

# **Optimal Clitic Positioning in Czech and Beyond**

Clitic syntax, morphology, and prosody from a constraint-based  
perspective

Dissertation  
zur Erlangung des philosophischen Doktorgrades  
an der Philosophischen Fakultät der Georg-August-Universität Göttingen

vorgelegt von  
Nina Adam  
aus Hanau

Göttingen 2024

# Contents

<b>List of Abbreviations</b>	<b>vi</b>
<b>1 Introduction</b>	<b>1</b>
1.1 Goal of this thesis . . . . .	1
1.2 Frameworks and central assumptions . . . . .	2
1.3 Sources for linguistic data . . . . .	2
1.4 Structure . . . . .	4
<b>2 Background I: Clitics</b>	<b>5</b>
2.1 Definition . . . . .	5
2.2 Classification . . . . .	7
2.2.1 Pro- vs. enclitics . . . . .	7
2.2.2 Special vs. simple clitics . . . . .	7
2.2.3 Phrasal vs. head-adjacent clitics . . . . .	8
2.3 Diachronic role in grammaticalisation . . . . .	9
2.4 Categorial affiliations . . . . .	9
2.5 Linguistic distribution . . . . .	10
2.6 Summary . . . . .	10
<b>3 Background II: Basics of Czech Grammar</b>	<b>11</b>
3.1 A remark on different varieties of Czech . . . . .	11
3.2 Phonology . . . . .	12
3.3 Morphology and morphosyntax . . . . .	12
3.3.1 Nominal categories . . . . .	12
3.3.2 Verbal categories . . . . .	13
3.4 Syntax . . . . .	14
3.4.1 Word order . . . . .	14
3.4.2 Pronouns, copulas, negation . . . . .	17
3.4.3 Binding and extraction . . . . .	17
3.5 Information structure . . . . .	18
3.5.1 Focus-background structure . . . . .	19
3.5.2 Topic-comment structure . . . . .	21
3.5.3 Givenness . . . . .	22
3.6 Summary . . . . .	23

<b>4</b>	<b>The Clitic Inventory of Czech</b>	<b>24</b>
4.1	Types of Czech clitics . . . . .	24
4.1.1	Pronouns . . . . .	25
4.1.2	Auxiliaries . . . . .	27
4.1.3	The conditional complementiser and negation . . . . .	27
4.1.4	Adverbs . . . . .	29
4.2	Properties of second-position clitics . . . . .	29
4.2.1	Inability to appear sentence-initially . . . . .	30
4.2.2	No hosting of other clitics . . . . .	31
4.2.3	Inability to appear in isolation or bear stress . . . . .	32
4.3	Relative ordering within the second position . . . . .	33
4.4	Co-occurrence restrictions in the cluster . . . . .	36
4.5	Summary . . . . .	37
<b>5</b>	<b>Placement of Czech 2P Clitics</b>	<b>38</b>
5.1	Placement in simple clauses . . . . .	38
5.1.1	Clitics in second position . . . . .	38
5.1.2	Clitics in third position . . . . .	42
5.1.3	Clitics in first position and prosodic dependency . . . . .	45
5.2	Placement in complex clauses: clitic climbing . . . . .	48
5.2.1	Domains for Czech clitic climbing . . . . .	48
5.2.2	Co-occurrence restrictions for clitic climbing . . . . .	51
5.3	Summary . . . . .	52
<b>6</b>	<b>Overview of Theories of Clitic Placement</b>	<b>53</b>
6.1	Syntactic approaches . . . . .	53
6.1.1	Veselovská 1995: Clitics in C . . . . .	54
6.1.2	Lenertová 2001: Clitics in Fin . . . . .	55
6.1.3	Discussion . . . . .	55
6.2	Prosodic approaches . . . . .	59
6.2.1	Radanović-Kocić 1996: Prosodic placement in BCMS . . . . .	59
6.2.2	Discussion . . . . .	60
6.3	Syntax-and-PF approaches . . . . .	62
6.3.1	Halpern 1995: Prosodic Inversion . . . . .	62
6.3.2	Franks & Bošković 2001: Phases and copy deletion . . . . .	63
6.3.3	Bošković 2004: Copy deletion instead of PF movement . . . . .	65
6.3.4	Franks 2017: Relinearisation due to prosodic deficiency . . . . .	66
6.3.5	Discussion . . . . .	66
6.4	Morphology-and-OT approaches . . . . .	68
6.4.1	Anderson 1992/2000/2005: Clitics as phrasal affixes . . . . .	68
6.4.2	Richardson 1997: Constraining Czech clitic placement . . . . .	68
6.4.3	Discussion . . . . .	70
6.5	Interim conclusions . . . . .	72

6.6	Accounting for clitic ordering . . . . .	73
6.6.1	Syntactic approaches . . . . .	73
6.6.2	Syntax-and-PF approaches . . . . .	74
6.6.3	Morphology-and-OT approaches . . . . .	75
6.6.4	Morphological approaches . . . . .	76
6.6.5	Discussion . . . . .	76
<b>7</b>	<b>Comparing Clitic and Non-Clitic Verbs</b>	<b>77</b>
7.1	Morphology of clitic and non-clitic auxiliaries . . . . .	78
7.1.1	Four views on the status of clitic auxiliaries . . . . .	78
7.1.2	Omission and reduction . . . . .	80
7.1.3	Colloquial allomorphy . . . . .	82
7.1.4	Frequentative formation . . . . .	83
7.1.5	Summary . . . . .	84
7.2	Syntax of clitic and non-clitic verbal elements . . . . .	84
7.2.1	Fronting to the left periphery . . . . .	85
7.2.2	Verbal crossing . . . . .	85
7.2.3	Interaction of participle and <i>-li</i> . . . . .	88
7.2.4	Summary . . . . .	90
7.3	Past vs. passive: the participle . . . . .	90
7.4	The relation of past and conditional clitic . . . . .	92
7.4.1	Morphological shape of the conditional clitic . . . . .	92
7.4.2	The locus of the past tense interpretation . . . . .	94
7.4.3	Summary . . . . .	95
7.5	Conclusions . . . . .	96
<b>8</b>	<b>An OT Account of Czech 2P Placement</b>	<b>97</b>
8.1	Optimality Theory: Underlying assumptions . . . . .	97
8.1.1	Basic ideas of Optimality Theory . . . . .	97
8.1.2	Generalised Alignment . . . . .	98
8.1.3	Syntax in OT . . . . .	99
8.1.4	Summary . . . . .	101
8.2	Placement in second position . . . . .	101
8.2.1	Basic cases . . . . .	101
8.2.2	Gradient evaluation and INTEGRITY . . . . .	103
8.2.3	A note on placement in first position . . . . .	104
8.2.4	Summary . . . . .	106
8.3	Clitic climbing and the domain of EDGEMOST . . . . .	106
8.3.1	Climbing out of defective clauses . . . . .	106
8.3.2	Blocked climbing out of CP complements . . . . .	110
8.3.3	Initial placement in the embedded clause . . . . .	114
8.3.4	Consequences for coordination . . . . .	115
8.3.5	Summary . . . . .	117

8.4	Why 2P? Motivating the constraints . . . . .	117
8.4.1	General considerations concerning 2P . . . . .	118
8.4.2	Billings 2002: Scope and suffixation . . . . .	118
8.4.3	Dotlačil 2007: Information-structural restrictions . . . . .	122
8.4.4	Summary . . . . .	123
8.5	The order of clitics within the cluster . . . . .	123
8.6	Notes on a factorial typology . . . . .	127
8.7	Conclusions . . . . .	129
<b>9</b>	<b>Analysing 3P Placement and Topicalisation</b>	<b>131</b>
9.1	Argument topicalisation . . . . .	131
9.1.1	The syntax of topicalisation . . . . .	131
9.1.2	Capturing 3P after topics . . . . .	132
9.1.3	Split topicalisation . . . . .	134
9.1.4	Summary . . . . .	136
9.2	Verbal and VP topicalisation . . . . .	136
9.2.1	Literature and corpus data . . . . .	137
9.2.2	Judgement data . . . . .	139
9.2.3	Capturing the data: verb movement and uniformity . . . . .	141
9.2.4	Follow-up: fronting with context . . . . .	147
9.2.5	A note on verb topicalisation . . . . .	150
9.2.6	Summary . . . . .	151
9.3	Conclusions . . . . .	151
<b>10</b>	<b>Extending the Analysis to Further Slavic Languages</b>	<b>153</b>
10.1	Slovak: Czech's closest sibling . . . . .	153
10.1.1	The clitic inventory . . . . .	154
10.1.2	Clitic placement: the basic pattern . . . . .	154
10.1.3	Ordering in the cluster . . . . .	158
10.1.4	Prosodic enclitics or not? . . . . .	159
10.1.5	Third-position placement . . . . .	160
10.1.6	Clitic climbing . . . . .	162
10.1.7	Conclusions . . . . .	164
10.2	Upper Sorbian: 2P with variability . . . . .	165
10.2.1	The clitic inventory . . . . .	166
10.2.2	Clitic placement: the basic pattern . . . . .	169
10.2.3	Ordering in the cluster . . . . .	173
10.2.4	Third-position placement . . . . .	174
10.2.5	Prosodic enclitics or not? . . . . .	176
10.2.6	Conclusions . . . . .	176
10.3	BCMS: prosody and integrity . . . . .	178
10.3.1	Clitic placement: the basic pattern . . . . .	178
10.3.2	Anderson's (2005) analysis of BCMS . . . . .	180

10.3.3	Constraints on prosodic constituency . . . . .	181
10.3.4	The effect of EDGEMOST . . . . .	182
10.3.5	NONINITIAL vs. SUFFIX . . . . .	187
10.3.6	The nature of the first position in BCMS . . . . .	189
10.3.7	Integrating integrity: 2W and 2D . . . . .	190
10.3.8	VP raising and clitic lowering . . . . .	194
10.3.9	Which $\iota$ P is the clitics' domain? . . . . .	196
10.3.10	Conclusions . . . . .	200
<b>11</b>	<b>Conclusions and Outlook</b>	<b>201</b>
11.1	Modelling clitic placement . . . . .	201
11.2	The nature of (West Slavic) clitics . . . . .	201
11.3	The interaction of syntax and phonology . . . . .	202
11.4	Transferability to interface theories . . . . .	203
11.5	On clitics as a diagnostic . . . . .	203
11.6	Objectives for further research . . . . .	204
<b>A</b>	<b>List of constraints</b>	<b>205</b>
<b>B</b>	<b>Acceptability judgement tasks</b>	<b>207</b>
<b>C</b>	<b>German Summary and affidavit</b>	<b>208</b>
	<b>Online resources</b>	<b>211</b>
	<b>References</b>	<b>212</b>

# List of Abbreviations

## Abbreviations used in the glosses

1, 2, 3	first, second, third person
ACC	accusative case
AUX	auxiliary
COMP	complementiser
COND	conditional
DAT	dative case
FREQ	frequentative
GEN	genitive case
IMP	imperative
INDEF	indefinite article
INF	infinitive
INS	instrumental case
IPF	imperfective
LOC	locative case
NOM	nominative case
PART	particle
PASS	passive
PF	perfective
PL	plural
POL	polite form
POSS	possessive reflexive
PRET	preterite tense
PRF	perfect tense
PRS	present tense
PTCP	participle
Q	question marker
REFL	reflexive
REL	relative pronoun
SG	singular
VOC	vocative

**Other abbreviations and symbols**

1P, 2P, 3P, 4P	first, second, third, fourth position (clitic placement)
2D	position after the first syntactic daughter of a sentence
2W	position after the first prosodic word of a sentence
BCMS	Bosnian-Croatian-Montenegrin-Serbian
CP	complementiser phrase
DPFR	default principle of focus realisation
FH	functional head
GA	Generalised Alignment
GCat	grammatical category
ipf.	imperfective
IP	inflectional phrase
i. p. m.	instances per million positions
lit.	literally
MP	Minimalist Program
NP	noun phrase
OT	Optimality Theory
PCat	prosodic category
PCC	person case constraint
pf.	perfective
PIC	phase impenetrability condition
TP	tense phrase
ιP	intonational phrase
φ	phonological phrase
ω	prosodic word



# Chapter 1

## Introduction

Second-position clitics are found in many of the world's languages, and they are a widespread and controversial subject of linguistic analysis. At first glance, capturing second-position cliticisation appears easy, if one does not take into account language-specific deviations from the general pattern. Thus, on one hand, a thorough investigation into concrete linguistic patterns is necessary if one wishes to capture the data. On the other hand, when developing a language-specific theory that aims to capture the entire data range, it is important to bear in mind that the cross-linguistic occurrence of second-position clitics should be based on universal principles.

### 1.1 Goal of this thesis

The descriptive part of this thesis is concerned with the following questions: Where is the clitic cluster positioned? What can precede it? More specifically, I will investigate under which circumstances clitics appear in third position, and whether clitic positioning is influenced only by syntactic or also by prosodic factors in Czech.

Clitics can provide valuable insight and give us important clues as to how the interaction between syntax, morphology, and prosody can be modelled. The theoretical part analyses the advantages and problems of different theories of clitic placement and, based on the Czech data, attempts to draw a picture of the role of different parts of the grammar in clitic positioning. Also, the nature of Czech clitics, especially in contrast to verbs, will be discussed. The ultimate goal is to develop a theory of what precisely Czech clitics are, and how their specific properties translate into a certain behaviour in the language's grammatical structure.

Whilst the focus of this work lies on Czech, other languages, especially Slavic ones, are also taken into consideration. This allows us to put the Czech results into a broader typological perspective, which in turn will help to achieve another goal of this thesis: to develop a theory of clitics that not only explains the Czech data, but also captures typological variation as a direct result of the nature of clitics and the constraints that apply to them. In particular, I will strive to contribute to answering the following questions: Why is the second position special? Why do we find clitics in this position in so many languages?

## 1.2 Frameworks and central assumptions

This dissertation was not a priori committed to a specific framework. My intention is to base theoretical decision not only on – doubtlessly important – issues such as internal coherence, but also and centrally on the data that is the focal point of this work. Yet, during the course of my investigations into clitic placement, I have found that a constraint-based analysis has many advantages, which will become clear in the course of this thesis. More specifically, I will adopt an optimality theoretic (OT) model. Within this model, many common assumptions of government and binding theory and the Minimalist Program (MP) will be adopted.

I will include insights not only of previous OT analyses to clitic placement, but also those of recent MP approaches that recognise the importance of allowing a well-defined interaction of the different grammar modules in order to adequately model clitic positioning. These approaches, too, make use of constraints, which however only apply after syntax, in the phonological component. In contrast, the present approach captures not only phonology, but also morphology and syntax by using ranked, violable constraints that are all evaluated in parallel.

With respect to prosody, I will adopt a matching account, which entails that prosodic phrasing derives from, but is not identical to, syntactic phrasing. In order to avoid confusion of prosodic and syntactic constituency, I will use Greek letters for prosodic units:  $\omega$  for the prosodic word,  $\varphi$  for the phonological phrase, and  $\iota P$  for the intonational phrase (cf. for example Selkirk, 1986; Nespor & Vogel, 2007).

## 1.3 Sources for linguistic data

For the Czech data presented in this study, I partly relied on judgements found in the relevant literature. In addition to this, I regularly consulted with native speakers of Czech and Slovak, both to verify the judgements stated in previous works, and to obtain new data. Some simple examples I constructed myself, but then always asked for feedback from native speakers, usually through instant messaging or in videophone calls. Data from other languages was taken from the relevant literature.

In order to obtain a systematic data set, I conducted several acceptability judgement surveys. Before the beginning of the Covid-19 pandemic in March 2020, these were carried out in person. I interviewed native speakers of Czech both in Göttingen and in Prague. During the pandemic, I switched to online surveys, using the online survey tool *L-Rex*, which is being developed and maintained by Alexej Starschenko and Marta Wierzba since 2019 for the specific purpose of linguistic acceptability rating experiments.

For participant recruitment, I used *Prolific*, a platform for recruiting and managing participants for online surveys in all kinds of research fields. Through *Prolific*, participants are paid for each study they participate in. The platform allows the researcher to pre-screen participants for a large amount of features. In the Czech acceptability tasks, only people with Czech as their primary and only first language who were located in the Czech republic could participate. *Prolific* has around 130 members who fit this profile, therefore this strict filtering was possible. For Slovak, only filtering for Slovak as the first language yielded fewer than 50 eligible participants. Therefore, further filtering through *Prolific* was not used; however, the participants could be

further selected at a later point based on the collected metadata, the procedure for which is described below.

Whilst the in-person interviews had the advantage that I could note additional comments from the speakers and also answer their questions, the online surveys were of course less time-consuming (regarding both execution and evaluation) and thus yielded the greater amount of data. For this reason, I continued using *Prolific* and *L-Rex* when contact and travel restrictions were lifted.

In either modality, the introductory text made it clear that the speakers were to rely on their own intuitions about their daily language, not on prescriptive grammatical correctness (*Důležitá je vaše osobní intuice v běžném používání jazyka, ne gramatická pravidla*. ‘What matters is your personal intuition in common language use, not grammar rules.’). I also randomised the items, used fillers, and combined several questionnaires into one, in order to prevent the participants from discovering a pattern and to avoid effects of fatigue, priming etc.

The major studies of central importance for this thesis had over 60 participants each, whereas some further acceptability surveys were rather small, with respect to both the number of participants and the number of different lexicalisations for a given variable type. They therefore have the character of preliminary studies which point in a certain direction, but which should certainly not be understood as the final say on a certain data problem. This also applies to the Slovak data, which is limited by the small amount of native speakers available on *Prolific*. I believe that acceptability studies should be amplified in future research, as they greatly help to clarify the status of forms and constructions in contemporary Slavic languages. For considerations regarding the procedures and sample sizes for judgment data collection, see Schütze and Sprouse (2013).

For each of the major experiments, I calculated participant z-scores in order to level out the individual rating strategies of the participants (for example that some might tend towards the more positive or negative end of the scale). Whilst the questionnaires were always fully anonymous, they featured a final page which asked the participants for metadata to evaluate their linguistic background. I only included results of speakers which grew up monolingually in the Czech/Slovak republic, with Czech/Slovak as their only early language. Appendix B lists all surveys conducted for this dissertation and contains a link to the Open Science Framework (OSF) repository where the materials, results, and instructions of the experiments can be found.

In order to obtain naturally occurring examples and to get an impression of the frequency of certain configurations in Czech, I also used the SYN2015 corpus, a synchronous representative and referential corpus of contemporary (2010-2014) written Czech, with a size of 120,748,715 positions. It is part of the Czech National Corpus Project (*Český národní korpus*, ČNK), which is managed jointly by the Institute of the Czech National Corpus and the Institute of Theoretical and Computational Linguistics at Charles University, Prague.

Sources are always given directly below the examples when they are taken from the research literature or from the SYN2015 corpus. When no source is indicated, examples are from my questionnaire studies, which will be clarified in the running text; or, in very simple cases, I constructed examples myself and verified them by consulting a native speaker.

## 1.4 Structure

The two chapters following this introductory part provide the background for the two main aspects of this thesis: Chapter 2 gives general information on common assumptions about the definition, classification, and distribution of clitics. Chapter 3 then offers some basic insight into the grammar of Czech on the major levels of linguistic organisation.

After this, the central object of study, Czech second-position clitics, will be presented. In chapter 4, their lexical and morphophonological properties are described. Chapter 5 then turns the focus to the syntax of these elements, sketching their placement both in simple and in complex clauses. This concludes the foundational, descriptive part of this thesis.

As a basis for the theoretic discussion in the second half of the present dissertation, chapter 6 provides an overview of existing theories of clitic placement, both from approaches situated in the government and binding/Minimalist Program frameworks and such that adhere to assumptions of Optimality Theory. Then, in chapter 7, a closer investigation of the morphological and syntactic properties will yield a clearer picture of the nature of Czech clitic auxiliaries, and also allow us to draw more general conclusions about the nature of clitics and of different verb types in the language.

The analysis of the core phenomenon of this thesis, second-position placement, will be elaborated in chapter 8, encompassing mono-clausal placement as well as clitic climbing, and offering a discussion of the general principles that lead to the existence of second-position clitics in Czech and many other languages. Chapter 9 then presents the second half of the syntactic analysis, by demonstrating how the interaction of clitic placement with topicalisation can be modelled in a constraint-based account.

Finally, in chapter 10, I will sketch how the presented approach can be extended to further Slavic languages: two West Slavic languages, namely Slovak and Upper Sorbian, and one South Slavic language, Bosnian-Croatian-Montenegrin-Serbian (BCMS). Chapter 11 provides a conclusion and discussion of the results, as well as suggestions for worthwhile subjects of further research.

## Chapter 2

# Background I: Clitics

Since research on second-position clitics has been conducted for many decades now, starting with Wackernagel's (1892) work on clitics in Indo-European, a large body of knowledge has already been gathered. It is therefore reasonable to present a brief overview of these findings before beginning the investigation of Czech clitic placement phenomena. The present chapter first discusses the definition of clitics in section 2.1 and their classification in section 2.2. I will also comment on the role clitics play diachronically in section 2.3. Section 2.4 then sketches their categorial affiliations, whereas their cross-linguistic distribution is described in section 2.5.

### 2.1 Definition

What are clitics? This simple question defies a simple answer. There are several aspects to what makes up a clitic, and a clear line between clitics and other elements cannot always be drawn. Spencer and Luís (2012) even come to the conclusion that “ultimately there are no clitics. There are clitic-associated properties that tend to cluster together in typical cases, but these properties are in principle independent of each other” (Spencer & Luís, 2012, 220). However, the fact that they write this in their 360-page introductory work on clitics shows that obviously, this makes the study of clitics nonetheless worthwhile – on the contrary, it is part of the peculiar challenge that investigations of clitic behaviour provide.

The identification of clitics is complicated by the fact that they can be confused with two kinds of elements: They resemble full words in some respects, and affixes in others (cf. Spencer & Luís, 2012, 3). In their paper on the English negative *-n't*, Zwicky and Pullum (1983, 502-4) contrast clitics with affixes. Also, Zwicky (1977) investigates the properties of clitics in a range of languages. Based on these two seminal works' findings, I summarise the differences between clitics and affixes and those between clitics and full words in table 2.1 below.

	Affixes	Clitics	Full words
<b>Phonology</b>	form $\omega$ with adjacent word usually unstressed		form their own $\omega$ stressed
<b>Morphology</b>	cannot appear in isolation morphologically simple		can appear alone more complex
	selective w. r. t. host arbitrary gaps in affix- word combinations cannot attach to material containing clitics	not selective no arbitrary gaps can attach to material containing clitics	
<b>Morphosyntax</b>	strictly ordered among themselves simple distribution		more freedom complex distribution

Table 2.1: Clitics compared to affixes and full words

This overview shows that phonologically, clitics are like affixes, whereas morphologically, clitics mostly differ from affixes. Comparison with full words often fails due to the fundamental difference that unlike clitics, full words have no hosts. Also, a comparison with respect to morphosyntax is difficult, because what falls in the syntactic domain with full words is part of morphology when looking at affixes – the place that clitics occupy in this respect is one of the complex issues that I will discuss within this thesis. Also, it remains to be established whether clitics pattern with affixes or words regarding categorial affiliation: whilst full words belong to different grammatical categories, affixes are *syncategorematic* (cf. Zwicky & Pullum, 1983, 502).<sup>1</sup> A further property that distinguishes clitics from affixes is not included in the table because it spans across different linguistic levels: affixed words tend to display morphophonological and semantic idiosyncrasies; clitic groups do not (cf. Zwicky & Pullum, 1983, 504).

Due to their intermediate status, the use of expressions like *neither ... nor* is very common when defining clitics. This has been made explicit for example by Franks (2016), who shows that the most striking properties of clitics are in fact defects, i. e. the lack of certain properties: they are defective prosodically, semantically and syntactically (cf. Franks, 2016, 92). An interesting consequence of the fact that clitics are not selective with respect their hosts is Marantz’s (1988, 253-4)’s observation that “clitics involve a mismatch between the bracketing or structure motivated on semantic and syntactic grounds and the bracketing or structure motivated on phonological grounds” (Marantz, 1988, 253-4). As a simple example he gives the English sentence in 1, where the clitic future auxiliary is syntactically part of the VP, but prosodically attaches to the subject pronoun to its left; bracketing in 1a indicates syntactic constituency, whereas it indicates prosodic grouping in 1b.

- (1) a.  $[[I] \text{ } [[\textit{will} \text{ } [\textit{go to Milwaukee}]]]]$ .  
 b.  $[[I'U] \text{ } [\textit{go to Milwaukee}]]$ .

(cf. Marantz, 1988, 253)

<sup>1</sup>This might not be true of derivational affixes: given that they change the category of the stem they attach to, it is plausible to assume that they carry this categorial information. Inflectional affixes, in contrast, might arguably lack category features.

It is important to note that all the properties described here are that of *prototypical* clitics. This means that within the languages of the world, we find deviations from the described patterns. One of the major goals of the present thesis is to draw a clearer picture of the nature of Czech clitics, which involves a comparison between commonly assumed characteristics of clitics in general, and Czech ones in particular, with the linguistic reality displayed in the data.

## 2.2 Classification

There exist several ways of classifying clitics. Since most of these notions are widely used, I will shortly sketch them here. The focus of this dissertation is on second-position (2P) clitics in Czech, which belong into the groups of special clitics, phrasal clitics, and Wackernagel clitics. They are also commonly described as enclitics, but, as we will see, their phonological affiliation is in fact less clear.

### 2.2.1 Pro- vs. enclitics

Clitics are traditionally classified with respect to their phonological direction of attachment: the two most widespread types are *proclitics*, which form a prosodic unit with the *following* word, and *enclitics*, which attach to the *preceding* word (cf. Spencer & Luís, 2012, 1).<sup>2</sup> The general term *clitic* will allow us to discuss Czech clitic elements without making premature statements on their prosodic affiliation; as section 5.1.3 will show, determining this affiliation is indeed complex for Czech 2P clitics.

### 2.2.2 Special vs. simple clitics

The classic distinction of special and simple clitics introduced by Zwicky (1977) is often referred to, but not always accurately or completely. In fact, Zwicky (1977) actually makes a three-way distinction between special clitics, simple clitics and bound words. Also, these terms have partially shifted in meaning during the past decades, so I cite his original wording here:

- **Special clitics:** “An unaccented bound form acts as a variant of a stressed free form with the same cognitive and with similar phonological makeup” (Zwicky, 1977, 3). These forms “often show special syntax” and “it is unlikely that the weak forms are related to the strong ones by phonological rules of any generality” (Zwicky, 1977, 4).
- **Simple clitics:** “A free morpheme, when unaccented, may be phonologically reduced, the resultant form being phonologically subordinated to a neighboring word. Cliticization of this sort is usually associated with stylistic conditions” (Zwicky, 1977, 5). “These have ordinary syntax (the reduced forms occur in the same positions as the full forms) and ordinary phonology” (Zwicky, 1977, 6).
- **Bound words:** Morphemes that are “always bound and always unaccented show considerable syntactic freedom, in the sense that they can be associated with words of a variety

---

<sup>2</sup>Further prosodic clitic types are mesoclitics, which attach to a word’s stem, and endoclitics, which split up the word root (cf. Spencer & Luís, 2012, 207). These interesting items are, however, much rarer, and do not occur in the languages discussed in this thesis.

of morphosyntactic categories. Frequently, such a bound word is semantically associated with an entire constituent while being phonologically attached to one word.” (Zwicky, 1977, 6).

All three classes are unaccented; both special and simple clitics have free counterparts, whilst bound words do not. Special clitics and bound words differ from simple clitics in having “special” syntax and phonology. As examples for special clitics, Zwicky (1977) offers clitic pronouns in Romance and Slavic (Zwicky, 1977, 3); for simple clitics English reduced pronouns in casual speech (Zwicky, 1977, 5); and for bound words Latin *-que*, Tagalog particles, and English possessive *-s* (Zwicky, 1977, 6).

Whilst Zwicky’s (1977) observations are highly valuable, these frequently employed notions also have their problems. First of all, they are often not used in their original sense, and sometimes it is unclear what they are supposed to refer to. The fact that Zwicky (1977) only considers unaccented elements does not conform with the modern view on clitics, which places much less emphasis on this lack of accent, and even counts some accented items into the heterogeneous clitic family. Instead, special clitics are usually investigated with respect to their syntactic idiosyncrasies, which may or may not relate to phonological weakness (cf. Billings, 2002, 58).

Also, whether or not an element has a full counterpart – as in the distinction between special clitics and bound words – is usually of minor interest in the study of clitics. For example, Czech auxiliary clitics have no free counterparts, whilst Czech pronouns do – but otherwise they behave in the same manner, so assigning them to different categories does not lead to any gains in knowledge. An additional problem is that the question whether an element has a full counterpart cannot always be answered straightforwardly, given that special clitics may be phonologically very distinct from their full forms. It appears that distinguishing different types of elements solely on this factor is not particularly useful (cf. Spencer & Luís, 2012, 43).

### 2.2.3 Phrasal vs. head-adjacent clitics

Billings (2002) distinguishes phrasal clitics from head-adjacent clitics. Phrasal clitics have a “relative promiscuity of attachment”, whereas head-adjacent clitics appear with the verb in clauses, and with the noun in nominal expressions (cf. Billings, 2002, 55-6). Whilst phrasal clitics may be subject to Wackernagel’s law, head-adjacent clitics may be subject to the Tobler-Mussafia law, both of which I will sketch in the following paragraphs.<sup>3</sup>

There are at least two types of Wackernagel, i. e. 2P clitics: they may follow the first (phonological) word of their sentence, or the first constituent. Halpern (1995) calls these former ones *2W clitics*, the latter ones *2D clitics* (with D referring to the first syntactic daughter of the clause). Some languages allow their clitics to occur both in 2W and 2D, for example BCMS, Luiseño, Ngiyambaa and Warlpiri.; in Czech, clitics are always in 2D (cf. Halpern, 1995, 16). Although Wackernagel originally made his observations on clitics that occur after the first full *word* (cf. Wackernagel, 1892, 342), today the term “Wackernagel clitic” refers to all clitics that

<sup>3</sup>A third, rather unusual, type that should be mentioned here are ditropic clitics, which belong to a certain phrase syntactically, but their phonological attachment is oriented in the opposite direction. These are found for example in Vedic (Götz Keydana, p. c.) or in Kwak’wala (cf. Spencer & Luís, 2012, 66). Yet, example 1 above can also be viewed as containing a ditropic clitic.



are tied to the second position, also those after the first phrase (cf. Spencer & Luís, 2012, 41). I will generally use the term “2P clitic” for these elements.

The Tobler-Mussafia law, which affects head-adjacent clitics, was stated by A. Tobler and A. Mussafia in the late nineteenth century (1875/1886) for medieval Romance: Object clitics appear postverbally only when they would otherwise be initial (cf. Spencer & Luís, 2012, 64). As we will see, in Czech, the Tobler-Mussafia effect plays no role.<sup>4</sup> A Slavic language which has head-adjacent clitics that are subject to the Tobler-Mussafia law is Bulgarian (cf. Franks & Bošković, 2001, 175).

The demarcation from affixes is sharper for phrasal than for head-adjacent clitics. This is due to the fact that the former can appear with different types of hosts, whereas the latter are restricted to a certain type of head, for example the verb (cf. Billings, 2002, 57). The Tobler-Mussafia effect is one of the criteria that can be used to identify a head-adjacent element as a clitic. Further criteria are the lack of arbitrary gaps and the ability to attach to elements that already contain clitics, as sketched in table 2.1.

### 2.3 Diachronic role in grammaticalisation

The positioning of clitics between full words and affixes is reflected in the role they play diachronically. They “represent various stages in the processes of grammaticalization” (Spencer & Luís, 2012, 3), i. e. they are the result of a gradual shift from full words on one side of the scale to affixes on the other. For example, Zwicky and Pullum (1983) show that English negative *-n't*, although once a simple clitic (which itself had obviously derived from the full word *not*), is now an inflectional suffix (cf. Zwicky & Pullum, 1983, 510).

Since this shift is gradual, different types of clitics can also represent different stages of grammaticalisation. For example, Zwicky (1977) remarks that “special clitics are often the remnants of an earlier system of simple clitics” (Zwicky, 1977, 6); and then, “what is a clitic at one stage is reinterpreted as a derivational or inflectional affix at the next” (Zwicky, 1977, 6). I will not discuss the historic development of Czech 2P clitics within this thesis. Yet, it is important to keep in mind that clitics often represent the temporal results of a gradual grammaticalisation process, and in consequence their properties are often vague and varying.<sup>5</sup>

### 2.4 Categorical affiliations

Given their relative proximity to affixes, most clitics are function words. In the verbal domain, they express tense, aspect, mood, negation or agreement, among others (cf. Spencer & Luís, 2012, 14). They can also mark evidentiality, and, at the clausal level, polar questions. In the nominal domain, clitics express for example case, definiteness or possession (cf. Spencer & Luís, 2012, 19-20). There are also clitics which have no direct contribution to a clause’s interpretation, but instead demarcate prosodic boundaries (cf. Skopeteas, 2010, 312).

<sup>4</sup>Czech does have a head-adjacent clitic, the sentential negator *ne-*, but it appears preverbally without exception; see section 4.1.

<sup>5</sup>On the other hand, it must also be noted that there are many phrasal clitics for which a non-clitic history is not attested, for example the pronominal clitics found in Old Indo-European languages such as Ancient Greek (Götz Keydana, p. c.).

Among the most well-known cases of clitics are the pronominal ones, which can typically express person, number, gender, and case (cf. Spencer & Luís, 2012, 24); pronominal object clitics are also widespread in the Slavic languages, in addition to clitic auxiliaries (cf. Franks & King, 2000, 8-9). I will not go into more detail here. It has become clear that the functions clitics can take are manifold, similar to those of affixes. The ones relevant to this thesis are the functions we find in the Slavic languages, and in particular in Czech – chapter 4 presents the clitic inventory of Czech in detail.

A question that will be of great interest in this thesis is whether clitics really belong to the grammatical categories they resemble, or whether, like (inflectional) affixes, they do not. For example, we will see in chapter 7 that Czech auxiliary clitics are in many ways distinct from non-clitic auxiliary verbs.

## 2.5 Linguistic distribution

Literature on clitics is notoriously multilingual. For instance, the introductory book by Spencer and Luís (2012) discusses examples from around 130 different languages and the monograph by Anderson (2005) features almost 100 languages. This is not surprising given the fact that clitics occur in a large proportion of the languages of the world. Zwicky (1977) assumes that “most languages – very possibly, all except those of the most rigidly isolating type – have morphemes that [...] are neither clearly independent words nor clearly affixes” (Zwicky, 1977, 1).

With respect to 2P clitics, Halpern (1995) offers the following examples for languages that have clitics that are tied to a second position (be it 2D or 2W): Pashto, Tagalog, Warlpiri, Luiseño, Ngiyambaa, Sanskrit, and the Slavic languages; this list is of course far from complete. Within the Slavic language group, 2P placement is found especially in BCMS, Czech, Slovak, and Slovenian, whereas Bulgarian and Macedonian mostly have verb-adjacent clitics (cf. Franks & King, 2000, 216).

## 2.6 Summary

This section has presented generally held assumptions about the properties, the subgroupings and the distribution of clitics. We have seen that the lines between clitics and full words, on the one hand, and affixes, on the other hand, are blurred – this is an aspect that makes the study of clitics both challenging and interesting. We have also seen that traditional classifications of clitics are not always helpful for modern investigations of their behaviour. For the current study, it suffices to say that we focus on phrasal 2P clitics, and within this group mainly on 2D clitics. Other common classifications, such as the distinction of pro- and enclitics, will be subject to scrutiny within the course of this thesis.

## Chapter 3

# Background II: Basics of Czech Grammar

The purpose of this chapter is to give the reader, especially if unfamiliar with Czech, an impression of the Czech language at different levels of linguistic organisation: morphology, syntax, phonology, and information structure. This serves to provide a background for understanding the numerous language examples within the present thesis. I will also introduce some common theoretical assumptions about Czech that provide a basis for my analysis. Some of the grammatical properties introduced here will also become relevant in later parts of this work. Since these will then be briefly introduced again at that point, readers who are familiar with Czech grammar may harmlessly scan or skip the following sections. Before immersing into these issues, I will specify the variants of the Czech language that will be discussed.

### 3.1 A remark on different varieties of Czech

At first, a clarification of what is meant with the term *Czech* in the present work is necessary. The Czech language is divided into four main dialectal groups: Bohemian, Central Moravian (Haná), Moravian-Slovak, and Silesian (cf. Short, 1993a, 528). The Bohemian group is the most widespread of these, and is also found in the largest city, the capital Prague. In fact, the Bohemian dialects are largely levelled out (cf. Vintr, 1991, 83). The majority of speakers I interviewed for this thesis is from areas where this variant is spoken (cf. section 1.3), and it is also the colloquial variant mostly discussed in the literature.

A factor that is even more relevant in Czech is the distinction between standard and spoken language. The standard language, called *spisovná čeština* “Standard Czech”, was created based on the Prague variant of Czech. It is highly regulated and found mostly in written language. It differs from the colloquially spoken language, *obecná čeština* “Common Czech”, in all aspects of linguistic organisation. Between these two, there are many intermediate forms, so a clear line cannot be drawn (cf. Short, 1993a, 455-6). Examples from the quoted literature are mostly from Standard Czech. When a form or construction is only found in the colloquial language, this is made explicit.

## 3.2 Phonology

Czech has five vowels, *a*, *e*, *i*, *o*, and *u*, which can all be either long or short. The grapheme *y* has the same quality as the grapheme *i*, but only the latter entails palatalisation of the preceding consonant. There are no reduced vowels; the epenthetic vowel, used for example with inflection, is *e*. The liquid consonants *l* and *r* can be syllable nuclei (cf. Sussex & Cubberley, 2006, 153-4).

Czech main word stress always falls on the leftmost syllable. Monosyllabic prepositions form a prosodic unit (i. e.  $\omega$ ) with the following word and thus carry main stress (cf. Vintr, 1991, 75). However, within  $\varphi$ 's, the “phonic centre is always placed on the last word” (Daneš, 1957, 141). For the entire  $\iota P$ , the rightmost  $\varphi$ -phrase is the most prominent one, and it is thus the default location of the sentence's focus (cf. Daneš, 1957, 141).<sup>1</sup> This systematic assignment of main prominence at the prosodic levels is schematically illustrated in 2. The effects of information structure on Czech word order and prosody will be discussed in section 3.5.

	$\zeta \zeta$	$\zeta \zeta \zeta \zeta$	$\zeta \zeta \zeta$	$\zeta \zeta \zeta$
(2)	$\omega$	( x )	( x )	( x )
	$\varphi$	( x )	( x )	( x )
	$\iota P$	( x )	( x )	( x )

## 3.3 Morphology and morphosyntax

Like all Slavic languages, Czech has a very rich morphology, employing both prefixation and suffixation in the nominal as well as in the verbal domain. In order to provide easily accessible and to-the-point examples, not all morphological categories will be marked in the glosses; for example, gender and aspect are usually not indicated, except when directly relevant to the discussion.

### 3.3.1 Nominal categories

Czech nominal morphology distinguishes three grammatical genders, and within the masculine, it also distinguishes animacy. The dual only survives idiosyncratically; only two number categories, singular and plural, are productive (cf. Sussex & Cubberley, 2006, 223). There are seven cases: nominative, genitive, dative, accusative, vocative, locative, and instrumental. Nominal categories, i. e. nouns, adjectives and pronouns, inflect in all these dimensions (cf. Sussex & Cubberley, 2006, 222). The Czech language has no articles, but the common language employs the demonstrative *ten* ‘this, that’ to mark definiteness on a noun phrase (cf. Sussex & Cubberley, 2006, 560).

---

<sup>1</sup>Daneš (1957) actually uses a different terminology: The “stress-unit” is the lowest level of prosodic organisation, and can be equated with  $\omega$ . One level above lies the “rhythm-unit”, which corresponds to  $\varphi$ . The highest level is the “utterance”, which appears to correspond more to  $\iota P$  than to *utterance* in its modern usage (cf. Daneš, 1957, 139-40). Also, he only distinguishes *rheme* and *theme*, whilst I assume a focus-background structure and a topic-comment structure, as elaborated in section 3.5 (cf. Daneš, 1957, 141).

### 3.3.2 Verbal categories

Verbs show number and person agreement, as well as tense and aspect; participles<sup>2</sup> additionally inflect for gender (cf. Short, 1993a, 464-5, 480-1). Commonly used tenses are present tense, past tense (perfect), and future tense. The tense system was simplified around 1400, due to the rise of grammatical aspect, resulting in the loss of the non-periphrastic tenses aorist and imperfect (cf. Vintr, 1991, 85).

The past tense is formed using the clitic auxiliary derived from the present tense of *být* ‘to be’ and the *l*-participle of the main verb. The *l*-participle is also used when forming the conditional, combined with the former aorist forms of *být*, now the clitic conditional auxiliary (cf. Vintr, 1991, 81). Examples for both past tense and conditional mood (in present tense) are given in 3.<sup>3</sup>

- (3) a. *Spala jsem.*  
 sleep.PTCP PASTAUX.1SG  
 ‘I slept.’
- b. *Spala bych.*  
 sleep.PTCP CONDAUX.1SG  
 ‘I would sleep.’

Future formation functions differently for the two different verbal aspects, imperfective and perfective. For imperfective verbs, the future tense is built using the non-clitic future auxiliary and the infinitive of the main verb. With perfectives, the simple, morphologically present tense form expresses future, as the contrast between 4 and 5 illustrates (cf. Naughton, 2005, 150-2).

- (4) Imperfective:
- a. *Kupujeme pivo.*  
 buy.1PL beer.ACC  
 ‘We are buying beer.’/‘We buy beer.’
- b. *Budeme kupovat pivo.*  
 AUX.FUT.1PL buy.INF beer.ACC  
 ‘We will be buying beer.’/‘We shall buy beer.’
- (5) Perfective:
- a. *Koupíme pivo.*  
 buy.1PL beer.ACC  
 ‘We will buy beer.’

<sup>2</sup>Short (1993a, 487) writes that the past tense participle, which always ends in *-l*, “should perhaps not be called a participle now at all.” This is an observation that will become relevant in the course of this thesis.

<sup>3</sup>Whilst theoretically possible, the conditional mood is rarely combined with the past tense. Speakers prefer to use the morphosyntactically simpler present conditional, and the past tense meaning is inferred pragmatically.

b. \**Budeme koupit pivo.*

AUX.FUT.1PL buy.INF beer.ACC

In addition to perfective and imperfective aspect pairs, a frequentative can be derived from imperfective verbs through use of the suffix *-va*. This thus yields three different aspectual categories for one lexical meaning, for example for the verb *give*: *dát* (pf.) – *dávat* (ipf.) – *dávávat* (freq.; cf. Vintr, 1991, 81). This distinction is going to be relevant later in the discussion, in section 7.1.4.

Concerning voice, Czech has a passive construction combining an auxiliary with a short form<sup>4</sup> passive participle. The passive participle agrees with the subject in gender and number, as does the above-mentioned *l*-participle. The auxiliary can be in present, future, or, as in example 6, past tense. This periphrastic passive developed through Latin influence in the 16<sup>th</sup> century (cf. Vintr, 1991, 85). In spoken Czech, the use of the morphological passive is however avoided. Instead, information-structural reordering is exploited, or a reflexive passive is used (cf. Naughton, 2005, 163).

(6) *Báseň byla napsána Seifertem v roce 1922.*  
poem be.PASTPTCP write.PASSPTCP Seifert.INS in year 1922

‘The poem was written by Seifert in 1922.’

(cf. Naughton, 2005, 163)

## 3.4 Syntax

The following subsections present a few basic facts about Czech syntax, and thus about the order, presence or absence, and movement of constituents. Since the syntax of the language is highly sensitive to information structure, more detail on this issue can be found in section 3.5.

### 3.4.1 Word order

Word order is relatively free in Czech. Non-clitic constituents can in principle appear in any position within the clause, though this is in fact highly regulated based on prosody and/or information structure, as described in section 3.5 below. Second-position clitics are thus exceptional in Czech in the sense that they occupy a strictly fixed position (cf. Short, 1993a, 495).

Despite this freedom, Czech is typically considered an SVO language, like most other Slavic languages, as can be seen for example in the *World Atlas of Language Structures* (WALS; cf. Dryer, 2013b), which bases its classification on the frequency of a specific order in texts (cf. Dryer, 2013a). In their analysis of a Czech corpus including roughly 30,000 clauses, out of which 6101 were transitive, Siewierska and Uhliřová (1998) find 63.1 % of SVO; the second most frequent order is OVS with only 14.6 %. SVO is thus found in almost two-thirds of transitive

<sup>4</sup>The long forms of the participle, i. e. those ending on a long vowel (here *-é*), are used as adjectives or predicates; the short forms are more rare in these contexts (cf. Naughton, 2005, 160).

sentences, and more than twice as frequent as the next most frequent order, thereby meeting the WALS criterion for being counted as dominant (cf. Siewierska & Uhliřová, 1998, 123).

Yet, the categorisation of Slavic languages as VO has also been disputed, for example by Szucsich and Haider (2015) and Haider and Szucsich (2022). They argue that properties that are typically associated with VO languages in the Germanic and Romance groups cannot be found in the Slavic languages. Alternatively, the authors propose that Slavic belongs to a third type of languages, where the directionality of VP (i. e. whether the object is merged to the left or to the right of the verb) is underspecified (cf. Szucsich & Haider, 2015, 96-7).

In reply to Haider and Szucsich (2022), Šimík and Jasinskaja (2022) investigate properties that are typically associated with VO languages in Czech and find that the language meets four out of these seven criteria. Thus, even if it is true that the Slavic languages in general deviate from the typical VO pattern, Czech as an individual language does so to a much lower degree.

There are two additional factors that are of importance in this discussion. One is a point raised by Fanselow (2020): It is far from clear that the features used as criteria for distinguishing between VO and OV base order are indeed related to the position of these elements – a promising alternative candidate for some of them is the position of the *subject*, i. e. whether there is a VP-external functional subject position or not (cf. Fanselow, 2020, 10). This issue still requires further investigation which goes beyond the scope of this dissertation.

A second point, however, is that in Czech, VO orders can be used in an all-new sentence, whereas OV orders cannot, as example 7 illustrates. The latter order requires a special information-structural motivation, i. e. the rightmost element must be in focus. This is also confirmed in an experimental study by Šimík and Wierzba (2017), where  $S > V > O > PP$  and  $S > V > PP > O$  score significantly better than  $O > S > V > PP$  and  $S > O > V > PP$  in all-new sentences (cf. Šimík & Wierzba, 2017, 689). The effects of information structure on word order are discussed in detail in section 3.5.

