable constituents, e.g. between a noun and a postnominal modifier. Depending on how permissive one's judgments are, this rule can be added to the grammar or not.

Either way, within our theory, the generalization concerning the clitic cluster placement can be stated as follows: the cluster encliticizes to the last phonological word 'inside' the first length one string of languages in a sentence.

5.3 Analysis

5.3.1 Preliminaries

First, we make a case for the enclitic interrogative complementizer *li* which occurs in polar questions not being treated as an independent lexical item. *li* always comes encliticized either onto a finite verb or the complementizer *da*. The verb+*li* or *da*+*li* must occur question-initially. So we can give a lexical entry for *da li* which forms polar interrogatives out of finite clauses and requires *da li* to occur on the left edge of the question. For the finite verb+*li* polar question forming strategy,

1	2	3	4	5	6
AUX except *je*	dative clitics	accusative clitics	genitive clitics	*se*	*je*

Table 5.1: The order of enclitics in the cluster.

we can develop a lexical rule that maps each finite verb to its interrogative version which comes in a package with *li*, with the verb+*li* sequence having to occur question-initially. We will return to *li* in the chapter on interrogatives.

We are going to treat enclitics as functions over 'gappy' sentences, similar to how we analyzed quantificational noun phrases. So, a clitic will target a sentence missing an expression that's exactly like that clitic, tectogrammatically and semantically, which is essentially standard in categorial grammar (see for example Nishida (1996), Morrill and Gavarró (1992) and Kraak (1998)).

Table 5.1 shows the required order of clitics in the enclitic cluster (excluding *li* as explained at the beginning of this section):

We're going to enforce the order in the cluster via tectogrammatical types in the **S** family. As discussed in Chapter 3, the **S** family of types is indexed by natural numbers and **K**. **K**={m,e,q} parameter reflects the kind of sentence we're dealing with—a main declarative clause, an embedded declarative clause, or a question (main or embedded).

Intuitively, for each sign whose tecto type is $\mathbf{S}_{k,n}$, n refers to the maximal number of clitics that could be placed in that sentence. For example, if a sentence is

ARGUMENT'S RESULT TYPE	$\mathbf{S}_{n>5}$	$\mathbf{S}_{n>4}$	$\mathbf{S}_{n>3}$	$\mathbf{S}_{n>2}$	$\mathbf{S}_{n>1}$	$\mathbf{S}_{n>0}$
RESULT TYPE	$\mathbf{S}_5$	$\mathbf{S}_4$	$\mathbf{S}_3$	$\mathbf{S}_2$	$\mathbf{S}_1$	$\mathbf{S}_0$
CLITICS	AUX except *je*	dative clitics	accusative clitics	genitive clitics	*se*	*je*

Table 5.2: Tectogrammatical types of enclitics.

associated with the type $\mathbf{S}_{k,6}$ it means that no clitics have been placed inside that sentence, and up to six clitics could still occur inside of it. If a sentence has the type $\mathbf{S}_{k,0}$ it means that no more clitics can be placed inside of it. Recall that finite verbs in general build sentences of type $\mathbf{S}_{\mathrm{m},6}$ since they can't know in advance how many clitics may wind up placed inside of that sentence at the end.

The enclitics combine with 'gappy' sentences in the left-to-right order of their appearance in the clitic cluster. As they do so, they systematically reduce the number parameter of the sentence type, thereby preventing clitics that must occur to their left to apply after them. This way the order in the cluster in enforced.[21]

Informally, the number parameter order reverses the order of various clitic slots in the enclitic cluster. Table 5.2 lists, for each clitic slot, the type of sentence that the clitic can combine with, and the type of sentence that results after a given clitic combines with that sentence.

[21]The idea to use natural number parameters on sentence types to order clitics comes from Morrill and Gavarró (1992), however, our implementation is somewhat different.

So, for example, a dative clitic can combine with a sentence whose number parameter is greater than 4, that is, inside of which at least 5 more enclitics can be placed. It outputs a sentence whose number parameter is exactly 4, meaning that after the dative clitic is placed in the sentence, exactly 4 more enclitics can also occur in that sentence.

5.3.2 Pronominal Clitics

Suppose we're trying to generate the sentence *Ana mu daje knjigu* 'Ana gives him a book', where *mu* 'him' is a dative enclitic. The non-clitic version of the same pronoun is *njemu* 'him'. Below we give the lexical entry for the non-clitic and the clitic version of this pronoun.[22]

(190) $\vdash \mathrm{NJEMU}_{\mathbf{z}} : \mathbf{z}; \mathbf{NP}_{\mathrm{dat,\,m,sg,3}}; x : \mathbf{e}$

(191) $\vdash \lambda_{Fw}.\exists_{vt}[(F\ \mathrm{e}_{\mathbf{s}}\ v) \wedge (\mathbf{fst}_{\mathbf{z}}\ v\ t) \wedge w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc}_{\mathbf{s}}(\mathbf{lst}_{\mathbf{s}}\ t)\#\mathrm{mu}$
$(\mathbf{tsr}_{\mathbf{s}}\ t))) \circ (\mathbf{rst}_{\mathbf{z}}\ v)] : (\mathbf{z} \rightarrow \mathbf{Z}) \rightarrow \mathbf{Z};$
$(\mathbf{NP}_{\mathrm{dat,m,sg,3}} \multimap \mathbf{S}_{\mathrm{m},n>4}) \multimap \mathbf{S}_{\mathrm{m},4}; \lambda_{G}.(G\ x) : (\mathbf{e} \rightarrow \mathbf{p}) \rightarrow \mathbf{p}$

Tectogrammatically, the dative clitic is looking for a sentence in which at least 5 more clitics can be placed, and which is missing a noun phrase of exactly the same type as the non-clitic version of this dative pronoun, $\mathbf{NP}_{\mathrm{dat,\,m,sg,3}}$.

Phenogrammatically, the clitic first feeds the sentence with the appropriate dative gap the empty string, $\mathrm{e}_{\mathbf{s}}$, which results in s set of strings of languages

[22] We are semantically representing clitics as variables of type $\mathbf{e}$. We recognize that this is inadequate, however, the details about pronominal meanings and binding are unfortunately beyond the scope of this dissertation.

corresponding to the possible pronunciations of the sentence without the dative argument. v is one such string of languages.

The clitic then 'looks' inside the first string of languages in each v. This is some set of strings, and t is one such string. The dative clitic then encliticizes onto the last phonological word that the string t is built out of.

Below we show how to construct the sentence *Ana mu daje knjigu* 'Ana gives him a book'. First, we give the lexical entry for the verb.

(192) $\vdash \lambda_{vxw}.\mathbf{PER}(w \circ \mathtt{DAJE}_{\mathbf{z}} \circ v \circ x) : \mathbf{z} \to \mathbf{z} \to \mathbf{z} \to \mathbf{Z};$
$\mathbf{NP}_{acc,f,sg,3} \multimap \mathbf{NP}_{dat,m,sg,3} \multimap \mathbf{NP}_{nom,f,sg,3} \multimap \mathbf{S}_{m,6};$
$\mathsf{give} : \mathbf{e} \to \mathbf{e} \to \mathbf{e} \to \mathbf{p}$

The accusative object is a bare noun which has to undergo [NC] and [Quant] to become a full-fledged quantificational noun phrase, which will then scope over a sentence with the appropriate gap. So we first introduce an accusative trace and combine it with the verb.

Similarly, since the dative argument is a clitic, we introduce a dative trace as well and proceed with the sentence construction. Finally, the verb can combine with the subject, but both the dative and the accusative trace are kept track of in the context. The result is the following sign.

(193) $x : \mathbf{z}; \mathbf{NP}_{dat,m,sg,3}; x : \mathbf{e},$
$y : \mathbf{z}; \mathbf{NP}_{acc,f,sg,3}; y : \mathbf{e} \quad \vdash \mathbf{PER}(\mathtt{ANA}_{\mathbf{z}} \circ \mathtt{DAJE}_{\mathbf{z}} \circ y \circ x) : \mathbf{Z}; \mathbf{S}_{m,6}; (\mathsf{give}\ y\ x\ \mathsf{ana}) : \mathbf{p}$

We can bind the traces in either order. However, if we first bind the dative trace and place the clitic in the sentence, we won't be able to then use the quantificational version of *knjigu* 'a book'. This is because quantificational noun phrases are looking for gappy sentences whose number parameter is 6. But after the dative clitic combines with the sentence, its number parameter is reduced to 4. Therefore, the tectogrammatical typing requires us to first bind the accusative trace and introduce *knjigu* 'a book' into the derivation, and then deal with the clitic. In general, quantificational noun phrases will have to be dealt with before any clitics, since those quantificational noun phrases may wind up being clitic hosts.

After we bind the accusative trace and scope *knjigu* 'a book', we get the following sign:

(194) $$x : \mathbf{z}; \mathbf{NP}_{\mathrm{dat,m,sg,3}}; x : \mathbf{e} \vdash \mathbf{PER}(\mathtt{ANA}_{\mathbf{z}} \circ \mathtt{DAJE}_{\mathbf{z}} \circ \mathtt{KNJIGU}_{\mathbf{z}} \circ x) : \mathbf{Z}; \mathbf{S}_{\mathrm{m,6}}; (\text{exists book})(\lambda_y.(\text{give } y\ x\ \text{ana})) : \mathbf{p}$$

The dative trace is still in the context. Now we can bind it and finally introduce the dative clitic. The result is the following sign.

(195) $$\vdash \lambda_w.\exists_{vt}[((\mathbf{PER}(\mathtt{ANA}_{\mathbf{z}} \circ \mathtt{DAJE}_{\mathbf{z}} \circ \mathtt{KNJIGU}_{\mathbf{z}})\ v) \wedge (\mathbf{fst}_{\mathbf{z}}\ v\ t) \wedge w = \mathbf{toZ}(\lambda_s.s = (\mathbf{snc}_{\mathbf{s}}(\mathbf{lst}_{\mathbf{s}}\ t)\#\mathtt{mu}\ (\mathbf{tsr}_{\mathbf{s}}\ t))) \circ (\mathbf{rst}_{\mathbf{z}}\ v)] : \mathbf{Z}; \mathbf{S}_{\mathrm{m,4}}; (\text{exists book})(\lambda_y.(\text{give } y\ x\ \text{ana})) : \mathbf{p}$$

Tectogrammatically, because of the number parameter, no more dative clitics, and no auxiliary clitics can be placed in this sentence. Phenogrammatically, the sign denotes a set of strings of languages in each of which the clitic has now been encliticized onto the last phonological word in the first string of languages. Here are some such strings of languages:

(196)
a. $\mathbf{toZ}(\lambda_s.s = (\mathtt{ana\#mu})_{\mathbf{s}}) \circ \mathtt{DAJE}_{\mathbf{z}} \circ \mathtt{KNJIGU}_{\mathbf{z}}$
b. $\mathbf{toZ}(\lambda_s.s = (\mathtt{daje\#mu})_{\mathbf{s}}) \circ \mathtt{ANA}_{\mathbf{z}} \circ \mathtt{KNJIGU}_{\mathbf{z}}$
c. $\mathbf{toZ}(\lambda_s.s = (\mathtt{knjigu\#mu})_{\mathbf{s}}) \circ \mathtt{DAJE}_{\mathbf{z}} \circ \mathtt{ANA}_{\mathbf{z}}$

d. etc.

Now we show how to generate the sentence *Ana mu ga daje* ‘Ana gives it to him’, which contains two clitics that have to occur in this order. This sentence can be pronounced only two ways, *Ana mu ga daje* or *Daje mu ga Ana*. Below we give the lexical entry for the accusative clitic.

(197) $\vdash \lambda_{Fw}.\exists_{vt}[(F \ \mathtt{e}_{\mathbf{s}} \ v) \wedge (\mathbf{fst_z} \ v \ t) \wedge w = \mathbf{toZ}(\lambda_s.s = (\mathbf{snc_s}(\mathbf{lst_s} \ t)\#\mathtt{ga}$
$(\mathbf{tsr_s} \ t))) \circ (\mathbf{rst_z} \ v)] : (\mathbf{z} \rightarrow \mathbf{Z}) \rightarrow \mathbf{Z};$
$(\mathbf{NP}_{\mathrm{acc,m,sg,3}} \multimap \mathbf{S}_{\mathrm{m},n>3}) \multimap \mathbf{S}_{\mathrm{m,3}}; \lambda_G.(G\ x) : (\mathbf{e} \rightarrow \mathbf{p}) \rightarrow \mathbf{p}$

Phenogrammatically, this accusative clitic works the same way as the dative one we considered earlier. Tectogrammatically, the difference is in the sentence number parameter. Because the accusative clitic reduces the number parameter of the sentence to 3, it is impossible to first place the accusative clitic and then the dative clitic. So we first have to bind the dative trace and place the dative clitic, and then bind the accusative trace and place the accusative clitic. The result is the following sign:

(198) $\vdash \lambda_x.\exists_{yuvt}[(\mathbf{PER}(\mathrm{ANA}_{\mathbf{z}} \circ \mathrm{DAJE}_{\mathbf{z}}) \ v) \wedge (\mathbf{fst_z} \ v \ t) \wedge$
$y = (\mathbf{toZ}(\lambda_s.s = (\mathbf{snc_s}(\mathbf{lst_s} \ t)\#\mathtt{mu} \ (\mathbf{tsr_s} \ t))) \circ (\mathbf{rst_z} \ v)) \wedge (\mathbf{fst_z} \ y \ u) \wedge$
$x = (\mathbf{toZ}(\lambda_s.s = (\mathbf{snc_s}(\mathbf{lst_s} \ u)\#\mathtt{ih} \ (\mathbf{tsr_s} \ u))) \circ (\mathbf{rst_z} \ y))] : \mathbf{Z};$
$\mathbf{S}_{\mathrm{m,3}}; (\mathsf{give}\ x\ y\ \mathsf{ana}) : \mathbf{p}$

The phenogrammatical term is complicated but it denotes a set which contains exactly the following two strings of languages:

(199) a. $\mathbf{toZ}(\lambda_s.s = (\mathtt{ana\#mu\#ga})_{\mathbf{s}}) \circ \mathrm{DAJE}_{\mathbf{z}}$
b. $\mathbf{toZ}(\lambda_s.s = (\mathtt{daje\#mu\#ga})_{\mathbf{s}}) \circ \mathrm{ANA}_{\mathbf{z}}$

We analyze genitive clitics similarly.

5.3.3 The Inherent Reflexive *se*

In this section, we analyze the inherent reflexive *se*, which occurs in the same slot in the clitic cluster as the true reflexive *se*. While the true reflexive *se* has a corresponding full form *sebe*, the inherent reflexive does not.