(7) Context: *What's new?*

- a. *Marek ztratil mobil.*  
 Marek.NOM lose.PTCP mobile.phone.ACC  
 ‘Marek lost his mobile phone.’
- b. *#Marek mobil ztratil.*  
 Marek.NOM mobile.phone.ACC lose.PTCP

Whilst Szucsich and Haider (2015) and Haider and Szucsich (2022), but also Fanselow (2020), raise important issues concerning the criteria used for the categorisation of languages as VO or OV, the frequency and neutrality of subject-verb-object order still make an interpretation of Czech as underlyingly SVO plausible. I will therefore adopt this view in the present thesis.

In addition to information structure, question formation also affects word order. Interrogative phrases in content questions usually appear in initial position.<sup>5</sup> Fronting of multiple

<sup>5</sup>In the present thesis, I avoid the expressions *wh*-question, *wh*-movement etc., for the simple reason that interrogative phrases do not begin with these letters in Czech. One might instead refer to them as *kd*-phrases (*kdo* ‘who’, *kde* ‘where’, *kdy* ‘when’), but this would not fit the entire class of interrogative elements, which also features *co* ‘what’ and *proč* ‘why’, among others.

interrogative phrases is possible, and their respective order is relatively free, as shown in 8. In polar (*yes/no*) questions, the verb is optionally fronted, as in 9; if it is not, intonation marks the sentence as a question (cf. Short, 1993a, 496).

- (8) a. *Kdo co doporučil komisi?*  
 who.NOM what.ACC recommend.PTCP committee.DAT  
 b. *Co kdo doporučil komisi?*  
 what.ACC who.NOM recommend.PTCP committee.DAT  
 ‘Who recommended what to the committee?’

(cf. Szucsich & Haider, 2015, 116)

- (9) *Zabil Petr Pavla?*  
 kill.ptcp Petr.NOM Pavel.ACC  
 ‘Did Peter kill Pavel?’

(cf. Short, 1993a, 496)

I will assume that apart from phenomena such as interrogative fronting and clitic movement, the core constituents of a clause do not undergo obligatory movements. For Russian, this has been argued for example by Junghanns and Zybatow (1997, 298-9), and given the extensive word order freedom in Czech, I assume that it can be applied here as well. Thus, unlike in English, the subject is not required to move to the specifier of TP, but can remain low, without an expletive element having to occupy this position; compare the Czech sentences in 10 below to their English translations.

- (10) a. *Štěkali psi.*  
 bark.PTCP dogs.NOM  
 ‘Dogs barked.’  
 b. *\*There barked dogs. / \*Barked dogs.*  
 c. *Přišel muž.*  
 come.PTCP man.NOM  
 ‘A man came.’  
 d. *There came a man. / \*Came a man.*

A necessary exception to non-obligatory movement is the finite verb. I will follow Dvořák (2010) in her assumption that the finite verb always moves from V to v. Thus, in contrast to the infinitive, it does not form a constituent with the direct object (cf. Dvořák, 2010, 11); it does not need to move higher though (cf. Veselovská, 1995, 57-61). The necessity of this assumption will become evident in section 9.2 with respect to topicalisation data.



### 3.4.2 Pronouns, copulas, negation

Like many of the Slavic languages, but not for example Russian, Czech is pro drop, in the sense that referential pronominal subjects are omitted when they are not emphatic (cf. Fehrmann & Junghanns, 2012, 65-6). Also in contrast to Russian, Czech nominal predicates require a copula (cf. Vintr, 1991, 82), which is, despite its similarity to the clitic auxiliary, not a 2P clitic – an issue which will be an important aspect of the present thesis. The different pro-drop and copula properties of Czech and Russian yield the double contrast between the two languages shown in example 11.

- (11) a. Czech:  
*Je nemocen.*  
 COP.3SG ill
- b. Russian:  
*Он болен.*  
 he ill  
 ‘He is ill.’

(cf. Vintr, 1991, 82)

For sentential negation, the negative prefix *ne-* is attached to the finite verb (exceptions regarding clitics will be discussed below). Czech is a negative concord language, i. e. negative indefinite pronouns occur together with sentential negation, yielding a simple negated interpretation, as shown in 12. In contrast to Russian, the genitive of negation has disappeared in Czech (cf. Vintr, 1991, 82).

- (12) *Nikdo nic neslyšel.*  
 nobody nothing NEG:hear.PTCP  
 ‘Nobody heard anything.’

(cf. Vintr, 1991, 82)

### 3.4.3 Binding and extraction

Binding of reflexives is highly subject-sensitive, i. e. reflexives are bound by the nominative element within their clause, as can be seen in 13a. This is also true within subordinated clauses, where the reflexive is bound by the subject of its own clause, even if it is not overtly present, as 13b shows (cf. Lenertová, 2004, 164). What applies to the reflexive possessive in these examples is also true for the clitic reflexive pronouns *se* and *si*.

- (13) a. *Martin<sub>1</sub> Petra<sub>2</sub> představil své<sub>1/\*2</sub> kolegyni.*  
 Martin.NOM Peter.ACC introduce.PTCP POSSREFL colleague.DAT  
 ‘Martin introduced Peter to his colleague.’

- b. *Matka<sub>1</sub> mu<sub>2</sub> zakázala PRO<sub>2</sub> dát ten dopis své<sub>2</sub>/<sub>\*1</sub> ženě.*  
 mother him forbid.PTCP give.INF that letter POSSREFL wife.  
 ‘Mother forbade him to give the letter to his wife.’ (not: ‘... to her wife.’)

(cf. Lenertová, 2004, 164)

In his detailed study of content questions in Russian, Polish, and Czech, Meyer (2004) shows that Czech permits both long interrogative movement and long topicalisation; examples are shown in 14. Yet, restrictions exist, for example with respect to the type of matrix verb or whether the embedded sentence is a subject or object clause (cf. Meyer, 2004, 190-3).

- (14) a. *Koho byste řekli, že jsem sehnal koho na záskok?*  
 who.ACC COND.2PL say.PTCP COMP AUX.1SG get.PTCP on replacement  
 ‘Whom would you say that I got as a replacement?’
- b. *K tomu nevím, co bych řekl k tomu.*  
 to that NEG:know.1SG what COND.1SG say.PTCP  
 ‘To that I don’t know what I would say.’

(cf. Meyer, 2004, 194)

### 3.5 Information structure

It is often said that Czech, like most other Slavic languages, displays great freedom of word order. This is only partially true. Many orderings are in principle grammatical, but strongly context-dependent. It is thus not syntactic dependency that regulates constituent order, but rather information structuring – hence they can be described as “discourse-configurational” (cf. Šimík & Wierzba, 2017, 672).

Since the left periphery is especially relevant for information-structure related movements, we need to clarify its composition. I will follow Fehrmann and Junghanns (2012) in assuming the structure as given in 15. Topicalisation occurs when the Fin head contains a topic feature; but also leftward moved foci can appear here, hence no designated topic and focus positions (as in the style of Rizzi, 1997) are assumed.

- (15) [CP C [FinP Fin [TP ... ]]]

Before describing the interaction between Czech word order and information structuring, it is necessary to clarify the relevant notions. Following Junghanns (2002c, 15) and Fehrmann and Junghanns (2012, 75), I will assume that information structure has two dimensions: topic-comment structure and focus-background structure. I will use *topic* in the sense of what the sentence is about, and *focus* as representing the important information (cf. Junghanns & Lenertová, 2007, 234-5). In a sentence, everything which is not topic is *comment*, and everything which is not focus is *background*. Focus-background and topic-comment can be represented by F

and T features assigned to syntactic nodes, which correlate with specific prosody (cf. Lenertová & Junghanns, 2007, 348). Note that neither topic nor focus can be defined solely with reference to givenness (cf. Lenertová & Junghanns, 2007, 348-9).

### 3.5.1 Focus-background structure

One of the most striking features of information-structurally determined ordering is what Lenertová and Junghanns (2007) call the “default principle of focus realisation” (DPFR), which states that “the focus exponent appears at the right periphery of the clause” (Lenertová & Junghanns, 2007, 349). This is true also for wide focus sentences as the ones in 16. However, as 17 shows, the DPFR only applies to non-contrastive, i. e. presentational, foci. Contrastive foci do not occupy a specific position in the clause.<sup>6</sup> The syllable carrying the main stress is marked by small capitals.

- (16) *Proč jede Karel na dovolenou zrovna tam?*  
 why go.3SG Karel on holiday exactly there

‘Why does Karel go exactly there on his vacation?’

- a. *Poradil mu to jeho TErapeut.*  
 advise.PTCP him.DAT it.ACC his therapist

‘His therapist advised him to do it.’

- b. *Slíbil to své ŽEně.*  
 promise.PTCP it.ACC his.DAT wife.DAT

‘He promised it to his wife.’

(cf. Lenertová & Junghanns, 2007, 347)

- (17) *Slyšel jsem, žeš ztratil peníze.*  
 hear.PTCP AUX.1SG, COMP:2SG lose.PTCP money.ACC

‘I’ve heard that you lost your money.’

- a. *Ne, KLÍče jsem ztratil.*  
 no keys.ACC AUX.1SG lose.PTCP

- b. *Ne, ztratil jsem KLÍče.*  
 no lose.PTCP AUX.1SG keys.ACC

‘No, I lost my keys.’

(cf. Lenertová & Junghanns, 2007, 350-1)

<sup>6</sup>Neither does *verum focus*, which is realised on (non-clitic) verbs or on verificational sentential adverbs (cf. Junghanns & Lenertová, 2007, 246).

But also when only considering presentational foci, (superficial) exceptions to the DPF<sub>R</sub> occur: one is the phenomenon of so-called “rightward backgrounding”, which produces left-peripheral minimal foci (cf. Junghanns & Lenertová, 2007, 244), see 18a. Another one is the optionality of a left-peripheral focus exponent – i.e. the constituent which carries the main stress – in all-focus clauses, see 18b. This renders the sentence more emphatic, but does not change focus-background structure (cf. Lenertová & Junghanns, 2007, 347).

- (18) a. What did you buy him?

*KNihu jsem mu koupila.*  
book.ACC AUX.1SG him.DAT buy.PTCP

‘I bought him a book.’

(cf. Lenertová & Junghanns, 2007, 351)

- b. Why does Karel want to go there?

*Své ženě to slíbil.*  
his.DAT wife.DAT it.ACC promise.PTCP

‘He promised it to his wife.’

(cf. Lenertová & Junghanns, 2007, 347)

Inthetic clauses – i.e. sentences with one-argument predicates, all-focus and no (overt) topic – both unmarked V>NOM and marked NOM>V occur. They are licit and interpreted as all-focus as long as the subject carries the focus accent (cf. Junghanns & Lenertová, 2007, 247-8). Here, too, the structure with the fronted focus exponent, sentence 19b, is the more emphatic one (cf. Lenertová & Junghanns, 2007, 352).

- (19) a. *Umřel ředitel.*

die.PTCP headmaster.NOM

- b. *Ředitel umřel.*

headmaster.NOM die.PTCP

‘The headmaster has died.’

(cf. Lenertová & Junghanns, 2007, 352)

Lenertová and Junghanns (2007) conclude that focus exponents can generally be fronted in maximally focused sentences (cf. Lenertová & Junghanns, 2007, 353). They show that this movement is syntactic, because it is blocked by strong islands like adjunct clauses, and assume that it is caused by a general emphasis operator. However, focus exponent fronting is also clearly restricted by prosodic factors – namely, it cannot cross stressed material (cf. Lenertová & Junghanns, 2007, 358-9). This is due to the prosody of all-focus sentences with a fronted focus exponent: after the focus accent, the contour flattens (cf. Lenertová & Junghanns, 2007, 356). Thus, focus exponent movement is only available when the other elements in the clause can be prosodically integrated (cf. Lenertová & Junghanns, 2007, 360).

### 3.5.2 Topic-comment structure

In contrast to foci, Czech aboutness topics typically appear at or near the left edge of the clause. Also note that, whereas every sentence has a focus, not every sentence has a topic. In sentences with canonical subject-verb order, the (non-emphatic) subject is typically background and topic (cf. Fehrmann & Junghanns, 2012, 83). An example for an internal topic is the initial accusative object of the sentence in 20. Internal topics are assumed to be higher than TP, but below C; in a structure as in 15, they are thus located in Spec-FinP. External topics are assumed to adjoin to CP and have a co-referential element within the clause, as 21 illustrates: the topic *plevel* is co-referential with the demonstrative pronoun *ten* (cf. Junghanns & Lenertová, 2007, 236). The structural annotations in the following examples are my own.

- (20) [<sub>FinP</sub> *Tuto reklamní plochu Vám pronajme:*] *reklamní*  
 this.ACC advertising.ACC space.ACC you.DAT rent.3SG advertising.NOM  
*agentura RENCAR a.s.*  
 agency.NOM RENCAR JSC

‘This advertising space can be rented from RENCAR agency.’

(cf. Junghanns & Lenertová, 2007, 235)

- (21) [<sub>CP</sub> *Plevel<sub>1</sub>*, [<sub>CP</sub> *ten<sub>1</sub> se nedá tak snadno vyhubit.*]]  
 weed.NOM that.NOM REFL NEG:let.3SG so easily eliminate.INF

‘Weed is not so easy to get rid of.’

(cf. Junghanns & Lenertová, 2007, 236)

Verb-subject inversion, however, is not an instance of verb topicalisation (or verb focus). Instead, it occurs when a subject is background (i. e. not focus), and comment (i. e. not topic); the subject then occupies its canonical position in Spec-TP. In this context, the verb moves out of the focused VP (cf. Fehrmann & Junghanns, 2012, 76-7). An example is given in 22, where Voříšek, a dog, has been mentioned before, but is not the topic.

- (22) [<sub>FinP</sub> *Leží*] [<sub>TP</sub> *Voříšek v mokré trávě...*]]  
 lie.3SG Voříšek.NOM in wet.LOC grass.LOC

‘Voříšek lies in the wet grass...’

(cf. Fehrmann & Junghanns, 2012, 74)

### 3.5.3 Givenness

As noted, topic-comment structure and focus-background structure must be kept apart from givenness. A well-established definition of givenness is that it is “knowledge which the speaker assumes to be in the consciousness of the addressee at the time of the utterance” (Chafe, 1976, 30). How different word orders (assuming neutral intonation, i. e. with the main accent on the last word) yield different interpretations of phrases as new or given is illustrated in example 23.

- (23) a. Basic word order (SVO):  
*Chlapec našel lízátko.*  
 boy.NOM found lollipop.ACC  
 → Possible interpretations:  
 boy = new, lollipop = new  
 boy = given, lollipop = new  
 boy = given, lollipop = given
- b. Derived word order (OVS):  
*Lízátko našel chlapec.*  
 lollipop.ACC found boy.NOM  
 → Possible interpretation:  
 boy = new, lollipop = given

(cf. Kučerová, 2012, 2)

What we see here is that movement of the object out of its base position yields a givenness interpretation that was not available before. In addition, the interpretations possible in the basic word order get lost. Movement is thus a way to mark the givenness status of constituents (cf. Kučerová, 2012, 2-3).

Kučerová (2012) proposes that the different interpretations regarding givenness in the above examples are due to a partition dividing an utterance into a given and a new part. This partition is caused by a so-called G(ivenness)-operator that marks everything in its scope, i. e. all elements that are syntactically higher, as given (cf. Kučerová, 2012, 5). In contrast, Šimík, Wierzba and Kamali (2014) assume that movement of given elements is simply a consequence of the fact that these elements avoid bearing nuclear stress (cf. Šimík et al., 2014, 430-1).

In an elaborate experimental study, Šimík and Wierzba (2017) show that givenness-induced constituent movement in Czech is indeed mainly mediated through prosody – a general bisection of the sentence in a given and a new part cannot be confirmed (Šimík & Wierzba, 2017, 700), although part of the data can be explained by the preference for a given-before-new ordering in Czech (Šimík & Wierzba, 2017, 693).<sup>7</sup>

<sup>7</sup>The authors investigate three West Slavic languages, Czech, Slovak, and Polish, and find that in all three languages, a ban on stressed given expressions is active. The strategies to accommodate this are however different: whilst Polish makes use of stress shift, Czech and Slovak prefer reordering of constituents (Šimík & Wierzba, 2017, 700).

### 3.6 Summary

The previous sections have presented some basic information about Czech phonology, morphology and syntax, and their interaction with information structure. The Czech language has a rich nominal and verbal morphology, yet a simplified tense system. We have seen that there is great word order variation driven by discourse-related factors such as givenness, topic-comment structure, and focus-background structure. The exploitation of this capacity for reordering is preferred over operations such as passive formation.

## Chapter 4

# The Clitic Inventory of Czech

The following sections serve the purpose of introducing the linguistic entities that make up the empirical scope of the present thesis, i. e. the types of Czech clitics that will be investigated. Section 4.1 provides some terminological clarifications: the classificatory systems of Czech clitics by Avgustinova and Oliva (1995) and by Junghanns (2002b) are presented and compared, which allows us to extract the unambiguous 2P clitics most relevant for this thesis. Then follows, in section 4.2, a summary of the properties that actually make these elements clitic. Section 4.3 describes the ordering of 2P clitics with respect to each other; deviations from the template are introduced in section 4.4. Section 4.5 then provides a concluding summary.

### 4.1 Types of Czech clitics

Czech has two main types of clearly special clitics in the sense of Zwicky (1977): clitic auxiliaries and clitic pronouns. Other clitic elements include clitic negation, a clitic complementiser and (potentially) clitic adverbs. Clearly simple clitics in the sense of Zwicky (1977) occur among prepositions and conjunctions – these do not display any syntactic idiosyncrasies (cf. Avgustinova & Oliva, 1995, 8).

Both Avgustinova and Oliva (1995) and Junghanns (2002b) provide a classification of Czech clitics, and I will lay out and compare their systems in the following paragraphs; whilst the systems are not identical, their basic insights into the organisation of the Czech clitic inventory are compatible and provide a helpful terminological basis.

Junghanns (2002b) differentiates between *lexical* and *phonological* clitics. The former are necessarily clitic, both prosodically and syntactically, and constitute a closed class. The latter can optionally be prosodically incorporated into their host; however, phonological clitics have no special syntactic properties (cf. Junghanns, 2002b, 139-40). Avgustinova and Oliva (1995) distinguish *pure* and *semi* clitics: the former always appear in second position, the latter do so only optionally. Whilst Junghanns's (2002b) lexical clitics also include the non-2P clitics *-li* and *ne-*, Avgustinova and Oliva (1995) list them as *verbal*, since their pure and semi clitics only include phrasal (2P) clitics (cf. Avgustinova & Oliva, 1995, 9-12). Table 4.1 shows how the Czech clitic inventory is grouped within the two systems.<sup>1</sup>

---

<sup>1</sup>The differences between the two systems are most striking concerning phonological/semi clitics. In this respect, Junghanns (2002b) objects to Avgustinova and Oliva's (1995) classification in several points. Since these



lexical		phonological	
pure	verbal		semi
Pronouns: <i>mi, tě, ti, mu, ho</i>	<i>ne-</i>	Other unstressed pronouns	
Reflexives: <i>se, si</i>	<i>-li</i>	Short adverbs	
Perfect auxiliaries: <i>jsem, jsi, jsme, jste</i>		Copular forms of <i>být</i>	Passive auxiliaries
Conditional auxiliaries: <i>bych, bys, by, bychom, byste</i>		Complementisers	
		Possessive pronouns	
		...	

Table 4.1: Classification of Czech clitics

Only the elements that are both lexical in Junghanns’s (2002b) sense and pure in Avgustinova and Oliva’s (1995) sense are unambiguously 2P clitics. They thus constitute the focus of this work. They contrast with phonological and semi clitics, which are not tied to the second position. In what follows, I elaborate on the different groups of lexical clitics in Czech, and will also provide a note on short adverbs.

Throughout most of this thesis, only the properties of lexical pure clitics will be relevant. However, whenever differentiation is necessary, I will employ the terminology of Junghanns (2002b): the elements that constitute the focus of this work are thus called *lexical* clitics, whilst only optionally cliticising elements are *phonological*.<sup>2</sup> In all examples within this thesis, only 2P lexical clitics will be marked by bold type, to aid their identification and to avoid loading the text with explanations of the clitic status of each element in every given example.

#### 4.1.1 Pronouns

Only accusative, genitive and dative singular pronouns are lexical clitics. In contrast, locative and instrumental pronouns often occur after prepositions, most of them are bisyllabic, and they have no weak counterparts, thus they can be stressed. Since Czech is a pro-drop language, nominative pronouns only ever surface when stressed. Thus, the accusative, genitive and dative pronouns are the only clitic personal pronouns. Additionally, in the third person, only the masculine/neuter forms are lexical clitics, whilst the feminine forms are not. All unambiguously clitic pronouns are shown in table 4.2 (Junghanns, 2002b, 143).

	ACC/GEN	DAT
1SG		<b>mi</b>
2SG	<b>tě</b>	<b>ti</b>
3SG.M/N	<b>ho</b>	<b>mu</b>
REFL	<b>se</b>	<b>si</b>

Table 4.2: Unambiguously clitic pronouns of Czech

types of clitics are, however, not the focus of this work, I will not review his arguments here (cf. Junghanns, 2002b, 127).

<sup>2</sup>Also note that Czech lexical clitics are *phrasal clitics* in Billings’s (2002) terminology, presented in section 2.2. Yet, phonological and semi clitics can also sometimes function as phrasal clitics. Whilst Billings’s (2002) classification is useful, the analysis of Czech 2P clitics requires more detail.

These are also the forms that Avgustinova and Oliva (1995) classify as pure pronominal clitics, i. e. that always show clitic behaviour. This means for example that they cannot be focused; most of these pronouns thus have strong counterparts which are used when for example contrastively stressing the pronoun.<sup>3</sup> However, some other pronouns do not have different clitic and non-clitic forms, but only one form, which can, but must not, behave like a clitic. They belong to Avgustinova and Oliva's (1995) semi clitics, and under this category they list the lexical elements shown in table 4.3 (cf. Avgustinova & Oliva, 1995, 9).

	ACC	GEN	DAT
1SG	mě	mě	mně
2SG			
3SG.F	ji	jí	jí
1PL	nás	nás	nám
2PL	vás	vás	vám
3PL	je	jich	jim

Table 4.3: Ambiguously clitic/non-clitic pronouns of Czech

Whilst the forms in table 4.2 always display the properties of clitics listed in section 4.2 below, such as inability to appear sentence-initially and to carry contrastive stress, the pronouns in table 4.3 may appear unstressed in second position in one sentence but be placed, for example, sentence-finally and be contrastively stressed in another. For this reason, the lexical clitics listed in table 4.2 are much better suited for the present investigation, as their status as clitics is always clear. As an example of a pronominal semi- (i. e. phonological) clitic which occurs inside the clitic cluster, Avgustinova and Oliva (1995) provide sentence 24, in which the phonological clitic *jí* is clearly part of the cluster, since it is flanked by pure (i. e. lexical) clitics.

- (24) *Jan se jí ho včera pevně rozhodl dát k narozeninám.*  
 Jan REFL her.DAT it.ACC yesterday firmly decide.PTCP give.INF to birthday

‘Jan yesterday firmly decided to give it to her for her birthday.’

(cf. Avgustinova & Oliva, 1995, 10)

Another pronominal element with both clitic and non-clitic uses is the neuter demonstrative pronoun *to*, which can replace the third person singular pronoun *ho*, and also serves as a sentential pronoun. Like the pronouns in table 4.3, it is thus a phonological clitic. In sentence 25, it is clearly de-stressed and directly follows the lexical clitics *jsem* and *mu*. However, whether it is actually part of the clitic cluster, or just follows it, cannot be determined.

- (25) *Včera jsem mu to dal.*  
 yesterday AUX.1SG him.DAT IT.ACC give.PTCP

‘Yesterday I gave it to him.’

<sup>3</sup>The corresponding strong forms are: *tě – tebe; ti – tobě; ho – jeho; mu – jemu; se – sebe; si – sobě*. Instead of *mí*, the semi clitic *mně*, included in table 4.3, is used (cf. Naughton, 2005, 78-9).

(cf. Veselovská, 1995, 287)

### 4.1.2 Auxiliaries

Clitic auxiliaries in Czech all derive from the verb *být*, ‘to be’. However, not all auxiliaries derived from this verb are equally clitic. The forms beginning in *j(s)-*, which are the present tense forms of *být*, are only lexical clitics when used as past auxiliaries. When instead they serve as passive auxiliaries or as copula, they are only optionally clitic; as noted above, passive auxiliaries are classed as semi clitics by Avgustinova and Oliva (1995),<sup>4</sup> and copulas as phonological clitics by Junghanns (2002b). The other lexical clitics of the *být* paradigm are the conditional forms (*bych* and so on), whilst the future auxiliary (*budu* and others) is non-clitic. All these forms are shown in table 4.4 below, with pure/lexical clitics set in bold print.

COP	PASS.AUX	PAST.AUX	COND.AUX	FUT.AUX
jsem (su)	jsem (su)	<b>jsem</b>	<b>bych (bysem)</b>	budu
jsi (seš)	jsi (seš)	<b>jsi</b>	<b>bys (bysi)</b>	budeš
je	je	–	<b>by</b>	bude
jsme	jsme	<b>jsme</b>	<b>bychom (bysme)</b>	budeme
jste	jste	<b>jste</b>	<b>byste</b>	budete
jsou	jsou	–	<b>by</b>	budou

Table 4.4: Copula and auxiliary forms of the verb *být* (2P clitics in bold, colloquial forms in brackets)

Note that even though the auxiliaries derived from the present tense forms of *být* (first three columns) are all isomorphic, the purely clitic past auxiliary displays a gap in the third person that is absent from the copula and passive auxiliaries. The forms in brackets are colloquial variants. Chapter 7 provides an in-depth analysis of the different groups within the *být* paradigm.

### 4.1.3 The conditional complementiser and negation

Just as the pronominal and auxiliary clitics discussed so far, the two clitics discussed in the present section are lexical clitics in the sense of Junghanns (2002b). Yet, they do not belong to

<sup>4</sup>The reason that Avgustinova and Oliva (1995) only classify the passive auxiliary, but not the copula, as semi-clitic, is the fact that the former, but not the latter, can occur inside the clitic cluster. In this case, illustrated by 26b, it appears in the same templatic position as the past auxiliary. This thus leads to a three-way distinction regarding auxiliaries and copulas: past and conditional auxiliaries are pure clitics, passive auxiliaries are semi-clitics, and copulas and future auxiliaries are not clitic (cf. Avgustinova & Oliva, 1995, 9-12).

- (26) a. *Tyto knihy mu jsou antikvářem nabízeny jako protihodnota.*  
 these books him.DAT AUX.3PL antiquarian.INS offer.PASS as countervalue
- b. *Tyto knihy jsou mu antikvářem nabízeny jako protihodnota.*  
 these books AUX.3PL him.DAT antiquarian.INS offer.PASS as countervalue

‘These books were offered to him by the antiquarian as countervalue.’

(cf. Avgustinova & Oliva, 1995, 9)

Avgustinova and Oliva’s (1995) pure clitics, as they are not in second position. This is evident for sentential negation, which will be discussed first. For the complementiser *-li*, the situation is less obvious. I will show below why it can be assumed that this clitic does not cluster with the other clitics in the second position of its clause.

Verbal negation in Czech is expressed through a verbal proclitic particle (cf. Avgustinova & Oliva, 1995, 11) which occurs left-adjacent to the verb regardless of the verb’s position in the clause. The clitic forms a prosodic word with its host; this can be seen from the fact that the main accent, which is on the first syllable of a word in Czech, falls on the proclitic (cf. Sussex & Cubberley, 2006, 181). Sentence 27c shows that the negative element is not subject to the Tobler-Mussafia effect – it always precedes the verb, even if this puts it at the left edge of the clause.

- (27) a. *Irina nejedla ty sušenky.*  
 Irina NEG:eat.PTCP those cookies  
 ‘Irina did not eat the cookies.’
- b. *Ty sušenky nejedla.*  
 those cookies NEG:eat.PTCP  
 ‘She did not eat the cookies.’
- c. *Nejedla ty sušenky.*  
 NEG:eat.PTCP those cookies

A Czech clitic that might be in the process of changing from a 2P phrasal clitic to a verbal clitic is the interrogative *-li*. It is only used in embedded polar questions or as a conditional (in the sense of ‘if’) in modern Czech. It attaches to a clause-initial verb;<sup>5</sup> any other kind of host renders the sentence “archaic poetic” (Short, 1993a, 498) or completely unacceptable. An example is provided by 28, where only the first variant, with *-li* attached to verb, is possible, whereas the object NP or the adverbial cannot host the clitic.

- (28) a. *Máte-li pochyby, zatelefonujte na informace.*  
 have.2PL-if doubts call.IMP.PL to information  
 ‘If you have any doubts, call the information.’
- b. *\*Pochyby-li máte...*  
 doubts-if have.2PL
- c. *\*Dnes-li máte pochyby...*  
 today-if have.2PL doubts

(cf. Toman, 1996, 508)

Since in such configurations, *-li* occurs in second position, it is often seen as the first element within the clitic cluster, and thus as a 2P clitic. However, this is problematic. In the following

<sup>5</sup>Toman (1996) specifies this: “In Modern Czech [...] *-li* can only follow a finite verb” (Toman, 1996, 508). In section 7.2.3, I am going to argue that it can also attach to a participle.

example, it occurs after the first word, whilst the reflexive clitic appears after the first constituent. Apparently, this sentence is judged as oldish, but becomes much worse when *-li* is attached to the entire first constituent (*lásce své-li* or *své lásce-li*; cf. Avgustinova & Oliva, 1995, 15).

- (29) *Lásce-li své se v životí budeš protiviti, žebrákem půjdeš*  
 love.ACC-if self's.ACC REFL.ACC in life will.2SG oppose.INF beggar.INS go.2SG  
*světem.*  
 world.INS

‘If you oppose your love in your life, you will go through the world as a beggar.’

(cf. Avgustinova & Oliva, 1995, 14)

Avgustinova and Oliva (1995, 11-13) conclude that, apart from such archaic examples as 29 above, in modern Czech, *-li* is a purely verbal enclitic which obeys the independent requirement that the combination verb+*li* must, like an interrogative phrase, appear first in the clause. There are several ways to capture this theoretically: either *-li* attaches to the verb early, and then they move to the left periphery as a unit (thus as a kind of complex complementiser), or *-li* is base-generated in the left periphery and attracts the verb.

I conclude that *-li* behaves differently from the 2P clitics we find among pronouns and auxiliaries. Like the negative particle *ne-*, it is a verbal clitic, but it additionally requires the verb it attaches to to be clause-initial, which is probably a reflex of its historic role of a 2P clitic. Since *-li* is therefore not a phrasal clitic (anymore), it will not feature in my analysis of 2P clitics and will also not be marked by bold print in the examples.

Alternatively, one could analyse both *ne-* and *-li* as affixes (cf. Spencer & Luís, 2012, 20) – after all, head-adjacent clitics cannot easily be distinguished from affixes because they lack the clitic-typical property of not selecting specific hosts (cf. section 2.1). This would also bring the sentential negator *ne-* in line with the homophonous prefix that attaches to adjectives and nouns to derive their antonym, as in *nejistý* ‘uncertain’ and *nevíra* ‘disbelief’.

#### 4.1.4 Adverbs

Other elements that can be clitic are the following adverbs: *tu* ‘here’, *tam* ‘there’, *však* ‘though’, *už* ‘already’, *prý* ‘allegedly’, *ale* ‘however’, *teda/tedy* ‘so (cf. Franks & King, 2000, 103)’, *asi* ‘probably’, *snad* ‘perhaps’, and probably others. Junghanns (2002b, 143) classifies them as phonological clitics which are placed according to semantic criteria; in the system of Avgustinova and Oliva (1995), they belong to the class of semi clitics. Logically, they are only unambiguously part of the 2P clitic cluster when they appear at its beginning or between other clitics, but not when they appear at the cluster’s end (as was also the case with *to* in example 25 above). Their prosodic neutrality is not always easily judged (cf. Franks & King, 2000, 104).

## 4.2 Properties of second-position clitics

The Czech lexical 2P clitics that I have identified in the previous subsections (and that are marked by bold print in tables 4.2 and 4.4) have several typical properties, which both aid their

identification and are vital for understanding their nature. These properties are syntactic and prosodic, and they directly derive from these elements' clitic status. They will be illustrated in detail in the following subsections. In addition, the auxiliary clitics differ morphologically from other auxiliaries in the *být* paradigm; these differences are examined in detail in chapter 7. As noted in section 2.2.2, Czech 2P clitics are a good example for the fact that distinguishing clitic types based on the presence or absence of a non-clitic counterpart is unproductive: whilst pronouns always have a full counterpart, auxiliaries do not – yet, both display the same syntactic and prosodic behaviour.

#### 4.2.1 Inability to appear sentence-initially

The most striking property of all 2P lexical clitics is the fact that they cannot occur at the beginning of a sentence (cf. Junghanns, 2002b, 134). This applies to pronouns as well as to auxiliaries, as examples 30 and 31 show, respectively. The c examples additionally show that the same restrictions do not apply to strong pronouns and to copulas.

- (30) a. *Pozvali ho.*  
 invite.PTCP him.ACC  
 ‘They invited him.’  
 b. \**Ho pozvali.*  
 him.ACC invite.PTCP

(cf. Franks & King, 2000, 5)

- c. *JEHO pozvali.*  
 him.ACC invite.PTCP  
 ‘They invited HIM.’

- (31) a. *Pozval jsem Petra na pondělí.*  
 invite.PTCP PASTAUX.1SG Petr.ACC to monday  
 ‘I invited Petr for Monday.’  
 b. \**Jsem pozval Petra na pondělí.*  
 PASTAUX.1SG invite.PTCP Petr.ACC to monday  
 c. *Jsem doma.*  
 COP.1SG at.home  
 ‘I am at home.’

(cf. Fried, 1994, 163-4)

## 4.2.2 No hosting of other clitics

The proclitic negative marker *ne-* never attaches to other clitics. Even when a lower verb can host the negative particle, as in 32d, this is only possible when this verb is not a lexical clitic. Negation can thus serve as a test for clitic-hood (cf. Toman, 1980, 309). Strictly speaking, this restriction only applies to auxiliary clitics, as *ne-* is a verbal clitic, as shown above; thus logically, *ne-* cannot attach to a pronoun anyway.

- (32) a. *Nespala jsem.*  
 NEG:sleep.PTCP PASTAUX.1SG  
 ‘I did not sleep.’
- b. \**Spala nejsem.*
- c. *Nebudu spát.*  
 NEG:FUTAUX.1SG sleep.INF  
 ‘I will not sleep.’
- d. *Nebyli byste jí to řekli.*  
 NEG:be.PTCP CONDAUX.2PL her.DAT it say.PTCP  
 ‘You would not have told her.’
- e. *Byli byste jí to neřekli.*  
 be.PTCP CONDAUX.2PL her.DAT it NEG:say.PTCP
- (for d and e: cf. Short, 1993a, 510)
- f. \**Byli nebyste jí to řekli.*  
 be.PTCP NEG:CONDAUX.2PL her.DAT it say.PTCP

In a similar vein, the pronominal 2P clitics can never co-occur with prepositions; only strong forms of clitics are allowed there (Junghanns, 2002b, 134; cf. Naughton, 2005, 78-9). An example is given in 33. Most prepositions can be characterised as simple clitics in Czech, as they prosodically attach to their complement (cf. Avgustinova & Oliva, 1995, 8): they thus carry the main stress at word-level, which is always on the first syllable in Czech. This means that they form a p-word together with the following word. Thus, the situation is analogous to the one found with *ne-* (cf. Sussex & Cubberley, 2006, 181).

- (33) a. *Pro tebe.*  
 for you.ACC
- b. \**Pro tě.*  
 for you.ACC

Combinations of pronouns and masculine accusative *-ň* exist in spoken language, but Franks and King (2000, 101) interpret these as single lexical items, i. e. inflected prepositions.

- (34) *Zaň se každý stydí.*  
 for:him REFL.ACC everyone be.ashamed.3SG

‘Everyone is ashamed of him.’

(cf. Franks & King, 2000, 5)

### 4.2.3 Inability to appear in isolation or bear stress

Lexical clitics like the first person auxiliary *jsem* cannot be used as one-word answers, as example 35c shows; instead, the main verb, i. e. the participle *napsal*, is used in such a context, as example 35b illustrates.

- (35) *Napsals to?*  
 write.PTCP:2SG it

‘Did you write it?’

- a. *Ano, včera jsem to napsal.*  
 yes yesterday AUX.1SG it write.PTCP

‘Yes, I wrote it yesterday.’

- b. *Ano, napsal.*  
 yes write.PTCP

‘Yes, I did.’

- c. \**Ano, jsem.*  
 yes AUX.1SG

(cf. Veselovská, 1995, 105)

As mentioned above, 2P clitics also cannot bear contrastive stress, as sentence 36b illustrates for pronouns. Instead, in order to express contrast on a pronoun, the non-clitic variant must be used, as shown in 36a. For auxiliaries, the situation is slightly different, because there, no non-clitic alternatives exist. Thus, for verum focus, the participle is stressed instead of the clitic auxiliary, similarly to the one-word answer above.

- (36) a. *Neposlal jsem ho JEMU/\*MU, ale jí.*  
 NEG:send.PTCP AUX.1SG it him.DAT but her.DAT
- b. \**Neposlal jsem MU ho, ale jí.*  
 NEG:send.PTCP AUX.1SG him.DAT it but her.DAT

I did not send it to him, but to her.’

(adapted from Veselovská, 1995, 272)



### 4.3 Relative ordering within the second position

When a sentence contains several 2P clitics, they appear adjacent to each other. This is commonly called a “clitic cluster”. I am using this term in a purely descriptive sense here, thus simply describing the fact that clitics accumulate in a certain part of the sentence.<sup>6</sup> The ordering of the clitics within the cluster is rather strictly regulated; a statement of the clitic sequence of a given language as in 37 and 38 is commonly called a “template”.

(37) Clitic template for Czech according to Zwicky (1977, 24):

AUX – REFL – FREE DATIVE – IO – DO

(38) Clitic template for Czech according to Franks and King (2000, 105):

LI – COND/AUX – NON-ARG.DAT/REFL – DAT – ACC/GEN

The two hierarchies make the same predictions, except for free/non-argumental datives: whilst Franks and King (2000) state that the reflexive may precede or follow the free dative, according to Zwicky (1977), only the former ordering is possible. Franks and King (2000) provide some examples with a free dative preceding a reflexive to prove their point; one is given in 39. Sentences with the opposite ordering, and which therefore fit both views, are however easier to find; one is shown in 40. The sentences also illustrate the fact that the non-argumental dative precedes the argumental one, as is typical in the Slavic languages (cf. Franks & King, 2000, 105).

(39) *On ti se mi ani neomluvil.*

he you.DAT REFL.ACC me.DAT NEG.even NEG:apologise.PTCP

‘(I’m telling) you, he didn’t even apologise to me.’

(cf. Fried, 1994, 173, fn. 23)

(40) *Já jsem se ti mu to neodvážila říci.*

I AUX.1SG REFL you.DAT him.DAT it.ACC NEG:dare.PTCP say.INF

‘(I’m telling) you, I did not dare say it to him.’

(cf. Veselovská, 1995, 273)

Another difference between the two templates is the inclusion of *-li*. However, I have already argued above that in fact, the complementiser does not occupy the same position as the other clitics, since it is not a 2P clitic. The embedded clause in example 41 thus contains three lexically 2P clitics, printed, as usual, in bold. Note that *to* here is not accusative, but nominative. It is the subject of the embedded sentence, hence the neuter agreement suffix *-o* on the participle. Since subject pronouns are never lexical clitics in Czech, *to* in this example is most probably not

<sup>6</sup>In a less neutral sense, “cluster” is often used to describe the assumption that the clitics all attach to the same syntactic node (cf. for example Bošković, 2001, 50).

part of the clitic cluster in the narrower sense. Then, neither is the adverb *tedy*. Short (1993a), who originally provided the example, counts both these elements as part of the clitic cluster, classifying them as “other pronouns”, and “particles”, respectively.

- (41) *Ptali se, nemělo-li by se mu to tedy říct.*  
 ask.PTCP REFL.ACC NEG:have.PTCP.N-if COND.3 REFL.ACC him.DAT it.NOM so  
 tell.INF

‘They asked whether he ought not then to be told.’

(cf. Short, 1993a, 495)

A range of short adverbs like *už* ‘already’ and *prý* ‘allegedly’ are, by some speakers, allowed to appear initially in the cluster, or even in other positions within it, as shown in 42 (cf. Avgustinova & Oliva, 1995, 25, fn. 32). In the SYN2015 corpus, such sequences are extremely rare: 15 different search combinations of clitic + *už* + clitic yielded only a total of five results.

- (42) a. *Jan už se jim ho rozhodl nedávat.*  
 Jan already REFL.ACC them.DAT it.ACC decide.PTCP NEG:give.INF  
 ‘Jan has already decided not to give it to them.’  
 b. <sup>?</sup>*Jan se už jim ho rozhodl nedávat.*  
 Jan REFL.ACC already them.DAT it.ACC decide.PTCP NEG:give.INF  
 c. <sup>?</sup>*Jan se jim už ho rozhodl nedávat.*  
 Jan REFL.ACC them.DAT already it.ACC decide.PTCP NEG:give.INF

(cf. Avgustinova & Oliva, 1995, 25, fn. 32)

Note that the hierarchy is governed by several factors. Although case plays a role, the reflexive clitics *se* and *si* must precede other pronominal clitics, regardless of case, as example 43 shows. It is also worth noting that the status of the reflexive (whether it is part of a *reflexivum tantum* or an argument of the verb) plays no role for its placement (cf. Spencer & Luís, 2012, 231).

- (43) a. *Představila jsem mu tě včera.*  
 introduce.PTCP AUX.1SG him.DAT you.ACC yesterday  
 ‘I introduced you to him yesterday.’  
 b. *Představila jsem se mu včera.*  
 introduce.PTCP AUX.1SG REFL.ACC him.DAT yesterday  
 ‘I introduced myself to him yesterday.’

(cf. Franks, 1998, 42)

The two hierarchies in 37 and 38 additionally differ in the sense that one refers to syntactic roles, whilst the other refers to case. Zwicky's (1977) template thereby neglects the fact that the morphophonologically identical accusative and genitive pronouns appear in the same position. I thus conclude that the template as given by Franks and King (2000) is the more adequate one, except for the inclusion of *-li*. The ordering of Czech clitics is thus best described as in 44 below.

(44) Clitic template for Czech:

COND/AUX – NON-ARG.DAT/REFL – DAT – ACC/GEN

In contrast to, for example, BCMS (cf. Bošković, 2004, 53), the clitic cluster in Czech cannot be split up, according to my informants: therefore, sentence 45 is not good in the b- and c-variants. Possibly in a more poetic style, such splits may occur very rarely, as example 46 from the SYN2015 corpus shows. Generally, it appears that cluster interruption virtually does not occur; also Toman (1999) judges “clitic cluster splitting, occasionally quoted, unacceptable” (Toman, 1999, 205).

(45) a. *Jak jsem ti řekla, Jan by mu nepomohl.*  
 as AUX.1SG YOU.DAT say.PTCP Jan COND.3 him.DAT NEG:help.PTCP

‘As I have told you, Jan would not help him.’

b. \**Jan by, jak jsem ti řekla, mu nepomohl.*  
 Jan COND.3 as AUX.1SG YOU.DAT say.PTCP him.DAT NEG:help.PTCP

c. \**Jan mu, jak jsem ti řekla, by nepomohl.*  
 Jan him.DAT as AUX.1SG YOU.DAT say.PTCP COND.3 NEG:help.PTCP

(46) *Spatřil jsem v dálce dům nade vše milovaný, (...) s nímž jsem,*  
 see.PTCP AUX.1SG in distance house above all beloved with which AUX.1SG  
*sirotek, se dávno rozloučil.*  
 orphan REFL.ACC long.ago separate.PTCP

‘In the distance I saw the house which I had loved above all, (...) to which I, an orphan, had said goodbye long ago.’

(SYN2015)

Also unlike in BCMS (cf. Bošković, 2001, 50), the Czech cluster does not allow for partial ellipsis. According to my Czech consultants, the only grammatical option in 47 is to fully repeat the entire cluster, as in 47a. This indicates that the cluster behaves as a unit that must remain intact.

(47) Context: *Do you think that Hana and Marek would give you their cat?*

- a. *Hana by mi ho darovala, a Marek by mi ho*  
 Hana COND.3 me.DAT him.ACC give.PTCP and Marek COND.3 me.DAT him.ACC  
*daroval také.*  
 give.PTCP also  
 ‘Hana would give him to me, and Marek would give him to me, too.’
- b. \**Hana by mi ho darovala, a Marek by mi také.*  
 Hana COND.3 me.DAT him.ACC give.PTCP and Marek COND.3 me.DAT also
- c. \**Hana by mi ho darovala, a Marek by ho také.*  
 Hana COND.3 me.DAT him.ACC give.PTCP and Marek COND.3 him.ACC also
- d. \**Hana by mi ho darovala, a Marek by také.*  
 Hana COND.3 me.DAT him.ACC give.PTCP and Marek COND.3 also

Generally, it is a matter of some debate how to account for clitic templates. While some see the template as a kind of independent construction plan for the clitic cluster, others assume that clitic ordering directly results from the syntactic hierarchy of the elements; still others hold intermediate views. Section 6.6 gives an overview of the different assumptions.

#### 4.4 Co-occurrence restrictions in the cluster

Before concluding this introduction to clitic ordering, I want to remark on some restrictions on the sequence of pronouns within the clitic cluster. Such restrictions are commonly called “person-case constraint” (PCC), and they can be found in many languages with clitic clusters (Franks & King, 2000, 106). For Czech, Sturgeon et al. (2012) identify two PCC’s, the “weak” and the “me-first” PCC, which can be unified to a single restriction, the “strictly descending PCC”. These three constraints are defined as in 48.

(48) PCC’s in Czech (cf. Sturgeon et al., 2012, 2):

- a. Weak PCC: If there is a 3rd person, it has to be the direct object.
- b. Me-first PCC: If there is a 1st person, it has to be the indirect object.
- c. ⇒ Strictly descending PCC: The argument with the higher person specification (where 1 is higher than 2 is higher than 3) has to be the indirect object.

The consequence of this is that in Czech, as shown in 49, after a second or third person dative clitic, only third person (49a), but not first or second person (49b), accusative clitics are possible.

- (49) a. *Pak mi/ti/mu/jí/nám/vám/jim ho/ji/je/to představil.*  
 then 1/2/3SG.M/F/1/2/3PL 3SG.M/F/3PL introduce.PTCP  
 DAT ACC

- b. \**Pak ti/mu/jí/vám/jim mě/tě/nás/vás představil.*  
 then 2/3SG.M/F/2/3PL 1/2SG/1/2PL introduce.PTCP  
 DAT ACC

(cf. Franks & King, 2000, 106)

A strategy to avoid a PCC violation is to reorder the clitics, so that accusative-dative order results, as in example 50a. In reality, however, many languages which have a weak PCC display great variation, both within speakers and for individuals. In a corpus study, Sturgeon et al. (2012) confirmed this for Czech. For example, they not only found 50a, where the PCC is respected, but also the sentence in 50b, where dative-accusative order prevails despite the PCC violation (cf. Sturgeon et al., 2012, 6-7).<sup>7</sup>

- (50) a. ... *já tě mu nedám!*  
 I you.ACC him.DAT NEG:give.1SG  
 ‘I won’t give you to him!’
- b. *Dám mu tě do pytle.*  
 give.1SG him.DAT you.ACC into bag  
 ‘I’ll give you to him in a bag.’