Recall that by certain verbs require the occurrence of *se* which contributes nothing in terms of meaning. We analyze such verbs as introducing a hypothesis via their lexical entries. For example, consider the sentence *Boji ga se Ana* ‘Ana is afraid of him’, where the relevant verb is *bojati se* ‘be afraid’. We give the following lexical entries for *boji*, *ga* and the inherent reflexive *se*:

(200) $x : \mathtt{z}; \mathbf{SE}; x : \mathbf{e} \vdash \lambda_{uv}.\mathbf{PER}(v \circ \mathtt{BOJI}_{\mathtt{z}} \circ u \circ x); \mathbf{NP}_{\text{gen, m,sg,3}} \multimap \mathbf{NP}_{\text{nom, f,sg,3}} \multimap \mathbf{S}_{\text{m},6}; \lambda_{uv}.(\text{is-afraid } u\ v) : \mathbf{e} \to \mathbf{e} \to \mathbf{p}$

(201) $\vdash \lambda_{Fw}.\exists_{vt}[(F\ \mathtt{e_s}\ v) \wedge (\mathtt{fst_z}\ v\ t) \wedge w = \mathtt{toZ}(\lambda_s.s = (\mathtt{snc_s}(\mathtt{lst_s}\ t)\#\mathtt{ga}\ (\mathtt{tsr_s}\ t))) \circ (\mathtt{rst_z}\ v)] : (\mathtt{z} \to \mathtt{Z}) \to \mathtt{Z};$
$(\mathbf{NP}_{\text{gen,m,sg,3}} \multimap \mathbf{S}_{\text{m},n>2}) \multimap \mathbf{S}_{\text{m},2}; \lambda_G.(G\ y) : (\mathbf{e} \to \mathbf{p}) \to \mathbf{p}$

(202) $\vdash \lambda_{Fw}.\exists_{vt}[(F\ \mathtt{e_s}\ v) \wedge (\mathtt{fst_z}\ v\ t) \wedge w = \mathtt{toZ}(\lambda_s.s = (\mathtt{snc_s}(\mathtt{lst_s}\ t)\#\mathtt{se}\ (\mathtt{tsr_s}\ t))) \circ (\mathtt{rst_z}\ v)] : (\mathtt{z} \to \mathtt{Z}) \to \mathtt{Z};$
$(\mathbf{SE} \multimap \mathbf{S}_{\text{m,n}>1}) \multimap \mathbf{S}_{\text{m},1}; \lambda_G.(G\ x) : (\mathbf{e} \to \mathbf{p}) \to \mathbf{p}$

The tectogrammatical type **SE** is precisely the type of the inherent reflexive *se*. In constructing the sentence, first withdraw the genitive hypothesis and place the genitive clitic, which results in the following sign:

(203) $x : \mathtt{z}; \mathbf{SE}; x : \mathbf{e} \vdash \lambda_w.\exists_{vt}[(\mathbf{PER}(\mathtt{ANA}_{\mathtt{z}} \circ \mathtt{BOJI}_{\mathtt{z}} \circ x)v) \wedge (\mathtt{fst_z}\ v\ t) \wedge w = \mathtt{toZ}(\lambda_s.s = (\mathtt{snc_s}(\mathtt{lst_s}\ t)\#\mathtt{ga}\ (\mathtt{tsr_s}\ t))) \circ (\mathtt{rst_z}\ v)] : \mathtt{Z};$
$\mathbf{S}_{\text{m},2}; (\text{is-afraid } y \text{ ana}) : \mathbf{p}$

While the verb introduces a hypothesis in its lexical entry, which is still in the context in the sign above, since *se* doesn’t contribute anything semantically the semantic variable in the hypothesis does not occur anywhere in the verb's semantic term. When the hypothesis is withdrawn, the semantic term is vacuously abstracted on:

(204) $\vdash \lambda_{xw}.\exists_{vt}[(\mathbf{PER}(\mathrm{ANA}_{\mathbf{z}} \circ \mathrm{BOJI}_{\mathbf{z}} \circ x)v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$
$w = \mathbf{toZ}(\lambda_s.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\mathrm{ga}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)] : \mathbf{z} \rightarrow \mathbf{Z};$
$\mathbf{SE} \multimap \mathbf{S}_{\mathrm{m},2}; \lambda_x.(\text{is-afraid } y \text{ ana}) : \mathbf{e} \rightarrow \mathbf{p}$

The clitic *se* simply gets rid of this vacuous abstraction. Below is the sign that represents in the grammar the sentence *Ana ga se boji* 'Ana is afraid of him':

(205) $\vdash \lambda_x.\exists_{yuvt}[(\mathbf{PER}(\mathrm{ANA}_{\mathbf{z}} \circ \mathrm{BOJI}_{\mathbf{z}})\ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$
$y = (\mathbf{toZ}(\lambda_s.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\mathrm{ga}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)) \wedge (\mathbf{fst_z}\ y\ u) \wedge$
$x = (\mathbf{toZ}(\lambda_s.s = (\mathbf{snc_s}(\mathbf{lst_s}\ u)\#\mathrm{se}\ (\mathbf{tsr_s}\ u))) \circ (\mathbf{rst_z}\ y))] : \mathbf{Z};$
$\mathbf{S}_{\mathrm{m},1}; (\text{is-afraid } y \text{ ana}) : \mathbf{p}$

The phenogrammatical term of this sign denotes a set that contains the following two strings of languages:

(206) a. $\mathbf{toZ}(\lambda_s.s = (\mathrm{ana\#ga\#se})_{\mathbf{s}}) \circ \mathrm{BOJI}_{\mathbf{z}}$
b. $\mathbf{toZ}(\lambda_s.s = (\mathrm{boji\#ga\#se})_{\mathbf{s}}) \circ \mathrm{ANA}_{\mathbf{z}}$

5.3.4 The True Reflexive *se*

First we will analyze the non-clitic reflexive *sebe*, then extend that analysis to the clitic reflexive.

The Non-clitic Reflexive

The reflexive *sebe* is an accusative case marked noun phrase, but has no gender or number features. In simple sentences, it is interpreted as coreferential with the closest subject. For example:

(207) a. Ana$_i$ voli sebe$_i$.
Ana$_{NOM,f,sg,3}$ love$_{sg,3}$ self$_{ACC}$
'Ana loves herself'

b. Ana$_j$ hoće da Marko$_i$ voli sebe$_{i/*j}$.
Ana$_{NOM,f,sg,3}$ want$_{sg,3}$ DA Marko$_{NOM,m,sg,3}$ love$_{sg,3}$ self$_{ACC}$
'Ana wants that Marko loves himself'

If *sebe* occurs in a subject controlled complement, it is interpreted as coreferential with the matrix subject.

(208) subject control

a. Ana$_i$ hoće da voli sebe$_i$.
Ana$_{NOM,f,sg,3}$ want$_{sg,3}$ DA love$_{sg,3}$ self$_{ACC}$
'Ana wants to love herself'

b. Marko$_i$ mora voljeti sebe$_i$.
Marko$_{NOM,m,sg,3}$ must$_{sg,3}$ love$_{inf}$ self$_{ACC}$
'Marko must love himself'

If *sebe* occurs in an object controlled complement, it is interpreted as coreferential with the matrix object; if *sebe* is the object controller, then it is interpreted as coreferential with the matrix subject.

(209) object control

a. Ana$_j$ nagovara Marka$_i$ da nominira sebe$_{i/*j}$.
Ana$_{NOM,f,sg,3}$ persuade$_{sg,3}$ Marko$_{ACC,m,sg,3}$ DA nominate$_{sg,3}$ self$_{ACC}$
'Ana is persuading Marko to nominate himself'

b. Ana$_i$ nagovara sebe$_i$ da nominira Marka.
Ana$_{NOM,f,sg,3}$ persuade$_{sg,3}$ self$_{ACC}$ DA nominate$_{sg,3}$ Marko$_{ACC,m,sg,3}$
'Ana is persuading herself to nominate Marko'

We give the following lexical entry schema for the reflexive *sebe*:

(210) $\vdash \lambda_{Gw}.(G\, \mathtt{SEBE}_{\mathbf{z}}\; w) : (\mathbf{z} \to \mathbf{z} \to \mathbf{Z}) \to \mathbf{z} \to \mathbf{Z};$ $\prod_{g,g':\mathbf{Gdr},n,n':\mathbf{Num},p,p':\mathbf{Prs}}(\mathbf{NP}_{\mathrm{acc},g,n,p} \multimap \mathbf{NP}_{\mathrm{nom},g',n',p'} \multimap \mathbf{S}_{\tau,6}) \multimap \mathbf{NP}_{\mathrm{nom},g',n',p'} \multimap \mathbf{S}_{\tau,6}]; \lambda_{Fz}.(F\, z\, z) : (\mathsf{e} \to \mathsf{e} \to \mathsf{p}) \to \mathsf{e} \to \mathsf{p}$
where τ is a metavariable that ranges over inf, m.

So, *sebe* takes as its argument a finite or infinitival verb phrase missing an accusative object and alters its meaning so that by the time the whole verb phrase combines with the subject the appropriate interpretation is obtained, and the reflexive is coreferential with the matrix subject. Phenogrammatically, the reflexive just lowers itself into its argument, much like a quantificational noun phrase, ensuring that it freely permutes with other clausal constituents, just like noun phrases in general.

Suppose we're trying to analyze the sentence *Ana voli sebe* 'Ana loves herself'. Below is the version of the lexical entry for *sebe* needed to generate the sentence *Ana voli sebe* 'Ana loves herself', and the result of its combination with the verb *voli*.

(211) $\vdash \lambda_{Gw}.(G\, \mathtt{SEBE}_{\mathbf{z}}\; w) : (\mathbf{z} \to \mathbf{z} \to \mathbf{Z}) \to \mathbf{z} \to \mathbf{Z}; (\mathbf{NP}_{\mathrm{acc,\ f,sg,3}} \multimap \mathbf{NP}_{\mathrm{nom,\ f,sg,3}} \multimap \mathbf{S}_{\mathrm{m},6}) \multimap \mathbf{NP}_{\mathrm{nom,\ f,sg,3}} \multimap \mathbf{S}_{\mathrm{m},6}; \lambda_{Fz}.(F\, z\, z) : (\mathsf{e} \to \mathsf{e} \to \mathsf{p}) \to \mathsf{e} \to \mathsf{p}$

(212) $\vdash \lambda_{w}.\mathbf{PER}\ (w \circ \mathtt{VOLI}_{\mathbf{z}} \circ \mathtt{SEBE}_{\mathbf{z}}) : \mathbf{z} \to \mathbf{Z}; \mathbf{NP}_{\mathrm{nom,f,sg,3}} \multimap \mathbf{S}_{\mathrm{m},6}; \lambda_{z}.(\mathsf{love}\ z\ z) : \mathsf{e} \to \mathsf{p}$

At this point, the verb phrase can combine with the subject resulting in the following sign, whose semantic term guarantees the correct interpretation whereby the object and the subject of the verb are coreferential.

(213) $\mathbf{PER}\ (\mathtt{ANA}_{\mathbf{z}} \circ \mathtt{VOLI}_{\mathbf{z}} \circ \mathtt{SEBE}_{\mathbf{z}}) : \mathbf{Z}; \mathbf{S}_{\mathrm{m},6}; (\mathsf{love\ ana\ ana}) : \mathsf{p}$

Now we show an example of the reflexive occurring in subject controlled verb phrases. Suppose we are trying to construct a representation of *Ana hoće da voli sebe* 'Ana wants to love herself'. We first combine the reflexive with the verb as in the example above, and then introduce a subject trace.

(214)

$$\dfrac{\begin{array}{c}\vdash \lambda_w.\mathbf{PER}\ (w \circ \mathtt{VOLI_z} \circ \mathtt{SEBE_z}); \\ \mathbf{NP}_{\mathrm{nom,f,sg,3}} \multimap \mathbf{S}_{\mathrm{m,6}}; \lambda_z.(\mathsf{love}\ z\ z)\end{array} \qquad \dfrac{}{x; \mathbf{NP}_{\mathrm{nom,f,sg,3}}; x \vdash x; \mathbf{NP}_{\mathrm{nom,f,sg,3}}; x}\,[\mathrm{Ax}]}{x; \mathbf{NP}_{\mathrm{nom,f,sg,3}}; x \vdash \mathbf{PER}\ (x \circ \mathtt{VOLI_z} \circ \mathtt{SEBE_z}); \mathbf{S}_{\mathrm{m,6}}; (\mathsf{love}\ x\ x)}\,[\multimap\mathrm{E}]$$

We then proceed to construct an embedded clause with a subject gap, as we do for subject control sentences in general.

(215)

$$\dfrac{\dfrac{\begin{array}{c}\vdash \lambda_{Xw}.\exists_v[(X\ v) \wedge w = (\mathtt{DA_z} \circ v)]; \\ \mathbf{S}_{\mathrm{m,6}} \multimap \mathbf{S}_{\mathrm{e,6}}; \lambda_q.q\end{array} \qquad \begin{array}{c}x; \mathbf{NP}_{\mathrm{nom,f,sg,3}}; x \vdash \mathbf{PER}\ (x \circ \mathtt{VOLI_z} \circ \mathtt{SEBE_z}); \\ \mathbf{S}_{\mathrm{m,6}}; (\mathsf{love}\ x\ x)\end{array}}{x; \mathbf{NP}_{\mathrm{nom,f,sg,3}}; x \vdash \lambda_w.\exists_v[(\mathbf{PER}\ (x \circ \mathtt{VOLI_z} \circ \mathtt{SEBE_z})\ v) \wedge w = (\mathtt{DA_z} \circ v)]; \mathbf{S}_{\mathrm{e,6}}; (\mathsf{love}\ x\ x)}\,[\multimap\mathrm{E}]}{\vdash \lambda_{xw}.\exists_v[(\mathbf{PER}\ (x \circ \mathtt{VOLI_z} \circ \mathtt{SEBE_z})\ v) \wedge w = (\mathtt{DA_z} \circ v)]; \mathbf{NP}_{\mathrm{nom,f,sg,3}} \multimap \mathbf{S}_{\mathrm{e,6}}; \lambda_x.(\mathsf{love}\ x\ x)}\,[\multimap\mathrm{I}]$$

The conclusion of the proof above can be an argument of a subject control verb. Next, *hoće* 'wants' combines with this embedded clause with a subject gap (216), and then the whole verb phrase combines with the subject, resulting in (217).