(cf. Sturgeon et al., 2012, 7)

Due to the just mentioned inter- and intra-speaker variation, the existence of PCC’s in Czech has sometimes been questioned. To clarify this issue, Sturgeon et al. (2012) also carried out an acceptability judgement task, where native speakers rated ditransitive sentences with clitic direct and indirect object. The results showed a clear decrease in acceptability both for sentences with a weak PCC violation and for sentences with a me-first PCC violation. The authors conclude from this that something like the strictly descending PCC is active in Czech, though not for all speakers (cf. Sturgeon et al., 2012, 5-6).

Franks (2017) argues that the common formulation of PCC’s as in 48 is misleading: since PCC violations can be avoided by reordering the clitics, PCC’s should better be stated in terms of linear order or structural height. For his reformulation of these constraints, see section 6.6.2.

## 4.5 Summary

The previous sections have shown that Czech lexical 2P clitics comprise reflexive and non-reflexive pronouns, past tense auxiliaries and conditional auxiliaries, whilst other prosodically cliticising elements are not tied to the second position. All Czech 2P clitics eschew bearing contrastive stress and appearing initially or in complete isolation. They also cannot support other clitics, i. e. neither negation nor prepositions. Within the clitic cluster, ordering is highly regulated, and the cluster is impermeable to any non-clitic elements. For pronominal clitics, exceptions from the templatic order may occur due to PCC constraints.

<sup>7</sup>The native speaker I asked about these sentences found all configurations in 49b acceptable and also clearly rejected re-ordered sentences such as the one in 50a. Apparently, for this consultant, templatic order is the only operative restriction, and no PCC is active.

## Chapter 5

# Placement of Czech 2P Clitics

Having described the clitic inventory of Czech, we can investigate the positions in which 2P clitics appear in this language. The present chapter is divided into two parts: The first one, section 5.1, describes the placement patterns within simple clauses, where we can observe 2P as well as 3P effects, and, under certain conditions, also placement in first position. The second part, section 5.2, addresses another widely discussed aspect of clitic positioning: clitic climbing, i. e. constructions where clitics appear outside of their original clause, in a higher matrix sentence.

### 5.1 Placement in simple clauses

Section 4.1 has presented the lexical elements that can be characterised as unambiguous 2P clitics. Firstly, this statement requires clarification: what exactly is the second position, i. e. how is the *first* position defined? Secondly, there are cases that appear to be better labelled “clitic third”, rather than “clitic second”. These phenomena form a central part of this thesis: What principles can third-position clitics in Czech be derived from? Are they deviations from the norm, or can we find an explanation that captures this “delayed clitic placement” as a natural consequence of the general clitic placement mechanism? I will also address the issue of clitics appearing initially, both in a syntactic and in a prosodic sense.

#### 5.1.1 Clitics in second position

Czech clitics appear in second position. This is however only a very broad description, since, as mentioned in the general introduction to clitics in section 2.2, the exact definition of “second position” is language-specific. In Czech, 2P clitics must be preceded by exactly one constituent. In the terminology of Halpern (1995, 15), Czech clitics are thus classified as “2D”, meaning that they follow the first syntactic daughter of their constituent (cf. Halpern, 1995, 16), i. e. the clause. The element preceding the clitics can apparently be of any syntactic type, for example subject, object, or adverbial, and does not have to be phrasal: it can also be a verbal head or a complementiser, as the variations in 51, featuring the lexically clitic conditional auxiliary, show.

- (51) a. *Anna by pomohla svému bratrovi.*  
Anna COND.3 help.PTCP POSS.DAT brother.DAT  
‘Anna would help her brother.’

- b. *Svému bratrovi by Anna pomohla.*  
 POSS.DAT brother.DAT COND.3 Anna help.PTCP  
 ‘Anna would help her brother.’
- c. *Bez váhání by pomohla svému bratrovi.*  
 Without hesitation COND.3 help.PTCP POSS.DAT brother.DAT  
 ‘Without hesitation, she would help her brother.’
- d. *Pomohla by svému bratrovi.*  
 help.PTCP COND.3 POSS.DAT brother.DAT  
 ‘She would help her brother.’
- e. ... že by Anna pomohla svému bratrovi.  
 COMP COND.3 Anna help.PTCP POSS.DAT brother.DAT  
 ‘...that Anna would help her brother.’

Unlike in, for example, BCMS, what Halpern (1995, 16) calls “2W”, i. e. clitic positioning after the first word, does not occur in Czech, thus 52b is ungrammatical; the clitic pronoun must follow the entire NP and cannot intervene between demonstrative and noun phrase. Yet, there are other cases where a clitic appears to be involved in the splitting of a constituent: for example, Lenertová (2001) presents the sentence in 53, where the noun *knížku* is focused and fronted to the left periphery, and thus separated from its adjectival modifier *krásnou*. The clitic auxiliary then follows this noun. This phenomenon will be discussed and analysed in section 9.1.3.

- (52) a. *Ten básník mi čte ze své knihy.*  
 that poet me.DAT read.3SG from POSS.GEN book.GEN  
 ‘That poet reads to me from his book.’
- b. \**Ten mi básník čte ze své knihy.*  
 that me.DAT poet read.3SG from POSS.GEN book.GEN

(cf. Halpern, 1995, 17)

- (53) *KNÍŽKU jsem vám přinesla krásnou.*  
 book.ACC AUX.1SG you.DAT bring.PTCP nice.ACC  
 ‘I’ve brought you a nice book!’

(cf. Lenertová, 2001, 299)

The NP preceding the clitic can be maximally complex, as long as it still forms one constituent. So, for example, a relative clause can be added, as illustrated in 54. Several adverbials can also be clitic hosts if they are “of the same type” (Avgustinova & Oliva, 1995, 4), i. e. if they collectively refer to a single time or place and thus form a single constituent, as in 55; in contrast, Czech clitics cannot follow two adverbials that do not form such a semantic unit, as the contrast between 56a and 56b shows.

- (54) a. *Knihy, které tady vidíte, se dnes platí zlatem.*  
 books REL here see.2PL REFL today pay.3SG gold.INS  
 ‘The books you can see here are paid for with gold today.’

(cf. Toman, 1996, 506)

- b. *Knihy se dnes platí zlatem.*  
 books REFL today pay.3SG gold.INS  
 ‘The books are paid for with gold today.’

- (55) *Každý týden v neděli dopoledne v devět hodin se stará paní svátečně  
 ustrojila a šla na mši.*  
 every week in Sunday morning in nine hours REFL.ACC old lady festively  
 dress.PTCP and go.PTCP on mass

‘Every week on Sunday morning at nine o’clock the old lady dressed up festively and went to the mass.’

(cf. Avgustinova & Oliva, 1995, 4)

- (56) a. *Dnes by ti je jistě prodali levněji.*  
 today COND.3 you.DAT them.ACC certainly sell.PTCP cheaper  
 ‘Today they certainly would have sold them to you at a cheaper price.’

(cf. Franks & King, 2000, 105)

- b. *\*Dnes jistě by ti je prodali levněji.*  
 today certainly COND.3 you.DAT them.ACC sell.PTCP cheaper

Given that interrogative phrases are constituents, it is not surprising that in multiple questions, the clitics typically follow the first interrogative phrase, as shown in 57. The order of interrogative phrases, on the other hand, is not fixed, so there can be variation with regard to which question word precedes the clitic (cf. Rudin, 1988, 175, and examples therein).

- (57) a. *Kdo ho kde viděl je nejasné.*  
 who.NOM him.ACC where see.PTCP is unclear  
 ‘Who saw him where is unclear.’

- b. *\*Kdo kde ho viděl je nejasné.*  
 who.NOM where him.ACC see.PTCP is unclear

(cf. Rudin, 1988, 466)



Whilst the constituent that hosts the clitics can be of any type, it must be part of the clause the clitics belong to. As Junghanns (2002b) shows, extra-clausal material cannot host the clitic: neither the vocative in 59 nor the adverbial clause in 60. This latter case shows that the syntactic type of an adverbial – whether it contains an inflected verb (as in 60a) or not (as in 51c, 55, and 56a) – is crucial for clitic placement.<sup>1</sup>

- (59) a. *Mamínko, omluvil jsem se babičce.*  
 mum.VOC apologise.PTCP AUX.1SG REFL grandma.DAT

‘Mum, I have apologised to grandma.’

- b. \**Mamínko, jsem se omluvil babičce.*  
 mum.VOC AUX.1SG REFL apologise.PTCP grandma.DAT

(cf. Junghanns, 2002b, 130)

- (60) a. *Až se vrátí, chtěla bych s ním mluvit.*  
 when REFL return.3SG want.PTCP COND.1SG with him.INS speak.INF

‘When he returns, I would like to talk to him.’

- b. \**Až se vrátí, bych chtěla s ním mluvit.*  
 when REFL return.3SG COND.1SG want.PTCP with him.INS speak.INF

(cf. Lenertová, 2004, 150)

Also elements such as *a* ‘and’, *ale* ‘but’, and *no* ‘well’ cannot host the clitics. Foreign language learners of Czech are frequently told that these words “do not count”; the reason is arguably that they, too, are extra-clausal. Examples are given in 61. A search in the SYN2015 corpus supports this: both the sequences *ale bych* and *a bych* only have 0.02 i. p. m. each, whereas the results for these sequences with an intervening word are 820 times higher for *ale* and 1450 times higher for *a*.

- (61) a. \**A mu nic nedávej.*  
 and him.DAT nothing NEG:give.IMP

‘And don’t give him anything.’

<sup>1</sup>German V2, for example, is indifferent to the syntactic type of the adverbial:

- (58) a. *Morgen möchte ich mit ihm sprechen.*  
 tomorrow want.COND.1SG 1SG with him.DAT speak.INF  
 ‘Tomorrow I’d like to speak to him.’  
 b. *Wenn er zurückkommt, möchte ich mit ihm sprechen.*  
 when he return.3SG want.COND.1SG 1SG with him.DAT speak.INF

‘When he returns, I’d like to speak to him.’

- b. \**No ti poslal dopis.*  
 well you.DAT send.PTCP letter  
 ‘Well, he sent you a letter.’

(cf. Fried, 1994, 162)

We can summarise this section by stating that Czech clitics appear after the first constituent. Sequences which do not form one constituent cannot precede the clitic cluster, and neither can constituents that form their own verbal projection or that are extra-clausal. So far, it thus appears that clitic placement can be described in very simple syntactic terms. The following section presents a deviation from this pattern.

### 5.1.2 Clitics in third position

There are cases when clitics can appear in third position. For example, the complementiser *že* permits intervening material: clitics can appear in third position when preceded by an element which is fronted for information-structural reasons, as is the case with *Petr* in 62b. Regarding this phenomenon’s frequency, a search in the SYN2015 corpus showed that the third-position effect after a complementiser is 31 times less frequent than the direct sequence complementiser-clitic.<sup>2</sup>

- (62) a. ... *že se Petr odstěhoval.*  
 COMP REFL Petr move.PTCP  
 b. ... *že Petr se odstěhoval.*  
 COMP Petr REFL move.PTCP  
 ‘...that Petr has moved.’

(cf. Fried, 1994, 159)

Examples where the information-structural status of the phrase in the second position is clear are provided by Lenertová (2001); the author draws them from Czech films and literature. Sentence 63a shows a topic preceding the clitics, and 63b a focus; capitals indicate focus intonation. Also note that in 63a, the topic is not an argument of the verb, and also not a NP, but an adverbial phrase.

- (63) a. *Já myslím, že tady na severu **se** *nemusíme oslovovat plnými*  
 I think.1SG COMP here in north REFL NEG:must.1PL address.INF full.INS  
*tituly.*  
 titles.INS*

‘I think that here in the north, we don’t have to address each other with full titles.’

<sup>2</sup>I searched for the direct sequence *že bych*, which had 57.59 i.p.m, and for the sequence *že X bych*, which had 1.86 i.p.m. Note that the intervening word X is unlikely to be a clitic, since auxiliaries come first in the clitic cluster. The conditional clitic was used because unlike the past tense clitic, it is not homophonous with any non-clitic or only phonologically clitic element.

- b. *Věřil byste, že i REVMA jsem ztratil?*  
 believe.PTCP COND.2PL COMP even rheumatism AUX.1SG lose.PTCP

‘Would you believe that I have even lost my RHEUMATISM?’

(cf. Lenertová, 2001, 300-1)

Other phrasal elements can thus also appear in second position, preceding the clitics, such as sentential adverbials, which then “receive a special emphasis”, according to Lenertová (2001, 301); additional examples are shown in 64. We can arguably subcategorise this under the complementiser+focus configurations.

- (64) a. *Myslím, že naopak jsme byli velmi jednotní.*  
 think.1SG COMP on.the.contrary AUX.1PL be.PTCP very unanimous

‘On the contrary: I think we were quite of the same mind.’

- b. *Řekl stručně, že v podstatě jsme byli prozrazeni.*  
 say.PTCP briefly COMP in fact AUX.1PL be.PTCP betray.PTCP

‘He recently said that in fact we were betrayed.’

(cf. Lenertová, 2001, 301)

However, not all elements are permitted in second position: a verb cannot intervene between complementiser and clitic, as 65 exemplifies. Yet, we will see that matters are more complicated than that: there is a slight difference in acceptability depending on the type of verb, and VP fronting to that position is also possible in some configurations. This will be illustrated and discussed in section 9.2. For the present purposes, it suffices to note that verb occurrence in second position is restricted.

- (65) \*... *že nedal by mu to.*  
 COMP NEG.give.PTCP COND.3 him.DAT it

‘...that he would not give it to him.’

(cf. Veselovská, 1995, 110)

There is another interesting case of clitic third which does not involve the presence of a complementiser: in content questions, a focused constituent can appear between the interrogative phrase and the clitic; this is shown in 66a. Example 66b demonstrates that without the stressed and fronted subject, the clitic regularly follows the first constituent of its own clause, which is the interrogative.<sup>3</sup> Note that interrogative phrases, as well as complementisers, are independently required to appear in the left periphery. The elements that appear first in clitic-third

<sup>3</sup>One intriguing problem of Czech syntax connected to this construction is the fact that a topic phrase cannot occupy the position between interrogative phrase and clitic (cf. Lenertová, 2001, 299). It is often assumed that Czech left-peripheral foci and topics occupy the same position, as they are in complementary distribution. The puzzle is: why can foci and topics not co-occur, as well as topics and interrogative phrases, but foci and interrogative phrases can? Whilst clearly meriting further investigation, this issue goes beyond the scope of the present thesis and I will consequently leave it aside.

configurations are thus both relevant for sentence type and independently required to occupy the clause-initial position.

- (66) a. *A co EMA by na to řekla?*  
 and what Ema COND.3 to this say.PTCP  
 ‘And what would EMA say to that?’

(cf. Lenertová, 2001, 299)

- b. *A co by Ema na to řekla?*  
 and what COND.3 Ema to this say.PTCP  
 ‘And what would Ema say to that?’

A potentially related area are multiple questions. The preceding subsection has shown that clitics typically appear after the first interrogative phrase. However, there are exceptions. In 67a, the multiple question does not ask for a pair-list answer, but just for a single pair. In 67b, the interrogative clause is an enquiry about the reference of *it* in the context; the questioner wants to make sure he understands the topic of the conversation correctly (cf. Meyer, 2004, 82-3).

- (67) a. ... *kdo koho si váží víc.*  
 who.NOM whom.ACC REFL.DAT value.3SG more  
 ‘... who values whom more.’
- b. Context: *Don't you know when it began?*  
*Co kdy se začalo?*  
 what.NOM when REFL.ACC begin.PTCP  
 ‘When what began?’

(cf. Meyer, 2004, 82-3)

For the elements that appear between complementiser and clitic, we have seen that they are always phrasal, and that they occupy the second position in the clause for information-structural reasons. Note that, as discussed in section 3.5, topics occupy the left periphery, but foci typically appear at the right edge, so 63b could also have the word order as given in 68.

- (68) *Věřil byste, že jsem ztratil i REVMA?*  
 believe.PTCP COND.2PL COMP AUX.1SG lose.PTCP even rheumatism  
 ‘Would you believe that I have even lost my RHEUMATISM?’

To summarise, the clitic-third configurations that have been illustrated in this section are composed as follows: The first position can be occupied by a complementiser (for example *že* ‘that’) or by an interrogative phrase (for example *co* ‘what’). The second position can be occupied by a topic, by a left-peripheral focus, or by a second interrogative phrase. The clitic

cluster then follows these information-structurally positioned elements and thus occurs third in the clause.<sup>4</sup>

### 5.1.3 Clitics in first position and prosodic dependency

Czech clitics are often characterised as enclitics, which require a host to their left due to their phonological weakness. There are, however, reasons to doubt this classification: first, the colloquial language allows clitics to appear in first position; second, even in the standard language we find cases where the clitic cannot be prosodically attached to the left. I will address these two phenomena in turn in this section.

Beginning with sentence-initial placement in the colloquial language, Czech clitics can sometimes appear initially, i. e. in first position. According to Toman (1996, 506), this is the result of truncation of certain short and semantically reduced words such as *tak* ‘so’ and *to* ‘it’. This latter case is illustrated in example 69, where Colloquial Czech allows for the b-variant.

(69) a. *To se uvidí.*  
it REFL see.3SG

b. *Se uvidí.*  
REFL see.3SG

‘One will see.’ (*lit.*: ‘It will be seen.’)

(cf. Toman, 1996, 506)

Another example is provided by Franks and King (2000), shown in 70. The spelling of the first person singular auxiliary as *sem* instead of *jsem* is a colloquial spelling, reflecting the fact that the initial *j* is almost always silent before *s*. Also the spelling of the verb *řek* without the participle suffix *-l* is a variant which reflects elision in colloquial pronunciation.

(70) *Sem ti to už přece řek.*  
AUX.1SG you.DAT it already though say.PTCP

‘I already told you that, though.’

(cf. Franks & King, 2000, 114)

In this example, however, deletion of one of the just mentioned words does not appear plausible; note that the pronominal *to* appears lower in the clause. Also Lenertová (2004) and Toman (1996) point out that truncation cannot explain all cases of first-position placement. Franks and King’s (2000) interpretation is that the clitic syntactically follows a silent pronominal subject *já* (cf. Franks & King, 2000, 114). Nevertheless, prosodically, the clitic has no host to its left, and thus cannot be enclitic.

<sup>4</sup>Lenertová (2001) also mentions the possibility that a topic occurs first, and the complementiser follows (cf. Lenertová, 2001, 301-2). This is a case of an external topic, as remarked upon in section 3.5.2. Since in this configuration, the topic presumably adjoins to CP, I do not analyse this as 3P placement, but as 2P placement with a preceding extra-clausal element.

In general, some speakers' grammars appear to grant the clitic auxiliaries much more freedom than others. Not only are both the conditional and the past auxiliary allowed in first position in these variants, clitics can also appear in a fourth position following a stressed constituent. An example is given in 71. Short (1993a, 495) notes that this freedom is not granted to all clitics equally: the reflexive *se* appears to be especially prone to what he calls "slippage". Whilst these colloquial phenomena will not be the focus of my analysis, I will shortly come back to this issue in section 8.2.3.

- (71) *Jistě namítnete, že to VÁM se zítra nestane.*  
 surely object.2PL COMP it you.PL.DAT REFL.ACC tomorrow NEG:happen.3SG

'You will surely object that to YOU, it will not happen tomorrow.'

(cf. Franks, 1998, 38)

However, this does not mean that in the standard language, we can simply analyse 2P clitics as enclitics: Franks and King (2000) claim that the clitic's direction of prosodic attachment is underspecified, i. e. that it can just as well attach to a host to their right. This is also assumed by Toman (1996), who provides example 72 to support this point. Another example, found in Toman (1986), is the sentence in 73. Importantly, these are accepted by all speakers, and not only in colloquial language, so they cannot just be attributed to a different grammatical system.

- (72) *Pomáhat mu, by dnes stálo mnoho.*  
 help.INF him.DAT COND.3 today cost.PTCP much

'Helping him would cost a lot today.'

(cf. Toman, 1996, 507)

- (73) *Ten doktor, co mu důvěruješ, se neholí.*  
 that doctor REL him.DAT trust.2SG REFL.ACC NEG:shave.3SG

'That doctor, whom you trust, doesn't shave.'

(cf. Toman, 1986, 125)

In example 72, the first clitic is followed by a pause, whilst the second clitic is *preceded* by one. Note we cannot assume that *mu* and *by* form a cluster in this sentence, as their order would then be reversed. Neither can *mu* host *by*, as *mu* is itself a clitic. It follows that *by* cannot be prosodically attached to the left. In 73, too, a clitic follows an intonational break: this is not only perfectly grammatical, in fact, the position right after the intonational break is the *only* grammatical position for the reflexive *se*, since it is the position after the sentence's first constituent. It thus appears that prosodically, Czech clitics can be either pro- or enclitic. This is an important observation, since the clitics' reputed encliticity is a crucial factor in many theories of clitic placement.

One might conclude from this that Czech clitics are not prosodically dependent at all. Support for this assumption could come from the existence of bisyllabic clitic auxiliaries: the first and second person plural conditional auxiliaries *bychom* (and its colloquial variant *bysme*) and *byste*. At first glance, these appear to be capable of bearing their own word stress. Yet, if one attributes the prosodic dependency of clitics to their incapacity to project prosodic feet, as argued for example by Franks (2017, 147), then the difference between mono- and bisyllabic clitics disappears – neither of them project prosodic feet.<sup>5</sup> Thus, bisyllabic clitics cannot be seen as evidence for clitics’ prosodic independence.

Other factors directly point to the prosodic weakness of clitics: as shown in section 4.2, Czech clitics cannot carry focus accent and they are not permitted to appear alone, as a single-word statement – this shows that Czech clitics do require *some* prosodic host. Also note in this context that what applies to monosyllabic auxiliaries just as well applies to bisyllabic ones.

On the other hand, clitics can even be preceded *and* followed by an intonational break at the same time – they must thus not necessarily be included into a prosodic phrase neither to their left nor to their right. This is shown for pronouns and auxiliaries by the sentences in 74, for which Junghanns (2003) confirmed the acceptability with intonational pauses both preceding and following the clitics with several native speakers. Note that, whilst these sentences are grammatical with the intonational breaks, they are not when the clitic is placed in a position other than the one following the first constituent, as 75 shows for the pronominal clitic in 74a.

- (74) a. *Důkazy, které předložila, ho, jak je vidět, nepřesvědčují.*  
 evidence REL present.PTCP him.ACC as is see.INF NEG:convince.3PL  
 ‘Obviously, the evidence that she has presented does not convince him.’
- b. *Knihy, které tady vidíte, se, jak tvrdí, dnes platí zlatem.*  
 books REL here see.2PL REFL.ACC as claim.3PL today pay.INF gold.INS  
 ‘They claim that the books you can see here are paid for with gold today.’
- c. *Tu televizi, kterou jsme u Vás našli, jste, jak víme, nekoupila v obchodě.*  
 this TV REL AUX.1PL at you find.PTCP AUX.2PL as know.1PL  
 NEG:buy.PTCP in shop  
 ‘We know that you didn’t buy the TV set that we found at your place at a shop.’
- d. *Ten člověk, který tu žije, by, jak tvrdí, mohl být terorista.*  
 this man REL here live.3SG COND.3 as claim.3PL can.PTCP be.INF terrorist  
 ‘They claim that the man who lives here might be a terrorist.’

(cf. Junghanns, 2003, 7)

<sup>5</sup>As Franks (2017) correctly notes, the observation that clitics do not project prosodic feet must not be conflated with their general ability to carry stress. Whilst they are never independently carriers of stress, they can carry the stress of a larger prosodic unit they incorporate into. For example, Macedonian has fixed ante-penultimate stress, and this stress can fall onto a clitic if it is the ante-penultimate syllable (cf. Franks, 2017, 147).

- (75) a. \***Ho** *důkazy, které předložila, jak je vidět, nepřesvědčují.*  
 him.ACC evidence REL present.PTCP as is see.INF NEG:convince.3PL
- b. ?\**Důkazy, které předložila, jak je vidět, ho nepřesvědčují.*  
 evidence REL present.PTCP as is see.INF him.ACC NEG:convince.3PL
- c. \**Důkazy, které předložila, jak je vidět, nepřesvědčují ho.*  
 evidence REL present.PTCP as is see.INF NEG:convince.3PL him.ACC

(cf. Junghanns, 2003, 7)

It thus seems that Czech clitics do not have a fixed direction of cliticisation. Moreover, they can appear without any prosodic host – if their placement in the syntactically second position does not provide them with one, they will do without. On the other hand, Czech lexical clitics are clearly prosodically weak in some way, as they cannot appear alone and cannot carry stress. It therefore appears that the syntactic behaviour of Czech clitics cannot be explained merely through their prosodic properties. This is also a point raised by Anderson (1992), who notes that there appears to be no connection between the properties of special clitics and prosodic deficiency: “prosodic dependence can be found either with or without special placement, and vice versa” (Anderson, 1992, 204).

A similar observation is made by Zimmerling (2012). He surveys a range of Slavic and non-Slavic languages and concludes that several features of 2P cliticisation that are attributed to their prosodic dependency are actually independent of one another. If prosodic deficiency caused clitics to appear after the first constituent, and not clause-initially or after a sentential host, we would expect to find only languages which either have none or all of these features – however, Zimmerling (2012) shows that all kinds of combinations occur (cf. Zimmerling, 2012, 9). This is strong evidence that prosodic deficiency cannot explain the placement of all 2P clitics – and clearly not of those found in Czech.

## 5.2 Placement in complex clauses: clitic climbing

To conclude our introductory description of Czech clitic placement, the following subsections sketch the phenomenon of clitic climbing, where a clitic or clitic cluster is not positioned within the clause of its predicate, but within a superordinate clause. Clitic climbing occurs in many languages, both with verbal clitics, as in Romance, and with 2P clitics, as in Slavic (cf. Spencer & Luís, 2012, 164-5). In Czech, clitic climbing is rather restricted (Franks & King, 2000, 117-20). I will describe the configurations in which climbing is possible, required, or barred in Czech in section 5.2.1, and also illustrate additional morphophonological restrictions on climbing in section 5.2.2. Section 5.3 then provides a concluding summary.

### 5.2.1 Domains for Czech clitic climbing

Generally, climbing out of finite clauses is ungrammatical in Czech; clitics remain within their own CP. This is demonstrated in example 76. The same also applies to infinitive clauses that clearly have a CP layer, i. e. which feature an interrogative element in Spec-CP (cf. Lenertová,



2004, 156), as can be seen in example 77. It is also important to note at this point that, given that climbing is restricted to infinitival clauses in Czech, climbing only applies to pronominal clitics, whereas auxiliary clitics are necessarily excluded from such constructions.

- (76) a. *Řekl, že mi ho můžete ukázat.*  
 say.PTCP COMP me.DAT him.ACC can.2PL show.INF  
 ‘He said that you can show him to me.’
- b. \**Řekl mi ho, že můžete ukázat.*  
 say.PTCP me.DAT him.ACC COMP can.2PL show.INF

(cf. Lenertová, 2004, 156)

- (77) a. *Ale nevím opravdu, jak ho zapisovat.*  
 but NEG.know.1SG really how him.ACC record.INF  
 ‘But I really do not know how to record him.’
- b. \**Ale nevím ho opravdu, jak zapisovat.*  
 but NEG.know.1SG him.ACC really how record.INF

(cf. Lenertová, 2004, 156)

Climbing obligatorily occurs inside a clause, in the sense that genitive clitics cannot remain within their NP; thus, 78b is ungrammatical. Clitics must appear in the second position of the entire clause, as shown in 78c (cf. Toman, 1986, 128). This is, however, not climbing *sensu stricto*, i. e. out of a clausal domain.

- (78) a. *Karel měl málo kapra.*  
 Karel have.PTCP little carp.GEN  
 ‘Karel had a bit of carp.’
- b. \**Karel měl málo ho.*  
 Karel have.PTCP little him.GEN
- c. *Karel ho měl málo.*  
 Karel him.GEN have.PTCP little  
 ‘Karel had a bit of it.’

(cf. Toman, 1986, 128)

I now turn to climbing where clitics do not appear in the clause they belong to. In multi-predicate structures, clitic climbing out of infinitival complements of modal verbs is obligatory, according to Veselovská (1995, 305). Thus, 79a is the only acceptable option for clitic placement, where the clitic occurs second within the entire clause. This is confirmed by my native speaker consultants. The same applies to complex clauses with phasic matrix verbs like *začít* ‘to begin’, as in 80. The syntactic status of these infinitive clauses – whether they are VPs, TPs, or CPs – will have to be established; see section 8.3.

- (79) a. *Včera mu to musel dát.*  
 yesterday him.DAT it.ACC must.PTCP give.INF  
 ‘Yesterday he had to give it to him.’
- b. \**Včera musel mu to dát.*  
 yesterday must.PTCP him.DAT it.ACC give.INF
- c. \**Včera musel dát mu to.*  
 yesterday must.PTCP give.INF him.DAT it.ACC

(cf. Veselovská, 1995, 305)

- (80) a. *Jan se začal smát.*  
 Jan REFL.ACC start.PTCP laugh.INF  
 ‘Jan started to laugh.’
- b. ?\**Jan začal smát se.*  
 Jan start.PTCP laugh.INF REFL.ACC
- c. ?\**Jan začal se smát.*  
 Jan start.PTCP REFL.ACC laugh.INF

(cf. Avgustinova &amp; Oliva, 1995, 14)

Veselovská (1995) reports that Czech clitics do not leave ECM structures that lack a CP layer, and provides 82 as an example. Since the clause must contain a PRO subject, which is located in IP/TP, Veselovská (1995) concludes that clitics cannot cross the IP boundary (cf. Veselovská, 1995, 303-4). Yet, it is not evident to me why 82 should be an ECM clause – after all, the PRO subject is not exceptionally case-marked, but simply receives case from the verb it is merged with.<sup>6</sup> For discussion, see section 8.3.

- (82) a. *Nutil Petra<sub>1</sub> PRO<sub>1</sub> dát mu tu knihu.*  
 force.PTCP Petr.ACC give.INF him.DAT that.ACC book.ACC  
 ‘He made Petr give him that book.’
- b. \**Nutil mu Petra dát tu knihu.*  
 force.PTCP him.DAT Petr.ACC give.INF that.ACC book.ACC

(cf. Veselovská, 1995, 303-4)

<sup>6</sup>Note that *nutil* ‘force’ takes two complements, a subject and a direct object:

- (81) *Proč to udělal? – Karel ho nutil.*  
 why it do.PTCP Karel.NOM him.ACC force.PTCP  
 ‘Why did he do this? – Karel made him.’

A real case of ECM in Czech can be found in example 83; here, the subject of the embedded clause (consisting of the infinitive *ležet*) receives accusative case from the matrix verb and therefore surfaces as a clitic object pronoun, *ho*, which climbs to the higher clause, forming a cluster with the matrix clause auxiliary clitic *jsem*.<sup>7</sup> Thus, climbing is blocked out of object control clauses like 82, but is actually possible out of ECM clauses like 83.

- (83) *V jaké poloze jste ho našel ho ležet?*  
 in which position aux.2PL him.ACC find.PTCP lie.INF  
 ‘Which position did you find him lying in?’

(cf. Junghanns, 2002a, 65)

Generally, Junghanns (2002b) notes that movement restrictions on clitics do not apply to non-clitics, i. e. that clitics are generally less free in their movement: embedded interrogative infinitives, infinitives embedded into predicate adverbial phrases, and right-extraposed subjects are islands for clitic climbing, but not for movement of full NPs (cf. Junghanns, 2002b, 131-2). Thus, clitic climbing and its absence cannot simply be captured by appealing to general extraction properties of the language. Section 8.3 will show how it can be incorporated into the analysis of clitic placement pursued in the present thesis.

### 5.2.2 Co-occurrence restrictions for clitic climbing

For the sake of completeness, I will shortly sketch some morphophonological co-occurrence restrictions with respect to clitic climbing. When a clitic climbs up to a position which already hosts an identical clitic, haplology occurs, i. e. the two clitics are obligatorily merged/one of the clitics is deleted, as example 84 shows. In general, two reflexive clitics are not permitted within the same cluster, so the lower clitic stays within its clause, as example 85 demonstrates.

- (84) a. *Jan se snažil se elegantně obléci.*  
 Jan REFL.ACC try.PTCP REFL.ACC elegantly dress.INF  
 b. *Jan se snažil elegantně obléci.*  
 Jan REFL.ACC try.PTCP elegantly dress.INF  
 ‘Jan tried to dress elegantly.’

(cf. Avgustinova & Oliva, 1995, 15)

- (85) a. *Stále se snažím získat si její přízeň.*  
 constantly REFL.ACC try.1SG win.INF REFL.DAT her favour  
 ‘I am constantly trying to win her favour.’  
 b. *\*Stále se si snažím získat její přízeň.*  
 constantly REFL.ACC REFL.DAT try.1SG win.INF her favour

<sup>7</sup>I would like to thank Irenäus Kulik for pointing me to this example.

- c. \**Stále si se snažím získat její přízeň.*  
 constantly REFL.DAT REFL.ACC try.1SG win.INF her favour

(cf. Avgustinova & Oliva, 1995, 16)

However, Avgustinova and Oliva (1995) claim that this restriction does not apply when the reflexives do not underlyingly form one cluster, and give the example in 86. They argue that here, the clitics are not a cluster, because otherwise, they would have to be ordered differently (cf. the clitic templates in section 4.3). But it is not evident to me that the restriction “dative >> accusative” also applies to reflexive clitics, which have their own slot in the template. Furthermore, I presented the sentence in 86 to two native speakers, who completely rejected it, as well as other sentences with reflexive sequences. One even judged them to be barely comprehensible. It appears that the restriction on two co-occurring reflexives is absolute for at least some native speakers.

- (86) *Naučil se si vždy znovu získat její přízeň.*  
 learn.PTCP REFL.ACC REFL.DAT always again win.INF her favour

‘He learned to win her favour always anew.’

(cf. Avgustinova & Oliva, 1995, 16)

Since such haplogies can be captured by simple phonological constraints, and are secondary to the general rules of clitic placement, I will not investigate them in more depth within this thesis. However, when investigating clitic climbing, these phenomena must be kept in mind, because using complex sentences with identical clitics or with two reflexives could potentially distort results.

### 5.3 Summary

This chapter has shown that typically, Czech clitics appear after the first constituent of their own clause. They may stand in third position when a constituent appears after the initial complementiser (or interrogative phrase) for information-structural reasons. First-position placement is limited to colloquial styles; yet, also the standard language has no strict *prosodic* requirement on clitic attachment, i.e. clitics are not always *enclitic* – the placement properties of Czech lexical clitics thus cannot simply be reduced to prosodic weakness.

With respect to complex clauses, we have seen that clitics climb out of infinitival complements of modal or phasic verbs, as well as out of ECM clauses. In contrast, they do not rise out of finite clauses; neither do they climb out of infinitival clauses with PRO subjects. Section 8.3 will provide an answer to the question how we can generalise over these two groups and thereby capture the behaviour of clitics in complex sentence constructions. I have also presented two rules of haplogy that are tied to climbing: two phonologically identical clitics may not co-occur within the cluster, and neither may two reflexive clitics.

## Chapter 6

# Overview of Theories of Clitic Placement

The previous chapter has shown that the mere statement “Czech clitics appear in second position” is not nearly enough to actually capture the facts. A definition of this second position – and thereby also of the first position – has to be given and (seeming) exceptions must be taken into account, too. As clitics have been of interest for different linguistic disciplines for decades now, many different theories have emerged, and they give different answers to the following questions: What are clitics? Why do they occupy a specific position? How can this position be defined? Can deviations be described as results from the general rule? Of course, not all approaches that concern themselves with clitic placement can be discussed here, so I will confine myself to the most relevant ones. Also, I mostly consider theories which either directly make reference to Czech or a closely related language, or which base their assumptions on a larger number of languages, including ones that display patterns similar to those found in Czech. The approaches presented in this chapter are arranged in four different groups: syntactic (section 6.1), prosodic (section 6.2), syntax-and-PF (section 6.3), and morphology-and-OT (section 6.4). After presenting and discussing representatives of these groups, I will provide an interim summary (section 6.5), before briefly turning to the issue of clitic ordering and how it is analysed in the respective approaches (section 6.6).

### 6.1 Syntactic approaches

Syntactic accounts all have in common that they assume clitic positioning to result from the clitics’ location within the syntax: they occupy a fixed head position which is located near the left edge, traditionally the C head. However, these theories also all allow phonology to play a role. It can trigger movement of elements in front of the clitic or it can determine the clitic’s potential hosts. In this sense, none of the approaches to clitic placement presented in this section functions entirely without interface effects. Nonetheless, the presented theories are all syntactic in the sense that the second position results from processes that take place *in the syntax*. Prosodic factors are assumed not out of the theories’ design, but out of the necessity to explain the data.

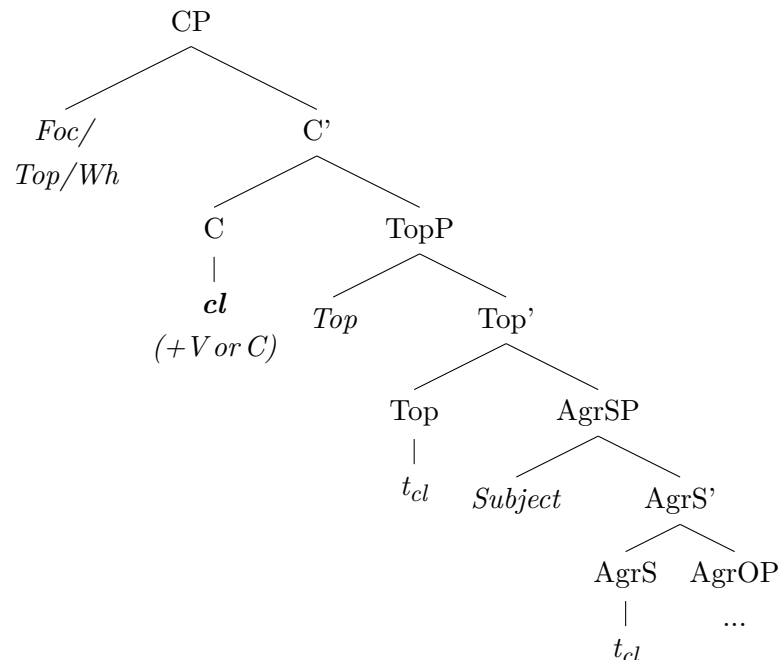
### 6.1.1 Veselovská 1995: Clitics in C

Veselovská (1995) investigates Czech syntax in general, and clitic placement in particular, in great detail. She assumes that clitics surface in C, an assumption also made in earlier proposals such as Čavar and Wilder (1994). Concerning the base position of auxiliary clitics, Veselovská (1995) argues that, as the past and conditional auxiliary clitics lack tense, aspect, and (in the third person) number, they are generated above TP; but, as they agree with the subject, they must be generated in AgrS (cf. Veselovská, 1995, 98-106).

Movement of the clitics from AgrS up to C is motivated as follows: The non-clitic finite verb does not need to rise to the highest verbal functional head (FH) because it fully reflects all FH's in its paradigm. Yet, for subject pro drop, all verbal FH's (from AgrO up to AgrS) must be strictly adjacent to the verb. Thus, no overt head may intervene. However, as stated above, clitic auxiliaries are generated in AgrS. In this position, they would intervene in the verbal FH chain. It follows that clitics must move to a higher position, i. e. to C (Veselovská, 1995, 91-2).

The element that precedes the clitics is either interrogative, or a focus, or a topic – or it is a verb, which is fronted into C when the clitic requires a host (cf. Veselovská, 1995, 107 and 123). This leads to strict adjacency of verb and clitic in this case (cf. Veselovská, 1995, 89 and 120). A similar constellation arises when the C head is filled by an overt complementiser. The positions she assumes for the Czech left periphery as well as the clitic's base and derived positions are illustrated in the tree structure in 87.

(87) The Czech left periphery and clitic positions according to Veselovská (1995, 15):



There is one necessary exception to C as the clitics' final landing site: the above described cases of clitic-third, where, in embedded sentences, a topic or focus phrase intervenes between complementiser (or interrogative phrase) and clitic. In this configuration, the clitic must remain in the Top head position. This is the reason for two topic positions in the structure in 87:

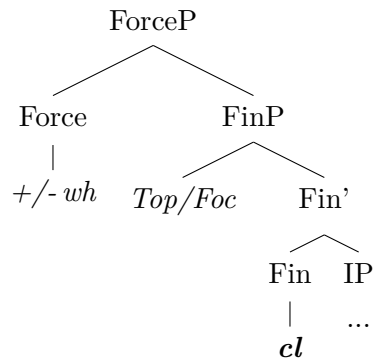
the higher position is for topics in matrix clauses, the lower position for topics in embedded sentences where the complementiser occupies C. It is not clear to me why the requirement that the clitics rise to the highest head position is lifted exactly under these circumstances.

The fact that the verb only moves up from AgrO to provide a host for the clitic prevents such ungrammatical configurations as \*C-V-Cl, \*Foc-V-Cl, \*Top-V-Cl etc. When a complementiser or an NP already precede the clitic, it has a host and thus there is no trigger for verb movement to the sentence-initial position.

### 6.1.2 Lenertová 2001: Clitics in Fin

In her 2001 paper, Lenertová is concerned with the structure of the Czech left periphery. She argues against a split left periphery in the style of Rizzi (1997), at least for Czech. In her analysis, the clitic is positioned in Fin, thus lower than in Veselovská's (1995) approach. The fact that the clitic never appears initially in a sentence and can under the above-mentioned circumstances appear in third position is captured through feature-checking: there is a feature in the C-domain that must be checked and thus triggers movement of an NP or a verb; this feature is obligatory in main clauses, but optional in embedded sentences (cf. Lenertová, 2001, 302), thus deriving the optionality of clitic-third configurations. In contrast to Veselovská's (1995) approach, the position of the clitic does not vary: it always occupies Fin. There is also only one position for topics, which is desirable as Czech does not permit multiple fronted topics (cf. Lenertová, 2001, 299). Lenertová's (2001) view of the Czech left periphery and the clitics' position in it is shown in 88.

(88) The Czech left periphery according to Lenertová (2001, 302):



When the verb moves to check the feature in the C-domain, it lands in Fin. This is not made explicit by the author but must be assumed with respect to the Head Movement Constraint. Fronted NPs land in the topic/focus position in Spec-FinP. Note that the approach does not assume that the clitic's prosodic deficiency triggers verb movement; instead, it is the above-mentioned C-domain feature.

### 6.1.3 Discussion

Many arguments against purely syntactic analyses of 2P clitics have been discussed in the literature, and many authors (for example Franks & Bošković, 2001 and Anderson, 1992) have

concluded that syntax on its own cannot be responsible for their placement. In this section, I will review the arguments made against syntactic analyses of 2P cliticisation and complement them with my own observations. This not only serves the purpose of convincing the reader that a purely syntactic account cannot be the solution we search for – many of the problems these accounts face also illuminate the nature of cliticisation, which is at its core an interface phenomenon.

First, the fact that clitics always appear second is often attributed to phonological deficiency, in the sense that an element, for example the verb, moves in front of the clitic to provide a phonological host for it. This means that the syntax must access phonology (cf. Richardson, 1997, 137), an assumption that is excluded in many theories, especially generative ones. Also, as was shown in section 5.1.3, Czech clitics do not necessarily require a host to their left: in some sentences, they do not have a prosodic host, even though they occupy the second position. These sentences, such as 73, repeated here as 89, are however perfectly grammatical.

- (89) *Ten doktor, co mu důvěruješ, se neholí.*  
 that doctor REL him.DAT trust.2SG REFL.ACC NEG:shave.3SG

‘That doctor, whom you trust, doesn’t shave.’

(cf. Toman, 1986, 125)

The alternative proposal, made by Lenertová (2001), that something always precedes the clitics because a feature in the C-domain must be checked, does not actually explain why clitics appear in second position, and also appears stipulated. What does it mean that this feature is optional in embedded sentences? And how does this optionality arise? Lenertová and Junghanns (2007) instead argue for an EPP feature which attracts the closest element to Spec-CP. Consequently, the verb only moves when no higher element, for example the subject, rises into the C-domain (Lenertová & Junghanns, 2007, 356).

Second, it remains unclear why clitics should move to a sentence-initial (head) position in the first place. The argument that their special movement properties are what makes them special clitics is not only circular, but also cannot explain why second-position clitics are so common cross-linguistically – the phenomenon should thus better be explained by independent principles. Whilst a common movement trigger, case, has been invoked to explain clitic movement, it is then unclear why other case-bearing phrases do not occupy the same positions as clitics.<sup>1</sup> Additionally, as has been shown by for example Lenertová (2004) with respect to clitic climbing, movement to the matrix clause occurs even when that clause very clearly cannot provide case for the clitic (cf. Lenertová, 2004, 159). This is illustrated by example 90, where an accusative clitic climbs to a passive matrix clause.

<sup>1</sup>When looking beyond the language sample of the present thesis, another interesting argument against feature-driven movement of clitics presents itself: As described by Skopeteas (2010), the Mayan languages not only have enclitics that spell out definiteness, person, and deictic categories, but also so-called “demarcative” clitics, which have no semantic value, but demarcate prosodic constituents (cf. Skopeteas, 2010, 307-9). Strikingly, interpretable clitics in these languages are placed according to the same principles as demarcative ones (cf. Skopeteas, 2010, 319). Since demarcative clitics arguably have no syntactic features, and no role in the syntax, their positioning cannot be related to such features – and thus, neither can the positioning of interpretable clitics.



- (90) *Ale tam ho byl donucen předat ho Lubinovi.*  
 but there it.ACC be.PASTPTCP force.PASSPTCP give.INF Lubin.DAT

‘But there he was forced to give it to Lubin.’

(cf. Lenertová, 2004, 159)

As explained in subsection 6.1.1, in Veselovská’s (1995) system, clitic movement to C is indirectly linked to the pro-drop parameter through the requirement that for pro-drop to be possible, no overt head may intervene between the verbal functional heads and the low verb itself. This is a strong claim that would require cross-linguistic investigation; note that the height of the non-clitic verb also plays a role, thus the connection is far from simple. For the Slavic languages at least, the languages with 2P clitics are mostly pro-drop (BCMS, Czech, Slovak, and Slovenian), but the Upper Sorbian colloquial language does not conform with this, having 2P clitics but no pro-drop (cf. Scholze, 2012, 204, 210). Conversely, not all Slavic pro-drop languages have 2P clitics, for example Polish and Bulgarian.

Another problem with Veselovská’s (1995) proposal is that it only explains the existence of 2P *auxiliaries*, not that of 2P *pronouns*. Being arguments of the head, the latter group is generated below V and thus would not intervene in the verbal chain. Whilst one could assume that they obligatorily move upwards, e. g. to AgrO in Veselovská’s (1995) system, this comes at the cost of stipulating obligatory object movement in a language where this is not otherwise attested (cf. section 3.4.1), and of assuming this only for an arbitrary subgroup of object pronouns.

Third, there is a considerable amount of data that shows that there cannot be one fixed sentential position for clitics. For example, in BCMS, the interpretation of adverbs like *pravilno* in 91 is dependent on their syntactic position: manner adverbs are adjoined to VP, whilst sentential adverbs are adjoined to TP. When a participle is fronted, as *odgovorio* in 91b, the sentential reading is not available any more. This suggests that the highest position a participle can rise to is below the sentential adverb’s position, i. e. below TP (cf. Bošković, 2001, 40).

- (91) a. *Jovan je pravilno odgovorio Mileni.*  
 Jovan AUX.3SG correctly answer.PTCP Milena.DAT

‘Jovan gave Milena a correct answer.’

‘Jovan did the right thing in answering Milena.’

- b. *Odgovorio je pravilno Mileni.*  
 answer.PTCP AUX.3SG correctly Milena.DAT

Only: ‘He gave Milena a correct answer.’

(cf. Bošković, 2001, 40)

Concerning the position of the clitic, the reasoning goes as follows: For the sentential reading in 91a to arise, the adverb must be adjoined to TP. The clitic in this example is thus *above* TP. In 91b, however, the participle is in a position below TP, as mentioned above. Crucially, the clitic must thus also be *below* TP. It follows that there cannot be one fixed syntactic position for

clitics (cf. Bošković, 2001, 41). In Czech, according to two of my consultants, there is a general preference for all types of adverbs to precede the participles, but placement after the participle is also possible, as shown in 92. Due to this lack of ordering preferences, the test can therefore arguably not be applied to Czech.

- (92) a. *Míla mi opatrně podala zkumavku.*  
 Míla me.DAT carefully pass.PTCP test.tube  
 b. *°Míla mi podala opatrně zkumavku.*  
 Míla me.DAT pass.PTCP carefully test.tube  
 ‘Míla passed me the test tube carefully.’  
 c. *Míla mi včera podala zkumavku.*  
 Míla me.DAT yesterday pass.PTCP test.tube  
 d. *°Míla mi podala včera zkumavku.*  
 Míla me.DAT pass.PTCP yesterday test.tube  
 ‘Míla passed me the test tube yesterday.’

Another argument in point concerns the fact that the element preceding the clitic(s) can be of any syntactic category. Most strikingly, whilst the category of the initial constituent is irrelevant, its prosodic realisation is not: when an element is a phonologically null *pro*, it cannot be a host, as 94 shows.<sup>2</sup>

- (94) a. *Maria ti pomáhá pověsit plakáty.*  
 Maria you.DAT help.3SG put.up.INF posters  
 ‘Maria will help you put up the posters.’  
 b. *pro Pomáhá ti pověsit plakáty.*  
 help.3SG you.DAT put.up.INF posters  
 ‘She will help you put up the posters.’  
 c. *\*pro Ti pomáhá pověsit plakáty.*  
 you.DAT help.3SG put.up.INF posters

None of these three variations involve any information-structural reordering of the sentence, and *Maria* must not necessarily be a topic or a focus. One can thus reasonably assume that the position which the subject *Maria* occupies in 94a is filled by a silent *pro* in both 94b and 94c. If the clitic occupied a fixed position in the sentence, it should be prosodically initial in these examples, but this renders the sentence ungrammatical, as 94c shows.

<sup>2</sup>This has also been noted for BCMS by Bošković (2001, 39):

- (93) *\*Je poljubio Anu.*  
 AUX.3SG kiss.PTCP Ana  
 ‘He kissed Ana.’

Fourth, in a sentence such as 94a, if the clitic were located at the left periphery, the pre-clitic NP would necessarily be topicalised. This is problematic, because 94a is grammatical in many different contexts, and is the neutral order in all-new contexts. A similar problem is observed for Vedic by Keydana (2011): One type of Vedic clitics attaches to the first phonological phrase ( $\varphi$ ) of their intonation phrase ( $\iota$ ). As argued by Nespor and Vogel (2007, 168), the syntactic input for this  $\varphi$  is the left periphery. Keydana (2011) points out the problem that this requires the left periphery to always be filled; otherwise the clitic would have no host. For Vedic this appears to be unproblematic, as it is for German, with the initial subject as a default topic (cf. Keydana, 2011, 124-5). It is not clear whether this is desirable for Czech, though. If possible, the left periphery should be reserved for information-structural or scope-related reordering.

Fifth, and finally, another argument against syntactic analyses is that the relative ordering of clitics cannot be captured, because it is dependent on phonological form, not on syntactic base position. For example, reflexive clitics, which can be direct objects, indirect objects or lexical, nevertheless share the same position in the cluster regardless of their syntactic role, as was shown in section 4.3 (cf. Richardson, 1997, 144). The different theoretic views on what underlies the ordering in the clitic cluster are sketched in section 6.6.

To summarise, syntactic approaches to clitic placement typically assume that clitics rise to a high position within the syntax. Additionally, some element must always precede the clitics, either for phonological reasons or due to the presence of some syntactic feature. Many problems of syntactic approaches are tied to the fact that it is difficult to find a motivation both for the movement of the clitic and for that of the pre-clitic element. Theories which assume that clitics have a fixed position in the syntax also have to face several types of data which contradict this assumption. All this does not mean that clitic placement is not at least partly syntactic – but to all appearances, its modelling requires mechanisms beyond that found in traditional purely syntactic analyses.

## 6.2 Prosodic approaches

For some languages, purely prosodic approaches to cliticisation have been proposed. These maintain that clitics are placed only with respect to phonological constituency. They do not occupy a certain position in the syntax, but instead appear in a position defined solely in prosodic terms. As we will see, this in fact appears to be the relevant mechanism for some languages – but not for all, as the case of Czech clitic placement clearly demonstrates.

### 6.2.1 Radanović-Kocić 1996: Prosodic placement in BCMS

For BCMS, clitic placement is often assumed to be entirely prosodic. Clitics in these languages are placed with respect to a purely prosodic constituent – which is partly built based on syntactic structure. An analysis along these lines is proposed for BCMS by Radanović-Kocić (1996). In her approach, whether an element is [+clitic] or not is only relevant for phonology, not for syntax. The feature [+clitic] is assigned to pronouns and auxiliaries, except when they carry phrasal stress or are preceded by an unstressed element (cf. Radanović-Kocić, 1996, 433); all

these elements must be placed after the first phonological phrase ( $\varphi$ ) of their intonational phrase ( $\iota P$ ; cf. Radanović-Kocić, 1996, 439-41).

The examples in 95 show how the approach works for BCMS, where prosodic structure has a great influence on clitic placement. Sentence 95a contains a restrictive relative clause, thus the speaker might have several sisters, and one of them is in Sarajevo. Here, the entire sentence forms one single  $\iota P$ , and the clitics appear after the first constituent, the subject. In contrast, 95b contains an appositive relative clause, so the speaker might have only one sister, about whom additional information is given. The appositive clause forms a separate  $\iota P$ ; the commas in this sentence coincide with intonational breaks that mark its edges. If the clitic is placed directly after this clause, as in 95c, it is initial in its  $\iota P$  and thus the sentence is ungrammatical.

- (95) a. *Ona moja sestra koja je u Sarajevu **vas se** sjeća.*  
 that my sister REL is in Sarajevo you REFL remember.3SG  
 ‘My sister who is in Sarajevo remembers you.’
- b. *Moja sestra, koja je u Sarajevu, sjeća **vas se**.*  
 my sister REL is in Sarajevo remember.3SG you REFL  
 ‘My sister, who is in Sarajevo, remembers you.’
- c. *\*Moja sestra, koja je u Sarajevu, **vas se** sjeća.*  
 my sister REL is in Sarajevo you REFL remember.3SG

(cf. Radanović-Kocić, 1996, 437)

This clearly shows that syntactic constituency is not the only decisive factor in BCMS for the definition of the second position. Another case that supports this point are heavy subjects, which can optionally be followed by a prosodic break. They consequently also optionally permit the delayed placement of clitics, as shown in 96.

- (96) a. *Kolutovi plavičastog dima **su se** penjali.*  
 circles bluish smoke AUX.3PL REFL rise.PTCP
- b. *Kolutovi plavičastog dima penjali **su se**.*  
 circles bluish smoke rise.PTCP AUX.3PL REFL
- ‘Bluish circles of smoke were rising.’

(cf. Radanović-Kocić, 1996, 435)

### 6.2.2 Discussion

Czech is clearly different from BCMS with regard to 2P clitic placement. For example, in Czech it is irrelevant for clitics whether a constituent includes an appositive relative clause or not, as the examples in 97 show. The intonational break in 97b does not lead to delayed clitic placement; instead, the clitic occurs directly after the first constituent, and thus after the break. Neither does a heavy subject, followed by a pause, induce delayed placement of the clitics; example 98b

is ungrammatical, because the clitic does not follow the first syntactic constituent. More data demonstrating that Czech clitics are insensitive to prosodic breaks was presented in section 5.1.3.

- (97) a. *Linda by chtěla cestovat do Londýna.*  
 Linda COND.3 want.PTCP travel.INF to London  
 ‘Linda would like to travel to London.’
- b. *Linda, která je Angličanka, by chtěla cestovat do Londýna.*  
 Linda REL is Englishwoman COND.3 want.PTCP travel.INF to London  
 ‘Linda, who is English, would like to travel to London.’
- c. *\*Linda, která je Angličanka, chtěla by cestovat do Londýna.*  
 Linda REL is Englishwoman want.PTCP COND.3 travel.INF to London
- (98) a. *Já a moje oblíbená teta Karolína jsme byli v Paříži.*  
 I and my favourite aunt Karolína AUX.1PL be.PTCP in Paris  
 ‘Me and my favourite aunt Karolína were in Paris.’
- b. *\*Já a moje oblíbená teta Karolína byli jsme v Paříži.*  
 I and my favourite aunt Karolína be.PTCP AUX.1PL in Paris

It is thus clear that whilst the prosodic approach appears adequate for BCMS, it cannot capture the Czech facts. Even though the two languages are rather closely related, clitic placement is obviously based on different principles: in BCMS, clitics are placed also with regard to prosodic constituents, whilst in Czech, the relevant constituents are syntactic. This is yet another way in which the concept of *second-position clitics* may take very different shapes in different languages.

Another notable difference between the two languages is that BCMS allows clitics to split up constituents in some cases, thus both sentences shown below in 99 are grammatical. In Czech, however, as example 100, first introduced as 52 in section 5.1.1, shows, placement of a clitic within an NP is not possible (but see section 9.1.3 for a different case of discontinuous constituents in Czech).

- (99) a. *Taj člověk je voleo Mariju.*  
 that man AUX.3SG love.PTCP Maria.ACC
- b. *Taj je člověk voleo Mariju.*  
 that AUX.3SG man love.PTCP Maria.ACC
- ‘That man loved Maria.’

(cf. Halpern, 1995, 16)

- (100) a. *Ten básník mi čte ze své knihy.*  
 that poet me.DAT read.3SG from his book  
 ‘That poet reads to me from his book.’
- b. *\*Ten mi básník čte ze své knihy.*  
 that me.DAT poet read.3SG from his book

(cf. Halpern, 1995, 17)

This might be related to the fact that Czech is generally not very permissive towards movements that split constituents (cf. Bošković, 2001, 154, fn. 59). Yet, it could also result from the different principles of clitic placement that have just been discussed. Whilst it is clear that the presented analysis of BCMS cannot be extended to Czech, it might be possible to find another approach under which clitic placement in both languages can be captured. This will be outlined in chapter 8 for Czech and in chapter 10.3 for BCMS.

### 6.3 Syntax-and-PF approaches

The fact that syntactic structure alone cannot explain the positioning of clitics does not necessarily entail that syntax has no role at all in clitic placement. To remedy the problems of syntactic approaches described in section 6.1.3, some authors have proposed that, like other auxiliaries, pronouns etc., clitics move in the syntax, but the final linearisation is influenced by prosodic requirements. Some earlier proposals in this direction have been made in the 1990ies, but syntax-and-PF approaches have especially gained traction in more recent publications.

#### 6.3.1 Halpern 1995: Prosodic Inversion

The approach to clitics developed by Halpern (1995) in the revised version of the author’s 1992 dissertation is basically a syntactic one: it assumes that clitics are syntactic entities. However, prosody as well as morphology play a role in clitic placement as well. The most well-known aspect of this is Prosodic Inversion, which entails that a “clitic may ‘trade places’ with a prosodic unit which is adjacent to it” (Halpern, 1995, 17). This possibility results from the mechanism of prosodic adjunction, which is defined as in 101.

(101) Prosodic adjunction of clitics

For a clitic X, which must have a prosodic host  $\omega$  to its left (respectively right):

- a. If there is a  $\omega$ , Y, comprised of material which is syntactically immediately to the left (right) of X, then adjoin X to the right (left) of Y.
- b. Else attach X to the right (left) edge of the  $\omega$  composed of syntactic material immediately to its right (left).

(cf. Halpern, 1995, 5)

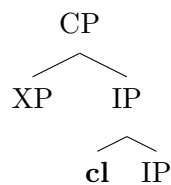
The last-resort option in 101b triggers a reversal of the linear order of a clitic and an adjacent prosodic word ( $\omega$ ). The central objective of this is to explain the placement of clitics within constituents, after the first word<sup>3</sup>. As mentioned in section 6.2, this occurs for example in BCMS, and it was illustrated there in example 99. Prosodic Inversion is only responsible for these configurations: note that it cannot reverse the order of a clitic and a following syntactic constituent, as prosodic adjunction only makes reference to prosodic words, not to syntactic constituency. Cases of clitics after the first constituent (“2D” in Halpern’s (1995) terminology) are the consequence of topicalisation of this initial constituent (cf. Halpern, 1995, 3-4).

In BCMS, both 2D and 2W placement can occur, though never in the same sentence at once. This is explained with topicalisation: if a topic phrase is moved in front of the clitics, they can p-adjoin to the  $\omega$  to their left, resulting in 2D placement. If, however, topicalisation does not take place, there is no material to the left of the clitics. In this context, Prosodic Inversion then occurs as a last-resort process, which leads to 2W placement (cf. Halpern, 1995, 5).

In this approach, syntax does not refer to neither phonological nor morphological information. Thus, the syntax treats clitics as syntactic entities and places them according to syntactic constituency. Phonology then treats them as prosodic material which is adjoined to another prosodic constituent (cf. Halpern, 1995, 2). On the other hand, the ordering of several clitics relative to each other is morphological: groups of clitics form a constituent whose internal ordering can be explained with the same mechanism that inserts inflectional affixes (cf. Halpern, 1995, 191).

An additional aspect in which the approach differs from other syntactic approaches is the syntactic position of clitics: they do not occupy a head position such as C, but are instead adjoined to the maximal projection of a constituent, for example to IP, similarly to adverbials (cf. Halpern, 1995, 57). Thus, for languages like Czech, Halpern (1995) proposes a structure as in 102.

(102) Adjunction of clitics according to Halpern (1995, 60):



### 6.3.2 Franks & Bošković 2001: Phases and copy deletion

Based on clitic data from Bulgarian and Macedonian, two closely related Eastern South Slavic languages, Franks and Bošković (2001) make an argument for multiple spell-out, tied to the assumption of phases. In contrast to Czech, clitics in these languages are verbal, i.e. they are always adjacent to the verb – nothing can intervene (cf. Franks & Bošković, 2001, 174). Typically, the clitics precede the verb, as example 103 shows.

<sup>3</sup>As mentioned before, Halpern (1995) calls this configuration “2W”, as opposed to “2D”, where the clitic follows the entire first constituent (cf. Halpern, 1995, 15).

(103) Bulgarian and Macedonian:

a. *Včera Petko mi go dade.*  
 yesterday Petko me.DAT it.ACC gave

‘Yesterday Petko gave it to me.’

b. \**Petko mi go včera dade.*  
 Petko me.DAT it.ACC yesterday gave

(cf. Franks & Bošković, 2001, 174)

In Bulgarian, however, there is an exception to this pattern: when the verb is initial in the sentence, the clitics follow instead of precede it, as 104b shows. This is the Tobler-Mussafia effect described in section 2.2. It is attributed to a difference in prosodic subcategorisation: Bulgarian clitics are always enclitic, which is why they cannot appear initially; this would be a PF violation. When the highest copy of a clitic would cause such a violation, a lower copy of the clitic is pronounced instead; this is in line with the “copy-and-delete” theory of movement (cf. Franks & Bošković, 2001, 176).

(104) a. Macedonian:

*Mi go dade Petko včera.*  
 me.DAT it.ACC gave Petko yesterday

b. Bulgarian:

*Dade mi go Petko včera.*  
 gave me.DAT it.ACC Petko yesterday

(cf. Franks & Bošković, 2001, 175)

However, this alone cannot explain the Bulgarian data shown in 105. First, sentence 105a illustrates that the conjunction *i* ‘and’ can host the clitics. Sentence 105b then illustrates the interaction of verbal clitics with the question marker *li*, which is also enclitic, but generated in C. Now, crucially, in 105c, we see what happens if both *i* and *li* occur in a sentence – the linearisation in 105c is identical to that in 105b; the verbal clitics follow the verb and *li*, as if the initial element, *i*, was not there.

(105) Bulgarian:

a. *I mi go dade Petko včera.*  
 and me.DAT it.ACC gave Petko yesterday

‘And Petko gave it to me yesterday.’

b. *Dade li ti go Petko včera?*  
 gave Q you.DAT it.ACC Petko yesterday

‘Did Petko give it to you yesterday?’



- c. *I dade li ti go Petko včera?*  
 and gave Q you.DAT it.ACC Petko yesterday  
 ‘And did Petko give it to you yesterday?’

(cf. Franks & Bošković, 2001, 178)

Franks and Bošković (2001) capture this with multiple spell-out. Following Chomsky (2000), they assume that CP is a phase, but IP is not. Since *li* is a complementiser, it introduces a CP, thus 105b and 105c both have a CP layer. In contrast, 105a does not. If spell-out occurs after CP is built, and encliticisation then takes place after spell-out, this means that in 105c encliticisation occurs *before* the conjunction *i* is merged – which means that *i* is “literally not there at the point when encliticisation takes place” (Franks & Bošković, 2001, 181). This is why a lower copy of the clitics must be pronounced to avoid a PF-violation at this point. The structure assumed for sentences 105b and 105c is shown in 106.

- (106) a. [<sub>CP</sub> [<sub>C</sub> [ti go dade] + li] **ti go dade** Petko včera] (cf. ex. 105b)  
 b. i [<sub>CP</sub> [<sub>C</sub> [ti go dade] + li] **ti go dade** Petko včera] (cf. ex. 105c)

(cf. Franks & Bošković, 2001, 180)

### 6.3.3 Bošković 2004: Copy deletion instead of PF movement

Bošković (2004) bases his analysis on BCMS, Bulgarian, and Macedonian data. He assumes that clitics are placed in the syntax, but that the second position effect results from PF requirements (Bošković, 2004, 38). In contrast to pure syntactic accounts, in this approach the clitics do not occupy one shared syntactic position (Bošković, 2004, 50). Also, PF movement (i. e. Prosodic Inversion) is rejected: as in Franks and Bošković (2001), PF restrictions instead prevent the pronunciation of the highest copy (Bošković, 2004, 44 and 63). As for the clitics’ precise syntactic nature, they are generated as morphological heads in phrasal positions and thus can undergo head movement (Bošković, 2004, 77, 81). Using the copy-and-delete mechanism and a clitic-specific PF requirement, the fact that clitics in BCMS do not occur directly after an intonational break, as shown in 107 (repeated from example 95 in section 6.2), is captured as sketched in 108.<sup>4</sup>

- (107) *Moja sestra, koja je u Sarajevu, sjeća vas se.*  
 my sister REL is in Sarajevo remember.3SG you REFL  
 ‘My sister, who is in Sarajevo, remembers you.’

(cf. Radanović-Kocić, 1996, 437)

<sup>4</sup>Bošković (2004) assumes that auxiliary clitics rise to AgrS, dative clitics to AgrIO, and accusative clitics to AgrDOP, whilst the main verb remains in VP, which is also where the pronominal clitics are generated; (cf. Bošković, 2004, 59).

(108) Delayed clitic placement following Bošković (2004):

1. Clitics are placed in their syntactic position:  
Moja sestra, koja je u Sarajevu, **vas se** sjeća ~~vas~~-se.
2. PF requires that BCMS clitics must not be initial in  $\iota P$ :  
Moja sestra, koja je u Sarajevu, ~~vas~~-se sjeća **vas se**.

### 6.3.4 Franks 2017: Relinearisation due to prosodic deficiency

Also in Franks's (2017) approach, clitics are placed in the syntax. Yet, they require prosodic support as they cannot project prosodic feet (cf. Franks, 2017, 147). Phonological phrases that constitute potential boundaries for clitic placement are  $\omega$ ,  $\iota P$ , and the utterance. Enclisis is thus the consequence of an inability to appear in some prosodically initial position (cf. Franks, 2017, 115-6). Syntactically, clitics are ambiguous between heads and phrases, an idea taken from Chomsky (1994).<sup>5</sup> The operations involved in clitic placement are summarised in 109.

(109) The mechanics of clitic placement according to Franks (2017, 216):

1. The clitics are placed in a certain syntactic position.
2. Spell-Out linearises them according to the Linear Correspondence Axiom (i. e. based on *c*-command, cf. Kayne, 1994).
3. Under prosodification, they adjoin to an adjacent  $\omega$ , following the relevant prosodic initiality restrictions, which may lead to re-linearisation.

### 6.3.5 Discussion

As demonstrated above, the Prosodic Inversion account of Halpern (1995) is designed to handle cases of 2W, found for example in BCSM, where clitics appear within constituents. Yet, it provides no additional explanation for languages such as Czech, where clitics are always placed after the first constituent, and thus in this respect faces the same problems as pure-syntax accounts, discussed in section 6.1.3. It also only appears to prosodic en- and proclitics, as can be seen in the definition of Prosodic Inversion in 101.

Halpern's (1995) analysis has also been criticised because it assumes PF movement: when the clitic does not find a suitable host to its left, it moves one position to the right in the prosodic structure. However, Bošković (2004) shows for BCMS that the elements which occupy the first position in 2W placement are actually syntactically independent. Given that PF movement should be assumed only for phenomena that are clearly not syntactic, he therefore rejects the notion of Prosodic Inversion (Bošković, 2004, 49).

The approaches by Bošković (2004) and Franks (2017) successfully capture clitic placement in South Slavic languages such as BCMS, where clitics are placed with respect to prosodic

---

<sup>5</sup>The mechanism which Franks (2017) employs is similar to copy-and-delete; instead of copies, though, he assumes multiattachment. This is an idea inspired by the computer science notion of pointers: trees are "recursive data structures wherein nodes are pointers either to other nodes (the recursive case) or to vocabulary items (the base case)" (cf. Franks, 2017, 48-50).

constituency, as was shown above. In Czech, however, the relevant domains are syntactic. The PF requirement in Czech would thus be: “clitics must not be initial within their CP”. This in consequence means that PF would have to access syntactic constituency, which it should not be able to do. The assumption of a PF component separate from syntax only makes sense if that component is not also occupied with syntactic information; it should thus only refer to phonological structures.

The only way PF can indirectly refer to syntactic structure is through multiple spell-out: in such an approach, syntax derivationally feeds phonology, with movement only in the syntax. Assuming that CP is a phase, i. e. a point at which the syntactic derivation is passed on to PF, it follows that the left CP edge can be indirectly visible also in PF operations (cf. Franks & Bošković, 2001, 179-80). We could therefore formulate a non-initiality restriction for Czech clitics that applies at PF without having to refer directly to syntactic constituency. The fact that clitic placement occurs each time a CP is built could also capture the clitic climbing facts described in section 5.2.1: it directly prevents clitics from appearing in a CP higher than their own. This idea will become relevant again in section 8.3.2.

Another way of explaining the clitics’ non-initiality would be the argument given for re-linearisation by Franks (2017), namely the requirement of prosodic support. This is, as we have seen in section 5.1.3, also problematic for Czech: whilst clitics in this language do need *some* phonological host, it does not necessarily have to be to their left.

Whilst the copy-and-delete approaches, especially when expanded by the assumption of phases, successfully solve the problem of purely syntactic theories that they blur the lines between syntax and phonology, another problem remains: assuming the deletion of higher copies of the clitics still means that clitics have to rise to a higher position to begin with – but, as we have seen in section 6.1.3, it is difficult to find a syntactic trigger for this movement.

Finally, there is an interesting claim that is made by many syntax-and-PF theories, for example Franks (2017): clitics are seen as syntactically ambiguous between heads and phrases (cf. Chomsky, 1994, 16). Whilst such a statement might appear questionably at first glance, clitics in fact provide good grounds for this view: First, Chomsky (1994) argues that pronominal clitics must be phrasal in their theta positions, but that they then undergo head movement and attach to an inflectional head (cf. Chomsky, 1994, 16). Second, even if one does not share this latter assumption, the Czech clitic system on its own already provides an argument for such a view: as we have seen, Czech clitics comprise both pronouns and auxiliaries, i. e. phrases as well as heads, and yet these two groups show exactly the same behaviour (cf. chapter 4).

To summarise: The syntax-and-PF approaches to clitic placement successfully remedy some serious problems of purely syntactic accounts by recognising that 2P cannot be explained solely in terms of syntactic processes, but is due to additional restrictions which can be located in the phonological component. They also provide a coherent view of how syntax and PF interact, and how syntactic structure can indirectly be visible in PF processes. Yet, they must maintain that before these processes take place, syntactic movement places them in a high position in the derivation.

## 6.4 Morphology-and-OT approaches

As has been shown above, placing clitics solely in the syntax has some severe problems. However, it has also become evident from reviewing prosodic approaches to clitic placement that in Czech, cliticisation is strongly dependent on syntactic constituency. How can this be captured in a theory without stipulating otherwise unmotivated movements? And from which principles can 2P effects be derived? An attempt to solve these problems is the A-Morphous Morphology framework developed by Anderson (1992) and later works. Here, morphology constitutes a unit of grammar in its own right, instead of just being part of the syntax. Clitics play a central role in this approach: they are seen as morphological entities which are not inserted in the syntax, but can nevertheless be placed with respect to syntactic structure. The mechanics of this placement are not derivational, but constraint-based. Using such clitic-specific constraints, Richardson (1997) develops an analysis of Czech 2P positioning.

### 6.4.1 Anderson 1992/2000/2005: Clitics as phrasal affixes

Anderson (1992) and later works by that author approach clitics from a different viewpoint than syntactic theories: he treats special clitics (i. e. clitics that display idiosyncratic syntactic behaviour) not as syntactic material, but as affixes added to phrases instead of words. They express the properties of their phrase and/or modifications to their phrase's semantics or discourse properties. One argument for this is that just like word-level affixes, special clitics are typically prosodically deficient (cf. Anderson, 1992, 198-200). This is also in line with the observation made in section 2.1 that clitics are phonologically very similar to affixes.

In order to model different clitic placement patterns, Anderson (2000) and Anderson (2005) use interacting constraints in an Optimality Theoretic (OT) framework. The constraints require clitics to align with certain edges, or they ban them from doing so. Since constraints are violable, and since one constraint is typically more "important" than the other, their effect is not absolute. The implications of these assumptions by Anderson (2000/2005) will be illustrated in the following subsection, when presenting Richardson's (1997) application of this approach to Czech.

Concerning the differences between 2D and 2W clitic placement that we have seen with respect to BCMS, Anderson (2000) derives them from different Integrity constraint rankings (cf. Anderson, 2000, 24). An Integrity Constraint is one that penalises the breaking up of a certain structure; since clitics are affixes, they can enter structures that syntax has no access to (cf. Anderson, 2000, 34). The role of Integrity will be illustrated for Czech in section 8.2.2, and for BCMS in chapter 10.3.6.

### 6.4.2 Richardson 1997: Constraining Czech clitic placement

The OT-based account of clitic placement as developed by Anderson (1996) is applied to Czech by Richardson (1997): Clitics are not verbal or pronominal heads, but phrasal affixes; there is thus no syntactic position for the clitic. Clitics (and affixes) are placed with respect to a certain domain. Crucially, each constraint can have its own domain, so mismatches occur. In Czech,

clitics need to be leftmost in IP, but must not be initial in CP; when IP and CP coincide, typical 2P is the result. The definition and ranking of the two constraints is as shown in 110.

- (110) a. NONINITIAL(CL;CP): A given clitic must not be initial in CP.  
 b. EDGEMOST(CL,l;IP): A given clitic should be as close to the left edge of IP as possible.  
 c. Ranking: NONINITIAL(CL;CP) >> EDGEMOST(CL,l;IP)

(cf. Richardson, 1997, 146-7; Anderson, 2005, 136-7)

The interaction of the two constraints can capture both 2P placement and clitic-third with fronted topics/foci in embedded clauses: When the subject is not topicalised, it remains within IP. The clitic is then placed at the left edge of IP, which satisfies the lower-ranked EDGEMOST constraint and, due to the presence of the complementiser in C, does not violate the higher-ranked NONINITIAL constraint either. This is shown for 111a in 112a. However, when the subject *is* topicalised, it moves out of IP, to the left periphery. When the clitic is now placed at the left edge of IP, it again violates neither the lower nor the higher constraint. It is preceded by two constituents because these are both above IP (cf. Richardson, 1997, 146-8).

- (111) a. ... *že se Petr odstěhoval.*  
           COMP REFL Petr move.PTCP  
 b. ... *že Petr se odstěhoval.*  
           COMP Petr REFL move.PTCP  
       ‘...that Petr has moved.’

(cf. Fried, 1994, 159)

- (112) a. [<sub>CP</sub> *že* [<sub>IP</sub> *se* Petr odstěhoval]]  
 b. [<sub>CP</sub> *že* Petr [<sub>IP</sub> *se* odstěhoval]]

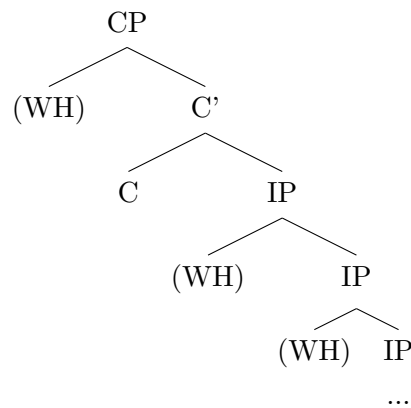
These are cases which might also be handled in the syntax, since they simply require the placement of the clitic in a position at IP’s left edge. However, as discussed above, verb-initial structures as in 113a are more problematic for syntactic accounts. In contrast, the interaction of the two constraints proposed by Richardson (1997) does not require the verb to move to host the clitic. Instead, the clitic is placed after the verb because it would otherwise be initial in the clause, violating the high-ranked NONINITIAL constraint, see 114a. A similar logic applies to initial, but not topicalised, subjects, as in 113b: they need not move to an IP-external position just to precede the clitic; see 114b (cf. Richardson, 1997, 148).

- (113) a. *Odstěhoval se.*  
 move.PTCP REFL  
 ‘He has moved.’
- b. *Petr se odstěhoval.*  
 Petr REFL move.PTCP  
 ‘Petr has moved.’

- (114) a. [CP [IP Odstěhoval se]]  
 b. [CP [IP Petr se odstěhoval]]

In his assumptions about Czech phrase structure, Richardson (1997) follows Rudin (1988). In her approach, subjects rise to the specifier of IP (cf. Richardson, 1997, 133). She also assumes that the Doubly-Filled Comp Filter is active in Czech, i. e. either Spec-CP or Comp must be empty (cf. Rudin, 1988, 487). Based on data with multiple interrogative phrases, she arrives at the structure for Czech shown in 115.

- (115) Czech phrase structure according to Rudin (1988, 478 and 486):



### 6.4.3 Discussion

The morphology-and-OT approaches by Anderson (1992/2000/2005) and Richardson (1997) avoid some of the major problems faced by the purely syntactic accounts: since they do not assume that clitics move in the syntax, they can capture the fact that they underlie different restrictions than regular syntactic entities like full NPs and main verbs. Thus, they can capture both the fact that clitics appear near the left edge of the clause, and the fact that they avoid being absolutely initial, with the use of entirely different mechanisms than are typically found in the syntax. In addition, the idea that clitics are not syntactic elements has a special appeal for Czech, where this group consists of both heads (auxiliaries) and phrases (pronouns), which nonetheless display the same peculiar behaviour.

However, an issue that remains open so far is how we can nevertheless ensure that clitic elements can fulfil the partly crucial roles that they evidently play in the syntax. Most strikingly,

how can we ensure that clitic pronouns occupy the slots they have in the verbal subcategorisation frame; or, the other way round, how do these pronominals receive their theta roles? And how are they assigned case? These are important questions that require an answer. At this point, I will only shortly remark that, if the OT-style constraints proposed by Richardson (1997) are integrated into a fully optimality-theoretic analysis in the sense of Prince and Smolensky (1993), then the issue of separate linguistic layers ceases to apply. In a parallel and monostratal OT architecture, it is unproblematic to capture the interactions of clitics with different levels of linguistic organisation. Chapter 8 will present a proposal for the analysis of Czech clitic placement in this spirit.

Before concluding this discussion, I want to remark on a technical detail of Richardson's (1997) account that requires closer attention: For a structure as in 112b, with a topicalised phrase (*Petr*) between complementiser and clitic, there must be a position below C but above IP that *Petr* can rise to (cf. Richardson, 1997, 148). In the structure proposed by Rudin (1988), given in 115, this does not exist. As in Lenertová's (2001) analysis, such a position should hold both foci and topics in order to explain their mutual exclusiveness. Yet, explicit topic and focus positions as in Italian (cf. Rizzi, 1997) are rejected for the Slavic languages by for example Lenertová (2001) and Fehrmann and Junghanns (2012). Lenertová (2001) argues that the approach by Rizzi (1997), who assumes CP to be split into Force, Top\*, Foc, Top\*, Fin, is not adequate for Czech, for several reasons: first, in Czech, multiple topics are not permitted; second, in the left periphery interrogative phrase+focus is possible, but not interrogative phrase+topic; and third, there are cases of topics preceding the complementiser (Lenertová, 2001, 298-302).

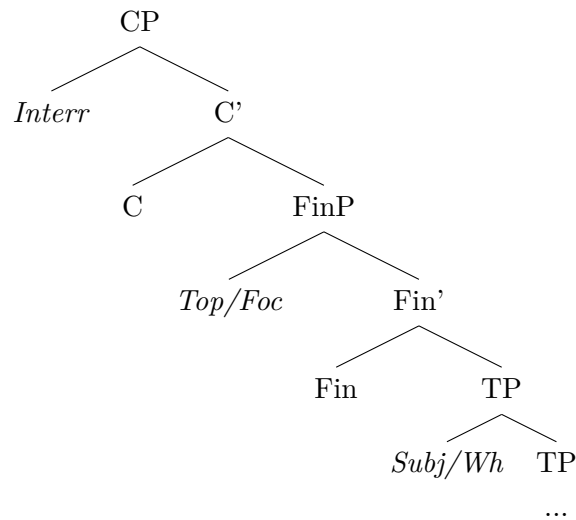
A Czech phrase structure which is thus both empirically motivated and adequate for the presented morphology-and-OT analyses could be the one given in 116 (cf. Fehrmann & Junghanns, 2012, 83). With this, the analysis of the above sentences in 111 and 113 now looks as in 117. The same results as illustrated above are obtained, but now there is a position for topicalised phrases, which actually was not available in 112b.<sup>6</sup>

---

<sup>6</sup>The structure does not exclude ungrammatical sequences of interrogative phrase and topic. Yet, as mentioned in footnote 3 of section 5.1.2, this cannot easily be remedied, since on the other hand, sequences of interrogative phrase and focus are allowed, whilst topic and focus are mutually exclusive. This is a puzzle which cannot be solved within the limits of this dissertation.

For the lower landing site of interrogative phrases in Spec-TP, see Meyer (2004).

(116) Updated phrase structure for Richardson's (1997) analysis:



- (117) a. [CP že [FinP [TP se Petr odstěhoval ]]] (cf. ex. 111a)  
 b. [CP že [FinP Petr [TP se odstěhoval ]]] (cf. ex. 111b)  
 c. [CP [FinP [TP Odstěhoval se ]]] (cf. ex. 113a)  
 d. [CP [FinP [TP Petr se odstěhoval ]]] (cf. ex. 113b)

## 6.5 Interim conclusions

It has become clear that neither pure syntactic nor pure phonological accounts can capture the Czech data; the former because syntactic mechanisms alone cannot derive the second position in which the clitics appear, as well as the position(s) before them; the latter because these approaches are designed for languages in which clitics are placed with respect to phonological, and not syntactic, constituency, and in which clitics are mostly strictly enclitic.

We therefore need a theory of clitic placement that allows us to combine the syntactic aspects of clitic placement with the non-syntactic, possibly phonological, aspects in a well-defined manner. This is possible both in the syntax-and-PF and the morphology-and-OT approaches. Whilst the first group claims that clitics have a position in the syntax that may be obscured by prosodic requirements, according to the second group, clitics are only later inserted by the morphology (Anderson, 2000).

In both accounts, 2P does not result from a simple rule that counts constituents and then places the clitic in second position. Instead, it is the effect of a restriction on clitics not to appear initially, combined with some mechanism or constraint that places them in a relatively high position. The two approaches offer two different explanations for this latter factor. According to syntax-and-PF accounts, clitics occupy a high position in the syntax, due to some featural requirement. But according to morphology-and-OT accounts, they need to be initial in a smaller domain, which is captured by a clitic-specific constraint.



In the present thesis, I will pursue an optimality-theoretic approach that aims to incorporate important insights from both theoretic branches. I will employ constraints like those proposed by Richardson (1997) that can capture Czech clitic placement in an accurate and elegant manner. At the same time, I will show that we can nevertheless maintain that clitics also play a role in other parts of the grammar, for example with respect to theta roles. For now, I am leaving open the question of whether clitics are morphological elements (i. e. phrasal affixes), or whether they can be treated in line with full NPs, verbs etc. I will come back to this issue in chapter 7.

## 6.6 Accounting for clitic ordering

The relatively fixed order of Czech clitics within the clitic cluster was demonstrated in section 4.3. There are different approaches to explaining the causes of that ordering – whilst some assume that it derives directly from syntactic position (6.6.1), others hold that the final surface order derives from phonology altering the syntactic output (6.6.2). Accounts that model clitic ordering independently from syntactic or phonological factors are on the one hand morphology-and-OT approaches, which apply the principle of ranked constraints to derive clitic order (6.6.3); and on the other hand purely morphological approaches, which hold that clitic ordering is simply the result of a morphological template (6.6.4).

### 6.6.1 Syntactic approaches

Syntactic accounts assume that not only the placement of the clitic cluster within the clause, but also the clitics' respective ordering derives entirely from their syntactic positions, and that no additional assumptions need be made for other parts of the grammar.

**Veselovská (1995)** holds that all Czech clitics, both verbal and pronominal ones, are heads that move in the syntax and cluster in C, as described in section 6.1.1. The ordering within the clitic cluster directly reflects the clitics' syntactic positions (cf. Veselovská, 1995, 314).

**Progovac (1996)** analyses the order of clitics within the cluster in BCMS, which can be sketched as shown in 118. She claims that clitic ordering in BCMS can be derived referring only to general syntactic rules, and that no manipulation of clitic ordering occurs outside of syntax.

(118) Clitic template for BCMS (cf. Progovac, 1996, 420):

*li* – AUX – DAT – ACC – *se*

The interrogative *li* appears first in the cluster because it is generated in C, and all other clitics right-adjoin to it. Auxiliaries are in Infl, which pronominal clitics must rise to, so auxiliaries precede pronominals. To explain why dative clitics precede accusative clitics, Progovac (1996) invokes Pesetsky's (1982) two-dimensional Path Condition, which is defined as in 119. This also explains the position of the reflexive *se*, assuming it is in the direct object position (cf. Progovac, 1996, 420-2). Thus, in this view, only one strategy for clitic placement, i. e. a syntactic one, is needed.

(119) Crossing Constraint – 2 dimensions (informal):

The lower an A'-bound trace occurs in a tree, the higher its binder must occur, in relation to other trace-A'-binder dependencies.

(cf. Pesetsky, 1982, 281)

### 6.6.2 Syntax-and-PF approaches

In contrast to purely syntactic accounts, syntax-and-PF approaches to clitic ordering hold that the order of clitics, although derived from their syntax, can be manipulated by additional requirements in PF.

**Bošković (2004)** discusses the order of clitics in BCMS (cf. the template in 118 above). He assumes that auxiliary clitics are structurally higher than pronominal clitics and that they do not share the same head position. Neither do the dative and the accusative clitic cluster together (cf. Bošković, 2004, 55-7). He also discusses the exceptional behaviour of the third person singular auxiliary clitic *je*, which in BCMS, in contrast to the other auxiliaries, appears *after* the pronominals in the cluster. He captures this by assuming that *je* is higher in the syntax than the pronominals, but there is a PF constraint that requires *je* to follow all other clitics; to achieve this, a lower copy is pronounced (cf. Bošković, 2004, 66). Thus, as with clitic placement described in section 6.3.3, Bošković (2004) also assumes for clitic ordering that it largely reflects syntactic positions, but that PF constraints can force the pronunciation of a lower copy.

**Franks (2017)** examines a range of Slavic languages, including Czech. Like Bošković (2004), he claims that the order of clitics generally derives from their syntax. For example, the precedence of dative before accusative clitics directly follows from syntactic structure, as does the precedence of *li* before auxiliaries (cf. Franks, 2017, 260). This ordering can then be manipulated by PF, which leads to deviations from the general rule, for example to PCC effects. In this context, he remarks that the traditional standard formulation of PCCs, as given for Czech in section 4.4, actually makes the wrong prediction, because they refer to direct and indirect objects, but not to linear order – yet, as we have seen, PCC violations can be avoided through the reordering of clitics, as example 48, repeated from section 4.4, illustrates. For this reason, Franks (2017) reformulates the PCC's with reference to linear order/structural height, as given in 120 (cf. Franks, 2017, 263-4).

(120) Reformulation of PCC's by Franks (2017, 265):

- a. Weak person-ordering constraint: In a combination of clitic pronouns, if there is a 3rd person, it has to be last/lowest.
- b. Me-first person-ordering constraint: In a combination of clitic pronouns, if there is a 1st person, it has to be first/highest.
- c. Strictly descending person-ordering constraint: In a combination of clitic pronouns, the argument with the “higher” person specification (1 >> 2 >> 3) has to be first/highest.

For the third-person singular auxiliary clitic *je* in BCMS, Franks (2017) assumes that its position can be captured in syntactic terms: it is lower than other auxiliaries, in T (whereas the other clitics are in Agr), because *je* has no featural content. This is also consistent with the situation in West Slavic; as we have seen for Czech in section 4.1, the third-person auxiliary clitic is null (cf. Franks, 2017, 225).

### 6.6.3 Morphology-and-OT approaches

As with respect to clitic placement, morphology-and-OT approaches also assume that the order *within* the clitic cluster is not determined by syntax, because clitics are phrasal affixes, and their placement as well as ordering is dependent on the interaction of violable constraints, and not on syntactic structure.

**Richardson (1997)** describes several problems posed by Czech clitic ordering with respect to purely syntactic accounts: First, ethical datives, which are not arguments, appear in the same position as argument datives – yet, their positions in syntax cannot be the same. Second, reflexive clitics can be accusatives, datives or lexical, but all share the same position; again, it is highly plausible to assume that they occupy different positions. Third, there are contracted forms of second person singular auxiliary and reflexive clitics (*jsi + se = ses*, *jsi + si = sis*) which cannot be derived syntactically from their underlying components (cf. Richardson, 1997, 144-6).

For these reasons, Richardson (1997) again follows Anderson (1992) and later works by that author in assuming that violable, ranked OT constraints regulate clitic ordering. These constraints refer to phonological realizations of the clitics, not to syntactic positions. They are more specific versions of the EDGEMOST constraint already defined in 110, and they all require that the element they refer to appear as far left in IP as possible; however, due to their different ranking, given in 121, some clitics will precede others.

- (121) EDGEMOST(Past/Cond,l,IP) >> EDGEMOST(Refl,l,IP)  
>> EDGEMOST(Dat,l,IP) >> EDGEMOST(Acc,l,IP)

(cf. Richardson, 1997, 146-8)

**Anderson (2000)** also claims that syntax cannot explain the relative order of clitics and that clitic placement is “not an instance of Lexical Insertion, but rather the overt manifestation of a phrasal analog to a Word Formation Rule”, which “modifies the PF shape of a phrase on the basis of its functional content” (Anderson, 2000, 15). Thus, what is often seen as an autonomous head is really part of the representation of a phrase’s properties.

One argument he gives for this that the shape of a clitic can be influenced by the one of a clitic that follows it, independently of general phonological rules in the language. This is difficult to explain if clitics are inserted from left to right. Additionally, a varying position (left vs. right) of the host does not lead to mirror-image clitic order; this means that clitics are introduced simultaneously, which can best be modelled in OT (cf. Anderson, 2000, 16-8). The order of two clitics is determined by the respective ranking of their corresponding EDGEMOST constraints. The general constraint format also employed by Richardson (1997) is given in 122.

- (122) EDGEMOST(e,E,D): e = element, E = edge, D = domain

(cf. Anderson, 2000, 19)

### 6.6.4 Morphological approaches

Finally, there are non-OT approaches that also claim that clitic ordering is determined morphologically, not syntactically.

**Fried (1994)** investigates Czech clitic placement and order. Based on the fact that the Czech clitic cluster cannot be interrupted by other material (cf. example 45 in section 4.3), she concludes that the clitic cluster behaves just like a phonological word and should therefore be analysed as one. She assumes a template with a total of eight slots (due to the fact that she also includes placement of *li* ‘if’ and *však* ‘however’), which the clitics enter before the entire template takes up its place in the sentence (cf. Fried, 1994, 171-3).

Further approaches that assume that clitic ordering is independent from the syntax and instead regulated by a morphological template are Halpern (1995) and Zimmerling (2012).

### 6.6.5 Discussion

The preceding subsections have shown that the general order of clitic elements within a cluster can partly be derived from their assumed positions in the syntax. Yet, many authors have also recognised that idiosyncrasies as, for example, PCC effects cannot be handled in purely syntactic accounts. Also phenomena like the separate placement of reflexives, regardless of their syntactic role, cast doubt on a purely syntactic analysis.