(216) $\lambda_{xw}.\exists_{zyv}[(\mathbf{PER}\ (\mathtt{VOLI_z} \circ \mathtt{SEBE_z})\ v) \wedge z = (\mathtt{DA_z} \circ v) \wedge (\mathbf{PER}(x \circ \mathtt{HO\acute{C}E_z})\ y) \wedge$
$w = y \circ \mathbf{toZ}(\mathbf{L}\ v)] : \mathbf{z} \rightarrow \mathbf{Z};$
$\mathbf{NP}_{\mathrm{nom,f,sg,3}} \multimap \mathbf{S}_{\mathrm{m,6}}; \lambda_z.(\mathsf{want}\ (\mathsf{love}\ z\ z)\ z) : \mathbf{e} \rightarrow \mathbf{p}$

(217) $\lambda_w.\exists_{zyv}[(\mathbf{PER}\ (\mathtt{VOLI_z} \circ \mathtt{SEBE_z})\ v) \wedge z = (\mathtt{DA_z} \circ v) \wedge (\mathbf{PER}(\mathtt{ANA_z} \circ \mathtt{HO\acute{C}E_z})\ y)$
$\wedge w = y \circ \mathbf{toZ}(\mathbf{L}\ v)] : \mathbf{Z}; \mathbf{S}_{\mathrm{m,6}}; (\mathsf{want}\ (\mathsf{love\ ana\ ana})\ \mathsf{ana}) : \mathbf{p}$

Next, we look at an example where the reflexive occurs in the object controlled complement. Suppose we are trying to represent the sentence *Ana nagovara Marka* *da nominira sebe* 'Ana is persuading Marko to nominate himself'. We combine the reflexive with the embedded verb, then construct an embedded sentence with a subject gap, resulting in the following sign.

(218) $\vdash \lambda_{xw}.\exists_{v}[(\mathbf{PER}\ (x \circ \mathtt{NOMINIRA}_{\mathbf{z}} \circ \mathtt{SEBE}_{\mathbf{z}})\ v) \wedge w = (\mathtt{DA}_{\mathbf{z}} \circ v)] : \mathbf{z} \rightarrow \mathbf{Z};$
$\mathbf{NP}_{\mathrm{nom,f,sg,3}} \multimap \mathbf{S}_{\mathrm{e,6}}; \lambda_{x}.(\mathsf{nominate}\ x\ x) : \mathbf{e} \rightarrow \mathbf{p}$

Consider the lexical entry for *nagovara* 'persuades' given in the previous chapter.

(219) $\lambda_{zFvw}.\exists_{xy}[(F \in_{\mathbf{S}} x) \wedge (\mathbf{PER}(v \circ \mathtt{NAGOVARA}_{\mathbf{z}} \circ z)\ y) \wedge w = y \circ \mathbf{toZ}(\mathbf{L}\ x)] :$
$\mathbf{z} \rightarrow (\mathbf{z} \rightarrow \mathbf{Z}) \rightarrow \mathbf{z} \rightarrow \mathbf{Z};$
$\mathbf{NP}_{\mathrm{acc,\ m,sg,3}} \multimap (\mathbf{NP}_{\mathrm{nom,m,sg,3}} \multimap \mathbf{S}_{\mathrm{e,6}}) \multimap \mathbf{NP}_{\mathrm{nom,f,sg,3}} \multimap \mathbf{S}_{\mathrm{m,6}};$
$\lambda_{yFx}.(\mathsf{persuade}\ y\ (F\ y)\ x) : \mathbf{e} \rightarrow (\mathbf{e} \rightarrow \mathbf{p}) \rightarrow \mathbf{e} \rightarrow \mathbf{p}$

We combine *nagovara* 'persuades' with its object *Marko*:

(220) $\lambda_{Fvw}.\exists_{xy}[(F \in_{\mathbf{S}} x) \wedge (\mathbf{PER}(v \circ \mathtt{NAGOVARA}_{\mathbf{z}} \circ \mathtt{MARKA}_{\mathbf{z}})\ y) \wedge w = y \circ \mathbf{toZ}(\mathbf{L}\ x)] :$
$(\mathbf{z} \rightarrow \mathbf{Z}) \rightarrow \mathbf{z} \rightarrow \mathbf{Z};$
$(\mathbf{NP}_{\mathrm{nom,m,sg,3}} \multimap \mathbf{S}_{\mathrm{e,6}}) \multimap \mathbf{NP}_{\mathrm{nom,f,sg,3}} \multimap \mathbf{S}_{\mathrm{m,6}};$
$\lambda_{Fx}.(\mathsf{persuade\ marko}\ (F\ \mathsf{marko})\ x) : (\mathbf{e} \rightarrow \mathbf{p}) \rightarrow \mathbf{e} \rightarrow \mathbf{p}$

Now we can combine *nagovara Marka* 'persuades Marko' with the controlled verb phrase, resulting in the following sign:

(221) $\lambda_{vw}.\exists_{xyz}[(\mathbf{PER}\ (\mathtt{NOMINIRA}_{\mathbf{z}} \circ \mathtt{SEBE}_{\mathbf{z}})\ z) \wedge x = (\mathtt{DA}_{\mathbf{z}} \circ z) \wedge$
$(\mathbf{PER}(v \circ \mathtt{NAGOVARA}_{\mathbf{z}} \circ \mathtt{MARKA}_{\mathbf{z}})\ y) \wedge w = y \circ \mathbf{toZ}(\mathbf{L}\ x)] : \mathbf{z} \rightarrow \mathbf{Z};$
$\mathbf{NP}_{\mathrm{nom,f,sg,3}} \multimap \mathbf{S}_{\mathrm{m,6}};$
$\lambda_{x}.(\mathsf{persuade\ marko}\ (\mathsf{nominate\ marko\ marko})\ x) : \mathbf{e} \rightarrow \mathbf{p}$

Finally, combining this verb phrase with the matrix subject *Ana*, we get the following sign:

(222) $\lambda_{w}.\exists_{xyz}[(\mathbf{PER}\ (\mathtt{NOMINIRA}_{\mathbf{z}} \circ \mathtt{SEBE}_{\mathbf{z}})\ z) \wedge x = (\mathtt{DA}_{\mathbf{z}} \circ z) \wedge$
$(\mathbf{PER}(\mathtt{ANA}_{\mathbf{z}} \circ \mathtt{NAGOVARA}_{\mathbf{z}} \circ \mathtt{MARKA}_{\mathbf{z}})\ y) \wedge w = y \circ \mathbf{toZ}(\mathbf{L}\ x)] : \mathbf{Z};$
$\mathbf{S}_{\mathrm{m,6}}; (\mathsf{persuade\ marko}\ (\mathsf{nominate\ marko\ marko})\ \mathsf{ana}) : \mathbf{p}$

Now we show an example where the reflexive is the object controller. Consider the sentence *Marko sebe smatra pristojnim* 'Marko considers himself polite'. Recall that we analyzed *smatra* 'considers' as taking an accusative noun phrase and an instrumental attributive adjective, ensuring that the accusative object and the adjective have the same gender and number. Below we repeat the lexical entry for the adjective given in the previous chapter, as well as the required tectogrammatical version of the lexical entry for the verb.

(223) $\lambda_v.\mathtt{PRISTOJNIM}_{\mathbf{z}} \circ v : \mathbf{z} \to \mathbf{z}; \mathbf{N}_{\mathrm{inst,m,sg}} \multimap \mathbf{N}_{\mathrm{inst,m,sg}}; \mathsf{polite} : \mathbf{e} \to \mathbf{p}$

(224) $\vdash \lambda_{vFw}.\mathbf{PER}(w \circ \mathtt{SMATRA}_{\mathbf{z}} \circ v \circ (F\, \epsilon_{\mathbf{S}})) : \mathbf{z} \to (\mathbf{z} \to \mathbf{z}) \to \mathbf{z} \to \mathbf{Z};$
$\mathbf{NP}_{\mathrm{acc,m,sg,3}} \multimap (\mathbf{N}_{\mathrm{inst,m,sg}} \multimap \mathbf{N}_{\mathrm{inst,m,sg}}) \multimap \mathbf{NP}_{\mathrm{nom,m,sg,3}} \multimap \mathbf{S}_{\mathrm{m,6}};$
$\lambda_{xPy}.(\mathsf{consider}(P\, x)\; y) : \mathbf{e} \to (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$

We cannot combine the reflexive *sebe* directly with this verb because of the type mismatch. However, we can introduce an accusative trace and combine *smatra* 'considers' with it, then combine the resulting sign with the instrumental adjective, and finally bind the accusative trace which results in the following sign.[23]

(225) $\vdash \lambda_{vw}.\mathbf{PER}(w \circ \mathtt{SMATRA}_{\mathbf{z}} \circ v \circ \mathtt{PRISTOJNIM}_{\mathbf{z}}) : \mathbf{z} \to \mathbf{z} \to \mathbf{Z};$
$\mathbf{NP}_{\mathrm{acc,m,sg,3}} \multimap \mathbf{NP}_{\mathrm{nom,m,sg,3}} \multimap \mathbf{S}_{\mathrm{m,6}}; \lambda_{xy}.(\mathsf{consider}(\mathsf{polite}\; x)\; y) : \mathbf{e} \to \mathbf{e} \to \mathbf{p}$

The reflexive can combine with this verb phrase, resulting in the following sign:

(226) $\vdash \lambda_{w}.\mathbf{PER}(w \circ \mathtt{SMATRA}_{\mathbf{z}} \circ \mathtt{SEBE}_{\mathbf{z}} \circ \mathtt{PRISTOJNIM}_{\mathbf{z}}) : \mathbf{z} \to \mathbf{Z};$
$\mathbf{NP}_{\mathrm{nom,m,sg,3}} \multimap \mathbf{S}_{\mathrm{m,6}}; \lambda_{z}.(\mathsf{consider}(\mathsf{polite}\; z)\; z) : \mathbf{e} \to \mathbf{p}$

[23] In general, it is a theorem of both intuitionistic (a) and linear (b) logic, which constitute the type system of our phenogrammtical and semantic component, and tectogrammatical component respectively, that for any formulas ϕ, ψ and χ:

(1) a. $\phi \to \psi \to \chi \vdash \psi \to \phi \to \chi$
b. $\phi \multimap \psi \multimap \chi \vdash \psi \multimap \phi \multimap \chi$

Once this verb phrase combines with the subject *Marko*, we will get a sign associated with the semantic term is $\vdash (\mathsf{consider}\ (\mathsf{polite\ marko})\ \mathsf{marko}) : \mathbf{p}$, just as desired.

The Clitic Reflexive

Moving onto the clitic reflexive *se*, since the clitics in general cannot be placed in the sentence until all the potential hosts have been accounted for, the clitic reflexive *se* cannot be placed in the sentence until after the subject is already there. However, once the subject has already combined with the verb phrase, we can no longer ensure that the sentence is interpreted correctly, with the object and the subject being coreferential. Therefore, in order to correctly compose a sentence in which the clitic reflexive clitic *se* occurs, we need to introduce a complicated hypothesis into the derivation so that the correct interpretation can be arrived at and the clitic can find an appropriate host, which may be the subject.

While in the derivation above we considered a specific tectogrammatical version of the reflexive, the reader should keep in mind that its 'basic' tectogrammatical type, listed in the lexicon, is dependently typed since *sebe* does not care about the gender, person or number of the subject that it's coreferential with. We will use the following type abbreviations, for legibility purposes:

(227) a. $\mathbf{REFL} =_{def}$
$\prod_{g:\mathbf{Gdr},n:\mathbf{Num},p:\mathbf{Prs}}[(\mathbf{NP}_{\mathrm{acc},g,n,p} \multimap \mathbf{NP}_{\mathrm{nom},g,n,p} \multimap \mathbf{S}_{\mathrm{m},6}) \multimap \mathbf{NP}_{\mathrm{nom},g,n,p} \multimap \mathbf{S}_{\mathrm{m},6}]$

b. $\mathbf{refl} =_{def} (\mathbf{e} \rightarrow \mathbf{e} \rightarrow \mathbf{p}) \rightarrow \mathbf{e} \rightarrow \mathbf{p}$

c. $\mathtt{r} =_{def} (\mathtt{z} \rightarrow \mathtt{z} \rightarrow \mathtt{Z}) \rightarrow \mathtt{z} \rightarrow \mathtt{Z}$

Below is the reflexive hypothesis we'd need to introduce to get *Ana se voli* 'Ana loves herself'.

(228) $\quad G : \mathtt{r}; \mathbf{REFL}_{\mathrm{f,sg,3}}; G : \textsf{\textbf{refl}} \vdash G : \mathtt{r}; \mathbf{REFL}_{\mathrm{f,sg,3}}; G : \textsf{\textbf{refl}}$

Below is the lexical entry for the clitic reflexive *se*.

(229)
$$\vdash \lambda_{Pw}.\exists_{vt}[(P\ (\lambda_{Qx}.Q\ \mathtt{e}_{\mathbf{s}}\ \ x)\ v) \wedge (\mathtt{\mathbf{fst}}_{\mathbf{z}}\ v\ t) \wedge$$
$$w = \mathtt{\mathbf{toZ}}(\lambda_{s}.s = (\mathtt{\mathbf{snc}}_{\mathbf{s}}(\mathtt{\mathbf{lst}}_{\mathbf{s}}\ t)\#\mathtt{se}\ (\mathtt{\mathbf{tsr}}_{\mathbf{s}}\ t))) \circ (\mathtt{\mathbf{rst}}_{\mathbf{z}}\ v)] :$$
$$(\mathtt{r} \to \mathtt{Z}) \to \mathtt{Z};$$
$$(\mathbf{REFL}_{\mathrm{f,sg,3}} \multimap \mathbf{S}_{\mathrm{m},n>1}) \multimap \mathbf{S}_{\mathrm{m,1}}; \lambda_{P}.P(\lambda_{Qz}.(Q\ z\ z)) : (\textsf{\textbf{refl}} \to \textsf{\textbf{p}}) \to \textsf{\textbf{p}}$$

In the semantic term above, $\vdash P : \textsf{\textbf{refl}} \to \textsf{\textbf{p}}$, $\vdash Q : \textsf{\textbf{e}} \to \textsf{\textbf{e}} \to \textsf{\textbf{p}}$, and so $\vdash \lambda_{Qz}.(Q\ z\ z) : (\textsf{\textbf{e}} \to \textsf{\textbf{e}} \to \textsf{\textbf{p}}) \to \textsf{\textbf{e}} \to \textsf{\textbf{p}}$.

In the phenogrammatical term, $\vdash P : \mathtt{r} \to \mathtt{Z}$, $\vdash Q : \mathtt{z} \to \mathtt{z} \to \mathtt{Z}$, and so $\vdash \lambda_{Qx}.Q\ \mathtt{e}_{\mathbf{s}}\ x : (\mathtt{z} \to \mathtt{z} \to \mathtt{Z}) \to \mathtt{z} \to \mathtt{Z}$. Suppose we are trying to generate the sentence *Ana se voli* 'Ana loves herself'. We introduce the reflexive hypothesis and combine the verb *voli* with it, resulting in the following sign.

(230)
$$G : \mathtt{r}; \mathbf{REFL}_{\mathrm{f,sg,3}}; G : \textsf{\textbf{refl}} \vdash G\ (\lambda_{xy}.\mathtt{\mathbf{PER}}(y \circ \mathtt{VOLI}_{\mathbf{z}} \circ x)) : \mathtt{z} \to \mathtt{Z};$$
$$\mathbf{NP}_{\mathrm{nom,f,sg,3}} \multimap \mathbf{S}_{\mathrm{m,6}}; G\ (\lambda_{xy}.\textsf{love}\ x\ y) : \textsf{\textbf{e}} \to \textsf{\textbf{p}}$$

Now this verb phrase combines with the subject, and then the reflexive hypothesis is withdrawn:

(231)
$$G : \mathtt{r}; \mathbf{REFL}_{\mathrm{f,sg,3}}; G : \textsf{\textbf{refl}} \vdash (G\ (\lambda_{xy}.\mathtt{\mathbf{PER}}(y \circ \mathtt{VOLI}_{\mathbf{z}} \circ x))\ \mathtt{ANA}_{\mathbf{z}}) : \mathtt{Z}; \mathbf{S}_{\mathrm{m,6}};$$
$$(G\ (\lambda_{xy}.\textsf{love}\ x\ y)\ \textsf{ana}) : \textsf{\textbf{p}}$$

(232)
$$\vdash \lambda_{G}.(G\ (\lambda_{xy}.\mathtt{\mathbf{PER}}(y \circ \mathtt{VOLI}_{\mathbf{z}} \circ x))\ \mathtt{ANA}_{\mathbf{z}}) : \mathtt{r} \to \mathtt{Z}; \mathbf{REFL}_{\mathrm{f,sg,3}} \multimap \mathbf{S}_{\mathrm{m,6}};$$
$$\lambda_{G}.(G\ (\lambda_{xy}.\textsf{love}\ x\ y)\ \textsf{ana}) : \textsf{\textbf{refl}} \to \textsf{\textbf{p}}$$

The clitic reflexive can combine with the sentence with a withdrawn reflexive hypothesis. The result, tectogrammatically, is $\mathbf{S}_{\mathrm{m},1}$. Below we show step-by-step reductions of the resulting phenogrammatical and semantic term.

(233) $\lambda_{Pw}.\exists_{vt}[(P\,(\lambda_{Qz}.Q\ \mathtt{e}_{\mathbf{s}}\ \ z)\ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\mathtt{se}$
$(\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)](\lambda_{G}.(G\,(\lambda_{xy}.\mathbf{PER}(y \circ \mathtt{VOLI}_{\mathbf{z}} \circ x))\ \mathtt{ANA}_{\mathbf{z}}))$

$\rightsquigarrow \lambda_{w}.\exists_{vt}[((\lambda_{G}.(G\,(\lambda_{xy}.\mathbf{PER}(y \circ \mathtt{VOLI}_{\mathbf{z}} \circ z))\ \mathtt{ANA}_{\mathbf{z}}))\ (\lambda_{Qz}.Q\ \mathtt{e}_{\mathbf{s}}\ \ z)\ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$
$w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\mathtt{se}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)]$

$\rightsquigarrow \lambda_{w}.\exists_{vt}[((\lambda_{Qz}.Q\ \ \mathtt{e}_{\mathbf{s}}\ \ z)\ (\lambda_{xy}.\mathbf{PER}(y \circ \mathtt{VOLI}_{\mathbf{z}} \circ x)\ \mathtt{ANA}_{\mathbf{z}})\ \ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge w =$
$\mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\mathtt{se}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)]$

$\rightsquigarrow \lambda_{w}.\exists_{vt}[((\lambda_{z}.(\lambda_{xy}.\mathbf{PER}(y \circ \mathtt{VOLI}_{\mathbf{z}} \circ x)\ \mathtt{e}_{\mathbf{s}}\ \ z)\ \ \mathtt{ANA}_{\mathbf{z}})\ \ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge w =$
$\mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\mathtt{se}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)]$

$\rightsquigarrow \lambda_{w}.\exists_{vt}[((\lambda_{z}.(\lambda_{y}.\mathbf{PER}(y \circ \mathtt{VOLI}_{\mathbf{z}})\ z)\ \ \mathtt{ANA}_{\mathbf{z}})\ \ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$
$w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\mathtt{se}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)]$

$\rightsquigarrow \lambda_{w}.\exists_{vt}[((\lambda_{z}.(\mathbf{PER}(z \circ \mathtt{VOLI}_{\mathbf{z}})\ \mathtt{ANA}_{\mathbf{z}})\ \ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$
$w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\mathtt{se}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)]$

$\rightsquigarrow \lambda_{w}.\exists_{vt}[(\mathbf{PER}(\mathtt{ANA}_{\mathbf{z}} \circ \mathtt{VOLI}_{\mathbf{z}})\ \ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$
$w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\mathtt{se}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)]$

(234) $\lambda_{P}.P(\lambda_{Qz}.(Q\ z\ z))(\lambda_{G}.(G\ (\lambda_{xy}.\mathsf{love}\ x\ y)\ \mathsf{ana}))$
$\rightsquigarrow \lambda_{G}.(G\ (\lambda_{xy}.\mathsf{love}\ x\ y)\ \mathsf{ana})(\lambda_{Qz}.(Q\ z\ z))$
$\rightsquigarrow \lambda_{Qz}.(Q\ z\ z)\ (\lambda_{xy}.\mathsf{love}\ x\ y)\ \mathsf{ana}$
$\rightsquigarrow \lambda_{z}.((\lambda_{xy}.\mathsf{love}\ x\ y)\ z\ z)\mathsf{ana}$
$\rightsquigarrow \lambda_{z}.(\mathsf{love}\ z\ z)\mathsf{ana}$
$\rightsquigarrow (\mathsf{love\ ana\ ana})$

5.3.5 Auxiliaries

There are three types of verbal clitics in Serbo-Croatian. The clitics of *htjeti* 'want, will' are mainly used to form periphrastic future tense. In the periphrastic future tense construction, they are subject control verbs, combining with either an infinitival verb phrase or an embedded declarative clause missing a nominative argument.

Another set of verbal clitics are the present tense clitics of *biti* 'be' which take a variety of predicative complements, including adjectives, noun phrases, passive participles, prepositional phrases, and, finally, past participles in the periphrastic past tense construction (see Chapter 4). Finally, the aorist clitics of *biti* 'be' are used to form the conditional mood and require a past participle complement.

The examples below show various types of sentences containing verbal clitics.

(235) clitics of *htjeti*

a. Ja ću pivo.
$\text{I}_{NOM,sg,1}$ $\text{will}_{sg,1}$ $\text{beer}_{ACC,n,sg}$
'I want a beer'

b. Ana će doći.
$\text{Ana}_{NOM,f,sg,3}$ $\text{will}_{sg,3}$ come_{inf}
'Ana will come'

c. Ana će da dođe.
$\text{Ana}_{NOM,f,sg,3}$ $\text{will}_{sg,3}$ DA $\text{come}_{sg,3}$
'Ana will come'

(236) present tense clitics of *biti*

a. Oni su studenti.
$\text{they}_{NOM,pl,3}$ $\text{are}_{pl,3}$ $\text{students}_{NOM,m,pl,3}$
'They are students'

b. Marko je pametan.
Marko$_{NOM,m,sg,3}$ is$_{sg,3}$ smart$_{NOM,m,sg}$
'Marko is smart'

c. Mi smo iz Sarajeva.
we$_{NOM,pl,1}$ are$_{pl,1}$ from Sarajevo$_{GEN,n,sg}$
'We are from Sarajevo'

d. Knjiga je pročitana.
book$_{NOM,f,sg}$ is$_{f,sg}$ read$_{pas,NOM,f,sg}$
'A/the book is/has been read'

e. Ti si došao.
you$_{NOM,sg,2}$ are$_{sg,2}$ arrived$_{ppl,m,sg}$
'You arrived'

(237) aorist clitics of *biti*

a. Ana bi htjela kupiti auto.
Ana$_{NOM,f,sg,3}$ would$_{sg,3}$ want$_{ppl,f,sg}$ buy$_{inf}$ car$_{ACC,m,sg}$
'Ana would like to buy a car'

b. Vi biste trebali više učiti.
you$_{NOM,pl,2}$ would$_{pl,2}$ need$_{ppl,m,pl}$ more study$_{inf}$
'You should study more'

All verbal clitics occur in the first slot in the clitic cluster, except for *je*, the third person present tense clitic of *biti*. We first analyze the slot 1 verbal clitics, and then return to *je*.

Slot 1 Auxiliaries

First we consider slot 1 auxiliaries in main declarative clauses. Precisely because they occur in the first slot in the clitic cluster, we can analyze them as taking an appropriate complement, and then the subject, and encliticizing onto the first word in the resulting string. This would be impossible if they occurred further to the right in the clitic cluster.

For the version of *će* which takes an infinitival verb phrase complement, we give the following lexical entry:

(238) $\vdash \lambda_{xyw}.\exists_{vt}[(\mathbf{PER}(y \circ x)\ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$
$w = \mathbf{toZ}(\lambda_s.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\text{će}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)] : \mathbf{z} \to \mathbf{z} \to \mathbf{Z};$
$\prod_{g:\mathbf{Gdr}}[(\mathbf{NP}_{\text{nom},g,\text{sg},3} \multimap \mathbf{S}_{\text{inf},6}) \multimap \mathbf{NP}_{\text{nom},g,\text{sg},3} \multimap \mathbf{S}_{\text{m},5}];$
$\lambda_{Fx}.(\text{FUT}\ (F\ x) : (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$

It is almost identical to its non-clitic counterpart tectogrammatically, except that its ultimate result type has the parameter 5, instead of 6. Phenogrammatically, it lets its complement infinitival verb phrase, and the subject which it eventually picks up, freely permute, and then it finds some host to encliticize onto.

The entire sentence *Maja će pročitati knjigu* 'Maja will read a book' is represented by the following sign:

(239) $\vdash \lambda_w.\exists_{vt}[(\mathbf{PER}(\text{MAJA}_\mathbf{z} \circ \text{PROČITATI}_\mathbf{z} \circ \text{KNJIGU}_\mathbf{z})\ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$
$w = \mathbf{toZ}(\lambda_s.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\text{će}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)] : \mathbf{Z}; \mathbf{S}_{\text{m},5};$
$((\text{exists book})\lambda_x.\text{FUT}\ (\text{read maja}\ x)) : \mathbf{p}$

The phenogrammatical term of this sign denotes a set of six strings of languages, with the clitic encliticized onto the last phonological word inside of the very first length-one string of languages. So we correctly predict that all of the following are possible:

(240) a. Maja će pročitati knjigu.
Maja$_{NOM,f,sg,3}$ will$_{sg,3}$ read$_{inf}$ book$_{ACC,f,sg}$
'Maja will read a/the book'

b. Maja će knjigu pročitati.
c. Knjigu će pročitati Maja.
d. Knjigu će Maja pročitati.
e. Pročitati će knjigu Maja.
f. Pročitati će Maja knjigu.

Two remarks are in order. First, in examples (e) and (f) above, in reality there are serious phonetic consequences when the infinitive host and the clitic get smushed into a single word, partly reflected in orthography. In Croatia, when the infinitive hosts a *htjeti* clitic, it is typically written as *Pročitat će Maja knjigu*, while elsewhere the standard is to write *Pročitaće Maja knjigu*. Either way, one syllable is lost, and depending on the coda of the penultimate syllable of the infinitive, even more dramatic changes occur. Here we just note this fact, but cannot adequately represent it in our framework.

Second, while in Chapter 4 we entertained a more restrictive generalization concerning infinitival verb phrases, whereby they had to remain contiguous and rightmost in the sentence, we are pretty confident that no native speakers would reject any of the examples above. For example, consider the fact that there are writing standards for the clitic being hosted by the infinitive which is but a part of the clitic's complement verb phrase. It seems that in general, at least when it comes to clitic placement, any constituent of the infinitival verb phrase can be the clitic host. Therefore, we claim that the lexical entry above in essentially universal.

We must give another lexical entry for *htjeti* which combines not with an infinitive but with an embedded clause with a nominative noun phrase gap to construct a main clause in future tense. Recall that embedded clauses generally have to remain contiguous and occur on the right edge of the matrix clause. This essentially leave only the subject, or some word in the noun phrase subject, as the possible host. We give it the following lexical entry:

(241) $\lambda_{Fvw}.\exists_{yt}[(\mathbf{PER}\,v\,y) \wedge (\mathbf{fst_z}\,y\,t) \wedge w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\,t)\#\text{će}$
$(\mathbf{tsr_s}\,t))) \circ (\mathbf{rst_z}\,y) \circ \mathbf{toZ}(\mathbf{k}\,(F \in_{\mathbf{s}} x))] : (\mathbf{z} \to \mathbf{Z}) \to \mathbf{z} \to \mathbf{Z};$
$\prod_{g:\mathbf{Gdr},n:\mathbf{Num}}[(\mathbf{NP}_{\text{nom},g,\text{sg},3} \multimap \mathbf{S}_{\text{e},n}) \multimap \mathbf{NP}_{\text{nom},g,\text{sg},3} \multimap \mathbf{S}_{\text{m},5}];$
$\lambda_{Fx}.\text{FUT}\,(F\,x) : (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$

Phenogrammatically, it combines with the embedded clause missing a subject. Then, it combines with the matrix subject and permutes it, then encliticizes onto the first word of the subject . If this matrix subject consists of an adjective and a noun, this means that either the noun or the adjective could be the clitic host. Of course, in a more restrictive grammar which doesn't allow noun phrase discontinuities in general, the permutation of the subject would have no effect. So the exact predictions of this lexical entry depend on how permissive one's grammar is with respect to noun phrases. Either way, however, the subject noun phrase, together with the attached clitic, must then precede the embedded clause complement. The meaning we construct is the same as in the case of an infinitival controlled complement.

The aorist clitics of *biti* can combine with past participles only. However, in combination with these clitics, they do not express past meaning at all. In our analysis of predicatives in the previous chapter, we assumed that past participles are a kind of predicative complement, and that they carry the past tense meaning. For example, we gave the following lexical entry for *spavao* 'slept'.