A further problem not addressed by any of these theories is the respective ordering of dative and accusative clitics: Dvořák (2010) convincingly argues that there are two types of ditransitive verbs, one group where the dative argument is generated higher than the accusative argument, and one group where the dative argument is generated first, and thus c-commanded by the accusative argument. Yet, in contrast to what would be predicted by purely syntactic theories, these different verb types do not display different clitic ordering.

Clitic ordering thus provides additional evidence that an adequate theory of 2P clitic positioning in Czech and beyond is one that permits the interaction of different linguistic levels. In section 8.5, I will sketch how we can incorporate the cluster-internal ordering into the constraint-based analysis of Czech clitic placement developed in this thesis.

## Chapter 7

# Comparing Clitic and Non-Clitic Verbs

In the course of this thesis, it has become clear that clitics in Czech display a behaviour quite different from that of non-clitics. In the case of clitic verbs, this is particularly interesting given the fact that both clitic and non-clitic auxiliaries are typically described as belonging to one verbal lexeme, *být* ‘to be’, and are in parts formally very similar, as illustrated in table 7.1, repeated from section 4.1. When trying to explain their distinct behaviour, the question arises whether this traditional view is correct and the clitic and non-clitic variants of *být* synchronically derive through a (de-)cliticisation process, or whether they are distinct lexical items. If the latter is true, then we must ask if clitic auxiliaries are verbal heads, or if they belong to a different category. This is an important question to answer if we wish to develop a theory of Czech clitic placement.

COP	PASS.AUX	PAST.AUX	COND.AUX	FUT.AUX
jsem (su)	jsem (su)	<b>jsem</b>	<b>bych (bysem)</b>	budu
jsi (seš)	jsi (seš)	<b>jsi</b>	<b>bys (bysi)</b>	budeš
je	je	–	<b>by</b>	bude
jsme	jsme	<b>jsme</b>	<b>bychom (bysme)</b>	budeme
jste	jste	<b>jste</b>	<b>byste</b>	budete
jsou	jsou	–	<b>by</b>	budou

Table 7.1: Copula and auxiliary forms of the verb *být* (2P clitics in bold, colloquial forms in brackets)

The majority of the discussion in the following sections is devoted to the differences between the clitic and non-clitic parts of the *být* paradigm, both morphologically (section 7.1), and syntactically (section 7.2). I then approach the special problem of what distinguishes the past and the passive auxiliary (section 7.3). Finally, I discuss the differences between the two clitic auxiliaries, the past and the conditional (section 7.4). Since both the clitic auxiliaries and the passive auxiliary combine with participles, the nature of these will also be central to the discussion. I will show that the past auxiliary has lost its verbal stem and has become affixal, which in consequence drives the past participle to develop into a finite verb. This is not the case

for the passive participle, which combines with the (only phonologically clitic) passive auxiliary. The conditional auxiliary will be shown to be a combination of two clitics.

## 7.1 Morphology of clitic and non-clitic auxiliaries

In the following subsections, I will first sketch four different approaches to the status of clitics with respect to non-clitics (section 7.1.1); in particular the copula, the passive auxiliary and the past auxiliary, i. e. the forms stemming from the present tense of *být*, since these are formally almost identical, as can be seen in table 7.1. The competing approaches will then be evaluated with respect to different kinds of morphological data: omission and reduction (section 7.1.2), colloquial allomorphy (section 7.1.3), and frequentative formation (section 7.1.4).

### 7.1.1 Four views on the status of clitic auxiliaries

Fried (1994) assumes that the different clitic and non-clitic present tense forms of *být* ‘to be’ are “clitic and non-clitic uses of the same lexical item” (Fried, 1994, 163), only distinguished through their prosodic subcategorisation frame, which is why she speaks of “prosodic allomorphy”. The first person singular copula/auxiliary, for example, can then either be prosodically free, or required to incorporate into a preceding prosodic word, as represented in 123.<sup>1</sup>

(123) Clitic and non-clitic allomorphs according to Fried (1994, 165):

- a. *jsem*<sub>1</sub>
- b. *jsem*<sub>2</sub> [[ ]<sub>ω</sub> \_ ]<sub>ω</sub>

In Junghanns’s (2002b) view, clitics belong to parts of speech just like non-clitics, but are distinguished from them through a feature [+clitic].<sup>2</sup> This feature is responsible for their syntactic behaviour. The feature specification for the non-clitic passive auxiliary is thus as in 124a, the one for the clitic past auxiliary as in 124b. In addition, clitics have a prosodic subcategorisation frame which causes them to form a prosodic word with their host, shown in 124c (cf. Junghanns, 2002b, 136). The crucial difference to Fried’s (1994) approach is that Junghanns (2002b) does not assume clitic and non-clitic elements to stem from one underlying lexical entry.

(124) Clitic and non-clitic features according to Junghanns (2002b, 136):

- a. Passive auxiliary: [+V, -N, +aux, -cl]
- b. Past auxiliary: [+V, -N, +aux, +cl]
- c. Prosodic subcategorisation of clitics: [[ ]<sub>ω</sub> \_ ]<sub>ω</sub>

Whilst both Fried (1994) and Junghanns (2002b) view clitic auxiliaries as belonging to the same category as non-clitic ones, Franks and King (2000) assume that “clitic auxiliaries are pure

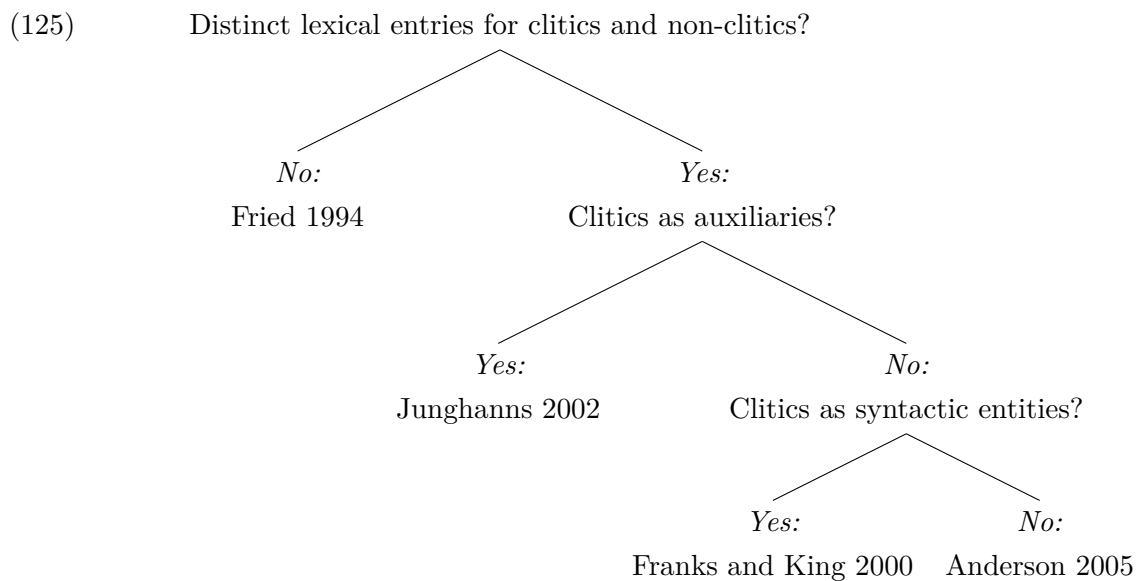
<sup>1</sup>This requirement is, as I have argued in section 5.1.3, problematic, since Czech 2P clitics do not need a prosodic host. Yet, I will leave this issue aside here, since it is not central to the current discussion.

<sup>2</sup>To be more precise, this feature characterises all *lexical* clitics, in his terminology, thus including all clitics that are tied to the second position. What I subsume here, for ease of presentation, under “non-clitics” are actually phonological clitics; these do not have the feature [+clitic] and are not tied to the second position (cf. Junghanns, 2002b, 135, and section 4.1 of this thesis).

realizations of person-number agreement features, whereas the nonclitic copulas consist of the verb ‘be’ plus these features” (Franks & King, 2000, 95).<sup>3</sup> The reduced nature of clitics is also central for Franks (2016): clitics are deficient semantically, prosodically and syntactically. Czech copulas have “a T element in addition to the Agr element of the auxiliary”, which means that clitics are deficient in their lack of tense (cf. Franks, 2016, 121-2).<sup>4</sup> Viewing auxiliary clitics as pure agreement morphology entails that they do not belong to the same category as non-clitic auxiliaries.

This view is not far from that of Anderson (2005). As described in section 6.4, the observation that clitics have strictly fixed positions in languages that otherwise have free word order leads him to conclude that “mechanisms other than those of the language’s basic syntax must be at work” (Anderson, 2005, 139). Clitics are instead placed by morphology: “The clitic is not itself a syntactic constituent of its phrasal domain, but rather a phonological marker of some feature(s) associated with the domain” (Anderson, 2005, 141), i. e. a phrasal affix.

To summarise, I have presented four different views of the lexical properties of clitic and non-clitic auxiliaries. Fried (1994) sees them as one lexical item with only prosodic allomorphy, Junghanns (2002b) assumes distinct entries for clitics and non-clitics, with both belonging to the class of auxiliaries, Franks and King (2000) and Franks (2016) see clitics as non-verbal feature realisations, and Anderson (2005) views them as phrasal affixes. The flowchart in 125 visualises the relations between these approaches and the questions we must answer to decide between them.



In order to answer these questions, the following subsections systematically examine the differences between the morphophonologically similar clitic and non-clitic forms presented in table 7.1. As already indicated in section 4.2, the clitic past tense auxiliary differs morphologically from the non-clitic copula and passive auxiliary in many respects, which can be grouped

<sup>3</sup>Confusingly, they also agree with Fried (1994) in her view of clitics and non-clitics as sharing one lexical entry, because the line between clitics and non-clitics does not coincide with that between auxiliaries and copulas (cf. Franks & King, 2000, 94). I do not know how this can be reconciled with the view expressed above.

<sup>4</sup>He also assumes, with Chomsky (1994), that pronominal clitics are ambiguous between heads and phrases, but this does not extend to auxiliary clitics (cf. Franks, 2016, 112).

into three categories: obligatory or optional omission and reduction, colloquial allomorphy, and formation of the frequentative. These will now be addressed in turn.

### 7.1.2 Omission and reduction

Turning to omission and reduction first, most striking is the obligatory gap in the third person singular and plural. Note that, since the past participle agrees in gender and number with the subject, the lack of an overt auxiliary in the past tense does not lead to ambiguity, as 126 shows. Additionally, but only in colloquial language, the first person singular and plural auxiliaries may also optionally be null, provided the subject is overtly expressed, as in example 127.

(126) a. *Spala.*  
 sleep.PTCP.F.SG  
 ‘She slept.’

b. *Spaly.*  
 sleep.PTCP.F.PL  
 ‘They slept.’

(127) a. *Já už spal.* (for: *Já jsem už spal.*)  
 I already sleep.PTCP  
 ‘I was already asleep.’

(cf. Toman, 1980, 307)

b. *My už spali.* (for: *My jsme už spali.*)  
 we already sleep.PTCP  
 ‘We were already asleep.’

The second person auxiliary, in contrast, cannot be omitted, but in the singular it can be contracted to *-s* in colloquial language.<sup>5</sup> It then attaches to the preceding element, as examples 128a and 128b show. Example 128c shows that with non-2P clitics such as the copula, this is not possible.

(128) a. *Tys přišel.* (for: *Ty jsi přišel.*)  
 you:2SG come.PTCP  
 ‘You came.’

<sup>5</sup>There appears to be a preference for *-s* to attach to an initial verb or pronoun, as in the examples in 128. According to a small acceptability judgement survey I conducted with eight speakers, NPs and adverbials are less accepted as hosts of the contracted auxiliary. For some speakers, it is also possible to place the clitic in third position, when the host, i. e. the second element in the clause, is the verb. For details, see the OSF link provided in appendix B.



- b. *Šels rychle?* (for: *Šel jsi rychle?*)  
 go.PTCP.2SG fast  
 ‘Did you walk fast?’
- c. *\*Tys učitel.* (for: *Ty jsi učitel.*)  
 you.2SG teacher  
 ‘You are a teacher.’

(cf. Toman, 1980, 306)

The different availabilities of omission and reduction are summarised in table 7.2. The gap in the third person fits the view that first and second forms have a morphological specification for person, whilst third forms do not (cf. Corbett, 1983), i.e. that the third person is actually a default that arises from a lack of person features. This is already noted by Benveniste (1946), who defines the third person as “the verb form which has as its function to express the non-person” (Benveniste, 1946, 4)<sup>6</sup>, which distinguishes it from first and second person. As evidence he cites the lack of a third person pronoun and the zero marking of third person on the verb; for the latter, he gives a multitude of examples from both Indo-European and non-Indo-European languages (cf. Benveniste, 1946, 5-6).

	clitic	non-clitic
1SG	<b>jsem</b> /∅	jsem
2SG	<b>jsi</b> /-s	jsi
3SG	∅	je
1PL	<b>jsme</b> /∅	jsme
2PL	<b>jste</b>	jste
3PL	∅	jsou

Table 7.2: Clitic and non-clitic present-tense forms of *být* ‘to be’ (2P clitics in bold)

However, as remarked by Toman (1980), the overall pattern, namely that the clitic past auxiliary can be omitted in the first person, must be omitted in the third person, but may not be omitted in the second person, appears odd (cf. Toman, 1980, 310). It is not compatible with Benveniste’s (1946) understanding of person, elaborated by Lazzeroni (1994), in which the first person opposes the second and the third in its *subjectivity* (“soggettività”), and the first and the second person oppose the third in their *personality* (“personalità”; Lazzeroni, 1994, 168).

On the other hand, the current stage of Czech matches the historic development of the loss of the past auxiliary in Russian, where the third person singular and plural could be omitted first, already around the 12<sup>th</sup> century. In later texts from the 13<sup>th</sup> to 17<sup>th</sup> centuries, 68 % of the third person were dropped, and 29 % of the first person, in contrast to only 3 % of the second person (cf. Kiparsky, 1967, 227-8). If one sees all persons as combinations of the features [ $\pm$  speaker]

<sup>6</sup>The translation from the original French is my own.

and [ $\pm$ addressee], then the zero forms are all intrinsically [ $-$ addressee], whilst [ $+$ addressee] forms resist omission more persistently in Russian and Czech.<sup>7</sup>

### 7.1.3 Colloquial allomorphy

For the first and second person singular, the colloquial language provides the variants *su* and *seš* for the (only prosodically clitic) auxiliaries/copulas *jsem* and *jsi*. This was indicated in table 7.1; examples are given in 129a-b for *su* and in 130a-b for *seš*.<sup>8</sup> These forms are not available for the lexically clitic past auxiliary, as 129c and 130c show.

(129) a. *Já su doma.* (for: *Já jsem doma.*)

I COP.1SG at.home

‘I am at home.’

b. *Já su pozván.* (for: *Já jsem pozván.*)

I AUX.1SG invite.PASSPTCP

‘I am invited.’

c. \**Já su spadl.* (for: *Já jsem spadl.*)

I AUX.1SG fall.PASTPTCP

‘I fell.’

(130) a. *Ty seš hlupák.* (for: *Ty jsi hlupák.*)

you COP.2SG fool

‘You are a fool.’

(cf. Toman, 1980, 306-7)

b. *Ty seš pozván.* (for: *Ty jsi pozván.*)

you AUX.2SG invite.PASSPTCP

‘You are invited.’

c. \**Ty seš spadl.* (for: *Ty jsi spadl.*)

you AUX.2SG fall.PASTPTCP

‘You fell.’

<sup>7</sup>With respect to Classical Greek, where only the second person forms of the verbs *say* and *be* are non-clitic, Kiparsky (1967) remarks “that precisely the second person singular requires an especially clear emphasis of the addressed person” (Kiparsky, 1967, 228; the translation from the original German is my own).

<sup>8</sup>The form of these allomorphs is derived from the first and second person singular verbal suffixes *-u* and *-eš*, and can thus be interpreted as a regularisation of the non-clitic present tense paradigm of *být*. The first person singular *su* is mainly found in Moravian dialects (cf. Franks & King, 2000, 95).

### 7.1.4 Frequentative formation

Finally, in order to express that an event is repeated or habitual, Czech employs a frequentative suffix *-va* (cf. Naughton, 2005, 165-6). Veselovská (2008) remarks that, while the passive auxiliary can support the frequentative suffix, the past auxiliary cannot. In her analysis, this is due to a difference in base position: whilst the passive auxiliary and the copula are generated in  $v^*$ ,<sup>9</sup> the clitic auxiliary is generated higher, in T.<sup>10</sup> Since the frequentative feature is located in  $v^*$ , the non-clitic can express iterativity, but the clitic cannot (cf. Veselovská, 2008, 559-60).

- (131) a. *Já jsem chválíval.*  
 I PASTAUX.1SG praise.FREQ.PASTPTCP  
 b. *\*Já bývám chválil.*  
 I be.FREQ.1SG praise.PASTPTCP  
 ‘I was repeatedly praising.’  
 c. *Já jsem chválíván.*  
 I PASSAUX.1SG praise.FREQ.PASSPTCP  
 d. *Já bývám chválen.*  
 I be.FREQ.1SG praise.PASSPTCP  
 ‘I am repeatedly being praised.’

(cf. Veselovská, 2008, 559)

Yet, as it turns out, the generalisation that the *-va* suffix only attaches to non-clitic verbs is insufficient: only the copula and the passive auxiliary can form a frequentative, as 132a-b demonstrate. In the other cases, including the non-clitic future auxiliary, the suffix attaches to the main verb, as shown in 132c-e. The generalisation thus appears to be that, within the *být* paradigm, frequentatives can only be formed from non-clitic present tense forms.<sup>11</sup>

- (132) a. *Já bývám doma.*  
 I be.FREQ.1SG at.home  
 ‘I am repeatedly at home.’  
 b. *Já bývám fotografován.*  
 I be.FREQ.1SG photograph.PASSPTCP  
 ‘I am repeatedly being photographed.’

<sup>9</sup>In the Minimalist Program, V and T are lexical, whilst C and  $v^*$  are functional heads (cf. Chomsky, 1995).

<sup>10</sup>She does not assume the T node to be the location of the tense feature, though (cf. Veselovská, 2008, 559). Earlier, she proposed that clitic auxiliaries must be generated *above* T because they lack tense (cf. Veselovská, 1995).

<sup>11</sup>Note that also in 132b, the frequentative suffix *can* attach to the participle; the difference is that in 132c-d, it *has to* (and in 132e it has to attach to the infinitive), because formation of a frequentative from the respective auxiliaries is impossible. Generally, all main verbs can form a frequentative, whether they are finite verbs, infinitives or participles.

- c. *Já jsem spávala.*  
 I aux.1sg sleep.FREQ.PASTPTCP  
 ‘I used to sleep.’
- d. *Já bych spávala.*  
 I COND.1SG sleep.FREQ.PASTPTCP  
 ‘I would sleep repeatedly.’
- e. *Já budu spávat.*  
 I FUT.1SG sleep.FREQ.INF  
 ‘I will sleep repeatedly.’

### 7.1.5 Summary

To summarise, this section has shown that clitic auxiliaries differ from non-clitic ones in a number of morphological aspects: they can be reduced or elided, they do not possess colloquial allomorphs that are available for non-clitic auxiliaries, and they do not support the frequentative suffix (but neither do non-clitic auxiliaries that are not derived from the present tense). What can we conclude from this regarding the above-presented theory types?

It turns out that Fried’s (1994) view that clitics and non-clitics are one lexical item and only differ prosodically cannot capture the morphological facts. Why should the clitic realisation of the auxiliary display such differences on the surface, if its only difference from the non-clitic was its phonological dependency? The great amount of morphological idiosyncrasies demonstrated above is a strong indicator that clitic and non-clitic auxiliaries are distinct lexical entities.

Regarding the other three types of approach, none of them appears to have trouble explaining the clitic auxiliaries’ special morphology. Junghanns’s (2002b) approach of viewing them as a separate class of auxiliary can explain why they have a different, and partly reduced, morphology. That the clitic past auxiliary cannot support the frequentative suffix, however, does not directly result from this account. The approaches by Franks and King (2000) and Anderson (2005) can explain reduction, elision and colloquial variants just as well. In addition, they have an answer to why the clitic cannot form a frequentative: if clitics are pure feature realisations, or affixes, they do not possess a verbal stem that the frequentative suffix can attach to. This does however not explain why the future auxiliary cannot form a frequentative either, and I must leave this question open for now.

## 7.2 Syntax of clitic and non-clitic verbal elements

We have seen several types of morphological data which suggest that clitic and non-clitic auxiliaries are lexically distinct items. Yet, all data shown so far is in accordance with either of the other three views (except from the uncertain case of frequentative suffixation). We now need to clarify whether clitic auxiliaries are verbs, or whether they constitute a different category. For this purpose, we will take a closer look at the syntactic behaviour of different types of clitic and

non-clitics verbs. First, section 7.2.1 will present the fronting patterns of different verb types. Section 7.2.2 then examines the crossing properties of clitic and non-clitic verbs. Finally, in section 7.2.3, we will see how the enclitic *-li* combines with different types of verbal elements.

### 7.2.1 Fronting to the left periphery

In an acceptability judgement study, I asked participants to rate sentences with verbal material occurring in pre-clitic position on a seven-point scale. The data will be discussed and analysed in detail in section 9.2, which also includes examples for all conditions. The sentences in 133 exemplify past participle fronting in a main clause for the verb alone (133a), and for the verb plus its object (133b).

- (133) a. *Posílali mu dopisy každý týden.*  
 send.PTCP him.DAT letters every week
- b. *Posílali dopisy mu každý týden.*  
 send.PTCP letters him.DAT every week
- ‘They were sending him letters every week.’

Figure 7.1 visualises the results, and strikingly shows that finite verbs and past participles pattern together, whilst infinitives and passive participles each display their own behaviour. The analysis I am going to propose for this in section 9.2 is that the finite verb leaves VP, which prevents joint topicalisation of verb and object. At this point, I will not go into the details of this. What is of concern for the present discussion is that past participles behave just like finite verbs with respect to fronting to the first position, both with and without direct object. This might suggest that the past participle *is* a finite verb – but then, the clitic past tense auxiliary cannot be a finite verb as well. This is in line with both Franks and King’s (2000) and Anderson’s (2005) views on the nature of clitics: if clitics are mere feature realisations or affixes, they are not verbs.

### 7.2.2 Verbal crossing

There appear to be restrictions on the type of verb which can host the clitic: if the VP consists of multiple verbs, only the highest one can function as a host. Examples 134a and 134b show that the clitic can be crossed by other verbs, whilst example 134c shows that non-clitic verbs cannot cross each other. The fact that the verbs cannot cross each other freely, but both can cross the clitic, indicates that there is a categorical difference between clitics and other verbs. Clitic verbs are thus not just verbs which occur in fixed positions, but differ from “regular” verbs in a more fundamental way (cf. Rivero, 1991, 327, 341).

- (134) a. *Koupil bych koupil knihy.*  
 buy.PTCP COND.1SG books
- ‘I would buy books.’

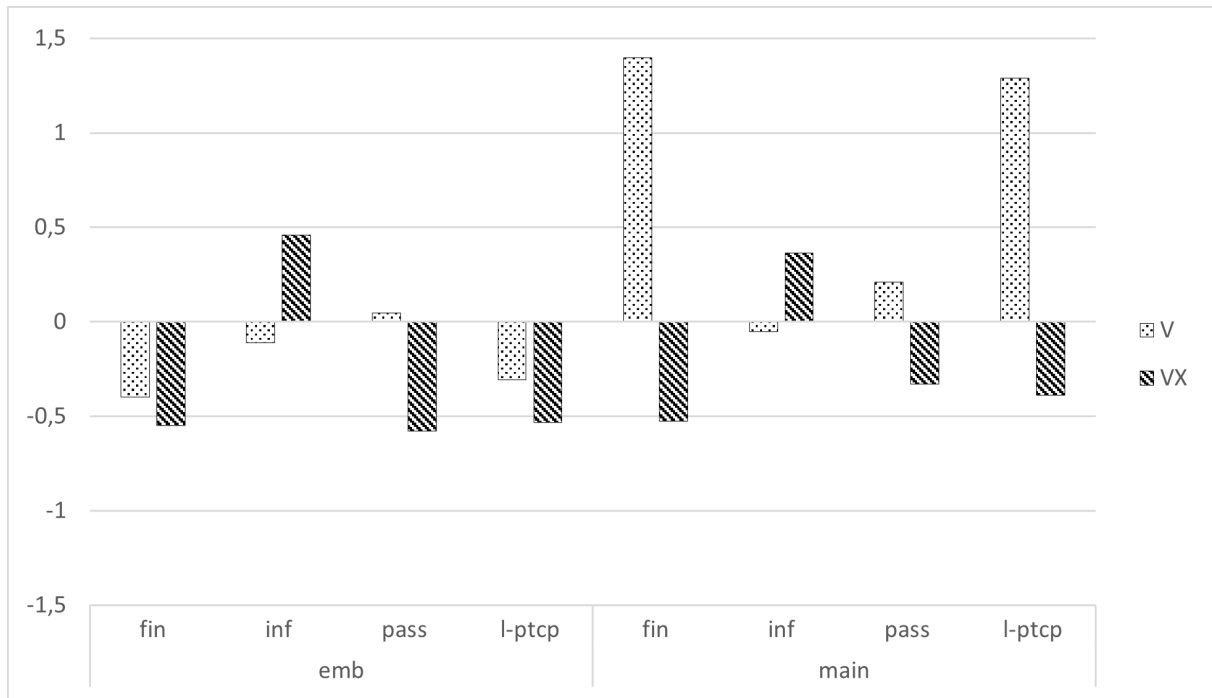


Figure 7.1: Acceptability judgements (z-scores) for fronting of different verb types with and without complements in embedded and main clauses (N=61)

- b. *Byl bych byl koupil knihy.*  
 be.PTCP COND.1SG buy.PTCP books
- c. \**Koupil bych byl koupil knihy.*  
 buy.PTCP COND.1SG be.PTCP books
- ‘I would have bought books.’

(cf. Rivero, 1991, 341, 327)

Rivero (1991) explains this based on the assumption that clitics are A-heads, whereas all other auxiliaries are A'-heads. The fact that *koupil* cannot be fronted in 134c results from it being the lower verb – 134c violates Relativised Minimality. In contrast, 134a and 134b do not violate Relativised Minimality, because clitic and non-clitic verbs are not of the same type (cf. Rivero, 1991, 327, 341).

Richardson (1997) objects that the only plausible factor that could lead to these elements' different affiliations is the clitics' prosodic deficiency. This, however, would mean that syntax accesses phonology. Richardson (1997) instead proposes, in line with Anderson (1992), that clitics are placed by morphology, which is the actual reason that the clitic past auxiliary does not behave like other finite verbs (cf. Richardson, 1997, 137).

Whilst Richardson's (1997) argument against the assumption that clitics and non-clitics constitute different kinds of heads does not hold up – after all, clitics and non-clitics do not only differ prosodically – a more crucial point here is the validity of the data: it is not as

uniform as assumed by either of these authors. First, crossing of non-clitic verbs *is* possible.<sup>12</sup> Under certain information-structural conditions, the order of non-clitic verbs can be reversed, for example when a focused modal is right-dislocated, as 135 shows (small capitals indicate main stress). Restrictions on verbal movement are thus not simply based on a clitic vs. non-clitic distinction.<sup>13</sup>

- (135) *Petr ty knihy číst CHTĚL, ale tys mu to rozmluvila.*  
 Petr those books read.INF want.PTCP but you:2SG him it dissuade.PTCP  
 ‘Petr WANTED to read those books, but you talked him out of it.’

Second, it should be noted that the ungrammaticality of example 134c might not be due to the verbs’ base-position, verb type etc. at all, but simply result from a garden-path effect: since both the past auxiliary (*byl*) and the main verb (*koupil*) are *l*-participles, the segment *Koupil bych* is interpreted as a complete verbal complex, in the present conditional. This is amplified by the fact that the past conditional is extremely rare in Czech. Instead, the present conditional is used to express all tenses, which are interpreted according to context. The native speakers I consulted confirmed that this is also their intuitive explanation for the sentence’s oddness.<sup>14</sup>

Also note that generally, it appears that verbs can only cross each other for information-structural reasons, whilst crossing of clitics happens independently. A suitable context for the sentence in 136b would be one where the final constituent, the direct object, is contrastively focused, as in 137, and where the infinitival main verb is already given in discourse.

- (136) a. *Chtěla koupit knihy.* → *neutral*  
 buy.INF want.PTCP books  
 b. *Koupit chtěla knihy.* → *reordered*  
 buy.INF want.PTCP books  
 ‘She wanted to buy books.’  
 c. \**Jsem koupila knihy.*  
 AUX.1SG buy.PTCP books  
 d. *Koupila jsem knihy.* → *neutral*  
 buy.PTCP AUX.1SG books  
 ‘I have bought books.’

<sup>12</sup>I would like to thank Uwe Junghanns for pointing this out to me.

<sup>13</sup>The sentences in 135 and 137 are the results of an online survey, in which I asked native speakers to provide plausible contexts for the core constructions (*Petr ty knihy číst chtěl* and *Koupit chtěla knihy*). The versions presented here were additionally confirmed by one of my consultants. For an alternative construction where the main verb precedes the modal due to leftward movement, *číst chtěl ty knihy*, no plausible contexts could be found.

<sup>14</sup>The rarity of the conditional past conditional is also reflected in the Czech National Corpus: here, sentences including a form of the conditional auxiliary plus the past participle of *být* only make up about 3.9% of all sentences with the conditional; the real ratio is probably even much lower, since combinations of conditional auxiliary and *být* participle can also just be the present conditional of ‘to be’.

- (137) *Koupit chtěla knihy, ale místo toho propila peníze v hospodě.*  
 buy.INF want.PTCP books but instead that.GEN drink.away.PTCP money in pub

‘She wanted to buy BOOKS, but instead she spent the money on drinks in the pub.’

The question is also whether one would want to assume verb movement in cases where verbs precede clitics anyway, as discussed for example in section 6.1.3. Whilst for a combination of modal verb + infinitive, it makes sense to assume that the modal is underlyingly higher (as is also evidenced by the fact that this order is by far more preferred), 2P needs not be seen as a fixed syntactic position at all. In the approach that I will assume for the modelling of clitic placement in the present thesis, the verb does not have to rise to the left periphery to precede the clitic cluster. This will be elaborated in section 8.2.

In consequence, crossing data such as those in 134 does not provide much evidence for the different nature of clitic verbs with respect to non-clitic ones, since the data is non-conclusive in two respects: First, it is not the case that non-clitic verbs can never cross each other; instead, most permutations are possible given the right information-structural conditions.<sup>15</sup> Second, the unacceptability of sentence 134c may actually result from a garden-path effect. Thus, crossing data gives us no additional information on whether the clitic is a verb or not.

### 7.2.3 Interaction of participle and *-li*

A possible argument against the claim that participles behave like finite verbs in Czech might be derived from the positioning of the verbal clitic *-li*, introduced in section 4.1.3. According to Toman (1996), it only attaches to finite verbs (cf. Toman, 1996, 508). If the participle is a finite verb as well, we would expect it to be a suitable host for *-li* just like (other) finite verbs. An initial corpus search for the combination participle+*li* yielded a total of 11.88 i. p. m., which amounts to only 4% of all occurrences of *-li* in the SYN2015 corpus. Two examples are given in 138.

- (138) a. *Byl-li však Bůh svědkem toho, co se stalo, pak  
 be.PTCP-if however God witness this.GEN what REFL happen.PTCP then  
 zná stejně dobře jako já pravdu.  
 know.3SG equally well as I truth.ACC*

‘If however God was witness to what happened, then He knows the truth as well as I do.’

<sup>15</sup>For differences between infinitives and participles with respect to movement, see section 9.2.3.



- b. *Chtěl-li jsem se něco dozvědět o Flaviovi,*  
 want.PTCP-if AUX.1SG REFL something find.out.INF about Flavius.LOC  
*nemohl jsem začít nikde jinde než na závodišti*  
 NEG:can.PTCP AUX.1SG begin.INF nowhere elsewhere than on racetrack.LOC  
*Circus maximus.*  
 Circus Maximus

‘If I wanted to find out anything about Flavius, I couldn’t start anywhere else but on the racetrack of Circus Maximus.’

(SYN2015)

To see whether these attested, yet at the same time quite infrequent constructions are accepted by native speakers, and whether *-li* displays any difference in behaviour towards different kinds of verbs, I conducted a small acceptability judgement survey with eight participants. They were asked to rate sentences beginning with a combination of verb+*li*. Four verb types were controlled for; in addition, the two items differed in aspect, to see whether it might have an influence. Examples for each verb type in the perfective aspect are provided in 139. The results, illustrated in figure 7.2, show a clear difference between finite verbs and past participles on the one hand, and infinitives and passive participles on the other hand, similarly to what was reported for verbal fronting in section 7.2.1. Aspect does not appear to have a consistent effect; the materials and results can be found in the OSF repository (cf. appendix B).

- (139) a. *Spravit-li umíš ty hodinky, budu dost překvapený.*  
 repair.INF-if can.2SG those watch will.1SG quite surprised  
 Intended: ‘If you can repair that watch, I will be quite surprised.’
- b. *Spravíš-li rádio, poslechneme si ten koncert.*  
 repair.2SG-if radio listen REF that concert  
 ‘If you repair the radio, we will listen to that concert.’
- c. *Spravila-li jsi kolo, uděláme si výlet.*  
 repair.PASTPTCP-if AUX.1SG bike make.1PL REFL trip  
 ‘If you have repaired the bike, we’ll go on a trip.’
- d. *Spraven-li je výtah, nemusíme jít pěšky.*  
 repair.PASSPTC-if is lift NEG:must.1SG go.INF on.foot  
 Intended: ‘If the lift is repaired, we don’t have to walk.’

Thus, the verb types which resist VP topicalisation appear to be the ones which can be combined with *-li*, possibly because they are more prone to rise to the left periphery (cf. section 9.2). In any case, compatibility of *-li* is another type of data where the behaviour of past participles and finite verbs is strikingly similar. And again, if the *l*-participle acts as the finite verb of its clause, then the clitic past auxiliary can be argued not to be a finite verb. Although

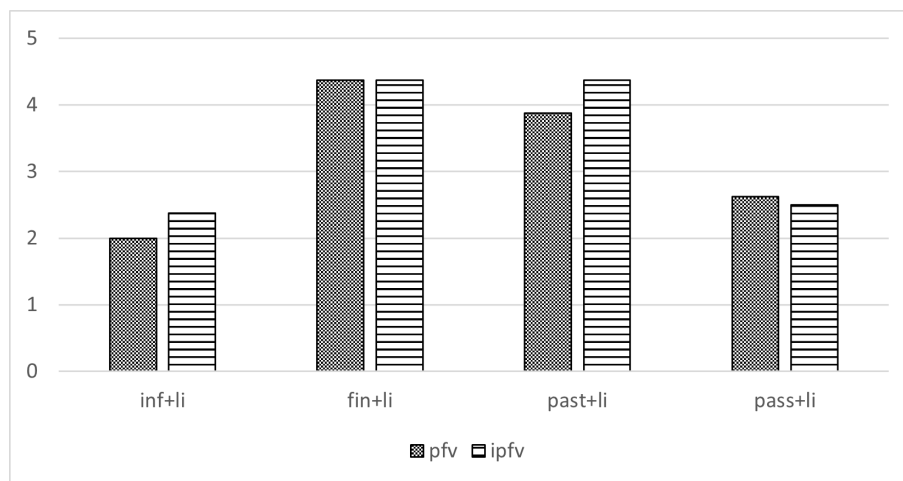


Figure 7.2: Acceptability judgements for combining different verb types with *-li* (N=8)

the data must be treated with care due to the small sample size and low number of test items, some tendencies are quite strong, and are in line with the previous results for VP topicalisation.

#### 7.2.4 Summary

To sum up, we have seen that only if we assume clitic and non-clitic forms of *být* to be fundamentally different, we can explain the peculiar morphological behaviour of Czech 2P clitics. Regarding the question of the clitics' precise nature, evidence from verbal fronting and from the behaviour of *-li* suggests that the auxiliary clitic might not be a verb, but instead some spell-out of features/an affix that combines with an inflected participle, which functions as the finite verb of its sentence.

### 7.3 Past vs. passive: the participle

In this section, I take a closer look at the differences between the past and the passive. As noted by Franks and King (2000), it is at first glance surprising that the line between clitics and non-clitics does not match the line between auxiliaries and copulas, but instead is drawn between the past and the passive auxiliary (cf. Franks & King, 2000, 94). We have seen in section 7.1 above that these two types of auxiliaries not only differ with respect to elision and allomorphy, but also in their ability to support the frequentative suffix. This might suggest that the clitic past auxiliary does not have a verbal stem that the suffix could attach to.<sup>16</sup> Support for this comes from the syntactic behaviour of the past participle, which suggests that it functions as the finite verb of its clause, whereas the past auxiliary does not. But why is it that past and passive are so fundamentally different in Czech?

Morphologically, the past participle (also *l*-participle) and the passive participle are very similar: they carry the same suffixes that appear on short-form adjectives (cf. Naughton, 2005,

<sup>16</sup>As noted above, the question remains why the non-clitic future auxiliary cannot form a frequentative either. The frequentative suffix apparently only attaches to present-tense auxiliaries; but it is not clear why that is the case.

55, 141), whereas the much more frequent long-form adjectives have a different set of suffixes (cf. Naughton, 2005, 52). The two sets of suffixes are shown in table 7.3. What all these forms have in common is that they inflect for number and gender, but not person, in contrast to Czech finite verbs.

short-form adjectives				
	M.ANIM	M.INAN	F	N
singular	-∅	-∅	-a	-o
plural	-i	-y	-y	-a
long-form adjectives				
	M.ANIM	M.INAN	F	N
singular	-ý	-ý	-á	-é
plural	-í	-é	-é	-á

Table 7.3: Inflectional suffixes of Czech participles and adjectives

However, this is not a contradiction to the proposal that the past participle displays a behaviour similar to that of finite verbs. Czech may in fact be developing into a similar direction as Russian, i. e. losing its past auxiliary and turning the *l*-participle into a finite verb (cf. Toman, 1999, 225; Kiparsky, 1967, 249). This is happening in the highly frequent past tense, but not in the passive, which is extremely rare in Czech (cf. Naughton, 2005, 163). Also, the passive participle still functions as a participle in Modern Czech: it can be used with long-form suffixes, for example as an attributive, as in 140a, or as a predicative, as in 140b. In the latter case, the line to a passive construction (shown in 140c) cannot always be drawn sharply. None of this applies to the past participle, which is only found in combination with clitics in past tense and conditional mood.

- (140) a. *otevřené dveře*  
 open.F.PL door.PL  
 ‘an open door’
- b. *Dveře byly otevřeny / otevřené.*  
 door.PL be.PASTPTCP.F.PL open.PASSPTCP.F.PL open.F.PL  
 ‘The door was open(ed).’
- c. *Dveře byly hned otevřeny a šli jsme dovnitř.*  
 door.PL be.PASTPTCP.F.PL promptly open.PASSPTCP.F.PL and go.PASTPTCP.M.PL  
 PASTAUX.1PL inside  
 ‘The door was promptly opened and we went inside.’

(cf. Naughton, 2005, 163)

Another interesting point is that the long forms of the passive participle fully inflect for case (cf. Naughton, 2005, 160), and even the short forms have retained an accusative (cf. Naughton,

2005, 55), whereas past participles reflect no case at all. Thus, as noted by Götz Keydana (p. c.), for both learners and speakers, the participial nature of the passive form is transparent, whereas that of the past form is not. Naughton (2005) does not even use the term “past participle”, but simply calls it “-l form” (cf. Naughton, 2005, 141).<sup>17</sup> Whilst historically, this form might have been a real participle, it is not any more. The only reason it is today still classed as such is its lack of person features – but this alone does not make a participle.<sup>18</sup>

## 7.4 The relation of past and conditional clitic

We have seen that the clitic auxiliaries differ from the non-clitic ones in fundamental ways. However, I have not yet addressed the issue of the relation between the two auxiliary clitics, the past and the conditional. The conditional auxiliary cannot undergo contraction or deletion as does the past auxiliary.<sup>19</sup> It is also morphophonologically more complex, which is not surprising, given that the conditional mood is more marked than the indicative. The following subsections investigate the differences between these two kinds of auxiliary clitics in detail.

### 7.4.1 Morphological shape of the conditional clitic

As observed by Toman (1980), the relatively new colloquial forms of the conditional auxiliary (which were given in brackets in table 7.1) can be viewed as combinations of a conditional stem with the past tense auxiliary, which would thus be a kind of suffix (cf. Toman, 1980, 310). As shown in table 7.4, the resulting (colloquial) paradigm of the conditional auxiliary is then completely regular: the conditional stem *by* is combined with the past auxiliary – note that its initial *j-* is typically silent. Toman (1980) remarks that this development “seems to indicate that the auxiliary *be* tends to obliterate its internal morphological structure, whereas the copula remains poly-morphemic” (Toman, 1980, 310), i. e. that the past auxiliary is drifting away from the copula by becoming mono-morphemic. We thus have additional support for our view that the past auxiliary lacks a verbal stem – on the other hand, we find that the conditional auxiliary, despite also being a 2P clitic, is still morphologically complex.

	PAST.AUX	COND.AUX
1SG	<b>jsem</b>	<b>bysem</b> ( <i>coll.</i> )
2SG	<b>jsi</b>	<b>bysi</b> ( <i>coll.</i> )
3SG	–	<b>by</b>
1PL	<b>jsme</b>	<b>bysme</b> ( <i>coll.</i> )
2PL	<b>jste</b>	<b>byste</b>
3PL	–	<b>by</b>

Table 7.4: Parallel past and conditional auxiliaries

<sup>17</sup>He also remarks that “some [passive] participles have become fully fledged adjectives, even to the extent that they form comparatives and superlatives” (cf. Naughton, 2005, 161).

<sup>18</sup>Also note that with clitic auxiliary contraction, discussed in section 7.1.2, person features may even be expressed on the past participle, e. g. *šels* ‘go.PTCP.2SG’.

<sup>19</sup>I should mention that if the conditional auxiliary were also deleted, the past tense would be indistinguishable from the conditional mood, since both auxiliaries combine with the past participle. Yet, since homophony of tenses, moods, and many other verbal features is widespread in possibly all languages, it cannot serve as an explanation for the clitics’ divergent behaviour.

Additional support for this view comes from reflexives. When they combine with a second person singular past auxiliary, the auxiliary itself vanishes and second person is instead expressed on the reflexive, as 141 shows. Interestingly, the same happens with a second person conditional auxiliary, but in this case, the conditional stem *by* persists, as can be seen in 142.<sup>20</sup>

- (141) a. *??Učila jsi se.*  
 learn.PASTPTCP AUX.2SG REFL  
 b. *Učila ses.*  
 learn.PASTPTCP REFL:2SG  
 ‘You learned.’

- (142) a. *??Učila bys se.*  
 learn.PASTPTCP COND.2SG REFL  
 b. *Učila by ses.*  
 learn.PASTPTCP COND REFL:2SG  
 ‘You would learn.’

Given that both the past and the conditional auxiliary combine with the past participle, we are facing a potential problem: if the conditional clitic has a verbal stem, i. e. is a finite verb, then what role does the past participle play here? At first glance, it thus appears that the conditional clitic contradicts our analysis of the past participle as a finite verb that combines with a suffixal auxiliary. An alternative is to interpret the conditional auxiliary not as a combination of verbal stem and clitic, but as a lexical combination of two clitics, one expressing the conditional, the other person and number, as shown in table 7.5. This is also the view Franks and King (2000) take in order to explain the existence of the clitic bisyllabic conditional plural forms *bychom/bysme* and *byste*.<sup>21</sup> Then, in both conditional and past tense sentences, the participle must act as the finite verb, because the auxiliaries have no verbal stem. Interestingly, this analysis provides a parallel between Czech and Polish, where conditional mood is expressed through the invariable clitic particle *by* (cf. Sussex & Cubberley, 2006, 369).

		1SG = <b>(j)sem</b>
		2SG = <b>(j)si</b>
		3SG = –
COND = <b>by</b>	+	1PL = <b>(j)sme</b>
		2PL = <b>(j)ste</b>
		3PL = –

Table 7.5: Composition of the conditional auxiliary

<sup>20</sup>According to my informants, this expression of person and number on the reflexive is obligatory, but they also note that not all speakers adhere to this. The question marks on the a-examples thus might indicate colloquial language, rather than generally reduced acceptability.

<sup>21</sup>The original motivation for this analysis might be called into question, since one may reject the requirement that clitics be monosyllabic (cf. Avgustinova & Oliva, 1995, 11). Anyway, Franks and King’s (2000) proposal receives support independently of this question here.

An additional piece of evidence that the conditional auxiliary clitic is not an independent verb comes from inflected complementisers: first, the resultative complementiser *aby* ‘so that’, shown in 143a for the singular, and in 143b for the plural; and second, the irrealis complementiser *kdyby* ‘if’, shown in 143c for the singular, and in 143d for the plural. Their forms can be viewed as combinations of *a/kdy* with the (non-colloquial) conditional auxiliary. Like the auxiliary on its own, these complex complementisers combine with the verb in the past (*l*-)participle (cf. Veselovská, 1995, 113).

- (143) a. *Chtěla, abych/ abys/ aby přišel.*  
 want.PTCP so.that:COND.1SG so.that:COND.2SG so.that:COND.3 come.PTCP.SG  
 ‘She wanted me/you(sg.)/him to come.’
- b. *Chtěla, abychom/ abyste/ aby přišli.*  
 want.PTCP so.that:COND.1PL so.that:COND.2PL so.that:COND.3 come.PTCP.PL  
 ‘She wanted us/you(pl.)/them to come.’
- c. *Kdybych/ kdybys/ kdyby znal...*  
 if:COND.1SG if:COND.2SG if:COND.3 come.PTCP.SG  
 ‘If I/you(sg.)/he knew...’
- d. *Kdybychom/ kdybyste/ kdyby znali...*  
 if:COND.1PL if:COND.2PL if:COND.3 come.PTCP.PL  
 ‘If we/you(pl.)/they knew...’

(cf. Veselovská, 1995, 113)

#### 7.4.2 The locus of the past tense interpretation

The fact that past and conditional share the same participle also has an interesting consequence for the origin of the past tense interpretation: it cannot be compositional, in the sense that the participle expresses anteriority, and the auxiliary provides the reference point (here: present). This is parallel to the argument made by Zwart (2017). He argues that the Dutch periphrastic past is not formed in syntax, but is part of the morphological paradigm, which also explains its lack of compositionality (cf. Zwart, 2017, 35-6).

For Czech, it appears clear that neither past nor conditional interpretation can arise compositionally from a combination of the participle plus the respective auxiliary: What feature precisely would this be that is shared by past and conditional?<sup>22</sup> Put differently: [+past] cannot be part of the participle because then the conditional would also necessarily have a [+past] interpretation; on the other hand, it cannot be part of the past auxiliary either because the latter can be dropped.