(242) $\vdash \mathtt{SPAVAO}_{\mathbf{z}} : \mathbf{z}; \mathbf{NP}_{\mathrm{nom,m,sg,3}} \multimap \mathbf{Prd}_{\mathrm{pl}}; \lambda_x.\mathsf{PST}(\mathsf{sleep}\ x) : \mathbf{e} \rightarrow \mathbf{p}$

To distinguish between conditional mood and past tense, we assume that participles are associated with lexical entries like the following one:

(243) $\vdash \mathtt{SPAVAO}_{\mathbf{z}} : \mathbf{z}; \mathbf{NP}_{\mathrm{nom,m,sg,3}} \multimap \mathbf{S}_{\mathrm{m,6}}; \lambda_x.(\mathsf{sleep}\ x) : \mathbf{e} \rightarrow \mathbf{p}$

Note that this sign has an ordinary declarative sentence as its result type, and that it's not associated with past meaning. Because of its phenogrammatical term and type, however, there is no danger of this participle combining with a subject and producing a non-sentence **Marko spavao*. This sign is a possible argument of an aorist clitic of *biti* building a sentence in conditional mood. We then give the following rule schema which converts conditional forming participles into past tense forming ones:

(244)

$$\frac{\vdash \phi : \mathbf{z}; \mathbf{NP}_{\tau,\tau',\tau'',\tau'''} \multimap \mathbf{S}_{\mathrm{m,6}}; \sigma : \mathbf{e} \rightarrow \mathbf{p}}{\vdash \phi : \mathbf{z}; \mathbf{NP}_{\tau,\tau',\tau'',\tau'''} \multimap \mathbf{Prd}_{\mathrm{pl}}; \lambda_x.\mathsf{PST}(\sigma\ x) : \mathbf{e} \rightarrow \mathbf{p}}\ [\mathrm{CPL}]$$

Consider the sentence *Maja bi kupila auto* 'Maja would buy a car'. We give the following lexical entry for *bi*, where W is some unanalyzed propositional operator:

(245) $\vdash \lambda_{xyw}.\exists_{vt}[(\mathbf{PER}(y \circ x)\ v) \wedge (\mathbf{fst}_{\mathbf{z}}\ v\ t) \wedge$
$w = \mathbf{toZ}(\lambda_s.s = (\mathbf{snc}_{\mathbf{s}}(\mathbf{lst}_{\mathbf{s}}\ t)\#\mathtt{bi}\ (\mathbf{tsr}_{\mathbf{s}}\ t))) \circ (\mathbf{rst}_{\mathbf{z}}\ v)] : \mathbf{z} \rightarrow \mathbf{z} \rightarrow \mathbf{Z};$
$\prod_{p:\mathbf{Prs}}[(\mathbf{NP}_{\mathrm{nom,m,sg},p} \multimap \mathbf{S}_{\mathrm{m,6}}) \multimap \mathbf{NP}_{\mathrm{nom,m,sg},p} \multimap \mathbf{S}_{\mathrm{m,5}}];$
$\lambda_{Px}.\mathsf{W}(P\ x) : (\mathbf{e} \rightarrow \mathbf{p}) \rightarrow \mathbf{e} \rightarrow \mathbf{p}$

The entire sentence is then represented by the following sign in the grammar:

(246) $$\vdash \lambda_{w}.\exists_{vt}[(\mathbf{PER}(\mathtt{MAJA}_{\mathbf{z}} \circ \mathtt{KUPILA}_{\mathbf{z}} \circ \mathtt{AUTO}_{\mathbf{z}})\ v) \wedge (\mathbf{fst}_{\mathbf{z}}\ v\ t) \wedge$$
$$w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc}_{\mathbf{s}}(\mathbf{lst}_{\mathbf{s}}\ t)\#\mathtt{bi}\ (\mathbf{tsr}_{\mathbf{s}}\ t))) \circ (\mathbf{rst}_{\mathbf{z}}\ v)] : \mathbf{Z}; \mathbf{S}_{\mathrm{m},5};$$
$$(\mathsf{exists\ car})(\lambda_{x}.\mathsf{W}(\mathsf{buy}\ x\ \mathsf{maja})) : \mathbf{p}$$

We predict that the sentence is pronounceable six different ways, corresponding to the six different orders of the participles, the subject and the object, with the clitic *bi* attached to the first word in each case.

Finally, we give a lexical entry for a representative of the clitic version of the copula. Recall from the previous chapter that non-aorist finite forms of the verb *biti* have the tectogrammatical type is $\prod_{g:\mathbf{Gdr},p:\mathbf{D}}[(\mathbf{NP}_{\mathrm{nom},g,\tau,\tau'} \multimap \mathbf{Prd}_{d}) \multimap \mathbf{NP}_{\mathrm{nom},g,\tau,\tau'} \multimap \mathbf{S}_{\mathrm{m},6}]$, where terms of type $\mathbf{D}$ specify the kind of the predicative phrase in question.

Excluding *je*, we can analyze present tense *biti* clitics on a par with other verbal clitics, whereby they combine with the predicative complement and then the subject, instead of being placed inside a sentence with an appropriate gap, as with pronominal clitics. We give the following lexical entry for *si*:

(247) $$\lambda_{xyw}.\exists_{vt}[(\mathbf{PER}(y \circ x)\ v) \wedge (\mathbf{fst}_{\mathbf{z}}\ v\ t) \wedge$$
$$w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc}_{\mathbf{s}}(\mathbf{lst}_{\mathbf{s}}\ t)\#\mathtt{si}\ (\mathbf{tsr}_{\mathbf{s}}\ t))) \circ (\mathbf{rst}_{\mathbf{z}}\ v)] : \mathbf{z} \to \mathbf{z} \to \mathbf{Z};$$
$$\prod_{g:\mathbf{Gdr},d:\mathbf{D}}[(\mathbf{NP}_{\mathrm{nom},g,\mathrm{sg},2} \multimap \mathbf{Prd}_{d}) \multimap \mathbf{NP}_{\mathrm{nom},g,\mathrm{sg},2} \multimap \mathbf{S}_{\mathrm{m},5}]$$
$$\lambda_{Px}.(P\ x) : (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$$

However, slot 1 clitics can only be analyzed as combining with an appropriate complement, and then the subject, in main declarative clauses. This analysis doesn't generalize to embedded declarative clauses or interrogative clauses.

When we analyze slot 1 clitics as combining with a complement and then the subject, we predict that they encliticize onto some phonological word contained in the complement or the subject. However, in embedded clauses the clitics obligatorily encliticize onto the complementizer *da*, in polar interrogatives onto the interrogative complementizer *da li*, and in constituent questions onto the fronted *wh* expression. The analysis of these verbal clitics presented above, however, predicts that the clitics do not encliticize onto the complementizers or the fronted *wh* expression, but occur further to the right of that initial element.

Here we focus on extending the analysis to the placement of enclitics in embedded declarative clauses. This analysis will give correct predictions for interrogatives as well, but we return to that issue in the next chapter.

The first thing that needs to be ensured is that the complementizer combines with a sentence before any clitics have been placed inside of it. This restriction is already built into our grammar because we associated the complementizer with the tectogrammatical type $\mathbf{S}_{\mathrm{m},6} \multimap \mathbf{S}_{\mathrm{e},6}$. The complementizer *da* ensures that its complement sentence has no enclitics already placed inside of by requiring that the number parameter of its argument and result type is 6.

The second issue is guaranteeing that the complementizer itself occurs leftmost in the embedded clause and that it is treated as a length one string of languages, because in our theory enclitics encliticize onto the last phonological word in the clause-initial length one string of languages. This is also already built into our grammar because the complementizer builds phenogrammatical terms consisting of the length one string of languages constructed out of *da* followed by the length one string of languages constructed out its complement clause.

Finally, to ensure that the slot 1 verbal clitics encliticize onto the complementizer, we must analyze them on a par with pronominal clitics, as combining with sentences with an appropriate gap.

We introduce the following type abbreviations:

(248) tectogrammatical type abbreviations

a. $\mathbf{IR} =_{def} \prod_{g:\mathbf{Gdr}}[(\mathbf{NP}_{\mathrm{nom},g,\tau,\tau'} \multimap \mathbf{S}_{\mathrm{inf},6}) \multimap \mathbf{NP}_{\mathrm{nom},g,\tau,\tau'} \multimap \mathbf{S}_{\mathrm{m},6}]$

b. $\mathbf{FR} =_{def} \prod_{n:\mathbf{N},g:\mathbf{Gdr}}[(\mathbf{NP}_{\mathrm{nom},g,\tau,\tau'} \multimap \mathbf{S}_{\mathrm{e},n}) \multimap \mathbf{NP}_{\mathrm{nom},g,\tau,\tau'} \multimap \mathbf{S}_{\mathrm{m},6}]$

c. $\mathbf{PR} =_{def} \prod_{d:\mathbf{D},g:\mathbf{Gdr}}[(\mathbf{NP}_{\mathrm{nom},g,\tau,\tau'} \multimap \mathbf{Prd}_{d}) \multimap \mathbf{NP}_{\mathrm{nom},g,\tau,\tau'} \multimap \mathbf{S}_{\mathrm{m},6}]$

(249) semantic type abbreviations

a. $\mathbf{pr} =_{def} (\mathbf{e} \to \mathbf{p}) \to \mathbf{e} \to \mathbf{p}$

Recall that *htjeti* 'will, want' clitics can combine with either an infinitival or a finite complement. For the infinitival-taking version of these clitics, the following hypothesis has to be introduced into the derivation:

(250) $G : \mathbf{z} \to \mathbf{z} \to \mathbf{Z}; \mathbf{IR}_{\tau:\mathbf{Gdr}}; G : \mathbf{pr} \vdash G : \mathbf{z} \to \mathbf{z} \to \mathbf{Z}; \mathbf{IR}_{\tau:\mathbf{Gdr}}; G : \mathbf{pr}$

The tectogrammatical type of the hypothesis has to be specified for particular person and number values, restricting the choice of a possible enclitic of *htjeti*. We give the following lexical entry for the 3rd person singular *će* which takes an infinitival complement.

(251) $\vdash \lambda_{Fw}.\exists_{vt}[(F\,(\lambda_{xy}.\mathtt{PER}(x \circ y))\ v) \wedge (\mathtt{fst_z}\ v\ t) \wedge$
$w = \mathtt{toZ}(\lambda_{s}.s = (\mathtt{snc_s}(\mathtt{lst_s}\ t)\#\text{će}\ (\mathtt{tsr_s}\ t))) \circ (\mathtt{rst_z}\ v)] :$

$$((\mathtt{z} \to \mathtt{z} \to \mathtt{Z}) \to \mathtt{Z}) \to \mathtt{Z};$$
$$\textstyle\prod_{k:\mathbf{K}}[(\mathbf{IR} \multimap \mathbf{S}_{k,6}) \multimap \mathbf{S}_{k,5}]; \lambda_{F}.(F\,(\lambda_{Px}.(\mathsf{FUT}\ P\ x))) : (\mathbf{pr} \to \mathbf{p}) \to \mathbf{p}$$

To construct the sentence *Ana će doći* 'Ana will come', we have to introduce an appropriate hypothesis, then construct a sentence as if it contained a subject control verb , although it doesn't. Then, we withdraw the subject control verb hypothesis and create something that *će* can combine with:

(252) $\vdash \lambda_{G}.(G\ \mathtt{DOĆI}_{\mathtt{z}}\ \mathtt{ANA}_{\mathtt{z}}) : (\mathtt{z} \to \mathtt{z} \to \mathtt{Z}) \to \mathtt{Z}; \mathbf{IR}_{\mathrm{f}} \multimap \mathbf{S}_{\mathrm{m},6}; \lambda_{G}.(G\ \mathsf{arrive}\ \mathsf{ana}) : \mathbf{pr} \to \mathbf{p}$

Once *će* combines with the sign below, tectogrammatically we get a main declarative clause whose number parameter is 5, i.e. the first enclitic slot has been filled. The resulting phenogrammatical term is somewhat complex; below we show all the reduction steps:

(253)
$$\lambda_{Fw}.\exists_{vt}[(F\ (\lambda_{xy}.\mathtt{PER}(x \circ y))\ v) \wedge (\mathtt{fst}_{\mathtt{z}}\ v\ t) \wedge$$
$$w = \mathtt{toZ}(\lambda_{s}.s = (\mathtt{snc}_{\mathtt{s}}(\mathtt{lst}_{\mathtt{s}}\ t)\#\text{će}\ (\mathtt{tsr}_{\mathtt{s}}\ t))) \circ (\mathtt{rst}_{\mathtt{z}}\ v)]$$
$$(\lambda_{G}.(G\ \mathtt{DOĆI}_{\mathtt{z}}\ \mathtt{ANA}_{\mathtt{z}}))$$

$$\rightsquigarrow \lambda_{w}.\exists_{vt}[(\lambda_{G}.(G\ \mathtt{DOĆI}_{\mathtt{z}}\ \mathtt{ANA}_{\mathtt{z}})(\lambda_{xy}.\mathtt{PER}(x \circ y))\ v) \wedge (\mathtt{fst}_{\mathtt{z}}\ v\ t) \wedge$$
$$w = \mathtt{toZ}(\lambda_{s}.s = (\mathtt{snc}_{\mathtt{s}}(\mathtt{lst}_{\mathtt{s}}\ t)\#\text{će}\ (\mathtt{tsr}_{\mathtt{s}}\ t))) \circ (\mathtt{rst}_{\mathtt{z}}\ v)]$$

$$\rightsquigarrow \lambda_{w}.\exists_{vt}[((\lambda_{xy}.\mathtt{PER}(x \circ y)\ \mathtt{DOĆI}_{\mathtt{z}}\ \mathtt{ANA}_{\mathtt{z}})\ v) \wedge (\mathtt{fst}_{\mathtt{z}}\ v\ t) \wedge$$
$$w = \mathtt{toZ}(\lambda_{s}.s = (\mathtt{snc}_{\mathtt{s}}(\mathtt{lst}_{\mathtt{s}}\ t)\#\text{će}\ (\mathtt{tsr}_{\mathtt{s}}\ t))) \circ (\mathtt{rst}_{\mathtt{z}}\ v)]$$

$$\rightsquigarrow \lambda_{w}.\exists_{vt}[(\mathtt{PER}(\mathtt{DOĆI}_{\mathtt{z}} \circ \mathtt{ANA}_{\mathtt{z}})\ v) \wedge (\mathtt{fst}_{\mathtt{z}}\ v\ t) \wedge$$
$$w = \mathtt{toZ}(\lambda_{s}.s = (\mathtt{snc}_{\mathtt{s}}(\mathtt{lst}_{\mathtt{s}}\ t)\#\text{će}\ (\mathtt{tsr}_{\mathtt{s}}\ t))) \circ (\mathtt{rst}_{\mathtt{z}}\ v)]$$

The resulting term denotes a set of two strings of languages, where either the clitic host is *Ana* or *doći*. If we had been placing the same clitic inside the embedded clause *da će Ana doći* 'that Ana will arrive', the sentence missing a subject control verb would combine with the complementizer, and then the hypothesis would be withdrawn, resulting in the following sign:

(254) $\vdash \lambda_{Gw}.\exists_{v}[(G\,\mathtt{DO\acute{C}I}_{\mathbf{z}}\,\mathtt{ANA}_{\mathbf{z}}\,v) \wedge w = (\mathtt{DA}_{\mathbf{z}} \circ v)] : (\mathbf{z} \rightarrow \mathbf{z} \rightarrow \mathbf{Z}) \rightarrow \mathbf{Z};$
$\mathbf{IR}_{\mathrm{f}} \multimap \mathbf{S}_{\mathrm{e},6}; \lambda_{G}.(G\ \mathsf{arrive\ ana}) : \mathbf{pr} \rightarrow \mathbf{p}$

After *će*, with the appropriately instantiated *k* parameter, combines with this sign, we wind up with an embedded clause whose number parameter is 5. Below we show the reduction of the resulting phenogrammatical term:

(255)
$$\lambda_{Fw}.\exists_{vt}[(F\,(\lambda_{xy}.\mathbf{PER}(x \circ y))\ v) \wedge (\mathbf{fst}_{\mathbf{z}}\ v\ t) \wedge$$
$$w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc}_{\mathbf{s}}(\mathbf{lst}_{\mathbf{s}}\ t)\#\acute{c}\mathtt{e}(\mathbf{tsr}_{\mathbf{s}}\ t))) \circ (\mathbf{rst}_{\mathbf{z}}\ v)]$$
$$(\lambda_{Gw'}.\exists_{v'}[(G\,\mathtt{DO\acute{C}I}_{\mathbf{z}}\,\mathtt{ANA}_{\mathbf{z}}\,v') \wedge w' = (\mathtt{DA}_{\mathbf{z}} \circ v')])$$

$$\leadsto \lambda_{w}.\exists_{vt}[(\lambda_{Gw'}.\exists_{v'}[(G\ \mathtt{DO\acute{C}I}_{\mathbf{z}}\ \mathtt{ANA}_{\mathbf{z}}\ v') \wedge w' = (\mathtt{DA}_{\mathbf{z}} \circ v')](\lambda_{xy}.\mathbf{PER}(x \circ y))\ v) \wedge (\mathbf{fst}_{\mathbf{z}}\ v\ t) \wedge w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc}_{\mathbf{s}}(\mathbf{lst}_{\mathbf{s}}\ t)\#\acute{c}\mathtt{e}\ (\mathbf{tsr}_{\mathbf{s}}\ t))) \circ (\mathbf{rst}_{\mathbf{z}}\ v)]$$

$$\leadsto \lambda_{w}.\exists_{vt}[(\lambda_{w'}.\exists_{v'}[((\lambda_{xy}.\mathbf{PER}(x \circ y))\mathtt{DO\acute{C}I}_{\mathbf{z}}\,\mathtt{ANA}_{\mathbf{z}}\,v') \wedge w' = (\mathtt{DA}_{\mathbf{z}} \circ v')]\ v) \wedge (\mathbf{fst}_{\mathbf{z}}\ v\ t)$$
$$\wedge w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc}_{\mathbf{s}}(\mathbf{lst}_{\mathbf{s}}\ t)\#\acute{c}\mathtt{e}\ (\mathbf{tsr}_{\mathbf{s}}\ t))) \circ (\mathbf{rst}_{\mathbf{z}}\ v)]$$

$$\leadsto \lambda_{w}.\exists_{vt}[(\lambda_{w'}.\exists_{v'}[(\mathbf{PER}(\mathtt{DO\acute{C}I}_{\mathbf{z}} \circ \mathtt{ANA}_{\mathbf{z}})\ \ v') \wedge w' = (\mathtt{DA}_{\mathbf{z}} \circ v')]\ v)$$
$$\wedge(\mathbf{fst}_{\mathbf{z}}\ v\ t) \wedge w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc}_{\mathbf{s}}(\mathbf{lst}_{\mathbf{s}}\ t)\#\acute{c}\mathtt{e}\ (\mathbf{tsr}_{\mathbf{s}}\ t))) \circ (\mathbf{rst}_{\mathbf{z}}\ v)]$$

$$\leadsto \lambda_{w}.\exists_{vv't}[(\mathbf{PER}(\mathtt{DO\acute{C}I}_{\mathbf{z}} \circ \mathtt{ANA}_{\mathbf{z}})\ \ v') \wedge v = (\mathtt{DA}_{\mathbf{z}} \circ v') \wedge (\mathbf{fst}_{\mathbf{z}}\ v\ t)$$
$$\wedge w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc}_{\mathbf{s}}(\mathbf{lst}_{\mathbf{s}}\ t)\#\acute{c}\mathtt{e}\ (\mathbf{tsr}_{\mathbf{s}}\ t))) \circ (\mathbf{rst}_{\mathbf{z}}\ v)]$$

This term also denotes a set of two strings of languages, where the first string is required to be constructed out of *da* onto which the auxiliary encliticized, followed by some permutation of the subject and the infinitive.

To construct a sentence in which there is a finite subject controlled verb phrase, such as *Ana će da vidi Marka* 'Ana will see Marko' we have to introduce the following hypothesis:

(256) $H : (\mathbf{z} \to \mathbf{Z}) \to (\mathbf{z} \to \mathbf{Z}); \mathbf{FR}_{6,\mathrm{f}}; G : \mathbf{pr} \vdash H : (\mathbf{z} \to \mathbf{Z}) \to (\mathbf{z} \to \mathbf{Z}); \mathbf{FR}_{6,\mathrm{f}}; G : \mathbf{pr}$

Eventually, the hypothesis is withdrawn, resulting in the following sign:

(257) $\vdash \lambda_H.(H\ (\lambda_{xw}.\exists_v[(\mathbf{PER}(x \circ \mathtt{VIDI}_{\mathbf{z}} \circ \mathtt{MARKA}_{\mathbf{z}})\ v) \land w = \mathtt{DA}_{\mathbf{z}} \circ v])\ \mathtt{ANA}_{\mathbf{z}}) : ((\mathbf{z} \to \mathbf{Z}) \to (\mathbf{z} \to \mathbf{Z})) \to \mathbf{Z}; \mathbf{FR}_{6,\mathrm{f}} \multimap \mathbf{S}_{\mathrm{m},6}; \lambda_G.G\ (\lambda_x.\text{see marko}\ x) : \mathbf{pr} \to \mathbf{p}$

We give the following lexical entry for *će* which occurs with a finite verb phrase complement:

(258) $\lambda_F.(F\ (\lambda_{Gxw}.\exists_{vt}[(\mathbf{PER}\ x\ v) \land (\mathbf{fst_z}\ v\ t)$
$\land w = \mathbf{toZ}(\lambda_s.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\text{će}(\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v) \circ$
$\mathbf{toZ}(\mathbf{k}\ (G\, \mathrm{e}_{\mathbf{s}}))])) : (((\mathbf{z} \to \mathbf{Z}) \to (\mathbf{z} \to \mathbf{Z})); \to \mathbf{Z}) \to \mathbf{Z}$
$\prod_{k:\mathbf{K}}[(\mathbf{FR} \multimap \mathbf{S}_{k,6}) \multimap \mathbf{S}_{k,5}]; \lambda_F.(F\ (\lambda_{Px}.(\text{FUT}\ P\ x)) : (\mathbf{pr} \to \mathbf{p}) \to \mathbf{p}.$

Tectogrammatically and semantically, the outcome is the same as with this clitic's counterpart which occurs with an infinitival verb phrase. Phenogrammatically, the situation is more complex because the clitic needs to turn the verb phrase into a length one string of languages and ensure that it occurs rightmost in the main clause, while encliticizing onto some phonological word in the subject. Below we show the full reduction of the phenogrammatical term obtained by combining this clitic with the sentence missing a subject control verb to construct a representation of the sentence *Ana će da vidi Marka* 'Ana will see Marko'.

(259) $\lambda_F.(F\ (\lambda_{Gxw}.\exists_{vt}[(\mathbf{PER}\ x\ v) \land (\mathbf{fst_z}\ v\ t) \land$
$w = \mathbf{toZ}(\lambda_s.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\text{će}(\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v) \circ \mathbf{toZ}(\mathbf{k}\ (G\, \mathrm{e}_{\mathbf{s}}))]))$

$$(\lambda_{H}.(H\,(\lambda_{x'w'}.\exists_{v'}[(\mathbf{PER}(x' \circ \mathtt{VIDI}_{\mathbf{z}} \circ \mathtt{MARKA}_{\mathbf{z}})\ v') \wedge w' = \mathtt{DA}_{\mathbf{z}} \circ v'])\ \mathtt{ANA}_{\mathbf{z}}))$$

$$\rightsquigarrow (\lambda_{H}.(H\,(\lambda_{x'w'}.\exists_{v'}[(\mathbf{PER}(x' \circ \mathtt{VIDI}_{\mathbf{z}} \circ \mathtt{MARKA}_{\mathbf{z}})\ v') \wedge w' = \mathtt{DA}_{\mathbf{z}} \circ v'])\ \mathtt{ANA}_{\mathbf{z}}))$$
$$(\lambda_{Gxw}.\exists_{vt}[(\mathbf{PER}\ x\ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$$
$$w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\acute{c}\mathtt{e}(\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v) \circ \mathbf{toZ}(\mathbf{k}\,(G\,\mathtt{e}_{\mathbf{S}}))])$$

$$\rightsquigarrow (\lambda_{Gxw}.\exists_{vt}[(\mathbf{PER}\ x\ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$$
$$w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\acute{c}\mathtt{e}(\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v) \circ \mathbf{toZ}(\mathbf{k}\,(G\,\mathtt{e}_{\mathbf{S}}))]$$
$$(\lambda_{x'w'}.\exists_{v'}[(\mathbf{PER}(x' \circ \mathtt{VIDI}_{\mathbf{z}} \circ \mathtt{MARKA}_{\mathbf{z}})\ v') \wedge w' = \mathtt{DA}_{\mathbf{z}} \circ v'])\ \mathtt{ANA}_{\mathbf{z}})$$

$$\rightsquigarrow \lambda_{xw}.\exists_{vt}[(\mathbf{PER}\ x\ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$$
$$w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\acute{c}\mathtt{e}(\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v) \circ$$
$$\mathbf{toZ}(\mathbf{k}\,(\lambda_{x'w'}.\exists_{v'}[(\mathbf{PER}(x' \circ \mathtt{VIDI}_{\mathbf{z}} \circ \mathtt{MARKA}_{\mathbf{z}})\ v') \wedge w' = \mathtt{DA}_{\mathbf{z}} \circ v'])\mathtt{e}_{\mathbf{S}})]\ \mathtt{ANA}_{\mathbf{z}}$$

$$\rightsquigarrow \lambda_{w}.\exists_{vt}[(\mathbf{PER}\ \mathtt{ANA}_{\mathbf{z}}\ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$$
$$w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\acute{c}\mathtt{e}(\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v) \circ$$
$$\mathbf{toZ}(\mathbf{k}\,((\lambda_{x'w'}.\exists_{v'}[(\mathbf{PER}(x' \circ \mathtt{VIDI}_{\mathbf{z}} \circ \mathtt{MARKA}_{\mathbf{z}})\ v') \wedge w' = \mathtt{DA}_{\mathbf{z}} \circ v'])\mathtt{e}_{\mathbf{S}})]$$

$$\rightsquigarrow \lambda_{w}.\exists_{vt}[(\mathbf{PER}\ \mathtt{ANA}_{\mathbf{z}}\ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$$
$$w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\acute{c}\mathtt{e}(\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v) \circ$$
$$\mathbf{toZ}(\mathbf{k}\,(\lambda_{w'}.\exists_{v'}[(\mathbf{PER}(\mathtt{VIDI}_{\mathbf{z}} \circ \mathtt{MARKA}_{\mathbf{z}})\ v') \wedge w' = \mathtt{DA}_{\mathbf{z}} \circ v']))]$$

For aorist clitics of *biti* which combine with past participle phrases, we give lexical entries like the following one for the 3rd person singular *bi*:

(260) $\vdash \lambda_{Fw}.\exists_{vt}[(F\,(\lambda_{xy}.\mathbf{PER}(x \circ y))\ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$
$w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\mathtt{bi}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)] :$
$((\mathbf{z} \to \mathbf{z} \to \mathbf{Z}) \to \mathbf{Z}) \to \mathbf{Z};$
$\prod_{k:\mathbf{K}}[((\mathbf{NP}_{\mathrm{nom},\tau,\mathrm{sg},3} \multimap \mathbf{S}_{\mathrm{m},6}) \multimap \mathbf{NP}_{\mathrm{nom},\tau,\mathrm{sg},3} \multimap \mathbf{S}_{k,6}) \multimap \mathbf{S}_{k,5}];$
$\lambda_{F}.F(\lambda_{Px}.\mathsf{W}(P\,x)) : (\mathbf{pr} \to \mathbf{p}) \to \mathbf{p}$

For clitic versions of the copula, we give lexical entries like the following one for the second person singular *si*:

(261) $\vdash \lambda_{Fw}.\exists_{vt}[(F\,(\lambda_{xy}.\mathbf{PER}(x \circ y))\ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$
$w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\mathtt{si}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)] :$

$$((\mathbf{z} \to \mathbf{z} \to \mathbf{Z}) \to \mathbf{Z}) \to \mathbf{Z};$$
$$\prod_{k:\mathbf{K}}[(\mathbf{PR} \multimap \mathbf{S}_{k,6}) \multimap \mathbf{S}_{k,5}]; \lambda_F.(F\,(\lambda_{Px}.P\,x)) : (\mathbf{pr} \to \mathbf{p}) \to \mathbf{p}$$

Phenogrammatically, these enclitics function exactly like *htjeti* clitics which take infinitival verb phrase complements.

We remain agnostic here as to whether to (i) allow two sets of verbal enclitic lexical entries into the grammar, one set which occurs main declarative clauses and the other, more complicated, which occurs in other types of clauses, or (ii) eschew the initial simple but not general enough analysis of slot 1 enclitics altogether, and just retain the most general, and the most complicated set of lexical entries which is clause universal.

Slot 6 Auxiliary *je*

The clitic *je* occurs in the rightmost slot of the clitic cluster and cannot be analyzed as combining with its complement and then the subject even in main declarative clauses, because that would preclude the possibility of other enclitics occurring in the same clitic cluster to the left of *je*. So we have to analyze it as combining with a sentence which is missing an expression exactly like *je*. Below is the schema for the hypothesis that has to be introduced:

(262) $G : \mathbf{z} \to \mathbf{z} \to \mathbf{Z}; \mathbf{PR}_{\tau:\mathbf{Gdr},\tau':\mathbf{D}}; G : \mathbf{pr} \vdash G : \mathbf{z} \to \mathbf{z} \to \mathbf{Z}; \mathbf{PR}_{\tau:\mathbf{Gdr},\tau':\mathbf{D}}; G : \mathbf{pr}$

We give *je* the following lexical entry schema:

(263) $\vdash \lambda_{Fw}.\exists_{vt}[(F\,(\lambda_{xy}.\mathtt{PER}(x \circ y))\,v) \wedge (\mathtt{fst}_{\mathtt{z}}\,v\,t) \wedge$
$$w = \mathtt{toZ}(\lambda_s.s = (\mathtt{snc}_{\mathtt{s}}(\mathtt{lst}_{\mathtt{s}}\,t)\#\mathtt{je}\,(\mathtt{tsr}_{\mathtt{s}}\,t))) \circ (\mathtt{rst}_{\mathtt{z}}\,v)] :$$
$$((\mathbf{z} \to \mathbf{z} \to \mathbf{Z}) \to \mathbf{Z}) \to \mathbf{Z};$$
$$\prod_{k:\mathbf{K}}[(\mathbf{PR} \multimap \mathbf{S}_{k,n>0}) \multimap \mathbf{S}_{k,0}]; \lambda_F.(F\,(\lambda_{Px}.P\,x)) : (\mathbf{pr} \to \mathbf{p}) \to \mathbf{p}$$

So, this lexical entry is completely analogous to other present tense *biti* clitics, except for the tectogrammatical number parameter, which *je* reduces to 0 as it is the rightmost clitic.