Another argument against compositionality is the lack of the auxiliary in the past tense third person, which has no effect on interpretation. If both periphrastic and synthetic forms

<sup>22</sup>Note that nevertheless, past and conditional can be homophonous in e.g. English, and (partly) German; yet, this seems to be conditioned phonologically.

are simply part of morphology, a mix of these two within one tense paradigm is unproblematic (cf. Zwart, 2017, 33). Also note that a purely compositional account cannot capture the fact that the *l*-participle can express past tense on its own, whilst the conditional auxiliary cannot express conditional mood on its own, as shown in 144.

- (144) a. *Byl* *v Praze*.  
 be.PASTPTCP in Prague  
 ‘He was in Prague.’
- b. *Byl by v Praze*.  
 be.PASTPTCP COND.3 in Prague  
 ‘He would be in Prague.’
- c. \**By v Praze*.  
 COND.3 in Prague
- d. \**V Praze by*.  
 in Prague COND.3

Continuing to follow Zwart’s (2017) argument, under the assumption that morphology is post-syntactic,<sup>23</sup> the positioning of the clitic auxiliary (or, in the case of Dutch, verb-second) must also apply post-syntactically, i. e. cannot be a consequence of syntactic processes (cf. Zwart, 2017, 44).

### 7.4.3 Summary

We have seen that the behaviour of past and conditional auxiliary can best be captured if we view them not as verbs, but as affixes/feature realisations without a verbal stem. Whereas the past auxiliary is a single clitic, the conditional auxiliary consists of two combined clitics. In this respect, the latter is a counterexample to the commonly assumed parallel between affixes and clitics (cf. table 2.1 in section 2.1): unlike affixes, the conditional clitic is not morphologically simple. It should however also be noted that the non-colloquial forms *bych* (1SG), *bys* (2SG) and *bychom* (1PL) cannot be analysed as two clitics as easily. We should best view them as morphologically opaque forms that spell out both conditional and person-number features (though *bys* might be a shortening of *by+(j)si*; then only the first person forms remain idiosyncratic).

We can also conclude that the term “past participle”, used in grammars and also in the present thesis, is twice misleading: It is not a real participle, but increasingly acts as a finite verb. And it does not express past tense, i. e. anteriority, either – otherwise, the conditional mood would necessarily entail a past tense interpretation, which it does not.

<sup>23</sup>The basic idea is that syntax only contains roots and grammatical features, and syntactic terminals are spelled-out through morphological paradigms, which provide inflected words (cf. Zwart, 2017, 32). Such a view cannot be taken in the OT analysis that I am going to map in the following chapters, since all constraints are processed in parallel here. Yet, whilst clitic placement cannot occur *after* syntax, it can take place *independently* from certain syntactic constraints nonetheless.

## 7.5 Conclusions

The overall picture that we have developed in the preceding sections for the different forms of *být* looks like this: Out of the three elements derived from the present tense of *být*, two are auxiliary verbs which only optionally cliticise phonologically: the copula and the passive auxiliary; for details and examples, see section 4.1.2. The past auxiliary, on the other hand, does not have a verbal stem and is tied to the second position. The other 2P clitic derived from *být*, the conditional auxiliary, is really a combination of an invariant conditional clitic and the past tense auxiliary clitic. The past participle that these two types of clitic auxiliaries combine with takes on the role of the finite verb. Finally, the future auxiliary is entirely non-clitic. These findings are in line with both Franks and King's (2000) analysis of clitics as pure feature realisations and Anderson's (1992) analysis of clitics as phrasal affixes.

An interesting aspect is the diachronic perspective on the Czech situation. It appears that with respect to the grammaticalisation of the past tense, Modern Czech occupies a space between Old Church Slavonic and Polish: whilst the former had a fully periphrastic perfect tense (cf. Huntley, 1993, 152), the latter has completely transformed the auxiliary into a (mobile) word affix (cf. Franks & King, 2000, 141). Czech is in an intermediate state, since its auxiliaries are reduced to mere expressions of features and the former participle is losing its participial status.<sup>24</sup> Another interesting point of comparison is Russian, which similarly has lost its past tense auxiliary and turned the past participle into a finite verb; it has dropped inflection of the conditional auxiliary and reanalysed it as a particle (cf. Toman, 1999, 225–6). However, since a diachronic analysis is beyond the scope of this book, I will not pursue this argument further.

Whilst this chapter has only treated clitic and non-clitic *verbs*, the analysis of clitic auxiliaries as feature-bearing elements without a stem also has implications for clitic *pronouns*. As their distribution is entirely parallel to that of the auxiliary clitics, we might extend our conclusions regarding those to the entire class of 2P clitics. The remaining chapters of this dissertation will pursue this line of thought and elaborate on how it can be integrated into an optimality-theoretic analysis.

---

<sup>24</sup>I would like to thank Götz Keydana for calling my attention to this aspect.



## Chapter 8

# An OT Account of Czech 2P Placement

This chapter presents an OT analysis of Czech clitic placement and the phenomena it interacts with. I will first sketch some general assumptions of OT in section 8.1. Then, I will substantiate and refine the approach to clitic placement by Richardson (1997) in several ways in section 8.2. Section 8.3 provides an analysis of clitic climbing out of embedded clauses, as well as multi-clausal structures that block climbing. I will also consider the reasons for the existence of 2P phenomena in section 8.4, and discuss an alternative OT approach to clitic placement by Billings (2002), based on both empirical and conceptual grounds. Section 8.5 briefly presents how our constraint-based analysis can incorporate the respective order of clitics within the clitic cluster. Finally, in section 8.6, I will sketch the predictions made by a factorial typology on the basis of the constraints employed in the analysis. Section 8.7 summarises and concludes this chapter.

### 8.1 Optimality Theory: Underlying assumptions

Before turning to the concrete analysis of Czech 2P placement in an OT framework, a few general ideas of Optimality Theory should be clarified. I will sketch some basic features of this model in section 8.1.1. Generalised Alignment, a constraint class crucial for the analysis of 2P clitics, will be introduced in section 8.1.2. Then, in section 8.1.3, I will approach some syntax-specific issues concerning OT.

#### 8.1.1 Basic ideas of Optimality Theory

The OT apparatus comprises two components, the Generator (Gen) and the Evaluator (Eval), out of which only the latter is responsible for the specific properties of languages. Gen generates all possible candidates for an input. It is usually assumed that this process is unrestricted, i. e. that Gen has *Freedom of Analysis* (cf. McCarthy & Prince, 1993, 9). Yet, for example Grimshaw (1997, 376) assumes that Gen is not entirely unconstrained, but restricted by the (inviolable) principles of X-bar, i. e. only X-bar-compatible structures can be generated. I will, however, not assume such a restriction here, since general structural constraints in Eval can instead prevent the occurrence of ill-formed representations. Thus, the descriptive as well as

explanatory burden is placed entirely on Eval – Gen, on the other hand, is deliberately uninteresting.

The candidates created by Gen are evaluated in Eval with respect to violable constraints that are ranked language-specifically. The principle that governs candidate evaluation is *strict domination*, meaning that “no amount of success on the weaker constraint can compensate for failure on the stronger one” (Prince & Smolensky, 2006, 126). Since constraint evaluation is a parallel, one-step process, OT is – unlike instantiations of the principles and parameters framework like government and binding and the Minimalist Program – non-transformational. An interesting property of OT is that all constraints are universal,<sup>1</sup> and that re-ranking of postulated constraints thus automatically yields a typology of possible systems (cf. Grimshaw, 1997, 404).

Figure 8.1 provides two abstract OT tableaux that illustrate the evaluation process and conventions of presentation: the tableaux are read from left to right, thus the leftmost constraint in each of them is the higher ranked. Asterisks indicate that a constraint is violated, and a following exclamation mark shows that this violation is fatal because another candidate fares better. The hand points to the winning candidate in each tableau. The two tableaux are representative of two language systems, where different candidates emerge due to a different constraint ranking.

	NO-A	NO-B
abc	*!	*
acc	*!	
☞ bcc		*

	NO-B	NO-A
abc	*!	*
☞ acc		*
bcc	*!	

Figure 8.1: Abstract OT tableaux with two different rankings, exemplifying two hypothetical language systems

Another strength of OT is that it allows us to capture concepts such as economy, which play an important role in, for example, the Minimalist Program, in an explicit, precise and computable way. A theory which is otherwise not constraint-based usually has no means of calculating the impact of such violable requirements, especially when several of these conflict. In OT, economy, captured through constraints such as STAY, is one of many constraints whose effects on the outcome are calculated through language-specific rankings (cf. Grimshaw, 1997, 417). This explains why economy, though universally effective in language, is not all-powerful, and we thus nonetheless find traces etc.

For a concise overview of the basic ideas of OT, the reader is referred to Prince and Smolensky (2006). Broekhuis (2008) offers a detailed comparison of Optimality Theory and the Minimalist Program.

### 8.1.2 Generalised Alignment

For the analysis that I will present in the following sections, the notion of *alignment* is crucial. More precisely, I will employ a family of well-formedness constraints termed “Generalised Align-

<sup>1</sup>This does not necessarily entail that the constraints are innate, i. e. encoded genetically. Universality might also result from “universal functional pressures like the architecture of the human vocal tract and perceptual system, or the task demands of communication” (Smolensky & Legendre, 2006, 43).

ment” (GA), which was originally formulated to capture reference to and matching of edges in phonology and morphology. The GA family is defined as in 145. It allows the alignment of one category’s edge with that of another. The categories may be either prosodic (PCat) or grammatical (GCat).<sup>2</sup>

(145)  $\text{ALIGN}(\text{Cat1}, \text{Edge1}, \text{Cat2}, \text{Edge2}) =_{\text{def}} \forall \text{Cat1} \exists \text{Cat2}$  such that Edge1 of Cat1 and Edge2 of Cat2 coincide.

Where:  $\text{Cat1}, \text{Cat2} \in \text{PCat} \cup \text{GCat}$ ;  $\text{Edge1}, \text{Edge2} \in \{\text{right}, \text{left}\}$

(cf. McCarthy & Prince, 1993, 2)

As noted above, OT assumes that the set of constraints that a grammar contains is universal, whereas the ranking of these constraints is language-specific. With respect to GA, the arguments in the template in 145 are also filled differently by learners of different languages, whereas the template itself is universal (cf. McCarthy & Prince, 1993, 7). GA, as a very general principle, can capture a great range of different phenomena, which in a derivational approach would have to be treated in entirely different ways (cf. McCarthy & Prince, 1993, 3-4).

The approach to clitic placement by Anderson (1992) and Richardson (1997), introduced in section 6.4, involves a constraint type called EDGEMOST, which is defined as in 146a. However, McCarthy and Prince (1993) note that EDGEMOST can be subsumed under GA as shown in 146b (cf. McCarthy & Prince, 1993, 14). Thus, EDGEMOST is simply a case of GA in which the edges of both categories are either both left or both right. The advantage of using EDGEMOST instead of ALIGN is thus only notational.

(146) a.  $\text{EDGEMOST}(\varphi, E, D)$ : The item  $\varphi$  is situated at edge E of domain D.

b.  $= \text{ALIGN}(\varphi, E, D, E)$

(cf. McCarthy & Prince, 1993, 14)

### 8.1.3 Syntax in OT

OT was first applied to phonology, but has been extended to syntax and morphology in various studies now.<sup>3</sup> I will first identify which input Gen receives and specify the nature of Gen’s output, i. e. the candidate set. Then I will present some basic assumptions we must make about the syntactic constraint set.

Turning first to the input to Gen, with respect to phonology it is an underlying representation. For syntax, it is an interpretation, i. e. a proposition: it includes heads and their arguments as well as information-structural roles such as focus and topic (cf. Smolensky, Legendre & Tesar,

<sup>2</sup>Generalised Alignment is conceptually close to edge-based theories of the syntax-phonology interface, especially to Selkirk (1986), who argues that prosodic structure is built based on syntactic structure by making reference only to the right or left edge of syntactic constituents (cf. Selkirk, 1986, 373).

<sup>3</sup>For example Grimshaw (1997), Anderson (2000), Billings (2002), Legendre, Raymond and Smolensky (2006), and Engels and Vikner (2006).

2006, 461-3). It is thus not a numeration<sup>4</sup>, as is assumed in MP; see also the discussion of these two alternatives in Broekhuis (2008, 50-6).

With respect to Gen's output, the candidate set, I assume with Grimshaw (1997) that candidates have the same meaning: "competing candidates have non-distinct logical forms, in a sense which must be made precise by further research, but which certainly must entail that they are truth functionally equivalent" (cf. Grimshaw, 1997, 375).<sup>5</sup> I will adopt this superficial assumption, as it suffices for the current purpose. It also has the conceptual advantage that it fits to OT's profile as a production grammar – the propositional content is the starting point, unlike in MP.

Turning now to the contents of the evaluator with respect to syntax, the structure that is assigned to a given input is evaluated by constraints. Unnecessary structure is penalised, for example through constraints such as STAY, which punishes traces/copies, or a constraint OBLIGATORY HEADS, which requires a projection to have a head (cf. Grimshaw, 1997, 401-2). But constraints are violable, and thus higher-ranked constraints might favour candidates which do not comply with these structural requirements.

Grimshaw (1997) argues that projections need not be labelled, since the properties of a projection simply result from its head. This also means that constraints must not – and in fact, cannot – refer to projection labels (cf. Grimshaw, 1997, 417-8).<sup>6</sup> However, we will see below that this radical, yet attractive view, cannot be applied to constraints that regulate clitic placement, which must refer to labels.

Eval not only comprises constraints on the structure of the output, but also syntactic faithfulness constraints, which "tie the success of an output candidate to the shape of the corresponding input; each faithfulness constraint asserts that an input and its output should be identical in a certain respect" (Prince & Smolensky, 2006, 128). Such constraints ensure that despite Gen's unrestrictedness, different inputs lead to different outputs – instead of the same maximally unmarked candidate winning regardless of the intended interpretation (cf. Smolensky et al., 2006, 523).

Whilst for phonology, candidates are evaluated in Eval with respect to an underlying representation, syntactic constraints refer to a given interpretation (cf. Smolensky et al., 2006, 462) – in each case, this is the input that Gen receives, as stated above. For the present analysis, however, this input interpretation is not directly relevant, since I do not employ faithfulness constraints (i. e. correspondences between input and output will not play a role), thus for the sake of simplicity, the interpretation is not given.

---

<sup>4</sup>A numeration is a set of lexical elements (words, features etc.); it must be empty at the end of the derivation (cf. Chomsky, 1995, 226).

<sup>5</sup>This may result directly from the input Gen receives; but given Gen's free analysis, it is also plausible that it can in principle generate candidates that deviate from the intended interpretation. Luckily, this is not a question that must concern us here – either such candidates never surface due to harmonic bounding (a notion referring to the situation when the violation profile of one candidate is a proper subset of that of another candidate: after evaluating a candidate with a harmony value of  $n$ , any candidate with a harmony value smaller than  $n$  will not be evaluated; cf. Smolensky et al., 2006, 497, and references therein), or they fail to pass certain faithfulness constraints (cf. the following subsection on Eval).

<sup>6</sup>"Every projection is just that, a projection, which has a grammatical category, may have a lexically realized head, may have acquired a head by movement, or may lack one completely. The properties of a projection are just a function of what happens to head it" (Grimshaw, 1997, 418).

### 8.1.4 Summary

To summarise, the OT model assumed here has the structure given in 147. The modelling of Czech cliticisation which I present in this chapter will, like any other OT analysis, focus on Eval, i. e. on capturing the data through identifying the responsible constraints and their respective ranking. I should also note that when I use terms such as “movement” in my analysis, this is strictly speaking only a metaphor. Given that I employ a full OT architecture, the analysis does not include transformations, i. e. there are no intermediate representations. However, an element can occupy several positions within one representation, its copies forming a chain.

(147) Interpretation → Gen → Candidates → Eval → Output

## 8.2 Placement in second position

In section 6.4, I have already sketched Richardson’s (1997) analysis of Czech clitic placement. I will essentially follow his argument and present the basic cases in section 8.2.1. Section 8.2.2 then offers a closer look at the evaluation of the constraints and the role of phrasal integrity. I will also comment on the integration of first-position placement into the present analysis in section 8.2.3. Section 8.2.4 provides a concluding summary.

### 8.2.1 Basic cases

Based on Anderson (1992) and later works by that author, Richardson (1997) proposes the two constraints in 148a-b and the constraint ranking in 148c to capture Czech clitic placement. The constraints are repeated from section 6.4.2.

- (148) a. NONINITIAL(Cl,CP) = A given clitic must not be initial in CP.  
 b. EDMOST(Cl,l,IP) = A given clitic should be as close to the left edge of IP as possible.  
 c. Ranking: NONINITIAL(Cl,CP) >> EDMOST(Cl,l,IP)

(cf. Richardson, 1997, 146-7)

I am assuming that all 2P clitics have a feature [+clitic], whereby the constraints in 148 identify them. In consequence, the constraints can only be violated by 2P clitics, and not by any other elements. Tableau 8.1 shows that NONINITIAL(Cl,CP) prevents ungrammatical clitic-first constructions as in 149a, and illustrates how of the remaining options, EDMOST(Cl,l,TP)<sup>7</sup> favours 2P constructions such as 149b over unacceptable candidates where the clitic is further to the right, as in 149c and 149d.

- (149) a. \**Si Šárka koupila knihy.*  
 REFL Šárka buy.PTCP books

<sup>7</sup>For the sake of coherence, I use *TP* instead of Richardson’s (1997) *IP*.

- b. *Šárka si koupila knihy.*  
 Šárka REFL buy.PTCP books  
 ‘Šárka bought herself books.’
- c. *\*Šárka koupila si knihy.*  
 Šárka buy.PTCP REFL books
- d. *\*Šárka koupila knihy si.*  
 Šárka buy.PTCP books REFL

	NONINI(Cl,CP)	EDGMO(Cl,I,TP)
a. [CP [TP <b>cl</b> NP V NP ]]	*!	
$\Rightarrow$ b. [CP [TP NP <b>cl</b> V NP ]]		*
c. [CP [TP NP V <b>cl</b> NP ]]		**!
d. [CP [TP NP V NP <b>cl</b> ]]		**!*

Table 8.1: OT tableau illustrating 2P placement with empty CP

Since CP and TP coincide in example 149, there is no position in which the clitic could satisfy both constraints at once. Thus, the ideal situation for a clitic is when there is overt material above TP, as in 150, where the clitic occurs directly after the interrogative phrase or complementiser. As the tableau in 8.2 illustrates, 2P placement in these contexts causes no violations of the two relevant constraints.

- (150) a. *Co si Šárka koupila?*  
 what REFL Šárka buy.PTCP  
 ‘What did Šárka buy herself?’
- b. *... že si Šárka koupila knihy.*  
 COMP REFL Šárka buy.PTCP books  
 ‘...that Šárka bought herself books.’

	NONINI(Cl,CP)	EDGMO(Cl,I,TP)
a. [CP <b>cl</b> XP/C [TP NP V ]]	*!	*
$\Rightarrow$ b. [CP XP/C [TP <b>cl</b> NP V ]]		
c. [CP XP/C [TP NP <b>cl</b> V ]]		*!
d. [CP XP/C [TP NP V <b>cl</b> ]]		**!*

Table 8.2: OT tableau illustrating ideal positioning with filled CP

As mentioned in section 6.4, a great advantage of this approach is that the verb does not have to move (i. e. to be topicalised) to host the clitic, which is especially undesirable in simple transitive sentences with a zero subject, as in 151. Instead, the clitic appears after the verb to

avoid the left edge of CP. This is illustrated in the tableau in 8.3. Note that the pattern is the same as with an initial NP as shown in tableau 8.1.

- (151) *Koupila si knihy.*  
 buy.PTCP REFL books  
 ‘She bought herself books.’

	NONINI(Cl,CP)	EDGMO(Cl,l,TP)
a. [CP [TP <b>cl</b> V NP ]]	*!	
b. [CP [TP V <b>cl</b> NP ]]		*
c. [CP [TP V NP <b>cl</b> ]]		**!

Table 8.3: OT tableau illustrating clitic placement after an initial verb

### 8.2.2 Gradient evaluation and Integrity

The tableaux 8.1 and 8.3 show that the non-parallel formulation of the two constraints is crucial: Since in these configurations, EDGEMOST is always violated when NONINITIAL is satisfied, multiple violations of EDGEMOST must be possible, i. e. the constraint must be gradient. In contrast, NONINITIAL must be absolute: either something is initial or not – otherwise the optimal candidate would be the one where the clitic is the furthest from the left edge.

Prince and Smolensky (1993) generally assume that EDGEMOST is gradient, more specifically that “degree of violation [is] measured by the distance of  $\varphi$  from the designated edge” (McCarthy & Prince, 1993, 15). But McCarthy and Prince (1993) state that whether Alignment is gradient or categorical “is a parameter to be set in the specification of individual constraints” (McCarthy & Prince, 1993, 56). If a constraint is gradient, the evaluation can simply refer to the notion of subset: “The optimal candidate among a set of Align violators is the one whose violation-string is contained in all other violation-strings, where the violation-string is the material separating the two specified edges” (McCarthy & Prince, 1993, 54).

Given the gradient evaluation of EDGEMOST, one would expect that Czech has 2W clitics, i. e. that clitics occur directly after the first word. Yet, as example 152, repeated from section 5.1.1, shows, this is not the case. In Czech, unlike in BCMS, clitics cannot break up the initial constituent.<sup>8</sup> Anderson (2000) attributes this to different rankings of members of the constraint family *Integrity*. An integrity constraint is one that penalises the breaking up of a certain structure (cf. Anderson, 2000, 24).<sup>9</sup> The constraint family is defined as in 153.

- (152) a. *Ten básník mi čte ze své knihy.*  
 that poet me.DAT read.3SG from his book  
 ‘That poet reads to me from his book.’

<sup>8</sup>Section 9.1.3 discusses cases of split topicalisation, which at first glance appear to present a counterexample – however, I will show there that it is not the clitic that breaks up the constituent, but information-structural (re-)ordering.

<sup>9</sup>A member of this constraint family that is high-ranked, or even undominated, in most languages, is INTEGRITY(Word) (cf. Anderson, 2000, 22).

- b. \**Ten mi básník čte ze své knihy.*  
 that me.DAT poet read.3SG from his book

(cf. Halpern, 1995, 17)

- (153) INTEGRITY(C) = A member of a category C may not be interrupted by phonological material that is not part of C.

Where:  $C \in \text{PCat} \cup \text{GCat}$

(cf. Anderson, 2000, 23)

In the case of Czech, the constraint family member INTEGRITY(XP) must be ranked above EDGEMOST(Cl,I,TP). The constraint domain *XP* includes lexical categories such as NP and PP, but not functional categories, to the effect that INTEGRITY(XP) still permits the clitics to occur for example inside VP, TP, or CP. The tableau in 8.4 shows the evaluation for sentence 152a. Even though every word preceding the clitics below the left TP edge causes a violation of EDGEMOST, not breaking up the initial phrase is more important, thus the 2D candidate wins over the one with 2W placement. The respective ranking of INTEGRITY(XP) and NONINITIAL cannot be determined, because the two never conflict; they are consequently on the same stratum (and hence no line is drawn between them in the tableau).

		NONI(Cl,CP)	ITGR(XP)	EDGM(Cl,I,TP)
☞	a. [TP [NP Ten básník] cl]			**
	b. [TP [NP Ten cl básník]]		*!	*

Table 8.4: OT tableau illustrating 2D vs. 2W placement in Czech

### 8.2.3 A note on placement in first position

Before concluding this analysis of Czech basic clitic placement, I want to remark on a phenomenon that I have not discussed so far, but that was sketched in section 5.1.3: clitics in first position. We have seen that *prosodic* first-position placement, which occurs regularly also in the standard language, is due to the fact that Czech clitics are not necessarily enclitic. They will thus follow a certain *syntactic* constituent, regardless of whether they follow a pause or not. This is already captured by the analysis presented in the preceding sections.

However, clitics that are also *syntactically* in first position, i.e. at the beginning of the sentence, cannot be captured in the present analysis, because the high ranking of NONINITIAL(Cl,CP) prevents their occurrence. Constructions where the clitic appears in absolute initial position in the sentence are only found in colloquial language; a relevant example is shown in 154, repeated from section 5.1.3. As noted there, assuming a truncation of short words like *tak* ('so') and *to* ('it') is not plausible in this case. The auxiliary clitic has a colloquial spelling, lacking the initial *j*-, and the participle spelling reflects colloquial pronunciation without the final *-l*.



- (154) *Sem ti to už přece řek.*  
 AUX.1SG you.DAT it already though say.PTCP

‘I already told you that, though.’

(cf. Franks & King, 2000, 114)

In a small acceptability judgement study with 16 participants, sentence 154 received a mean rating of 3.8 (with 5 being the best possible rating and 1 the worst). Two other sentences with the first person past auxiliary in first position, which did not have the colloquial spelling, received a slightly lower average rating of 3.0. These intermediate ratings are partly the result of a high standard deviation: for example, sentence 154 was rated with 5 by eight of the speakers, and with 1 by two of them. It is plausible to assume that the sentences are generally accepted colloquially, but that some participants, despite being asked to refer to their every-day common language, applied normative standards in their evaluations.<sup>10</sup>

Syntactic initial placement might reflect a change in the grammar of Czech: possibly for some speakers, the affected elements have lost their clitic status and thus acquired the positional freedom of non-clitic constituents. This is supported by the fact that speakers who accept first-position placement (1P) also accept clitics that appear lower in the clause, as in example 155a, repeated from section 5.1.3. Here, the reflexive clitic is the *fourth* constituent (4P), appearing after a focused non-clitic pronoun. This sentence received an average rating of 2.8;<sup>11</sup> another similarly structured token, sentence 155b, was rated slightly better, with an average of 4.1. Out of the 16 speakers, 13 were consistent in their rating of 1P and 4P; only with three speakers, the judgments for 1P and 4P had a difference higher than one.

- (155) a. <sup>?</sup>*Jistě namítnete, že to VÁM se zítra nestane.*  
 surely object.2PL COMP it you.PL.DAT REFL.ACC tomorrow NEG:happen.3SG

‘You will surely object that to YOU, it will not happen tomorrow.’

(cf. Franks, 1998, 38)

- b. *Víš, že už i PAVEL mi poslal odpověď na pozvání?*  
 know.2SG COMP already even Pavel me.DAT send.PTCP reply to invitation

‘Do you know that EVEN PAVEL has already sent me a reply to the invitation?’

If Czech clitics are becoming independent words in these grammars, then this goes against the often assumed grammaticalisation process as described by Zwicky (1977): “What is a clitic at one stage is reinterpreted as a derivational or inflectional affix at the next” (Zwicky, 1977, 6). One reason for this might be the general prosodic independence of Czech 2P clitics as described in section 5.1.3. It appears plausible that this autonomy leads speakers to analyse these elements

<sup>10</sup>The participant sample consisted of speakers both with a Bohemian and a Moravian linguistic background, but this factor was controlled for and did not yield different results.

<sup>11</sup>Interestingly, with example 155a, speakers from Bohemia and Moravia deviated in their judgements: whereas it received 3.5 points from Bohemian participants, it was judged with only 2.1 points by Moravian speakers. A similar difference occurred with another 4P sentence in the task, yet not with example 155b.

as full words instead of affixes, which then in consequence causes them to also grant the clitics the positional freedom of other full words. But note that these alternative grammars do not merely fall into the colloquial reign, nor are they entirely novel: as Junghanns (2021, 177-8) shows, 4P constructions can be found in texts that are clearly standard language, and over 100 years into the past.

Based on the facts just given, I assume that speakers who accept 1P and 4P constructions then employ a different grammar than when they adhere to the patterns described in this thesis. That grammar would thus require a different model where constraints are either ranked differently, or where the elements under discussion are not specified as clitics, so constraints like NONINITIAL and EDGEMOST do not apply to them. Crucially, I do not assume that speakers who also allow 1P simply have optionality between 1P and 2P placement – the placement of clitics in these alternative grammars is instead plausibly subject to other constraints which affect non-clitic constituents. Whether this is borne out has to be established by further research into clitic placement in different variants of Czech.

#### 8.2.4 Summary

This section has shown that with only a few very simple constraints, 2P clitic placement in Czech can be captured. We have also seen that the formulation of NONINITIAL as absolute and that of EDGEMOST as gradient is crucial. That clitics do not enter the first constituent is due to Anderson's (2000) general constraint INTEGRITY(XP).

### 8.3 Clitic climbing and the domain of EdgeMost

The following subsections show how the present approach can be extended in order to account for clitic climbing (section 8.3.1) and the blocking thereof (sections 8.3.2 and 8.3.3), as well as the consequences this has for analysing coordination structures in Czech (section 8.3.4.). The data presented here mostly stems from an acceptability judgement study that I carried out with 63 participants, who were asked to rate sentences with clitics in varying positions: second in the embedded clause, first in the embedded clause, and second in the matrix clause.

#### 8.3.1 Climbing out of defective clauses

According to Spencer and Luís (2012), verbs which trigger clitic climbing usually belong to one of the following three groups: phasic verbs (for example *begin*, *finish*), verbs of desire/belief (for example *want*, *believe*), and modals (for example *must*, *may*, *seem*; cf. Spencer & Luís, 2012, 166). This is also the case in Czech. An example of climbing with a phasic verb is given in 157, and one of climbing with a modal in 156; both are repeated from section 5.2.1.

- (156) *Včera mu to musel dát mu to.*  
 yesterday him.DAT it.ACC must.PTCP give.INF

‘Yesterday he had to give it to him.’

(cf. Veselovská, 1995, 305)

- (157) *Jan se začal smát se.*  
 Jan REFL.ACC start.PTCP laugh.INF

‘Jan started to laugh.’

(cf. Avgustinova & Oliva, 1995, 14)

The acceptability of climbing in these sentences was confirmed in the acceptability judgement study. I tested instances of all the above-mentioned groups: *začít* ‘begin’, *mocť* ‘can’, *muset* ‘must’, and *chtít* ‘want’. For each verb, there were two lexicalisations, thus a total of eight items. The different conditions are exemplified in 158 with a verb of desire. The results were the same for all matrix verbs: only the variant with the clitic in second position of the main clause was accepted (158c); both 1P (158a) and 2P (158b) in the embedded clause received low ratings. The mean z-scores are displayed in figure 8.2.

- (158) a. *Lucie už dlouho chce [si koupit nové kolo].*  
 Lucie already long.ADV want.3SG REFL.DAT buy.INF new bike  
 b. *Lucie už dlouho chce [koupit si nové kolo].*  
 Lucie already long.ADV want.3SG buy.INF REFL.DAT new bike  
 c. *Lucie si už dlouho chce [koupit nové kolo].*  
 Lucie REFL.DAT already long.ADV want.3SG buy.INF new bike

‘Lucie has been wanting to buy herself a new bike for a long time now.’

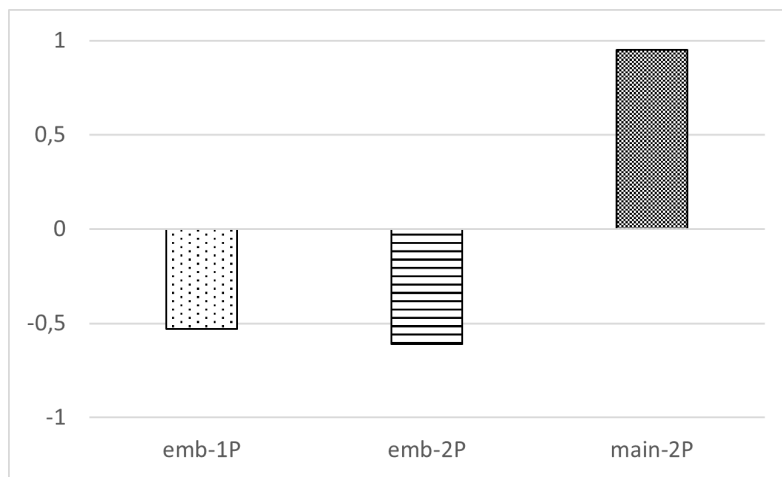


Figure 8.2: Acceptability judgements (z-scores) for different positions of clitics in a clause complex: climbing to main-2P or placement in emb-2P or emb-1P (N=63)

We could capture this by assuming, for example with Dotlačil (2007), that the complements of these verb types are not CPs, but only vPs. Then climbing would be a result of the high-ranked  $\text{EDGEMOST}(\text{Cl}, \text{l}, \text{TP})$  constraint, whilst the higher-ranked  $\text{NONINITIAL}(\text{Cl}, \text{CP})$  prevents placement in first position. However, there are problems with this claim. First, *chtít* ‘want’ is arguably a subject control verb, thus the embedded infinitival clause must have a PRO subject

that carries the agent theta role of *koupit* ‘buy’, and that occupies the specifier of TP. Second, also infinitive clauses with a tense adverb, and thus arguably with a TP layer, allow clitic climbing in Czech, as example 159 shows. The pronoun *ho* is the object of the embedded infinitive *ignorovat*, which can be modified by a temporal adverbial – yet, clitic climbing occurs (cf. Lenertová, 2004, 157).

- (159) *Místo toho se ho rozhodl [TP na moment / příště ignorovat ~~ho~~].*  
 instead it.GEN REFL him decide.PTCP for moment next.time ignore.INF

‘He decided instead to ignore him for a moment/next time.’

(cf. Lenertová, 2004, 157)

A slight alteration of the EDGEMOST constraint enables us to integrate these data, though: apparently, clitics strive for the *highest* TP in their clause. Thus, EDGEMOST must simply refer to this maximal TP within the overall CP, and can be restated as in 160.<sup>12</sup>

- (160) EDGEMOST(Cl,l,MaxTP)  
 = ALIGN(Cl,l,TP<sub>i</sub> such that ALIGN(TP<sub>i</sub>,l,CP,l), l)  
 = A clitic must occur as close to the left edge of the highest TP in its clause complex as possible.

This is also a good point to consider other constraints which interact with clitic climbing. First of all, the new version of EDGEMOST could simply be satisfied by placing the clitic directly within the higher clause, without any copy inside the embedded clause. This is of course undesirable, because pronouns are typically subcategorised for by the verb and must appear within their clause for theta assignment/checking. This is ensured by the constraint THETA, which is defined in 161 and integrates the theta criterion into the analysis; it is a plausible candidate for a universally undominated constraint (cf. Grimshaw, 1997, 387). Note that here we only need to consider pronominal clitics, and not auxiliary clitics, because the latter never climb, since they only occur within finite clauses, i. e. CPs.

- (161) THETA = An argument must be theta-marked.

(cf. Grimshaw, 1997, 387)

Second, having multiple copies of an element is not entirely free of charge: the constraint STAY, as formulated in 162 captures this economy principle and therefore prevents clitics (and other elements) from rising higher unnecessarily.

<sup>12</sup>The second line of this definition can be read as follows: “Align the left edge of a clitic with the left edge of a TP that aligns with the left edge of CP.” Technically, this constraint is also violated when the left edge of TP does not align with the left edge of CP. However, this is arguably always outweighed by other constraints that require certain elements to appear above TP, and thus for reasons of simplicity, I will ignore this aspect here. An alternative solution which would arguably yield the same results was proposed by Stavros Skopeteas (p. c.): instead of referring to the maximal TP, EDGEMOST could also refer to the *complement of C*. Whilst this would be the more elegant solution, it has the complication that when an element is fronted into the left periphery, there must be a Fin projection above TP, and in these contexts the *complement of Fin* would thus be the relevant domain. If one assumes that FinP is always present, this is not problematic, but I will stay agnostic on this issue here – the reader may however keep it in mind as an equivalent solution.

(162) STAY = Trace is not allowed.

(cf. Grimshaw, 1997, 373)

As noted above, THETA is high-ranked, or in all likelihood undominated, in all languages. STAY, on the other hand, is obviously ranked below EDGEMOST in Czech – otherwise, clitic climbing as well as the general leftward movement of clitics would never occur. Thus, the respective ranking of clitic-specific and general constraints must be as given in 163.

(163) THETA; NONINITIAL(Cl,CP) >> EDGEMOST(Cl,l,MaxTP) >> STAY

The tableau in 8.5 illustrates how the evaluator selects the candidate where the clitic has risen from its base position in the embedded vP to the edge of the higher TP. Candidate a has the clitic at the CP edge, which is fatal. Candidate c has the clitic at the lower TP edge, which causes a fatal violation of EDGEMOST(Cl,l,MaxTP). Candidate b, in contrast, has the clitic at its ideal position: since CP is filled by some phrase XP, the edge of the highest TP can be occupied without violating NONINITIAL(Cl;CP). In all candidates a-c a STAY violation occurs; candidate d shows that avoiding this violation by directly placing the clitic at the highest TP edge, without a lower copy in vP, causes a violation of THETA because the clitic must be theta-marked by its predicate in the verbal projection. At the same time, a higher copy must satisfy the requirement to be initial in the matrix TP. Thus, candidate b is optimal.<sup>13</sup>

	THETA	NOIN(CP)	EM(l,MxTP)	STAY
a. [CP <b>cl</b> XP [TP [TP [vP e <sub>H</sub> ]]]]		*!	*	*
b. [CP XP [TP <b>cl</b> [TP [vP e <sub>H</sub> ]]]]				*
c. [CP XP [TP [TP <b>cl</b> [vP e <sub>H</sub> ]]]]			*!	*
d. [CP XP [TP <b>cl</b> [TP [vP]]]]	*!			

Table 8.5: OT tableau illustrating clitic climbing to the highest TP

Before concluding the analysis of clitic climbing out of defective clauses, I must note that, according to, among others, Junghanns (2002a, 82), clitic climbing is either optional or preferred, but never obligatory. As an example he gives 164, where the clitic remains in second position in the embedded clause in 164a. Yet, I could not replicate this in my study, where this sentence was one of the two items with *chtít* ‘want’. Whereas the b-sentence received a mean z-score of 0.75, the a-sentence without climbing got a negative z-score of -0,75 on average (and similarly, the variant with the clitic in first position in the embedded clause got a mean z-score of -0.51). Thus, as with the entire data set reported above, informants also mostly reject clitic placement inside the defective clause for this particular example.

<sup>13</sup>There might be other overt constituents in a-d, in addition to XP; since they do not matter for the evaluation, I am leaving them out for reasons of space and readability. Also note that since the respective ranking of THETA and NONINITIAL(Cl,CP) cannot be determined, no line is drawn between these columns.

- (164) a. *Asi chtěla usušit ho pomalu.*  
 probably want.PTCP dry.INF it.ACC slowly
- b. *Asi ho chtěla usušit pomalu.*  
 probably it.ACC want.PTCP dry.INF slowly
- ‘Probably she wanted to dry it slowly.’

(cf. Junghanns, 2002a, 82)

On the other hand, Junghanns (2002a) argues that discourse factors play a crucial role in licensing clitic climbing.<sup>14</sup> It is thus possible that low clitic placement is dependent on very specific information-structural circumstances. Such a situation could actually be modelled without any additional assumptions in the present analysis: a constraint which refers to the presence of some information-structural feature would have to be ranked higher than EDGE-MOST(Cl,1,MaxTP). Then ideal clitic placement – in this case: climbing – is overridden by an independent, information-structurally motivated requirement.

If the lack of clitic climbing in sentences such as 164a is indeed conditioned by information structural factors, then we are dealing with a similar case as presented by clitic-third with topicalised elements, as described in section 8.2.1 above: there is no real optionality, but distinct contextual conditions lead to distinct outcomes. The alternative, that there is a free choice between two clitic positions, is difficult to model in an OT framework – but it is no less problematic in derivational models, as Kulik (2019) points out: if clitic movement, including clitic climbing, is feature-driven, then a feature either obligatorily triggers movement, or it does not (cf. Kulik, 2019, 10).

### 8.3.2 Blocked climbing out of CP complements

As sketched in section 5.2.1, clitic climbing is generally barred out of finite clauses in Czech, which is illustrated by example 165. It also does not occur with object control clauses such as 166: the clitic does not rise to the matrix clause, but occurs second in the embedded clause.<sup>15</sup> Similar data is offered by Veselovská (1995, 303-4), with analogous judgements. The parallel behaviour of these two multi-clausal constructions is easily captured if we assume that both finite clauses and object control clauses are CPs.<sup>16</sup> The structure of sentence 166a is then as in 167. Note that climbing is possible with subject control matrix verbs such as *chtít* ‘to want’, as was discussed in section 8.3.1.

<sup>14</sup>More precisely: “If a clitic functions as a distinct element of the entire sentence’s background, and is thus separated from the rest of the embedded sentence, whatever its discourse function, then this clitic moves to the highest accessible background position of the matrix clause, if adequate structural conditions are met” (Junghanns, 2002a, 84; the translation from the original German is my own).

<sup>15</sup>As argued already in section 6.1.3, the fact that the clitic does not attach to a null element like *pro* or PRO is another indicator that phonology, i. e. linearisation, is the basis for clitic placement, not syntactic position.

<sup>16</sup>Since interrogative phrase as well as topic extraction are both even possible out of complement clauses introduced by an overt complementiser (cf. Meyer, 2004, 190-4), these phenomena cannot be used to confirm the CP status of object control clauses. I am thus deriving it from the behaviour of clitics, which is evidently parallel to that in clauses with a complementiser.

- (165) a. *Řekl, že mi ho můžete ukázat.*  
 say.PTCP COMP me.DAT him.ACC can.2PL show.INF

‘He said that you can show him to me.’

- b. *\*Řekl mi ho, že můžete ukázat.*  
 say.PTCP me.DAT him.ACC COMP can.2PL show.INF

(cf. Lenertová, 2004, 156)

- (166) a. *Přinutil vrátného dát mu klíč.*  
 force.PTCP doorman.ACC give.INF him.DAT key

‘He made the doorman give him the key.’

- b. *\*Přinutil mu vrátného dát klíč.*  
 force.PTCP him.DAT doorman.ACC give.INF key

- (167)  $[_{CP} [_{TP} \text{Přinutil vrátného}_i [_{CP} [_{TP} \text{PRO}_i \text{ dát } \mathbf{mu} \text{ klíč } ]]]]$  (= 166a)

However, our current constraint set gives the wrong result, as tableau 8.6 shows. The frowning face indicates that the marked candidate is the winner despite not being the variant found in the data. This is because  $\text{EDGEMOST}(Cl, l, \text{MaxTP})$  allows multiple violations; thus, the further a clitic is from the highest TP (MaxTP), the more violations occur. In consequence, the candidate with the lower clitic position has more  $\text{EDGEMOST}$  violations than the one with the clitic in the higher CP. We also cannot fix the problem by reordering  $\text{STAY}$  and  $\text{EDGEMOST}$  – this would amount to saying that the clitic never moves to TP’s edge, so it would have to be base-generated there, which would cause a  $\text{THETA}$  violation (the constraint is not featured in the tableau for simplicity). Thus, since  $\text{STAY}$  is, and must be, a relatively low-ranked constraint, it cannot prevent climbing out of CP.

	NI(CP)	EM(l, MxTP)	STAY
a. $[_{CP} [_{TP} \text{V NP } [_{CP} [_{TP} \text{PRO V cl}]]]]]$		**!*	
⊖ b. $[_{CP} [_{TP} \text{V cl NP } [_{CP} [_{TP} \text{PRO V e}]]]]]$		*	*

Table 8.6: OT tableau illustrating wrongly predicted climbing out of a CP

In addition, there are even configurations where climbing out of CP would put the clitic in a position where both clitic-specific constraints are satisfied: if the higher clause allows the clitic to sit at TP’s left edge without being initial in CP, but the lower clause cannot offer this (because its left periphery is empty), then again the optimal candidate is the one where the clitic climbs out of the lower and to the higher CP. However, this is not what happens, as sentences like the one in 168 show.

- (168) a. *Myslím*, [<sub>CP</sub> *že* [<sub>TP</sub> *přinutil* *vrátného* [<sub>CP</sub> [<sub>TP</sub> *dát* ***mu*** *klíč*]]]]  
 think.1SG COMP force.PTPC doorman.ACC give.INF him.DAT  
 key  
 ‘I think that he made the doorman give him the key.’
- b. \**Myslím*, [<sub>CP</sub> *že* [<sub>TP</sub> ***mu*** *přinutil* *vrátného* [<sub>CP</sub> [<sub>TP</sub> *dát* *klíč*]]]]  
 think.1SG COMP him.DAT force.PTPC doorman.ACC give.INF  
 key

	NI(CP)	EM(1,MxTP)	STAY
a. [ <sub>CP</sub> Comp [ <sub>TP</sub> V XP [ <sub>CP</sub> [ <sub>TP</sub> V <b>cl</b> ]]]]		*!***	
⊗ b. [ <sub>CP</sub> Comp [ <sub>TP</sub> <b>cl</b> V XP [ <sub>CP</sub> [ <sub>TP</sub> V e]]]]			*

Table 8.7: OT tableau illustrating wrongly predicted climbing out of a CP

Fortunately, we already do have a constraint which can prevent clitics from climbing out of CP, NONINITIAL(CI,CP); but so far, we have not taken a closer look at *how* the clitic can actually leave the lower CP. But if we assume that it must, like other elements, obey a non-derivational analog of the phase impenetrability condition (PIC), then we get the ban on climbing out of CP for free: if phase impenetrability is ranked higher than EDGEMOST(CI,1,MaxTP), then a clitic which climbs out of its CP to the higher TP necessarily violates either PIC or NONINITIAL(CI,CP). These two constraints are in direct conflict and can never be satisfied simultaneously by a clitic.

Chomsky’s (1999) definition of the phase impenetrability condition is given in 169. Since CP is a phase, it follows from the PIC that elements that move out of a CP must do so through the left edge of that CP. We can formulate a non-derivational version of this condition which is compatible with Eval’s parallel computation as in 170.<sup>17</sup>

- (169) Phase Impenetrability Condition = In a phase  $\alpha$  with the head H, the domain of H is not accessible to operations at ZP, but only H and its edge.

$$[ZP Z \dots [HP \alpha [ H YP ]]]$$

(cf. Chomsky, 1999, 10)

- (170) CP-BARRIER = \*... X<sub>i</sub> ... [<sub>CP</sub> ... [<sub>YP</sub> X<sub>i</sub> ... ] ... ]  
 = An element which has a copy within CP must also have a copy at the left edge of that CP.

<sup>17</sup>Note that phases are not compatible with OT’s one-step parallel evaluation. Yet, assumptions similar to the PIC have existed before the emergence of phase theory, for example van Riemsdijk’s (1978) *Head Constraint*, developed based on his study of extraction out of prepositional phrases and preposition stranding in Dutch and English (cf. van Riemsdijk, 1978, 158-60). For this reason, I am naming the constraint CP-BARRIER, thus referring to the pre-phaseal notion.