To construct a sentence with a pronominal clitic and *je*, such as *Ana ga je vidjela* 'Ana saw him', we have to introduce an accusative noun phrase trace, and a *je* trace:

(264) $$G; \mathbf{PR}; G, x; \mathbf{NP}_{\mathrm{acc,m,sg,3}}; x \vdash (G\,(\mathtt{VIDJELA}_{\mathbf{z}} \circ \mathtt{x})\ \mathtt{ANA}_{\mathbf{z}});$$
$$\mathbf{S}_{\mathrm{m,6}}; (G\,(\mathsf{PAST}(\mathsf{see}\ x))\ \mathsf{ana})$$

We first bind the accusative trace, and combine the resulting sign with the accusative clitic, which gives the following:

(265) $$G; \mathbf{PR}; G \vdash \lambda_{w'}.\exists_{v't'}[((G\,\mathtt{VIDJELA}_{\mathbf{z}}\ \mathtt{ANA}_{\mathbf{z}})\ v') \wedge (\mathbf{fst_z}\ v'\ t') \wedge$$
$$w' = \mathbf{toZ}(\lambda_{s'}.s' = (\mathbf{snc_s}(\mathbf{lst_s}\ t')\#\mathtt{ga}\ (\mathbf{tsr_s}\ t'))) \circ (\mathbf{rst_z}\ v')] : \mathbf{Z};$$
$$\mathbf{S}_{\mathrm{m,3}}; (G\,(\mathsf{see}\ x)\ \mathsf{ana}) : \mathbf{p}$$

Now we bind the *je* trace and combine *je* with the resulting sign. Below we show the step by step reduction of the resulting phenogrammatical term.

(266) $$\lambda_{Fw}.\exists_{vt}[(F\,(\lambda_{xy}.\mathbf{PER}(x \circ y))\ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$$
$$w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\mathtt{je}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)]$$
$$(\lambda_{Gw'}.\exists_{v't'}[((G\,\mathtt{VIDJELA}_{\mathbf{z}}\ \mathtt{ANA}_{\mathbf{z}})\ v') \wedge (\mathbf{fst_z}\ v'\ t') \wedge$$
$$w' = \mathbf{toZ}(\lambda_{s'}.s' = (\mathbf{snc_s}(\mathbf{lst_s}\ t')\#\mathtt{ga}\ (\mathbf{tsr_s}\ t'))) \circ (\mathbf{rst_z}\ v')])$$

$$\rightsquigarrow \lambda_{w}.\exists_{vt}[((\lambda_{Gw'}.\exists_{v't'}[((G\,\mathtt{VIDJELA}_{\mathbf{z}}\ \mathtt{ANA}_{\mathbf{z}})\ v') \wedge (\mathbf{fst_z}\ v'\ t') \wedge$$
$$w' = \mathbf{toZ}(\lambda_{s'}.s' = (\mathbf{snc_s}(\mathbf{lst_s}\ t')\#\mathtt{ga}\ (\mathbf{tsr_s}\ t'))) \circ (\mathbf{rst_z}\ v')])$$
$$(\lambda_{xy}.\mathbf{PER}(x \circ y))\ v) \wedge (\mathbf{fst_z}\ v\ t) \wedge$$
$$w = \mathbf{toZ}(\lambda_{s}.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\mathtt{je}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)]$$

$$\rightsquigarrow \lambda_{w}.\exists_{vt}[(\lambda_{w'}.\exists_{v't'}[(((\lambda_{xy}.\mathbf{PER}(x \circ y)\ \mathtt{VIDJELA}_{\mathbf{z}}\ \mathtt{ANA}_{\mathbf{z}})\ v') \wedge (\mathbf{fst_z}\ v'\ t') \wedge$$
$$w' = \mathbf{toZ}(\lambda_{s'}.s' = (\mathbf{snc_s}(\mathbf{lst_s}\ t')\#\mathtt{ga}\ (\mathbf{tsr_s}\ t'))) \circ (\mathbf{rst_z}\ v')])\ v) \wedge$$

$(\mathbf{fst_z}\ v\ t) \wedge$
$w = \mathbf{toZ}(\lambda_s.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\mathtt{je}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)]$

$\rightsquigarrow \lambda_w.\exists_{vt}[(\lambda_{w'}.\exists_{v't'}[(((\mathbf{PER}(\mathrm{VIDJELA_z} \circ \mathrm{ANA_z})\ v') \wedge (\mathbf{fst_z}\ v'\ t') \wedge$
$w' = \mathbf{toZ}(\lambda_{s'}.s' = (\mathbf{snc_s}(\mathbf{lst_s}\ t')\#\mathtt{ga}\ (\mathbf{tsr_s}\ t'))) \circ (\mathbf{rst_z}\ v')])\ v)$
$\wedge (\mathbf{fst_z}\ v\ t) \wedge w = \mathbf{toZ}(\lambda_s.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\mathtt{je}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)]$

$\rightsquigarrow \lambda_w.\exists_{vtv't'}[(\mathbf{PER}(\mathrm{VIDJELA_z} \circ \mathrm{ANA_z})\ v') \wedge (\mathbf{fst_z}\ v'\ t') \wedge$
$v = \mathbf{toZ}(\lambda_{s'}.s' = (\mathbf{snc_s}(\mathbf{lst_s}\ t')\#\mathtt{ga}\ (\mathbf{tsr_s}\ t'))) \circ (\mathbf{rst_z}\ v') \wedge (\mathbf{fst_z}\ v\ t)$
$\wedge w = \mathbf{toZ}(\lambda_s.s = (\mathbf{snc_s}(\mathbf{lst_s}\ t)\#\mathtt{je}\ (\mathbf{tsr_s}\ t))) \circ (\mathbf{rst_z}\ v)]$

Essentially, the only two phonological words *Ana* and *vidjela* can freely permute. The accusative clitic encliticizes onto the first phonological word, resulting in a string of languages. *je* then encliticizes onto the first phonological word in that string of languages, which is either *Ana* or *vidjela* with *ga* encliticized onto it. So the clitics stack exactly as desired, and the whole sentence is predicted to be pronounceable two different ways, *Ana ga je vidjela* or *Vidjela ga je Ana*.

It is well known that in Serbo-Croatian, in the presence of the clitic *se*, *je* is typically not pronounced. This is true both for the inherent reflexive and the true reflexive *se*. For example:

(267) a.

Ana	se	vidjela.
$\text{Ana}_{NOM,f,sg,3}$	REFL	$\text{see}_{ppl,f,sg}$

‘Ana saw herself’

cf. *Ana se je vidjela.*

b.

Ana	ga	se	bojala.
$\text{Ana}_{NOM,f,sg,3}$	$\text{him}_{ACC,m,sg,3}$	SE	$\text{be-afraid}_{ppl,f,sg}$

‘Ana was afraid of him’

cf. *Ana ga se je bojala.*

Sentences which contain both clitics are not necessarily altogether unacceptable, but we certainly want to at least allow the possibility of *je* disappearing in the presence of *se*.

We can easily account for this phenomenon because of the number parameter of sentences. Basically, the clitic *je* 'knows' whether its complement sentence contains a slot 5 clitic (*se*) or not. In the former case, its argument sentence will have the number parameter 1, but in the latter case its parameter will be greater than 1. So we give the following lexical entry for *je* which occurs in a sentence that already contains *se*, which causes *je* not to be pronounced at all.

(268) $\vdash \lambda_{Fw}.(F\ (\lambda_{xy}.\mathbf{PER}(x \circ y))\ w) : ((\mathbf{z} \rightarrow \mathbf{z} \rightarrow \mathbf{Z}) \rightarrow \mathbf{Z}) \rightarrow \mathbf{Z};$
$\prod_{k:\mathbf{K}}[(\mathbf{PR} \multimap \mathbf{S}_{k,1}) \multimap \mathbf{S}_{k,0}]; \lambda_F.(F\ (\lambda_{Px}.P\ x)) : (\mathbf{pr} \rightarrow \mathbf{p}) \rightarrow \mathbf{p}$

For example, in constructing the sentence *Ana se vidjela* 'Ana saw herself', after the reflexive is placed and the *je* trace is bound, we have the following sign:

(269) $\vdash \lambda_{Gw'}.\exists_{v't'}[((G\,\mathrm{VIDJELA}_{\mathbf{z}}\ \mathrm{ANA}_{\mathbf{z}})\ v') \wedge (\mathbf{fst_z}\ v'\ t') \wedge$
$w' = \mathbf{toZ}(\lambda_{s'}.s' = (\mathbf{snc_s}(\mathbf{lst_s}\ t')\#\mathrm{ga}\ (\mathbf{tsr_s}\ t'))) \circ (\mathbf{rst_z}\ v')] :$
$(\mathbf{z} \rightarrow \mathbf{z} \rightarrow \mathbf{Z}) \rightarrow \mathbf{Z}; \mathbf{PR} \multimap \mathbf{S}_{\mathrm{m},1}; \lambda_G.(G\ (\lambda_x.\mathrm{PAST}(\mathrm{see}\ x\ x))\ \mathrm{ana}) : \mathbf{pr} \rightarrow \mathbf{p}$

Combining this sign with the version of *je* which is not going to be pronounced we get the following:

(270) $\vdash \lambda_w.\exists_{v't'}[((\mathbf{PER}\ (\mathrm{VIDJELA}_{\mathbf{z}} \circ \mathrm{ANA}_{\mathbf{z}})\ v') \wedge (\mathbf{fst_z}\ v'\ t') \wedge$
$w = \mathbf{toZ}(\lambda_{s'}.s' = (\mathbf{snc_s}(\mathbf{lst_s}\ t')\#\mathrm{se}\ (\mathbf{tsr_s}\ t'))) \circ (\mathbf{rst_z}\ v')] : \mathbf{Z};$
$\mathbf{S}_{\mathrm{m},0}; \mathrm{PAST}(\mathrm{see\ ana\ ana}) : \mathbf{p}$

The sentence expresses the right meaning, and is predicted to be pronounceable two different ways, *Ana se vidjela* and *Vidjela se Ana*.

5.3.6 1C and 1W placement

Preliminaries

In this section we explicitly address the 1C and 1W placement of enclitics. Recall that 1C placement refers to encliticization to the last phonological word of the initial constituent, and 1W refers to encliticization to the first phonological word of the initial constituent. However, not everything counts as a 'constituent' with respect to 1C placement. Below we enumerate different possibilities of enclitic cluster placement, and mark each possibility in terms of whether the permissive or the conservative version of our grammar predicts that placement or not. These are pretty much all the types of constituents which allow 1C placement. Clauses and verb phrases, including passive, infinitival and participial verb phrases, cannot in general host enclitics. The differences between the permissive and conservative versions of our grammar stem from the differences concerning noun phrase and prepositional phrase composition. Since we analyze enclitics as attaching to the last word in the initial length one string of languages, extending the grammar comes down to specifying in a more fine grained way what can count as that initial length one string of languages in a clause.

The problem of generalizing these two grammars to cover clitic placement possibilities comes down to two issues: (i) allow noun phrases and prepositional

TYPE	placement	permissive grammar	conservative grammar
adverbial phrase	Vrlo brzo je Ana došla. Vrlo je brzo Ana došla. ’Ana arrived very fast’	✓ ✗	✓ ✗
prepositional adverbials and predicatives	U velikom gradu je ona živjela. U velikom je gradu ona živjela. ‘She lived in a big city’	✗ ✓	✓ ✗
adjective + noun	Pametan momak je došao. Pametan je momak došao. ‘A smart young man arrived’	✗ ✓	✓ ✗
noun + postnominal modifier	Djevojka iz Beograda je lijepa. Djevojka je iz Beograda lijepa ‘The girl from Belgrade is pretty’	✓ ✗	✓ ✗
adjective + noun + postnominal modifier	Pametan momak iz Beograda je došao. Pametan je momak iz Beograda došao. ‘A smart young man from Belgrade arrived’	✓ ✗	✓ ✗
quantificational determiner + noun	Svaka djevojka je lijepa. Svaka je djevojka lijepa. ‘Every girl is pretty’	✗ ✓	✓ ✗

Table 5.3: Interim summary evaluation of grammar with respect to enclitic cluster placement.

phrases to fuse into strings of languages of length one in the permissive grammar, and (ii) allow 'reaching into' the clause initial length one string of languages and to split it into a length two string of languages, the first one built out of the initial phonological word. The latter is needed in both versions of the grammar.

Fusing noun phrases and prepositional phrases

This is needed for the permissive grammar only, to allow enclitics to attach to the last phonological word in an initial multi-word noun phrase. All phrasal noun phrases are quantificational, either because they contain a quantificational determiner, or because they underwent the [Quant] rule (see Chapter 3). So we need to write a rule that will target quantificational noun phrases and alter their phenogrammatical terms appropriately.

(271)

$$\frac{\vdash \phi : (\mathbf{z} \to \mathbf{Z}) \to \mathbf{Z}; (\mathbf{NP}_{\tau,\tau'\tau'',3} \multimap \mathbf{S}_{\mathrm{m},6}) \multimap \mathbf{S}_{\mathrm{m},6}; \sigma : (\mathbf{e} \to \mathbf{p}) \to \mathbf{p}}{\vdash \lambda_F.(F\,(\mathbf{toZ}\ \mathbf{k}(\phi\ \lambda_x.\mathbf{PER}\ x))) : (\mathbf{z} \to \mathbf{Z}) \to \mathbf{Z}; (\mathbf{NP}_{\tau,\tau'\tau'',3} \multimap \mathbf{S}_{\mathrm{m},6}) \multimap \mathbf{S}_{\mathrm{m},6};\ \sigma : (\mathbf{e} \to \mathbf{p}) \to \mathbf{p}}\ [1Z]$$

We illustrate how the rule works with a concrete example. Consider the noun phrase *lijepa djevojka* 'a pretty girl' after it has undergone the [Quant] rule. Focusing on the phenoterm only, since the tectogrammatical and the semantic component of the sign remain unchanged, here is how the [Z1] rule alters it:

(272)

$$\frac{\vdash \lambda_G.(G\,\mathtt{LIJEPA}_{\mathbf{z}} \circ \mathtt{DJEVOJKA}_{\mathbf{z}}) : (\mathbf{z} \to \mathbf{Z}) \to \mathbf{Z}}{\vdash \lambda_F.(F\,(\mathbf{toZ}\ \mathbf{k}(\lambda_G.(G\,\mathtt{LIJEPA}_{\mathbf{z}} \circ \mathtt{DJEVOJKA}_{\mathbf{z}})(\lambda_x.\mathbf{PER}\ x)))) : (\mathbf{z} \to \mathbf{Z}) \to \mathbf{Z}}\ [1Z]$$

The conclusion of this proof reduces to:

(273) $\vdash \lambda_F.(F\,(\mathtt{toZ}\;\mathtt{k}(\mathtt{PER}\;(\mathtt{LIJEPA}_{\mathtt{z}} \circ \mathtt{DJEVOJKA}_{\mathtt{z}})))) : (\mathtt{z} \to \mathtt{Z}) \to \mathtt{Z}$

So the noun phrase has been turned into a length one string of languages. Should it occur clause initially after it combines with the verb phrase, we will get 1C enclitic placement, because the enclitics would attach to the initial length one string of languages. With the addition of this rule to the permissive grammar, we get 1C placement of clitics for all noun phrases. Noun phrases, of course, do not have to undergo this rule, so 1W placement is still possible.