Implementing the PIC into a constraint-based syntactic analysis is desirable also apart from clitic climbing, since it is the basis for successive-cyclic movements as found for example with interrogative extractions. In addition, as will be discussed in section 8.4.3, it is plausible to generally ban clitics from CP’s left edge for information-structural reasons, and thus we can capture the non-initiality of 2P clitics as well as their restriction to their own CP based on the interaction of just two, independently motivated, constraints. The tableaux in 8.8 and 8.9 show correct results for climbing out of finite CPs and infinitive control clauses with the addition of the CP-BARRIER constraint.<sup>18</sup>

	NOIN (CP)	CP- BARR	EDGM (MxTP)	STAY
☞ a. [CP [TP V [CP že [TP cl]]]]			**	
b. [CP [TP V cl [CP že [TP el]]]]		*!	*	*
c. [CP [TP V cl [CP el že [TP el]]]]	*!		*	*

Table 8.8: OT tableau illustrating blocked climbing out of a finite CP

	NOIN (CP)	CP- BARR	EDGM (MxTP)	STAY
☞ a. [CP [TP V NP [CP [TP PRO V cl]]]]			***	
b. [CP [TP V cl XP [CP [TP PRO V el]]]]		*!	*	*
c. [CP [TP V cl NP [CP el [TP PRO V el]]]]	*!		*	*

Table 8.9: OT tableau illustrating clitic climbing out of a control clause

Dotlačil (2007) makes the important point that it is really the presence or absence of a CP *layer* that is relevant for clitic climbing – and not the intervention of a *C head*. He gives the examples in 171, where the a-sentence has an (overt) complementiser, *zda*, but the b-sentence does not (*jakou historku* being an interrogative phrase in Spec-CP). Yet, both sentences are not acceptable with clitic climbing (cf. Dotlačil, 2007, 79). This is further support for the current analysis, where constraints refer to the edges of projections, here specifically the CP edge.

- (171) a. \**Já mu nevím [CP zda říct pravdu].*  
 I him NEG.know.1SG whether say.INF truth  
 ‘I do not know whether to tell him the truth.’
- b. \**Já mu nevím [CP jakou historku říct].*  
 I him NEG.know.1SG what story say.INF  
 ‘I do not know what story to tell him.’

(cf. Dotlačil, 2007, 79)

<sup>18</sup>As becomes clear from the two tableaux in 8.8 and 8.9, we have to assume that NONINITIAL refers to all copies, whereas EDGEMOST only refers to the highest one.

### 8.3.3 Initial placement in the embedded clause

Veselovská (1995) classifies clauses like 166 as ECM (cf. Veselovská, 1995, 303-4), which is surprising – nothing is exceptionally case-marked in these configurations. Reflexives can serve as a test here, because they should be bound by a nominative PRO, but not by an ECM subject (cf. Lenertová, 2004, 164; see also section 3.4). Sentence 172 is exactly such a case: the reflexive can only refer to the main clause object, which must thus have a coindexed nominative PRO in the embedded clause. A proper case of ECM, in which climbing is possible, was presented in section 5.2.1. If we assume that ECM clauses are smaller than CPs, and therefore allow for case assignment out of the matrix clause, they can be naturally included into the present analysis.

- (172) *Doporučil jim<sub>1</sub> [PRO<sub>1</sub> se<sub>1</sub> umýt].*  
 recommend.PTCP them.DAT REFL.ACC wash.INF

‘He recommended them to wash themselves.’ Not: ‘...to wash him.’

(cf. Short, 1993a, 515)

However, sentence 172 presents a problem for the present analysis: note that here, the clitic is actually initial within the embedded CP! We cannot interpret the two clitics as a cluster in this example, because then the reflexive would have to precede the dative; also, climbing out of object control clauses is, as we have seen, not possible. Thus, we really do have a clitic at the left edge of a CP in the above example. To verify the existence of such constructions, I included two sentences like 172 into the acceptability questionnaire, with the same variations of clitic positioning as described above for defective clauses: 1P in the embedded clause, 2P in the embedded clause, and 2P in the matrix clause. And, indeed, participants accepted both 1P and 2P positioning within the embedded clause – in contrast to the versions with climbing, which were rejected. The sentences in 173 give an example for each of the three conditions, and figure 8.3 displays the mean results.

- (173) a. *Doporučila přátelům [CP se naučit programovat].*  
 recommend.PTCP friend.DAT.PL REFL.ACC learn.INF program.INF
- b. *Doporučila přátelům [CP naučit se programovat].*  
 recommend.PTCP friend.DAT.PL learn.INF REFL.ACC program.INF
- c. *Doporučila se přátelům [CP naučit programovat].*  
 recommend.PTCP REFL.ACC friend.DAT.PL learn.INF program.INF

‘She recommended her friends to learn how to programme.’

Note that it is highly implausible to analyse 173a as 3P in the matrix clause: first, because then 173c would have to be acceptable, too, since it features 2P placement within that same matrix clause; and second, because the highly restricted conditions for 3P placement, described in section 5.1.2, are not met. We are thus facing a dilemma here: either there is a CP barrier (as the binding of the reflexive by PRO suggests), then sentence 173a should not be acceptable,

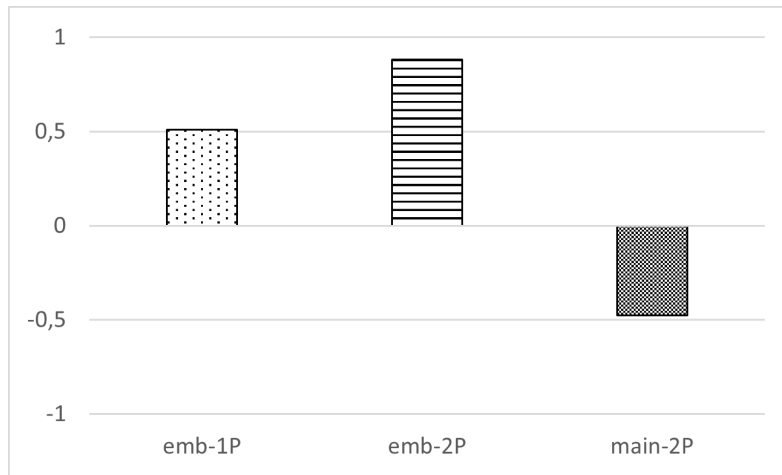


Figure 8.3: Acceptability judgements (z-scores) for different positions of clitics with a control clause: climbing to main-2P or placement in emb-2P or emb-1P (N=63)

because the clitic sits at precisely this barrier – but it is. Or there is no CP barrier; then sentence 173c should be the only acceptable option – but it is exactly the other way around.

Given that the first explanation, with an embedded CP, is the much more plausible one, due to both the binding of the reflexive and the fact that climbing out of the clause is impossible, as 173c shows, I conclude that in sentence 173a, the reflexive clitic *se* does sit at the left edge of the embedded CP. Possibly, a lower left CP boundary is not as “strong” as the highest one with respect to the violation of NONINITIAL(CI,CP), at least for some of the speakers.<sup>19</sup> On the other hand, speakers who allow 1P placement in the embedded clause would then, according to the present analysis, also have to allow clitic climbing – this is, alas, not the case.

### 8.3.4 Consequences for coordination

In his discussion of BCMS clitic placement, Bošković (2001) argues against the assumption that clitics are located in C, based on coordination structures as in 174. If clitics were located in this position, the first conjunct would have to be a CP, and consequently, the second conjunct, *razbio ga*, would have to be a CP as well. But he claims that it is much more plausible that the coordinated phrases are smaller (i. e. VPs, or AgrOPs etc.; cf. Bošković, 2001, 48-9); since otherwise, one would have to assume the deletion of an auxiliary in the second clause, which should not be allowed because it involves deletion of a part of an  $X^0$ , and because it violates the Head Condition, which demands that “no constituent can be deleted that is c-commanded by an overt  $X^0$  in its conjunct at S-Structure” (Bošković, 2001, 49).

- (174) [<sub>CP</sub> Ivan **je** kupio auto] i [<sub>CP</sub> razbio [<sub>CP</sub> **ga je**]].  
 Ivan AUX.3SG buy.PTCP car and ruin.PTCP it

‘Ivan bought a car and ruined it.’

(cf. Bošković, 2001, 48)

<sup>19</sup>With respect to intra-speaker variation, all three possible combinations occurred: some participants preferred 2P over 1P, some preferred 1P over 2P, and yet others had not preference.

In Czech, analogous constructions are possible, as 175a shows. This might not be surprising, as I have provided a range of arguments against the view that clitics are located in C throughout this thesis. Yet, Bošković's (2001) claim also constitutes a problem for our current OT model of Czech clitic placement: If *trénoval ho* was not a CP, but, say, a TP, then the lowest amount of  $\text{EDGEMOST}(\text{Cl}, 1, \text{MaxTP})$  violations would turn the balance, and, as with clitic climbing out of defective clauses, we would expect 175b to be the winning candidate. Even if we were to prevent such movement out of a conjunct clause by employing a more general constraint, we would not get the correct result: if the second conjunct is not a CP, then the clitic should be able to appear at its left edge; the unacceptability of 175c shows that this is not the case.

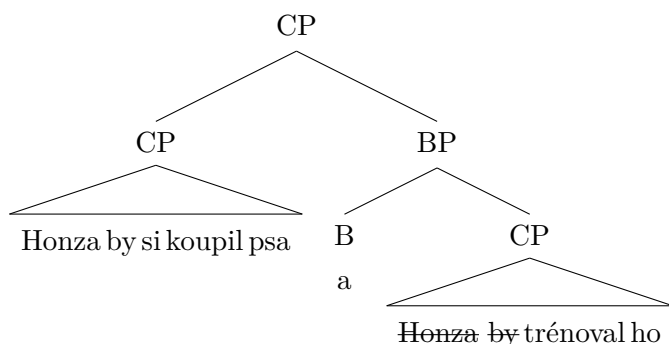
- (175) a. *Kdyby měl čas, Honza by si koupil psa a [?P trénoval ho].*  
 if have.PTCP time Honza COND.3 REFL buy.PTCP dog and train.PTCP  
 him.ACC

'If he had time, Honza would buy a dog and train it.'

- b. \**Kdyby měl čas, Honza by si ho koupil psa a [?P trénoval].*  
 if have.PTCP time Honza COND.3 REFL him.ACC buy.PTCP dog and  
 train.PTCP
- c. \**Kdyby měl čas, Honza by si koupil psa a [?P ho trénoval].*  
 if have.PTCP time Honza COND.3 REFL buy.PTCP dog and him.ACC  
 train.PTCP

I thus conclude that the second conjunct is a CP and that therefore, the clitic does not appear at its left edge – and, in consequence, it also does not climb out of it, due to the interaction of  $\text{NONINITIAL}(\text{Cl}, \text{CP})$  and  $\text{CP-BARRIER}$  described in section 8.3.2 above. The structure of the coordinated CPs as given in 176 follows Munn (1993), who holds that  $\text{BoolP}$  adjoins to the first conjunct (cf. Munn, 1993, 13).

- (176) Coordination of two CPs:



If we assume a structure as in 176 and the constraint  $\text{EDGEMOST}(\text{Cl}, 1, \text{MaxTP})$ , we can capture clitic placement: the clitic rises to the edge of the highest TP within its CP, but not

higher, so it does not leave its conjunct. The respective tableau for this evaluation is shown in 8.10. I thus agree with Bošković (2001) in that the BCMS data in 174 provides additional evidence that clitics do not cluster together in the C head. But nevertheless, if we accept the validity of  $\text{NONINITIAL}(\text{Cl}, \text{CP})$  – which is consistent with the data discussed so far – then at least in Czech, clausal conjuncts are CPs; for this reason, clitics do not appear at their left edge.

	NI (CP)	CP- BAR	EDGM (MxTP)	ST AY
a. $[_{\text{CP}} [_{\text{CP}} [_{\text{TP}} \text{NP V}]] [_{\text{BP}} \text{a} [_{\text{CP}} [_{\text{TP}} \text{V cl}]]]]$			****	
b. $[_{\text{CP}} [_{\text{CP}} [_{\text{TP}} \text{NP cl V}]] [_{\text{BP}} \text{a} [_{\text{CP}} [_{\text{TP}} \text{cl V cl}]]]]$	*!		*	**
c. $[_{\text{CP}} [_{\text{CP}} [_{\text{TP}} \text{NP V}]] [_{\text{BP}} \text{a} [_{\text{CP}} [_{\text{TP}} \text{cl V cl}]]]]$	*!		***	*

Table 8.10: OT tableau illustrating clitic placement with CP conjuncts

### 8.3.5 Summary

The previous subsections have shown how to capture climbing phenomena and the lack of climbing within the present analysis. For clitic climbing, only a slight modification of one of Richardson’s (1997) two clitic-specific constraints was necessary:  $\text{EDGEMOST}$  must refer to the *highest* TP within the clitic’s overall CP. The impossibility of clitic climbing out of CPs in Czech follows from the interaction of the other clitic-specific constraint,  $\text{NONINITIAL}(\text{Cl}, \text{CP})$  with the optimality-theoretic equivalent of phase impenetrability; no additional assumptions need to be made. Grimshaw’s (1997) general constraints  $\text{THETA}$  and  $\text{STAY}$  ensure that pronominal clitics receive their theta role in their base positions and that elements do not move unnecessarily. I have also argued that clitic placement gives us a hint concerning the size of clausal conjuncts in Czech: the fact that clitics eschew their left edge shows that they are best analysed as CPs.

Another interesting phenomenon that this section has addressed is optionality in clitic climbing: first, the optional remaining of clitics within infinitival complements, described for example by Junghanns (2002a; cf. example 164), which however could not be replicated in the acceptability survey; and second, the perplexing variation between 2P and 1P placement in embedded infinitival CPs (cf. example 173). This is an area which deserves much more attention in future research, and it points towards an interesting deviation of Czech clitics from generally expected clitic behaviour, as it was sketched in table 2.1 in section 2.1: they apparently do not behave like affixes in having an entirely “simple distribution”.

## 8.4 Why 2P? Motivating the constraints

The constraints used in the present analysis can all be acquired by a child based on the input it receives, with  $\text{NONINITIAL}$  and  $\text{EDGEMOST}$  adhering to the universal constraint format of GA. However, they do not explain *why* the input has this particular shape, i. e. why clitics appear in second position – but this issue is also far from solved outside of OT. In this section, I will first review a few general ideas on the occurrence of 2P clitics, and then discuss whether within OT, independently motivated constraints can be formulated to help shed light on this issue.

### 8.4.1 General considerations concerning 2P

Wanner (1996) makes the observation that the first position in a sentence typically has very high prominence, which is why it typically holds foci, topics, and stage setters. The second position then forms a kind of “valley”; clitics are often pronouns, auxiliaries and other material that does not convey prominent or new information, and the second position thus naturally attracts such background material (cf. Wanner, 1996, 538). However, since the second position can just as well contain full and, in some languages, even focused NPs, this does not seem a sufficient explanation.

Anderson (2005), who views clitics as phrasal affixes, draws an analogy between clitic placement in sentences and affix placement in words. In both contexts, he claims, grammatical material is placed at or near the edges. When discussing the striking difference in frequency between 2P clitics and their mirror-image, penultimate clitics, he speculates that “the identification (and stability) of left edges is important in itself for morphosyntactic parsing, while no corresponding significance is attached to right edges” (Anderson, 2005, 142). He also makes an argument specifically for Czech: the motivation for the clitics’ non-initiality is that in Czech, grammatical material is not placed at the beginning of the word, i. e. the language is suffixal. Clitics, too, obey this restriction (cf. Anderson, 2005, 137-8). Thus, clitics gravitate towards the left edge, but do not appear in first position. Yet, it is unclear why this should not apply to pronouns etc. in general, at least in suffixal languages – why do they only appear in 2P in some languages, but can occupy any later position in a sentence in others?

The fact that clitics appear near the left edge might be interpreted as reflecting their information-structural status: Since 2P clitics in Czech are pronouns and auxiliaries, one might argue that they represent given information (or at least are not focused) and therefore appear that far to the left. This is also argued by Junghanns (2002a): “Clitics as background material par excellence strive as far upwards in the hierarchic syntactic structure/as far leftwards in the linear order of sentence constituents as possible” (Junghanns, 2002a, 84)<sup>20</sup>. However, Short (1993a, 495) notes that clitics are not affected by information-structural ordering. Otherwise, their appearance at the right edge of very short sentences, regardless of the information-structural status of the first constituent, would not be expected. Clitics thus behave differently from other types of elements regarding information-structure-related ordering.

The generalisation we can make at this point is thus that 2P clitics have both a fixed prosody (they cannot bear independent stress) and a fixed information-structural role (they cannot be foci) – and at the same time, they also have a strictly fixed position with respect to other elements. Further than that, we simply do not have a comprehensive explanation why clitics are placed in second position at this point.

### 8.4.2 Billings 2002: Scope and suffixation

Billings (2002) makes an illuminative attempt to formulate independently motivated constraints that not only correctly *model* 2P placement, but also *explain* it. Instead of *EDGEMOST*, he proposes the constraint *SCOPE* as defined in 177. He uses the notion of scope in a sense much

<sup>20</sup>The translation from the original German is my own.

wider than the traditional one – i. e. it does not apply solely to operators. Clitics take scope in the sense that they have “some morphosyntactic property that is relevant to an entire phrasal domain” (Billings, 2002, 72).<sup>21</sup>

(177) SCOPE = Elements precede the domain over which they take scope.

(cf. Billings, 2002, 72)

The constraint leads to the exact same predictions as EDGEMOST(Cl,l,MaxTP), and at first glance appears less stipulated. Yet, the very loose notion of scope that Billings (2002) uses actually voids it of any meaning: it appears arbitrary which elements take scope and which ones do not, and thus any explanatory power is lost. I therefore argue that the advantage of SCOPE over EDGEMOST is only apparent, and since SCOPE additionally confounds the behaviour of clitics with that of operators, I conclude that EDGEMOST is more adequate.

To explain why clitics, despite being drawn to the left edge, do not appear initially, Billings (2002) proposes the constraint SUFFIX as defined in 178. Clitics thus need to follow a prosodic word ( $\omega$ ) because they are marked as suffixes.<sup>22</sup> This is in line with the traditional idea that 2P clitics are prosodically weak and therefore lean on the preceding word (for example Zwicky, 1977; Halpern, 1995).<sup>23</sup>

(178) SUFFIX = Morphemes marked as suffixes must follow some prosodic word.

(cf. Billings, 2002, 72)

SUFFIX does not make the same prediction as NONINITIAL: whilst the former requires clitics to attach prosodically to a preceding  $\omega$ , the latter only requires the clitic not to be initial in a certain domain – whether it is supported prosodically does not matter. Crucially, SUFFIX goes against Anderson’s (1992) observation that there is no connection between the properties of special clitics and prosodic deficiency, since “prosodic dependence can be found either with or without special placement, and vice versa” (Anderson, 1992, 203-4). And indeed, Czech is such a case: Czech clitics appear in second position whether they find a preceding host there or not, as cases like the reflexive clitic *se* in 179, first introduced in section 5.1.3 (and discussed in detail there), show.

(179) *Ten doktor, co mu důvěruješ, se neholí.*  
 that doctor REL him.DAT trust.2SG REFL.ACC NEG:shave.3SG

‘That doctor, whom you trust, doesn’t shave.’

<sup>21</sup>He does not state which property this is; but he refers to Anderson (1996), who holds that clitics are phrasal affixes because they represent a phrase’s subject and object features (cf. Billings, 2002, 69).

<sup>22</sup>One possible objection against this constraint is that clitics are not affixes. I will not discuss this issue at this point, but note that such objections can also be avoided if one formulates it instead as ENCLITIC; then suffixes and clitics can be kept apart, whilst the general argument remains the same.

<sup>23</sup>The formulation of this constraint might be slightly misleading, because it actually says nothing directly about prosodic attachment. Yet, in Billings’s (2002) discussion of the constraint, it becomes clear that in contrast to NONINITIAL, SUFFIX is in fact about prosodic encliticisation: it “requires morpholexical subcategorization to be parsed prosodically” and “is similar in spirit to the position taken by Franks (2000: 37) regarding his constraint PROSODIC SUPPORT” (Billings, 2002, 73).

(cf. Toman, 1986, 125)

The tableau in 8.11 shows how the two constraints by Billings (2002) make the wrong prediction for example 179: the evaluator favours the candidate where the clitic suffixes to the preceding relative clause (candidate b, marked by a frowning face). Yet, this is not what we find; the clitic does not obliterate the right edge of the RC; speakers utter it with a pause after the appositive. If anything, the clitic attaches prosodically to the following element, the verb (cf. Avgustinova & Oliva, 1995, 18-9), as in candidate c. The equal sign in the candidates' representations indicates the direction of prosodic attachment.

	SUFFIX	SCOPE/ EDM(MxTP)
a. [CP [TP <b>cl</b> = NP [RC] V]]	*!	
☹ b. [CP [TP NP [RC] = <b>cl</b> V]]		**
c. [CP [TP NP [RC] <b>cl</b> = V]]	*!	**
d. [CP [TP NP [RC] V = <b>cl</b> ]]		***!

Table 8.11: OT tableau using SUFFIX for Czech: wrong winner

However, when we repeat the evaluation with NONINITIAL(Cl,CP) instead of SUFFIX, we find that so far, our account simply has nothing to say about the position of the prosodic host with respect to the clitic. Therefore, in tableau 8.12, we have two winning candidates: the clitic is in the correct position, but both alternatives for attachment are equally optimal.

	NONINI (CP)	SCOPE/ EDM(MxTP)
a. [CP [TP <b>cl</b> = NP [RC] V]]	*!	
☹ b. [CP [TP NP [RC] = <b>cl</b> V]]		**
☹ c. [CP [TP NP [RC] <b>cl</b> = V]]		**
d. [CP [TP NP [RC] V = <b>cl</b> ]]		***!

Table 8.12: OT tableau using NONINITIAL for Czech: two winners

How can this be remedied? Remaining within Billings's (2002) argumentation, a clitic could satisfy both SCOPE and SUFFIX by simply attaching to the last word of the preceding clause – and thus appear in syntactic first position. This is prevented by an alignment constraint, defined in 180, which requires the left edge of a clause to be a  $\iota$ P left edge. The effect is that constructions like the one in 181 are not possible in Czech, where ALIGN(Clause,l, $\iota$ P,l) is high-ranked; whereas similar constructions are possible for example in Tagalog, where this constraint is ranked below SCOPE and SUFFIX (cf. Billings, 2002, 81-3).

(180) ALIGN(Clause,l, $\iota$ P,l) = A clause's leading edge must coincide with the leading edge of a  $\iota$ P.

(cf. Billings, 2002, 72)



- (181) \**Už* *víš*, [<sub>CP,IP</sub> =*si* *že Šárka koupila letadlo*]?  
 already know.2SG REFL.DAT COMP Šárka buy.PTCP airplane

‘Intended: Do you already know that Šárka has bought herself an airplane?’

However, this constraint actually has nothing to say about the Czech example in 179, because it refers to left edges only, whereas we are dealing with a right clause edge next to the clitic in this sentence. We thus need a more general constraint for keeping both left and right edges of clauses prosodically intact. A category of constraints commonly used for this purpose are matching constraints; the one relevant for the present purpose is defined as in 182.

- (182) MATCHCLAUSe: A clause in syntactic constituent structure must be matched by a corresponding *t*-phrase in prosodic representation.

(cf. Féry, 2017, 86)

What happens now when we rank this constraint above SCOPE/EDGEMOST can be seen in the tableaux in 8.13 and 8.14 below. With SUFFIX, candidate b (which is in the correct position, but has the wrong prosodic attachment) is no longer optimal, because it violates MATCHCLAUSe. The erroneously optimal candidate now is d, where the clitic appears lower in the clause in order to have an element in front of it that it can suffix to. In contrast, when we employ the NONINITIAL constraint, we get the correct result: out of the two candidates in the best syntactic position, b and c, MATCHCLAUSe selects the one where the clitic keeps the edge of the relative clause prosodically intact and instead prefixes to the following element. I therefore conclude that SUFFIX cannot be operative for Czech clitics.<sup>24</sup>

	MATCH CLAUSe	SUFFIX	SCOPE/ EDM(MxTP)
a. [ <sub>CP</sub> [ <sub>TP</sub> <b>cl</b> = NP [ <sub>RC</sub> ] V]]		*!	
b. [ <sub>CP</sub> [ <sub>TP</sub> NP [ <sub>RC</sub> ] = <b>cl</b> V]]	*!		**
c. [ <sub>CP</sub> [ <sub>TP</sub> NP [ <sub>RC</sub> ] <b>cl</b> = V]]		*!	**
☹ d. [ <sub>CP</sub> [ <sub>TP</sub> NP [ <sub>RC</sub> ] V = <b>cl</b> ]]			***

Table 8.13: OT tableau using SUFFIX and MATCH for Czech: wrong winner

	MATCH CLAUSe	NONINI (CP)	SCOPE/ EDM(MxTP)
a. [ <sub>CP</sub> [ <sub>TP</sub> <b>cl</b> = NP [ <sub>RC</sub> ] V]]		*!	
b. [ <sub>CP</sub> [ <sub>TP</sub> NP [ <sub>RC</sub> ] = <b>cl</b> V]]	*!		**
☞ c. [ <sub>CP</sub> [ <sub>TP</sub> NP [ <sub>RC</sub> ] <b>cl</b> = V]]			**
d. [ <sub>CP</sub> [ <sub>TP</sub> NP [ <sub>RC</sub> ] V = <b>cl</b> ]]			***!

Table 8.14: OT tableau using NONINITIAL and MATCH for Czech: correct winner

<sup>24</sup>The same holds for the alternative, ENCLITIC; cf. fn. 22. Also, it should be noted that in Czech, non-stressed syllables are not reduced, so there is no phonetic evidence for enclisis. How should a learner then know that the clitic is not independent? Arguably, a mere lack of clitic-first data might not be enough for a learner to form such a hypothesis.

Based on the arguments just given, I thus do not adopt Billings's (2002) SCOPE and SUFFIX for Czech clitics, and instead adhere to the constraints listed in section 8.2.4. The explanatory advantage of SCOPE is, as we have seen, only apparent, and SUFFIX makes the wrong predictions for Czech, where clitics are not strictly suffixal. On the other hand, we have seen that the clitics' direction of attachment can be regulated by general constraints like ALIGN(Clause,<sub>i</sub>P) or MATCHCLAUSE, which favour the prosodic demarcation of clause edges.

For the sake of argument, I have assumed here that Czech clitics incorporate into some neighbouring element, though not necessarily to their right. However, the data presented in section 5.1.3 has shown that they can also appear without any prosodic host. The question is whether this is a last resort option, or whether Czech clitics never en- or pro-cliticise, and are thus prosodically completely independent. In the absence of more fine-grained prosodic data, I am leaving this issue open. What is crucial here is that a suffix-like status clearly does not drive clitic placement in Czech, and that it can – and must – be modelled independently from prosody.

Note, however, that the constraints proposed by Billings (2002) might be operative in other parts of Czech grammar. For example, as we have seen in section 4.1.3, the negative proclitic *ne-* forms a prosodic unit with the verb that it attaches to. It is not a 2P clitic and does not permit prosodic breaks between itself and its host. The negative element *ne-* is thus a plausible candidate for an element which is subject to a constraint PREFIX or PROCLITIC. Also, we have already seen that clitics in BCMS are subject to prosodic restrictions. The role of SUFFIX for this language is discussed in section 10.3.5.

### 8.4.3 Dotlačil 2007: Information-structural restrictions

The preceding sections have shown that identifying independent factors that are responsible for the clitic-specific restrictions we observe in the data is not trivial. In this respect, Dotlačil's (2007) view on clitic climbing is highly interesting. He compares Czech clitic climbing with general movement restrictions in the language, and arrives at the following generalisation: If a phrase goes through the CP edge, it must be interpreted either as a contrastive topic or as a focus. But, as we have seen, clitics are incompatible with such an interpretation. Thus, if the phrase is a clitic, the sentence becomes illicit (cf. Dotlačil, 2007, 92).

Dotlačil's (2007) approach resembles the present one in tying restrictions on clitic climbing to the more general ban from the left CP edge. But it goes even further in arguing that this ban is information-structurally motivated. This has direct implications for the constraint we employ to capture this restriction, NONINITIAL(Cl,CP). It may actually have its source in a more general restriction on placing information-structurally “weak”, or defective, elements in salient positions, which could also be incorporated as a constraint into the present OT analysis. Since the focus of this dissertation is on clitics, I am not formulating this very general constraint here, and I will keep using NONINITIAL(Cl,CP) in the remainder of this thesis. Yet, the reader should keep in mind that it can be traced back to a far more general umbrella constraint that has consequences in all of Czech syntax.

The result is that we are not dealing with two clitic-specific constraints any more, but only one: EDGEMOST(Cl,l,MaxTP). It is thus a property specific to Czech clitics to strive towards

the highest TP – but the fact that they avoid the left CP edge (and therefore also do not climb) is simply due to their inability to be information-structurally salient.

#### 8.4.4 Summary

With respect to the reasons that underlie the existence of 2P clitics in Czech, the best explanation I can give at this point is as follows: First, clitics are subject to  $\text{NONINITIAL}(\text{Cl}, \text{CP})$  – or actually a more general version of this constraint – due to the first position’s information-structural status. Czech 2P clitics cannot be foci or contrastive topics and thus cannot occupy or pass through the CP edge. Second, why they appear close to the left TP edge, i.e. are subject to  $\text{EDGEMOST}$ , is more obscure; probably Anderson’s (2005) idea that grammatical material is drawn leftwards is on the right track. Further than that, we cannot explain clitic 2P placement at the moment – what we can and should aim to achieve, though, is a coherent model that accurately predicts their peculiar behaviour.

Whilst I have shown that Billings’s (2002) constraints cannot help shed light on the origins of 2P for Czech, his  $\text{SUFFIX}$  constraint, or some equivalent to it, is arguably relevant for true enclitics, i.e. clitics that always need to attach prosodically to an element to their left. For Czech 2P clitics, I assume that they are not specified as suffixes in the lexicon, and therefore this constraint has no bearing on them.

### 8.5 The order of clitics within the cluster

In my analysis, I have so far treated the Czech clitic cluster as a unit, which is justified as it behaves like one with respect to its placement in the sentence. It also cannot be split and does not allow for partial ellipsis (cf. section 4.3). I will now turn to the internal organisation of the cluster, i.e. the respective ordering of the clitics. I have already described the Czech pattern in section 4.3, from which I repeat the template that best captures it in 183.

(183) Clitic template for Czech:

$\text{COND/AUX} - \text{NON-ARG.DAT/REFL} - \text{DAT} - \text{ACC/GEN}$

Section 6.6 has presented different groups of approaches to clitic ordering and concluded that syntax on its own might be able to handle parts of the templatic pattern, but not deviations from it as represented for example by person-case constraints (PCC’s). Also, the fact that reflexives occupy their own slot even when they represent accusative or dative arguments is not derivable from syntax. Such points lead Anderson (1996) to assume a purely morphological model of clitic ordering, in which each clitic type has its own  $\text{EDGEMOST}$  constraint. The relative ranking of these constraints then produces the templatic order.

Following these assumptions, for Czech, Richardson (1997) proposes the constraints and their ranking as given in 184 (repeated from section 6.6.3).<sup>25</sup> Note that these are only more specific versions of the general  $\text{EDGEMOST}(\text{Cl}, \text{l}, \text{MaxTP})$  constraint we have employed so far. If we split it up and rank the resulting specific constraints as proposed by Richardson (1997), we

<sup>25</sup>I have slightly altered Richardson’s (1997) wording of the constraints by replacing *IP* with *MaxTP*, and *Past/Cond* with *Aux*.

get a group of constraints which captures the placement of the cluster in the clause as well as the ordering within that cluster at once.<sup>26</sup>

- (184)  $\text{EDGEMOST}(\text{Aux},l,\text{MaxTP}) \gg \text{EDGEMOST}(\text{Ref},l,\text{MaxTP})$   
 $\gg \text{EDGEMOST}(\text{Dat},l,\text{MaxTP}) \gg \text{EDGEMOST}(\text{Acc},l,\text{MaxTP})$

(cf. Richardson, 1997, 146-8)

In addition to this basic template, Czech also displays PCC effects, as was shown in section 4.4. Sturgeon et al. (2012) investigate these empirically and propose a constraint-based analysis which features two groups of EDGEMOST constraints, defined in 185. The first group,  $\text{EDGEMOST}(\pi)$ , captures the ranking of pronouns with different person specifications; the second group,  $\text{EDGEMOST}(\delta)$ , captures the ranking of pronominal case, thus it encompasses  $\text{EDGEMOST}(\text{Dat},l,\text{MaxTP})$  and  $\text{EDGEMOST}(\text{Acc},l,\text{MaxTP})$  introduced above. Note that both constraint types are only relevant for non-reflexive pronouns, since Czech reflexives do not exhibit effects of person or case.

- (185) a.  $\text{EDGEMOST}(\pi)$ : The clitic with person specification  $\pi$  is at the left edge of the clitic cluster.  
 b.  $\text{EDGEMOST}(\delta)$ : The clitic bearing the case  $\delta$  is at the left edge of the clitic cluster.

(cf. Sturgeon et al., 2012, 11-2)

PCC effects now arise due to the fact that observing the person hierarchy is more important than observing the case hierarchy, which leads to the overall ranking of person- and case-specific constraints as given in 186 (cf. Sturgeon et al., 2012, 12). I also include the constraints referring to auxiliaries and reflexives, both ranked above the five constraints that refer to non-reflexive pronouns, thus reflecting that the latter appear last in the cluster. For the sake of coherence, all constraints refer to the left edge of the maximal TP, instead of the left edge of the cluster as in 185.

- (186) Ranking of all constraints relevant for clitic ordering:  
 $\text{EDGEMOST}(\text{Aux},l,\text{MaxTP}) \gg \text{EDGEMOST}(\text{Ref},l,\text{MaxTP})$   
 $\gg \text{EDGEMOST}(1,l,\text{MaxTP}) \gg \text{EDGEMOST}(2,l,\text{MaxTP})$   
 $\gg \text{EDGEMOST}(3,l,\text{MaxTP}) \gg \text{EDGEMOST}(\text{Dat},l,\text{MaxTP})$   
 $\gg \text{EDGEMOST}(\text{Akk},l,\text{MaxTP})$

Let us now see how these constraints handle templatic effects, based on example 187, first introduced in section 4.3. Here, the PCC is not relevant, because there is only one non-reflexive pronominal clitic. As the tableau in 8.15 shows, candidates where the auxiliary is not first in the cluster (c and d) already fail at the highest constraint. The decision between the remaining

<sup>26</sup>Since Richardson (1997) assumes that free (or: ethical) datives occupy the same slot as indirect objects, he does not capture them separately in his constraints (cf. Richardson, 1997, 144). Yet, we could easily include them in accordance with Zwicky's (1977) template by introducing a constraint for free datives between the one for reflexives and the one for datives. For the sake of simplicity, I will ignore free datives in this discussion, though.

two candidates (a and b), which both have the auxiliary in first position, then falls to the second constraint. When the reflexive is second in the cluster, only one additional violation of *EDGEMOST(Refl)* arises, but when it is third, this causes two violations of this constraint, so candidate b loses to candidate a, which is in consequence the optimal one.

- (187) *Představila jsem se mu včera.*  
 introduce.PTCP AUX.1SG REFL.ACC him.DAT yesterday  
 ‘I introduced myself to him yesterday.’

(cf. Franks, 1998, 42)

	AUX	REFL	1	2	3	DAT	ACC
☞ a. [TP V jsem se mu]	*	**			***	***	
b. [TP V jsem mu se]	*	***!			**	**	
c. [TP V se mu jsem]	**!*	*			**	**	
d. [TP V mu jsem se]	**!	***			*	*	

Table 8.15: OT tableau illustrating clitic ordering

Next we take a look at the derivation of PCC effects, i. e. the reordering of two pronominal clitics. Example 188, repeated from section 4.4, illustrates such a case: the order of dative and accusative clitics is reversed, because otherwise the weak PCC would be violated, as third person would precede second. This is directly captured by our constraint ranking, as the evaluation in the tableau in 8.16 shows: since the person hierarchy is ranked above the case hierarchy, person effects can lead to a reversed order of pronominal clitics. The alternative ordering, *mu tě* (candidate b), is rejected due to the relative height of the constraint *EDGEMOST(2)*.

- (188) ... *já tě mu nedám!*  
 1SG you.ACC him.DAT NEG:give.1SG  
 ‘I won’t give you to him!’

(cf. Sturgeon et al., 2012, 7)

	AUX	REFL	1	2	3	DAT	ACC
☞ a. [TP NP tě mu]				*	**	**	*
b. [TP NP mu tě]				**!	*	*	**

Table 8.16: OT tableau illustrating clitic ordering with PCC effect

We thus see that Czech clitic ordering as well as deviations from that pattern due to the PCC are comprehensively captured in the presented approach. However, section 4.4 has also shown that not all speakers adhere to the PCC. How can these alternative grammars be covered in the analysis? The answer is simple: those speakers do not have the full set of constraints as in 186,

but only the reduced set, as already defined in 184, without the person-specific constraints. This also means that in Czech, the *person-case* constraint actually has nothing to do with *case* – the only relevant feature is *person*. This has also been emphasised by Franks (2017), who consequently reformulates the PCC’s and renames them *person-ordering constraints* (cf. Franks, 2017, 265, and section 6.6.2 of this thesis). Both his corrections as well as the present account thus also contribute to a clearer understanding of the nature of these restrictions on clitic ordering.<sup>27</sup>

An example of a sentence where the speaker ignores the PCC and solely adheres to the templatic order is given in 189, also repeated from section 4.4. Since person does not play a role in these grammars, the candidate where the dative precedes the accusative prevails over the re-ordered one, due to the ranking of  $\text{EDGEMOST}(\text{Dat})$  over  $\text{EDGEMOST}(\text{Acc})$ . The tableau in 8.17 illustrates this.

(189) *Dám mu tě do pytle.*  
 give.1SG him.DAT you.ACC into bag

‘I’ll give you to him in a bag.’

(cf. Sturgeon et al., 2012, 7)

	AUX	REFL	DAT	ACC
a. [TP V tě mu]			**!	*
☞ b. [TP V mu tě]			*	**

Table 8.17: OT tableau illustrating clitic ordering without PCC effect

There are two potential objections to the present analysis that I still need to address. The first might be that in the no-PCC grammar, three constraints have simply been left out. This appears to be a violation of the OT principle that all constraints are universal, and that grammars only differ with respect to the ranking of these constraints. Yet, this would be a misinterpretation of the nature of  $\text{EDGEMOST}$ , which is a subtype of the Generalised Alignment constraint family. As explained in section 8.1.2, the arguments of alignment constraints (i. e. categories and their edges) are inserted language-specifically by the learner, based on the data they observe; what is universal about alignment is the general constraint format (cf. McCarthy & Prince, 1993, 7). Clitics are a good example for why this assumption is plausible: In a language that has no clitics, why should a learner formulate clitic-specific constraints? And in a language that *does* have clitics, but where their person specification is irrelevant, why should one assume constraints that refer to person values of the clitics? Thus, assuming a smaller clitic constraint set for speakers that do not adhere to the PCC is well-grounded and entirely compatible with basic assumptions of OT.

<sup>27</sup>Whilst I will follow Franks’s (2017) proposal in calling these restrictions *person-ordering constraints* in the remainder of this thesis, I will adhere to the traditional abbreviation *PCC*, as it is a long-established name for this group of phenomena.

A second point of criticism is expressed by Franks (2017): he argues that OT approaches “simply recast the template as a stipulated ranking of constraints targeting specific lexical items”, instead of trying “to derive clitic ordering from more general aspects of language structure” (Franks, 2017, 259-60). And indeed, one can argue that at least the ordering of auxiliaries, dative pronouns and accusative pronouns – captured by the ranking  $\text{EDGEMOST}(\text{Aux}) \gg \text{EDGEMOST}(\text{Dat}) \gg \text{EDGEMOST}(\text{Akk})$  – reflects these elements’ respective base positions in the syntax, and so we are missing out on an opportunity to include independent motivation for clitic ordering into our analysis.

OT allows us to incorporate such insights through order preservation constraints (cf. Broekhuis, 2008, 46). I therefore propose a constraint which ensures that the basic order of clitic elements is reflected in the ordering within the clitic cluster. The constraint, which I will simply call  $\text{CLITICORDER}$ , is defined in 190. However, we still need to capture the fact that auxiliaries precede reflexives, whilst pronouns follow them, and we need to allow for PCC effects. This means that we can only replace  $\text{EDGEMOST}(\text{Dat})$  and  $\text{EDGEMOST}(\text{Acc})$  with  $\text{EDGEMOST}(\text{Pron})$ . The other constraints must remain in place, and our new order preservation constraint must be ranked below them. The complete set of constraints and their ranking is thus as given in 191.<sup>28</sup>

(190)  $\text{CLITICORDER}$  = If the foot of the chain of  $\alpha$  precedes the foot of the chain of  $\beta$ , then the head of the chain of  $\alpha$  precedes the head of the chain of  $\beta$ .

Where:  $\alpha, \beta \in \{x \mid x \text{ is a lexical clitic}\}$ <sup>29</sup>

(191) Ranking of all constraints relevant for clitic ordering (new):

$\text{EDGEMOST}(\text{Aux}, l, \text{MaxTP}) \gg \text{EDGEMOST}(\text{Refl}, l, \text{MaxTP})$

$\gg \text{EDGEMOST}(1, l, \text{MaxTP}) \gg \text{EDGEMOST}(2, l, \text{MaxTP})$

$\gg \text{EDGEMOST}(3, l, \text{MaxTP})$

$\gg \text{EDGEMOST}(\text{Pron}, l, \text{MaxTP}); \text{CLITICORDER}$

I conclude that our constraint-based analysis also permits us to capture the effects of syntactic base positions on the order within the cluster. Yet, there are many aspects of this ordering that are idiosyncratic, and cannot be explained through syntactic principles. Therefore, an OT account is best-suited to incorporate all facets that determine the precedence relations in the cluster in a precise and uniform way.

## 8.6 Notes on a factorial typology

An important aspect of Optimality Theory is its capacity to derive typological predictions from the re-ranking of constraints. Whereas this thesis is mainly concerned with one language, Czech, it is nonetheless interesting to investigate the predictions made by the proposed constraints, not

<sup>28</sup>For Czech, there is indeed evidence that suggests that the base position of the dative is higher than that of the accusative argument (cf. Sturgeon et al., 2012, 11). On the other hand, Dvořák (2010) argues that this is only true for one group of ditransitive verbs, whereas in a second group, the two objects are base-generated in reverse order. If she is correct,  $\text{CLITICORDER}$  would not apply, and instead the ordering of clitics in Czech would be fully templatic.

<sup>29</sup>I. e. it has the feature [+clitic], cf. section 8.2.1.

only to explore plausible typological patterns, but also to further investigate the functions and interactions of these constraints.

The three central constraints  $\text{NONINITIAL}(\text{Cl}, \text{CP})$ ,  $\text{EDGEMOST}(\text{Cl}, \text{l}, \text{MaxTP})$ , and  $\text{INTEGRITY}(\text{XP})$  that we have employed for the current analysis can be ranked in six ( $3 \times 2 \times 1$ ) different ways,<sup>30</sup> listed in table 8.18. 2D placement arises whenever  $\text{EDGEMOST}$  is ranked below  $\text{NONINITIAL}$  and  $\text{INTEGRITY}$ . The ordering of these two latter constraints does not make a difference, because they never conflict. 2W placement results when  $\text{EDGEMOST}$  is ranked below  $\text{NONINITIAL}$ , but  $\text{INTEGRITY}$  is lower than  $\text{EDGEMOST}$  – in this case, the clitics may split a constituent.<sup>31</sup>

The remaining three cases are not likely to be assumed by a language learner. In all these,  $\text{EDGEMOST}$  is ranked above  $\text{NONINITIAL}$ . Thus, the clitics simply always end up at the edge specified by  $\text{EDGEMOST}$  – whether they then violate  $\text{NONINITIAL}$  or not does not matter. Also, when  $\text{EDGEMOST}$  is satisfied, the clitic automatically does not occur inside a constituent, so  $\text{INTEGRITY}(\text{XP})$  is not relevant either. If confronted with such 1P data, the learner will consequently simply assume a (non-gradient)  $\text{EDGEMOST}$  constraint, and then move on (cf. Billings, 2002, 96).

$\text{NONINITIAL}$	>>	$\text{INTEGRITY}(\text{XP})$	>>	$\text{EDGEMOST}$	→ 2D
$\text{INTEGRITY}(\text{XP})$	>>	$\text{NONINITIAL}$	>>	$\text{EDGEMOST}$	
$\text{NONINITIAL}$	>>	$\text{EDGEMOST}$	>>	$\text{INTEGRITY}(\text{XP})$	→ 2W
$\text{EDGEMOST}$	>>	$\text{NONINITIAL}$	>>	$\text{INTEGRITY}(\text{XP})$	→ 1P
$\text{EDGEMOST}$	>>	$\text{INTEGRITY}(\text{XP})$	>>	$\text{NONINITIAL}$	
$\text{INTEGRITY}(\text{XP})$	>>	$\text{EDGEMOST}$	>>	$\text{NONINITIAL}$	

Table 8.18: Possible constraint rankings and their outcome

The reason for this asymmetrical behaviour of the possible orderings of  $\text{NONINITIAL}$  and  $\text{EDGEMOST}$  is that the former constraint is more tolerant:  $\text{NONINITIAL}$  excludes only one option, namely the one where an element is initial in a specified domain.  $\text{EDGEMOST}$ , on the other hand, is much more restrictive, because it *requires* exactly one specific option; thus, if it is ranked high, it selects this option and the other constraint has no effect any more.

Thus, although the three constraints that are central to the present analysis can be combined in six possible ways, they only yield three different patterns, and only two of these are actually best captured with these three constraints – therefore, the only re-ranking that yields an interesting contrast is the positioning of  $\text{INTEGRITY}(\text{XP})$  with respect to  $\text{NONINITIAL}$  >>  $\text{EDGEMOST}$ ; this is the contrast between 2D and 2W placement. Further variation with respect to clitic positioning is then due to how exactly these constraints' arguments are filled in, i. e. what their domains are.

<sup>30</sup>I am not discussing the more general constraints  $\text{MATCHCLAUSE}$ ,  $\text{STAY}$ , and  $\text{THETA}$  here.

<sup>31</sup>Billings (2002) develops a factorial typology with similar conclusions based on his constraints, which I have presented in section 8.4.2.



## 8.7 Conclusions

This chapter has shown how the optimality-theoretic concept of violable, ranked constraints can be used to model clitic 2P placement in Czech, both within simple clauses and within complex clauses, where clitic climbing occurs in the absence of a CP layer. Drawing on work by Richardson (1997), Grimshaw (1997) and Anderson (2000), placement of Czech clitics after the first constituent was modelled using just five constraints. A sixth constraint, CP-BARRIER, implements the insights of phase impenetrability and thereby derives the ban on climbing out of CP. A seventh constraint, MATCHCLAUSE, was employed to model the flexible prosodic attachment of Czech clitics (if they undergo any prosodic incorporation at all).