As for prepositional phrases, we can simply add the predicative and adverbial prepositional phrase lexical entries from the conservative grammar to the permissive grammar.

Splitting initial constituents

The second extension involves splitting a clause initial length one string of languages into a length one string of languages constructed out of the initial phonological word and the rest of that initial string of languages. This is needed for both the permissive and the conservative grammar.

This would allow the enclitic cluster to be hosted by an initial phonological word which under normal circumstances cannot be separated from the rest of its constituent, such as a noun or an adjective in the presence of a postnominal modifier in the constituent noun phrase.

Recall that Progovac (1996) maintains that only phonological that are normally separable from the remainder of their constituent can host clitics. So, for example, she doesn't accept examples where the noun and the postnominal modifier sequence is split by the enclitic cluster. If one wishes to maintain these judgments, then the following rule can simply be omitted from the grammar.

If one, however, wants to allow enclitics to attach to phonological words which normally can't be separated from the remainder of their constituent, the following rule needs to be added to the grammar:

(274)

$$\frac{\Gamma \vdash \phi : \mathtt{Z}; \mathbf{S}_{\mathrm{m},6}; \sigma : \mathtt{p}}{\Gamma \vdash \lambda_w . \exists_{vta}[(\phi\ v) \wedge (\mathtt{fst}_{\mathtt{z}}\ v\ t) \wedge (\mathtt{fst}_{\mathtt{s}}\ t) = a \wedge w = \mathtt{toZ}(\lambda_s . s = a_{\mathtt{s}}) \circ \mathtt{toZ}(\lambda_s . s = \mathtt{rst}_{\mathtt{s}}\ t)] : \mathtt{Z}; \mathbf{S}_{\mathrm{m},6}; \sigma : \mathtt{p}} \text{[1W]}$$

This rule can apply to any declarative main clause which hasn't had any clitics placed inside of it already. The reason it applies to main clauses only is that in embedded clauses the complementizer obligatorily hosts the enclitics, so the rule is only relevant for main clauses. The rule doesn't alter the semantics or the tectogrammatical type of the sign. Phenogrammatically, it constructs a different set of strings of languages than the input one by reaching into the first string of languages, extracting the first phonological word out of it, then constructing a length one string of languages out of that initial phonological word. Now, the enclitic cluster can encliticize into the first phonological word, that is, we get unrestricted 1W placement.

5.4 Conclusion

In this chapter, we gave our theory of encliticization in Serbo-Croatian. The general approach involved treating enclitics as functions looking for sentence with the right kind of gap, and then attaching themselves to the last word in the first string of languages of their argument. In contrast to ordinary phenogrammatical combination of expressions which works at the relatively high level of languages or strings of languages, clitics attach to their hosts at a deeper level, that of phonological words. So, the grammar correctly represents the fact that encliticization builds new phonological words. At the same time, because of how deeply they attach to their hosts, no subsequent ordinary phenogrammatical combination can rip apart the new phonological word created by encliticization.

The number parameter of the tectogrammatical family of sentence types was crucial for enforcing the ordering of enclitics in the clitic cluster. Because the number parameter keeps a very precise record of which slots in the clitic cluster have been filled, we were able to also account for the phenomenon of the clitic *je* disappearing in presence of *se*. This phenomenon is not definable semantically or tectogrammatically since it pertains to both the inherent and the true reflexive, and simply depends on whether the penultimate slot in the enclitic cluster is occupied or not. This is precisely what our analysis of this phenomenon depends on as well.

Finally, while we did have to introduce non-logical rules to get the details of the clitic cluster placement right, it is worth mentioning that in our grammar, there is no difference between 1C and 1W placement of clitics. In both cases, the clitics work the same way, encliticizing onto the last phonological word in the first length one string of languages. The real problem was correctly picking out the class of expressions which can be treated as initial length one strings of languages in a clause, and we more or less did that.

Chapter 6: Interrogatives

Chapter 7: Conclusion

Bibliography

Baldridge, Jason. 2002. *Lexically SpeciÞed Derivational Control in Combinatory Categorial Grammar*. Ph.D. thesis, University of Edinburgh.

Bernardi, Raffaella. 2002. *Reasoning with Polarity in Categorial Type Logic*. Ph.D. thesis, University of Utrecht.

Bošković, Željko. 1997. Superiority effects with multiple *wh* fronting in Serbo-Croatian. *Lingua* **102**: 1-20.

Bošković, Željko. 2001. *On the Nature of the Syntax-Phonology Interface*. Oxford: Elsevier.

Bošković, Željko. 2002. On multiple *wh* fronting. *Linguistic Inquiry* **33**: 351-383.

Bošković, Željko. 2004. Clitic placement in South Slavic. *Journal of Slavic Linguistics* **12**.

Bošković, Željko. 2009. Unifying first and last conjunct agreement. In *Natural Language and Linguistic Theory* **27**.

Browne, Wayles. 1974. On the problem of enclitic placement in Serbo-Croatian. In R. Brecht and C. Chvany, eds., *Slavic Transformational Syntax*. Ann Arbor: University of Michigan Press.

Chomsky, Noam. 1957. *Syntactic Structures*. The Hague: Mouton.

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

Chomsky, Noam. 1981. *Lectures on Government and Binding*. Holland: Foris Publications.

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

Curry, Haskell. 1961. Some logical aspects of grammatical structure. In Roman Jakobson, ed., *Structure of Language and Its Mathematical Aspects*.

Dowty, David. 1996. Toward a minimalist theory of syntactic structure. In Bunt, H. and A. van Horck, eds., *Discontinuous Constituency*. Mouton de Gruyter.

Franks, Steven and T. H. King. 2000. *A Handbook of Slavic Clitics*. Oxford University Press.

Gazdar, G., E. Klein, G. Pullum and I. Sag. 1985. *Generalized Phrase Structure Grammar*. Harvard University Press.

Godjevac, Svetlana. 1999. An autosegmental/metrical analysis of Serbo-Croatian intonation. In *OSU Working Papers in Linguistics* **54**.

Godjevac, Svetlana. 2000. *Word Order, Intonation and Focus Projection in Serbo-Croatian*. PhD thesis, The Ohio State University.

de Groote, Philippe. 2001. Towards Abstract Categorial Grammar. *Proceedings of ACL*, 2001.

de Groote, Philippe and Sarah Maarek. 2007. Type-theoretic extensions of Abstract Categorial Grammars. In R. Muskens, ed., *Proceedings of Workshop on New Directions in Type-Theoretic Grammars*.

Halpern, Aaron. 1995. *On the the Placement and Morphology of Clitics*. CSLI.

Hana, Jiří. 2007. *Czech Clitics in Higher Order Grammar*. PhD thesis, The Ohio State University.

Hoffman, Beryl. 1992. A CCG approach to free word order languages. In *Proceedings of ACL '92*.

Kathol, Andreas. 2000. *Linear Syntax*. Oxford University Press.

Kordić, Snježana. 2010. *Jezik i Nacionalizam*. Zagreb: Rotulus Universitas, Durieux.

Kraak, Esther. 1998. A Deductive Account of French Object Clitics. In *Syntax and Semantics: Complex Predicates in Nonderivational Syntax* 30:271-312.

Kubota, Yusuke. 2010. *(In)SSexibility of Constituency in Japanese in Multi-Modal Categorial Grammar with Structured Phonology*. PhD thesis, The Ohio State University.

Kubota, Yusuke and Carl Pollard. 2009. Phonological interpretation into preordered algebras. In *Proceedings of MoL11*.

Lambek, Joachim. 2001. Type grammar as pregroups. In *Grammars*, Vol. 4.

Lambek, Joachim and Philip J. Scott. 1986. *Introduction to Higher Order Categorical Logic*. Cambridge University Press.

Leko, Nedžad. 1999. Functional categories and the structure of DP in Bosnian. In Dimitrova- Vulchanova, M. and L. Hellan, eds., *Topics in South Slavic Syntax and Semantics*. JohnBenjamins.

Mihaliček, Vedrana. 2008. On multiple *wh* questions in Serbo-Croatian. Manuscript.

Mihaliček, Vedrana. 2010a. The order of *wh* words in multiple *wh* extraction in Serbo-Croatian. In Smirnova, A., V. Mihaliček and L. Ressue, eds., *Formal Studies in Slavic Linguistics*. Cambridge Scholars.

Mihaliček, Vedrana. 2010b. English questions in a linear logic based grammar. Manuscript.

Mihaliček, Vedrana. 2010c. (Not so) free word order in Lambda Grammar: The case of BCMS. In *Proceedings of FASSBL 2010*.

Mihaliček, Vedrana and Carl Pollard. To appear. Distinguishing phenogrammar from tectogrammar simplifies the analysis of interrogatives. In *Proceedings of Formal Grammar 2010*.

Moortgat, Michael. 1996. Generalized quantification and discontinuous type constructors. In H. Bunt, and A. van Horck, eds., *Discontinuous Constituency*. Mouton de Gruyter.

Moortgat, Michael. 1997. Categorial type logics. J. van Benthem and A. ter Meulen, eds., *Handbook of Logic and Language*. Amsterdam: Elsevier.

Morrill, Glyn. 1994. *Type Logical Grammar: Categorial Logic of Signs*. Dordrecht: Kluwer Academic Publishers.

Morrill, Glyn, Mario Fada and Oriol Valentin. 2007. Nondeterministic discontinuous Lambek calculus. *Proceedings of the Seventh International Workshop on Computational Semantics*, IWCS7, Tilburg.

Morrill, Glyn and Anna Gavarró.1992. Catalan Clitics. A. Lecomte, ed., *Word Order in Categorial Grammar*, pp. 211-232, Clermont-Ferrand. Editions Adosa.

Morrill, Glyn and Teresa Solias. 1993. Tuples, discontinuity and gapping in Categorial Grammar. In *Proceedings of the European Chapter Association for Computational Linguistics*, EACL93, Utrecht, 287-297.

Muskens, Reinhard. 2003. Language, lambdas, and logic. In Geert-Jan Kruijff and Richard Oehrle, eds., *Resource Sensitivity in Binding and Anaphora, Studies in Linguistics and Philosophy*. Kluwer.

Muskens, Reinhard. 2005. Sense and the computation of reference. *Linguistics and Philosophy* **28**:4.

Muskens, Reinhard. 2007a. Intensional models for the theory of types. *The Journal of Symbolic Logic* **72**:1.

Muskens, Reinhard. 2007b. Separating syntax and combinatorics in categorial grammar. *Research on Language and Computation* **5**:3.

Nishida, Chiyo. 1996. Second-Position Clitic Pronouns in Old Spanish and Categorical Grammar. In Aaron Halpern and Arnold Zwicky, eds., *Second Position Clitics and Related Phenomena*. Stanford, Ca.: Center for the Study of Language and Information. pp. 333-373.

Oehrle, Richard. 1994. Term-labeled categorial type systems. *Linguistics and Philosophy* **17**:6.

Penn, Gerald. 1999a. A Generalized-Domain-Based Approach to Serbo-Croatian Second Position Clitic Placement. In Bouma, G. et al., eds., *Constraints and Resources in Natural Language Syntax and Semantics*. CSLI.

Penn, Gerald. 1999b. Linearization and WH-Extraction in HPSG: Evidence from Serbo- Croatian. In Borsley, R. and A. Przepiorkowski, eds., *Slavic in HPSG*. CSLI.

Pollard, Carl. 2004. Type-Logical HPSG. In G. Jäger, P. Monachesi, G. Penn, and S. Wintner, eds., *Proceedings of Formal Grammar 2004*. Nancy, France: European Summer School in Language, Logic, and Information, pp. 107-124.

Pollard, Carl. 2008. Hyperintensions. *Journal of Logic and Computation* **18**:2.

Pollard, Carl. 2011. The calculus of responsibility and commitment. In Lecomte, A. and S. Tronçon, eds., *Ludics, Dialogue and Interaction*. *LNCS* 6505. Berlin: Springer.

Pollard, Carl and Ivan A. Sag. 1987. *Information-Based Syntax and Semantics, vol. 1*. CSLI Lecture Notes #13. CSLI.

Progovac, Ljiljana. 1996. Clitics in Serbian/Croatian: Comp as the second position. In Halpern A. and A. Zwicky, eds., *Approaching Second. CSLI Publications*.

Progovac, Ljiljana. 2005. *A Syntax of Serbian Clausal Architecture*. Bloomington: Slavica.

Radanović-Kocić, Vesna. 1996. Placement of Serbo-Croatian Clitics: A prosodic approach. In Halpern A. and A. Zwicky, eds., *Approaching Second*. CSLI Publications.

Reape, Mike. 1993. *A Formal Theory of Word Order: A Case Study in West Germanic*. Ph.D. thesis, University of Edinburgh.

Reape, Mike. 1996. Getting things in order. In Bunt, H. and A. van Horck, eds., *Discontinuous Constituency*. Mouton de Gruyter.

Rudin, Catherine. 1988a. On multiple questions and multiple *wh* fronting. *Natural Language and Linguistic Theory* **6**: 445-501.

Rudin, Catherine. 1988b. Multiple questions in South Slavic, West Slavic and Romanian. *The Slavic and East European Journal* **32**: 1-24.

Schütze, Carson. 1994. Serbo-Croatian second position clitic placement and the phonology- syntax interface. In *MIT Working Papers in Linguistics*. Cambridge: MIT Press.

Smith, Elizabeth. 2010. Correlational Comparison in English. PhD thesis, The Ohio State University.

Stjepanović, Sandra. 2003. Multiple Wh-fronting in Serbo-Croatian matrix questions and the matrix sluicing construction. In Boecx, C. and K. Grohmann, eds., *Multiple* wh *Fronting*. Philadelphia: John Benjamins.

Thomason, Richmond. 1980. A model theory for propositional attitudes. *Linguistics and Philosophy* **4**:1.

Vermaat, Willemijn. 2005. *The Logic of Variation. A Cross-Linguistic Account of Wh-Question Formation*. Ph.D. thesis, Utrecht Institute of Linguistics OTS, Utrecht University.

Yu, Kristine. 2009. The prosody of second position clitics and focus in Zagreb Croatian. In *Proceedings of FASL 17*.

Zec, Draga and S. Inkelas. 1990. Prosodically constrained syntax. In Inkelas, S. and D. Zec, eds., *The Phonology-Syntax Connection*. CSLI and the University of Chicago Press.

Zlatić, Larisa. 1997. *The Structure of the Serbian Noun Phrase*. PhD thesis, University of Texas at Austin.