I have also argued that whereas the NONINITIAL constraint cannot be traced back to a requirement for a host to the left of the clitic in Czech, it might, as proposed by Dotlačil (2007), actually be the product of a more general ban on information-structurally weak elements from the left edge of CP. This means that what causes the unusual syntactic behaviour of Czech clitics is not their prosody (at least not directly), but their inability to fulfil prominent information-structural roles. It also means that we are left with only one clitic-specific constraint; all other constraints reflect more general principles of grammar.

The final formulation of the constraints is thus as given in 192a-f; their ranking for Czech, as far as it can be determined based on the data discussed here, must be as in 192g. Note that the original formulation of EDGEMOST and NONINITIAL by Anderson (1992) and Richardson (1997) obscures the fact that the constraints are part of the well-established constraint family of generalised alignment. On the other hand, they are less technical and easier to read, so I will keep them as labels for the alignment constraints as specified in 192a-b.

(192) Relevant constraints and their ranking for Czech:

- a.  $\text{NONINITIAL}(\text{Cl}, \text{CP}) = * \text{ALIGN}(\text{Cl}, \text{l}, \text{CP}, \text{l})$   
= A clitic must not occur at the left edge of CP.
- b.  $\text{EDGEMOST}(\text{Cl}, \text{l}, \text{MaxTP}) = \text{ALIGN}(\text{Cl}, \text{l}, \text{MaxTP}, \text{l})$   
= A clitic must occur as close to the left edge of the highest TP in its clause complex as possible.
- c.  $\text{INTEGRITY}(\text{XP}) =$  An XP may not be interrupted by phonological material that is not part of XP.
- d.  $\text{STAY} =$  Trace is not allowed.
- e.  $\text{THETA} =$  An argument must be theta-marked.
- f.  $\text{CP-BARRIER} = * \dots X_i \dots [ \text{CP} \dots [ \text{YP } X_i \dots ] \dots ]$   
= An element which has a copy within CP must also have a copy at the left edge of that CP.
- g.  $\text{MATCHCLAUSE} =$  A clause in syntactic constituent structure must be matched by a corresponding  $\iota$ -phrase in prosodic representation.
- h. Ranking:  
 $\text{THETA}; \text{NONINITIAL}(\text{Cl}, \text{CP}); \text{CP-BARRIER};$

MATCHCLAUSE; INTEGRITY(XP)  
>> EDGEMOST(Cl,l,MaxTP) >> STAY

We have also seen that by splitting up EDGEMOST into more specific constraints that refer to subclasses of clitics, the internal ordering of the cluster can be captured. Whereas most of this ordering is idiosyncratic, an additional order preservation constraint, CLITICORDER, can integrate the effects of syntactic hierarchy into our analysis.

Before turning to the analysis of “late” clitic placement, i. e. 3P, I would like to comment on Grimshaw’s (1997) claim, mentioned in section 8.1.3, that syntactic constraints must not refer to labels such as CP, TP etc. Her analysis of auxiliary inversion and *do*-support in English indeed does not require any labelling of syntactic structure. Yet, the preceding sections of the present thesis have shown that clitics in Czech gravitate towards *specific* syntactic edges, whilst avoiding others. Therefore, I do not see how their behaviour could be modelled without reference to labels.

## Chapter 9

# Analysing 3P Placement and Topicalisation

The previous chapter has shown how clitics come to appear second within a clause, following whatever constituent comes first. With “delayed” clitic placement, i. e. 3P, there are however restrictions on the element preceding the clitic, and the present chapter investigates these restrictions. In section 5.1.2, I have already pointed out that clitic-third in Czech is always a consequence of topicalisation in embedded sentences: fronted elements occupy the position between the complementiser or interrogative phrase and the clitic. Section 9.1 will now demonstrate that in our constraint-based analysis of Czech clitic placement, cases of single-constituent topicalisation can be included free of charge. Then I turn to a more complex issue in section 9.2: the positioning of verbs and entire VPs in front of the clitic. I will show how we can capture some puzzling asymmetries with respect to verbal and VP topicalisation with only a few independently motivated constraints. Section 9.3 then summarises and concludes this chapter.

### 9.1 Argument topicalisation

The following subsections examine how argument topicalisation can be captured within a constraint-based approach. For the sake of simplicity, I will only look at NP arguments – yet, the same analysis can most likely also capture other types of arguments, for example PPs. In addition, non-argumental elements such as adverbials can also undergo topicalisation, as was shown in section 5.1.2. I assume that the same constraints apply in these cases, too.

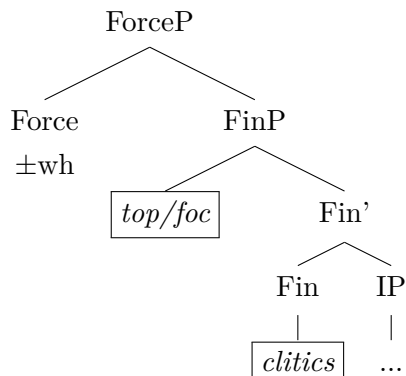
#### 9.1.1 The syntax of topicalisation

As sketched in section 6.4, Richardson (1997) also offers an analysis of optional 3P. Capturing true optionality within the OT apparatus employed here would be difficult; however, in the case of Czech 3P, optionality is only apparent – whether a clitic occurs in third or second position in an embedded clause depends on whether a constituent is topicalised (i. e. fronted to the left periphery) or not. Before taking a closer look at how this analysis works, we must clarify our notions of topicalisation.

To begin, I must point out that not only topics, but also left-peripheral foci can participate in 3P configurations, as shown in section 5.1.2. However, it is generally assumed that topics and left-peripheral foci in Czech occupy the same syntactic position, since they are in complementary distribution (cf. Lenertová, 2001, 299). I will thus not differentiate between the two in the following analysis. Instead, since foci can be realised at the right clause edge as well as in the left periphery (cf. section 3.5), I use topics for all judgement tasks presented later in this chapter, since a topic’s canonical position is unambiguously at the left clause edge.

Regarding the syntactic position of topicalised phrases, as already sketched in section 6.4.3, Lenertová (2001) argues against a split left periphery as postulated by Rizzi (1997) for Czech. First, because the post-complementiser position can also contain a range of other constituent types, such as foci but also information-structurally unmarked elements (cf. Fanselow & Lenertová, 2011, 173), which arguably would not fit into a specific topic projection. Second, in contrast to Italian, multiple topics as well as a combination of fronted topic and fronted focus are not possible in Czech. Instead, Lenertová (2001) proposes that all kinds of “topicalised” elements move to the specifier of FinP (cf. Lenertová, 2001, 299-302). The structure she assumes is shown in 193.

(193) Structure of the Czech left periphery as assumed by Lenertová (2001):



Whilst the precise nature of the projection hosting topics and foci is not crucial for the present purposes, I will yet follow Lenertová’s (2001) basic assumption that topicalised elements – meaning all fronted elements, thus also including fronted foci etc. – occupy a specifier position directly below C (or Force), and above IP/TP (see also Junghanns & Lenertová, 2007, 236), thus these elements are immediately adjacent to the complementiser. As has already become clear, I will not follow her analysis with respect to clitic positioning.

### 9.1.2 Capturing 3P after topics

I now turn to the analysis of clitic placement with respect to topics. When a topicalised element is present, as for example *Petr* in 194a, the clitic can never precede it because the topic position is above TP – thus  $\text{EDGE}_{\text{MOST}}(\text{CL}, \text{l}, \text{MaxTP})$  is unviolated in the post-topic position, and so is  $\text{NONINITIAL}(\text{CL}, \text{CP})$ . This is illustrated in the tableau in 9.1.

- (194) a. ... *že Petr si koupil knihy.*  
 COMP Petr REFL buy.PTCP books  
 ‘... that Petr bought books (for himself)?’
- b. ... *že si Petr koupil knihy.*  
 COMP REFL Petr buy.PTCP books

	NONINI(CP)	EDGMO(MaxTP)
a. [CP Comp [FinP <b>cl</b> NP <sub>Top</sub> [TP V ]]]		*!
☞ b. [CP Comp [FinP NP <sub>Top</sub> [TP <b>cl</b> V ]]]		
c. [CP Comp [FinP NP <sub>Top</sub> [TP V <b>cl</b> ]]]		*!

Table 9.1: OT tableau illustrating clitic-third after a topic

On the other hand, embedded sentences without a fronted topic, as exemplified by 194b, do not display clitic-third. Here as well, the clitic can find an ideal position where neither NONINITIAL nor EDGEMOST are violated, due to the presence of the complementiser; see the tableau in 9.2. Put differently: The presence of the complementiser makes it possible for the clitic to be at the left edge of TP without violating NONINITIAL. This does not change with the presence or absence of a topic.

	NONINI(CP)	EDGMO(MaxTP)
☞ a. [CP Comp [TP <b>cl</b> NP V ]]		
b. [CP Comp [TP NP <b>cl</b> V ]]		*!
c. [CP Comp [TP NP V <b>cl</b> ]]		*!*

Table 9.2: OT tableau illustrating clitic-second in the absence of a topic

The same effect arises with interrogative expressions: just as a complementiser in the C-head, an interrogative phrase in the specifier of C allows the clitic to occupy TP’s left edge without being clause-initial. This is illustrated in example 195, repeated from section 5.1.2: the clitic typically directly follows the interrogative phrase, but a topic or, in this case, a left-peripheral focus, can appear between the two. Constraint evaluation thus proceeds in parallel to what is shown in the tableaux in 9.1 and 9.2. Note that the conjunction *a* ‘and’ is extra-clausal and thus never affects clitic placement, as discussed in section 5.1.1.

- (195) a. *A co EMA by na to řekla?*  
 and what Ema COND.3 to this said  
 ‘And what would EMA say to that?’

(cf. Lenertová, 2001, 299)

- b. *A co by Ema na to řekla?*  
 and what COND.3 Ema to this said  
 ‘And what would Ema say to that?’

We can also analyse multiple questions, where only certain information-structural conditions allow both interrogative phrases to precede the clitics, in the same way. An example, also repeated from section 5.1.2, is provided by 196. I assume that the special single-pair reading of this question is triggered by the movement of the second interrogative phrase into the Fin projection.

- (196) ... *kdo koho si váží víc.*  
 who.NOM whom.ACC REFL.DAT value.3SG more  
 ‘... who values whom more.’

(cf. Meyer, 2004, 82)

### 9.1.3 Split topicalisation

As we have seen, we can describe Czech clitic placement as 2D, meaning that the clitics appear after the first entire constituent of the clause. Therefore, I have argued for a high ranking of INTEGRITY(XP) for Czech in section 8.2.2. Yet, one phenomenon might cast doubt on this generalisation: cases of split topicalisation as described, for example, by Lenertová (2001). She provides example 197, first introduced in section 5.1.1, and explains that such constructions involve an element which carries the main stress of a focused constituent (here: *krásnou KNÍŽKU* ‘nice BOOK’). This stressed element is fronted to initial position, thereby being separated from the rest of its constituent; the result looks like a case of 2W placement. She notes that this is found “especially in spoken Czech” (Lenertová, 2001, 299).

- (197) *KNÍŽKU jsem vám přinesla krásnou.*  
 book.ACC AUX.1SG you.DAT bring.PTCP nice.ACC  
 ‘I’ve brought you a nice book!’

(cf. Lenertová, 2001, 299)

In order to verify the acceptability of such constructions and to see whether they can also participate in 3P constructions, i. e. in embedded clauses, I carried out a small acceptability judgement study among eight native speakers. The results can be seen in figure 9.1; the OSF link to the materials and results can be found in appendix B. Examples for both the control constructions with neutral word order and the split sentences with topicalised noun are given in 198.

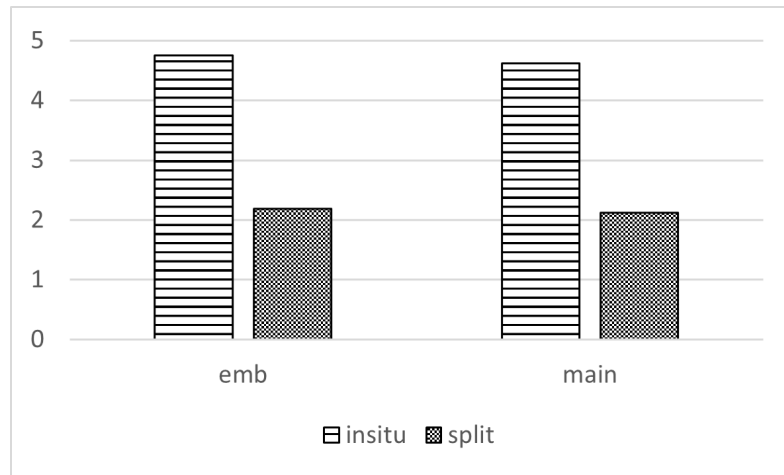


Figure 9.1: Acceptability judgements for split topicalisation: in-situ order vs. split with fronted noun (N=8)

- (198) a. *V parku jsem viděla strakatou krávu!*  
 in park AUX.1SG see.PTCP pied.ACC cow.ACC  
 ‘I’ve seen a pied cow in the park!’
- b. *Krávu jsem v parku viděla strakatou!*  
 cow.ACC AUX.1SG in park see.PTCP pied.ACC
- c. *Chtěla bych oznámit, že jsem v parku viděla strakatou krávu.*  
 want.PTCP COND.1SG report.INF COMP AUX.1SG in park see.PTCP pied.ACC  
 cow.ACC  
 ‘I’d like to report that I’ve seen a pied cow in the park.’
- d. *Chtěla bych oznámit, že krávu jsem v parku viděla strakatou.*  
 want.PTCP COND.1SG report.INF COMP cow.ACC AUX.1SG in park see.PTCP  
 pied.ACC

Surprisingly, the split topicalisation sentences received very low ratings, as figure 9.1 shows. Sentence 197, which was included as part of the filler set, fared slightly better, with an average rating of 3.0 and two speakers giving it a four or higher. It is possible that we are again facing the problem that speakers tend to give colloquial language a low rating because they are used to evaluate language based on normative rules. Yet, the results also show that some speakers accept these constructions, and that they can appear in main and embedded clauses, thus the second element in 3P constructions can be a partial constituent.<sup>1</sup>

<sup>1</sup>It is also possible that the result is partly coincidental, because the number of participant was very low: only eight speakers filled in the questionnaire. This was partly on purpose, because, as reported below, they were also asked to provide contexts for the sentences in free text, which entails a much more laborious evaluation.

In addition to the acceptability judgements, I also asked the participants to provide plausible contexts for main clause split topicalisation sentences like 197 and 198b. The results were largely unanimous: for example, most speakers formulated 199, or something very similar, as a context for sentence 198b. This shows that they perceive the sentences as most adequate in a context where the fronted noun is given and the sentence-final adjective is in focus.

- (199) *Jakou krávu jsi viděla v parku?*  
 what.ACC cow.ACC AUX.2SG see.PTCP in park  
 ‘What (kind of) cow did you see in the park?’

Regarding the question of potential 2W placement in Czech, it does not appear plausible to assume that the clitics split up an intact phrase in these examples. As Lenertová’s (2001) description of the phenomenon already implies, the separation of the noun from its adjective takes place for information structural or prosodic reasons, not due to clitic placement. This is also evident from the fact that not only the clitic intervenes between the two parts of the NP, but also other material. It thus appears that some information-structural or prosodic constraint is ranked higher than INTEGRITY, but that clitic-specific EDGEMOST constraints are still ranked lower. I will not attempt to define this high-ranked constraint here, as this would require a much deeper investigation of split topicalisation. What we can say at this point is that both in 2P and 3P configurations, clitics can appear after partial constituents if these are split up for independent reasons; thus, 2W clitic placement does not exist in Czech.

#### 9.1.4 Summary

The previous subsections have shown that topicalisation in the broader sense, i. e. the movement of (parts of) constituents to the left periphery, naturally yields third-position clitic placement in embedded clauses, because the ideal position for a clitic is at the left edge of TP, and therefore below any constituent that is in the clause’s left periphery, and thus above TP.

## 9.2 Verbal and VP topicalisation

In section 5.1.2, an interesting asymmetry was mentioned: in clitic-second configurations, the element preceding the clitic can be phrasal or a verbal head, such as an *l*-participle, but in clitic-third configurations, the element occupying the position between complementiser and clitic cannot be a participle. The relevant example given there, 65, is repeated here as 200. This observation deserves further investigation: Are all types of verb disallowed in this position? And what is the basis for this restriction? On the following pages, I will first draw a more fine-grained picture of which elements can occupy the position between complementiser and clitic, and which ones cannot. Then I will present an analysis to capture these patterns.

- (200) \*... *že nedal by mu to.*  
 COMP NEG.give.PTCP COND.3 him.DAT it  
 ‘...that he would not give it to him.’



(cf. Veselovská, 1995, 110)

### 9.2.1 Literature and corpus data

Regarding *l*-participles between complementiser and clitic, in the SYN2015 corpus, a few instances of this structure can be found. Three examples are given in 201, with *zemřela* in the relevant position in 201a, *opustil* in 201b, and *věřila* in 201c. Yet, all were judged either ungrammatical or archaic-poetic by my consultants<sup>2</sup> and are therefore marked with a question mark below. Sentence 201c was additionally used as a filler in an acceptability judgement task with 64 participants, where it received a mean rating of 3.93 on a seven-point scale where 7 was the best possible value.

- (201) a. <sup>?</sup>*Věřte mi, matko má, že zemřela bych radši, než takhle chodit ven.*  
 believe.IMP me mother my COMP die.PTCP COND.1SG rather than thus go.INF  
 out  
 ‘Believe me, my mother, that I would rather die than go out like this.’
- b. <sup>?</sup>*Radost nemám, že opustil jsem služby tvé kvůli ženám.*  
 joy NEG:have.1SG that left AUX.1SG services your because.of women  
 ‘I am not glad that I left your services for the sake of women.’
- c. <sup>?</sup>*Má chyba byla, že věřila jsem v krásné sny.*  
 my mistake be.PTCP COMP believe.PTCP AUX.1SG in beautiful dreams  
 ‘My mistake was that I believed in beautiful dreams.’

(SYN2015)

Infinitives appearing between complementiser and clitic could also be found in the corpus in very low numbers; for example *pamatovat* in 202a and *dýchat* in 202b. These structures were all fully accepted by my consultants; the fronting of the infinitive is sometimes perceived as emphasis on this element. Sentence 202a was also used as a filler in an acceptability judgement task with 64 participants, where it received an average rating of 5.41 on a seven-point scale where 7 was the best possible rating.

- (202) a. *Věděla, že pamatovat si bude právě tohle.*  
 know.PTCP COMP remember.INF REFL will.3SG just that  
 ‘She knew that she would remember just that.’

<sup>2</sup>These sentences also all contain other features that are archaic or poetic, namely the form of the possessive pronouns (*má* in 201a and 201c and *tvé* in 201b) and the order of noun and possessive pronoun in 201a and 201b (*matko má* and *služby tvé*). Also note that all examples are with clitic auxiliaries; with clitic pronouns, none were found.

- b. *Nezapomeňte, že dýchat byste měli “do břicha”.*  
 NEG.forget.IMP COMP breathe.INF COND.2PL have.PTCP into belly

‘Don’t forget that you should breathe “into your belly”’

(SYN2015)

According to Avgustinova and Oliva (1995), in main clauses, partially fronted VPs can occupy the position before the cluster, but only when the verb is either an infinitive, as in 203a, or a passive participle, as in 203b. The (past) *l*-participle, in contrast, can only occupy the position before the clitic alone, not together with its object, regardless of whether the clitic auxiliary is overt or not, as 203c and 203d show. Sentence 203d was also included as a filler in an acceptability judgement task with 64 participants, and received an average rating of 2.07 on a seven-point scale on which 7 was the best possible rating.

- (203) a. *Posílat dopisy ti budu pravidelně každý týden.*  
 send.INF letters you.DAT will.1SG regularly every week  
 ‘I shall send letters to you regularly every week.’
- b. *Nucen k takovým činům jste nebyl nikdo.*  
 force.PTCP to such actions AUX.2PL NEG:be.PTCP no-one  
 ‘None of you was forced to such actions.’
- c. *\*Posílal dopisy jsem ti pravidelně každý týden.*  
 send.PTCP letters AUX.1SG you.DAT regularly every week  
 Intended: ‘I sent letters to you regularly every week.’
- d. *\*Kupoval knihy si včera jenom Petr.*  
 buy.PTCP books REFL yesterday only Petr  
 ‘Only Petr bought books (for himself) yesterday.’

(cf. Avgustinova & Oliva, 1995, 27-8)

The corpus only contains very few instances of VPs, i. e. verb plus object, between complementiser and clitic. An example is 204: here, *snižovat cenu* is the infinitival, clausal subject of the verb *dařit*. The sentence was accepted by most of my informants, but some found it slightly dubious. When used as a filler sentence in an acceptability judgement task with 64 participants, it received an average rating of 5.73 on a seven-point scale on which 7 was the best possible rating.

- (204) *Myslíme, že snižovat cenu se bude dařit mnohem rychleji, než dnes tvrdí skeptici.*  
 think.1PL COMP decrease.INF price REFL will.3SG fare.INF much faster than  
 today claim.3PL sceptics

‘We think that lowering the price will go much faster than sceptics say today.’

(SYN2015)

It thus appears that in modern Czech, the position between complementiser and clitics can be occupied by an infinitive, but not by an *l*-participle. VP fronting to that position also occurs, but is probably highly restricted. Generally, the very low frequency of all these constructions in the corpus does not permit us to draw conclusions about their acceptability for speakers, and a more systematic investigation into speakers' judgements is advisable. Therefore, in the following subsection, I present such acceptability judgement data to clarify the status of these constructions.

### 9.2.2 Judgement data

In an acceptability judgement study, I investigated the possibility of fronting different types of verbs with or without an argument NP into the pre-clitic position. 64 monolingual native speakers of Czech who currently live in the Czech republic and use Czech as their primary language were recruited via *Prolific*. Three participants had to be excluded from the analysis because their response behaviour in the test items suggested that they had not understood the task. The experiment had a 2x4x2 design; 16 different items were constructed, which the participants saw in a pseudo-randomised, Latin-square distribution. The design is schematised in table 9.3, with reference to example sentences (item 1) that are given in 205-212. The lexical clitic in all sentences was the indirect object pronoun *mu*; in order to make sure that the verbs allow for VP fronting with the direct object in general, they were chosen from the list of Czech “dat-acc verbs” (in contrast to “acc-dat verbs” with a low dative) in Dvořák (2010).

	embedded clause		main clause	
	V	V+NP	V	V+NP
finite verb	205a	205b	209a	209b
infinitive	206a	206b	210a	210b
passive ptcp	207a	207b	211a	211b
past ptcp	208a	208b	212a	212b

Table 9.3: Design with example references for fronting of different verb types with and without complements in embedded and main clauses

(205) Fronting of a finite verb in an embedded clause:

- a. *Myslím, že posílají mu dopisy každý týden.*  
 think.1SG COMP send.3PL him.DAT letters every week
- b. *Myslím, že posílají dopisy mu každý týden.*  
 think.1SG COMP send.3PL letters him.DAT every week

‘I think that they are sending him letters every week.’

(206) Fronting of an infinitive in an embedded clause:

- a. *Myslím, že posílat **mu** budou dopisy každý týden.*  
 think.1SG COMP send.INF him.DAT will.3PL letters every week
- b. *Myslím, že posílat dopisy **mu** budou každý týden.*  
 think.1SG COMP send.INF letters him.DAT will.3PL every week
- ‘I think that they will be sending him letters every week.’

(207) Fronting of a passive participle in an embedded clause:

- a. *Myslím, že posílány **mu** byly dopisy každý týden.*  
 think.1SG COMP send.PASS him.DAT were letters every week
- b. *Myslím, že posílány dopisy **mu** byly každý týden.*  
 think.1SG COMP send.PASS letters him.DAT were every week
- ‘I think that letters were being sent to him every week.’

(208) Fronting of an *l*-participle in an embedded clause:

- a. *Myslím, že posílali **mu** dopisy každý týden.*  
 think.1SG COMP send.PTCP him.DAT letters every week
- b. *Myslím, že posílali dopisy **mu** každý týden.*  
 think.1SG COMP send.PTCP letters him.DAT every week
- ‘I think that they were sending him letters every week.’

(209) Fronting of a finite verb in a main clause:

- a. *Posílají **mu** dopisy každý týden.*  
 send.3PL him.DAT letters every week
- b. *Posílají dopisy **mu** každý týden.*  
 send.3PL letters him.DAT every week
- ‘They are sending him letters every week.’

(210) Fronting of an infinitive in a main clause:

- a. *Posílat **mu** budou dopisy každý týden.*  
 send.INF him.DAT will.3PL letters every week
- b. *Posílat dopisy **mu** budou každý týden.*  
 send.INF letters him.DAT will.3PL every week
- ‘They will be sending him letters every week.’

(211) Fronting of a passive participle in a main clause:

a. *Posílány mu byly dopisy každý týden.*  
send.PASS him.DAT were letters every week

b. *Posílány dopisy mu byly každý týden.*  
send.PASS letters him.DAT were every week

‘Letters were being sent to him every week.’

(212) Fronting of an *l*-participle in a main clause:

a. *Posílali mu dopisy každý týden.*  
send.PTCP him.DAT letters every week

b. *Posílali dopisy mu každý týden.*  
send.PTCP letters him.DAT every week

‘They were sending him letters every week.’

The results are visualised in figure 9.2. They show that with respect to pre-clitic positions, the absence or presence of a complementiser as well as the verb type are relevant. The sentences with the fronted finite verb in a main clause served as a control condition since due to the lack of an overt subject, they display canonical word order and clitic placement. Arguably for the same reason, the fronted *l*-participle in a main clause fared similarly good. Infinitives and passives yielded only intermediate results. In embedded clauses, all verb types lead to decreased judgements when fronted into the position between complementiser and clitic, with only a slight advantage of infinitives and passives over finite verbs and *l*-participles.

With regard to fronting of verb and object (or, in the case of passive sentences, verb and subject), only infinitives yield positive results, albeit not as high as the control conditions. Overall, the data suggests that the four verb types can be reduced to three groups: whilst infinitives and passives each display their own behaviour, finite verbs and *l*-participles behave alike in all conditions.

### 9.2.3 Capturing the data: verb movement and uniformity

The results are mostly in line with what Avgustinova and Oliva (1995) report for the data in 203 above. Yet they show that with an unambiguous 2P clitic such as *mu*, fronting of the passive participle with an argument is not necessarily grammatical. On the other hand, it must be noted that the type of complement is not the same (subject in the experiment vs. PP in their data). To investigate this effect, a small follow-up study was conducted where a passive participle was fronted with three different types of complements/adjuncts: the structural subject, the agent (in instrumental case), and a prepositional adverbial phrase. The control clauses without VP fronting already had a rather low mean z-score of 0.05, possibly due to the fact that the morphological passive is quite rare in Czech, and often perceived as cumbersome – the passivisation strategy typically used does not involve a passive participle, but a reflexive verb (cf. Sussex & Cubberley, 2006, 368-9). It is therefore not clear whether the uniformly low

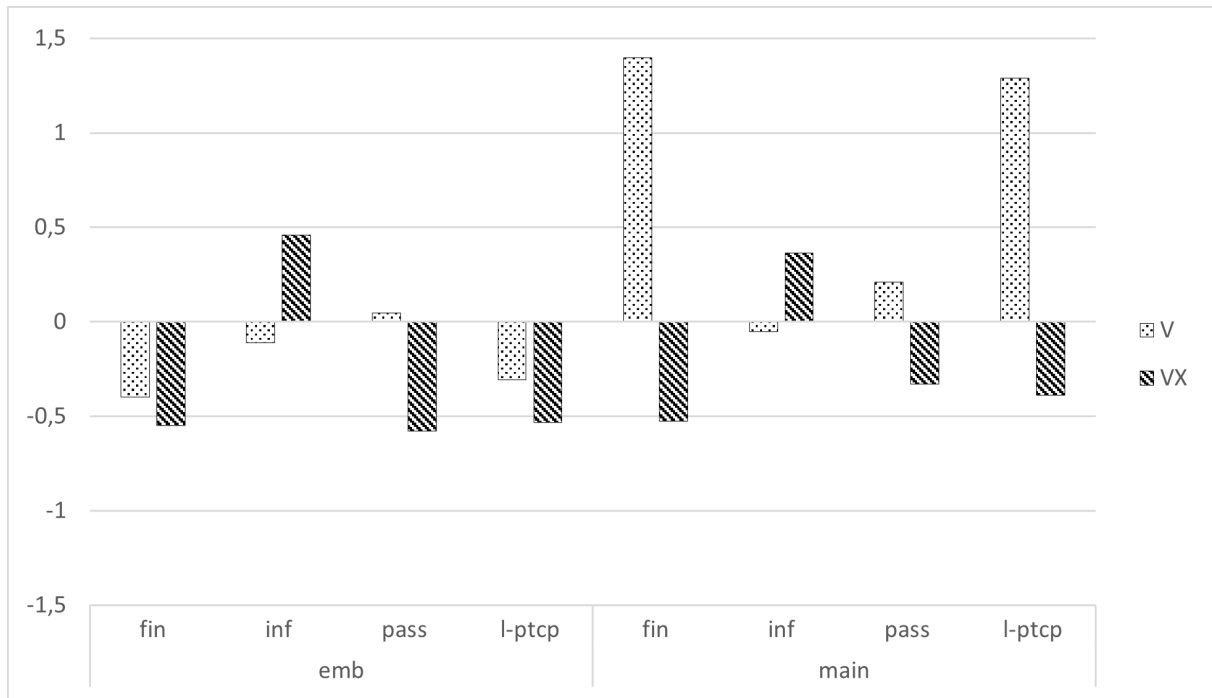


Figure 9.2: Acceptability judgements (z-scores) for fronting of different verb types with and without complements in embedded and main clauses (N=61)

judgements for all passive VP fronting sentences are simply an effect of several marked structures adding up, or indeed indicative of some kind of ban on this fronting operation. The experiment and its results can be found under the OSF link provided in appendix B.

Regarding the fronting of verb plus complement (the “VX” columns in figure 9.2), Veselovská (1995) invokes an “adjacency requirement” on finite verb and clitic when the verb is fronted: “When the verb is presumably fronted (into C) in Czech, CLs follow the verb and a strict PF adjacency requirement holds for the V+CL cluster” (Veselovská, 1995, 89).<sup>3</sup> In this spirit, Veselovská (2008) offers a syntactic explanation for the different properties of clitic and non-clitic auxiliaries: whilst the clitic past and conditional auxiliaries are generated in T (“in the clausal domain”), the passive auxiliary is generated in  $v^*$  (cf. Veselovská, 2008, 559-60). The author does not deal with different types of verbal material as clitic hosts, but if we apply her analysis to the data presented here, the following argument arises:

Non-lexical clitics such as the passive auxiliary occupy a low syntactic position and consequently can be preceded by both verbal and phrasal material. In contrast, clitics like the past auxiliary are generated in T and, as claimed in Veselovská (1995), adjoin to the highest head position in the clause (which is always C in her analysis). This position can hold a verb or a complementiser, and these two elements can thus precede the verb. This leads directly to the above-mentioned adjacency requirement, i. e. that nothing can intervene between verb or complementiser and clitic.

<sup>3</sup>There might a relation between this syntactic requirement and the prosodic pattern observed by Avgustinova and Oliva (1995, 18-9), who note that Czech clitics are most prone to attach prosodically to a neighbouring verb, be it to their left or to their right.

However, the acceptability of fronting phrasal material into a position between the complementiser or verb and the clitic, as in 204, 210b, and, marginally, 206b, represents a clear counterargument to strict adjacency, which relies on the idea that clitics are located in C: if they were, they would adjoin to the complementiser *že* or the verb, and we would not expect to find any intervening material that is clearly phrasal. Thus, we have come across further counter-evidence to the idea that clitics move to C, adding to what was sketched in section 6.1.3. If clitics do not move to C, then Veselovská's (1995/2008) explanation cannot be correct. Also, we cannot assume a purely phonological adjacency requirement either, given the just mentioned data with intervening material.

Instead, I propose that the difference is due to the fact that finite verbs move out of VP, whilst infinitives do not. The results for “fin” and “l-ptcp” in figure 9.2 also show that past participles pattern with finite verbs, and are clearly distinct from the other two groups, infinitives and passive participles. To capture this, I assume that the past participle *is* an inflected verb which undergoes head movement out of VP, whilst the infinitive is not. Veselovská (1995) offers a range of arguments for the assumption that the *l*-participle, like the finite verb, moves out of VP. She takes the lexicalist view that inflection is base-generated, thus she does not need raising to T (cf. Veselovská, 1995, 17). Her arguments against uniform raising of V to T, and for raising of V to *v*, are the flexible positioning of the verb with respect to both subjects and objects, as well as with respect to VP-adverbs (cf. Veselovská, 1995, 57-61). Another argument for movement of the *l*-participle is that it carries negation, whilst for example the passive participle does not (cf. Veselovská, 2008, 557).

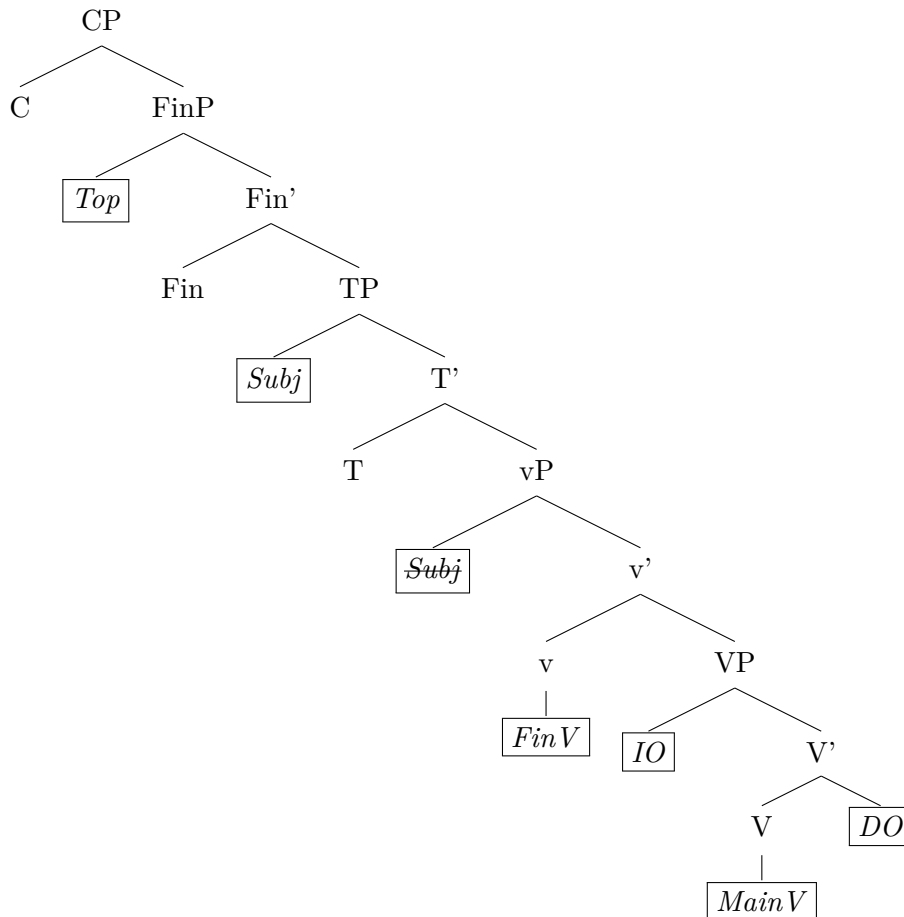
Also Franks and King (2000) briefly mention the different behaviour of past participles (i. e. *l*-participles) on the one hand and infinitives and passive participles on the other, and they similarly conclude that the *l*-participle moves out of VP (Franks & King, 2000, 113, fn. 20). In addition to all that, in chapter 7, I have presented several arguments for viewing the *l*-participle as a finite verb, among others the fact that the past and conditional auxiliaries have no verbal stem of their own, and the *l*-participle's loss of case marking. Thus, I believe that categorising *l*-participles as finite verbs is well-founded.

So how exactly does this approach capture the data? First, I assume that what occupies the first position in 210b and 206b (the sentences with fronted infinitive plus object) is a topicalised (i. e. fronted) VP. In the current analysis, clitic placement after a VP topic is captured in the same way as any other kind of topicalisation: topicalised phrases are above TP, thus the clitic's ideal position is below that phrase, directly at the left edge of TP. Second, if finite verbs and *l*-participles leave VP via head movement (for example to *v*, cf. Veselovská, 2008, 560), then the verb and its object cannot be topicalised as one phrase. Thus, joint topicalisation of finite verb and object, as exemplified in sentences 205b, 208b, 209b, and 212b is not possible; this is independent from clitic placement.

The syntactic structure I assume is made explicit in 213. Again, I follow Fehrman and Junghanns (2012) in their analysis of the left periphery in Slavic. They also show, based on evidence from adverb placement, that the subject (when backgrounded) moves to Spec-TP. But note that this does not amount to assuming obligatory subject movement to TP as found in English; arguments against an active subject EPP in Czech were presented in section 3.4.1. The

verb is attracted by Fin only in VSI structures (cf. Fehrman & Junghanns, 2012, 80-3). I also assume a split VP, and I follow Veselovská's (1995) view that (apart from VSI orders) the finite verb does not leave vP, as argued above (cf. Veselovská, 1995, 57-61). The order of merge for indirect and direct object might be reversed for certain types of ditransitive verbs; see Dvořák (2010).

(213) Syntactic structure for Czech:



To capture the relevant syntactic principle in OT, I assume a constraint, defined in 214, that implements the Chain Uniformity Condition by Chomsky (1995, 253).<sup>4</sup> It thus requires a chain to be uniform with respect to syntactic status. Since the finite verb undergoes head movement out of VP, it cannot be involved in phrasal movement to the left periphery later in the derivation, because this would involve a change in phrase structure status from minimal to maximal.<sup>5</sup>

<sup>4</sup>The Chain Uniformity Condition is not to be confused with either the Uniformity Condition (regarding overt and covert syntax) or the Uniformity Principle (regarding variation between languages). See also Chomsky (1995, 383, fn. 32).

<sup>5</sup>We could partially fix the problems of Veselovská's (1995/2008) account by combining it with the assumptions of the UNICHAIN approach. Infinitive plus object would then be topicalised as one phrase (possibly, then, to SpecCP). Yet, if we must let Uniformity in through the back door anyway, we can also simply fully adopt this approach and dispose of strict adjacency altogether.

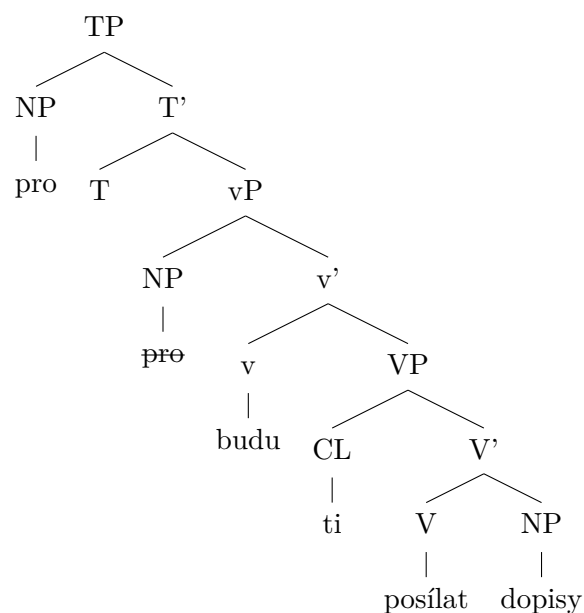


(214) UNICHAIN = A chain is uniform with regard to its phrase structure status S.

Where:  $S \in \{\text{maximal, minimal, neither}\}$

The trees in 216 and 217 show the relevant structures with infinitives and finite verbs, respectively. The a-sentences do not involve topicalisation to emphasise the structural position of the main verb. So, in 216a, the infinitive *posílat* remains within VP. This entire VP can then, in 216b, be topicalised to Spec-FinP. Note that the clitic is not in its surface position here. In contrast, the finite verb *posílám* in 217a moves out of VP, to v (Veselovská, 2008, 560). For this reason, it cannot participate in movement of a maximal projection to FinP, because this would violate UNICHAIN, and 217b is therefore ungrammatical.<sup>6</sup>

(216) a. Structural positions in a clause with infinitive main verb (future tense; without Fin projection):

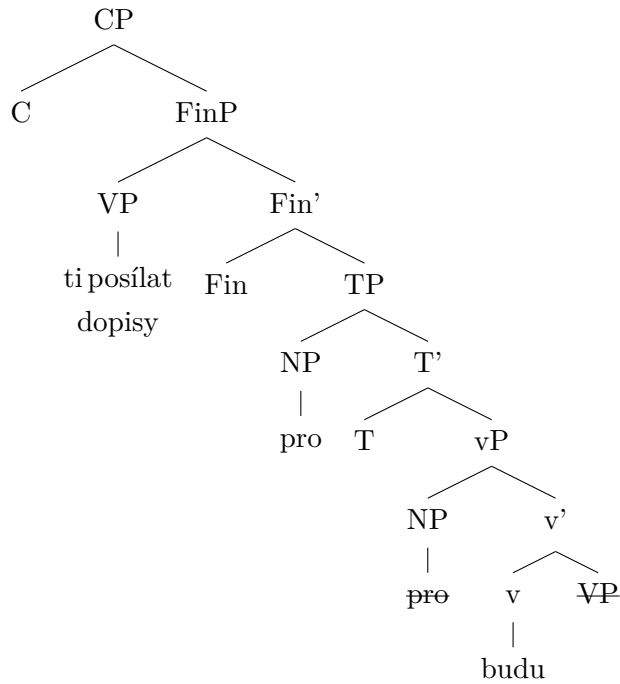


<sup>6</sup>Note that, for the analysis to accurately capture the presented data, we must assume a constraint that prohibits vP fronting. Otherwise, *posílám dopisy* could still be fronted to the left periphery in 217a, simply by fronting the entire vP. This would entail fronting of the subject, which in the present example is silent. However, overt subjects that arguably stay within vP, like indefinites, show that fronting of the entire vP is not possible, since examples like 215 are clearly rejected by speakers:

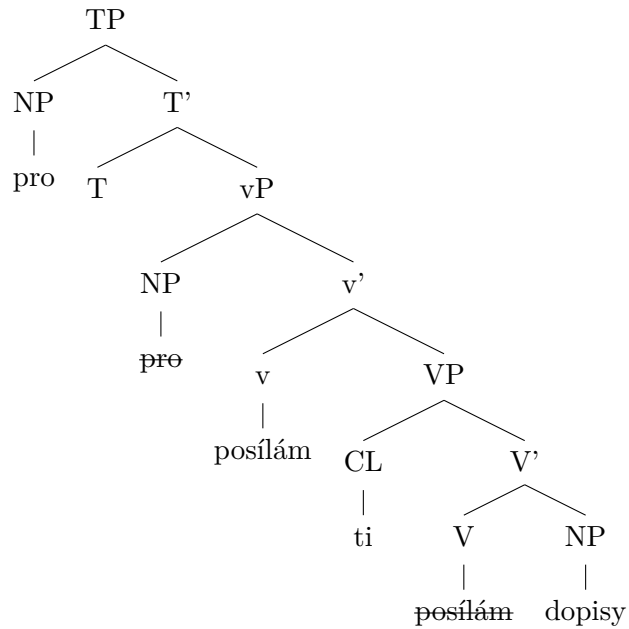
(215) \**Někdo pomáhat ti bude.*  
 someone help.INF you.DAT will.

‘Someone will help you.’

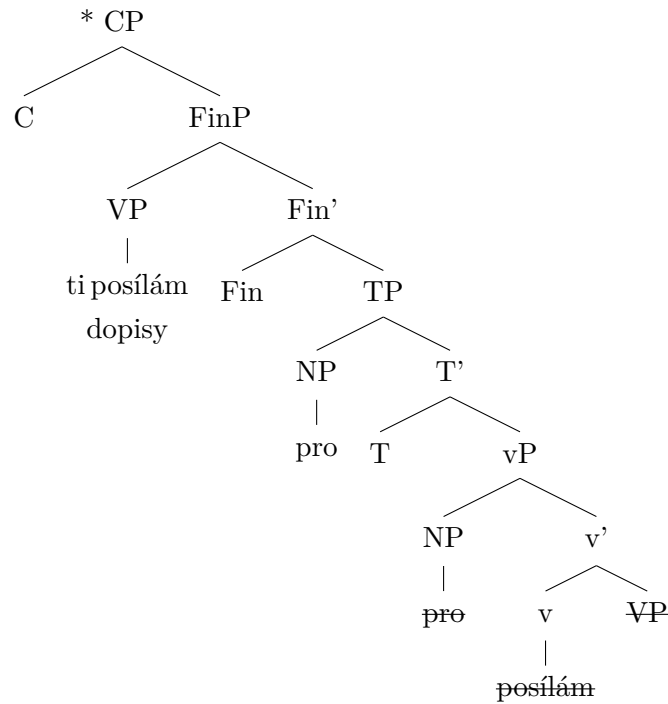
- b. Structural positions in a clause with infinitive main verb (future tense; full structure, without final clitic alignment):



- (217) a. Structural positions in a clause with finite main verb (present tense; without Fin projection):



- b. Structural positions in a clause with finite main verb (present tense; full structure, without final clitic alignment) – UNICHAIN is violated:



#### 9.2.4 Follow-up: fronting with context

The question now is: When the VP, i.e. the verb and its object, is marked as the topic, but UNICHAIN prevents their joint topicalisation, what happens instead? My hypothesis is that in this case, only the verb is fronted, because this preserves the underlying order. The relevant constraint, HEAD-COMPLEMENT, is proposed by Broekhuis (2008) and defined as in 218a. Following Costa (1998), I also assume a constraint TOP-FIRST, defined in 218b, according to which topics appear sentence-initially. However, unlike Costa (1998), I formulate it as an alignment constraint, since this permits a precise evaluation by the OT mechanism.<sup>7</sup> I also assume that information-structure related features like topic and focus are included in the input (cf. Song, 2012, 186). The ranking of the relevant constraints for Czech is shown in 218c.

- (218) a. HEAD-COMPLEMENT = A head precedes all terminals originally dominated by its complement (cf. Broekhuis, 2008, 357).
- b. TOP-FIRST = ALIGN(Topic,l,FinP,l)  
= An element marked as a topic must appear at the left edge of FinP.
- c. UNICHAIN >> TOP-FIRST >> HEAD-COMPLEMENT

Thus, the prediction is that when the VP topic is split, verbs are the exponents of VP-topichood, not objects. In a small follow-up acceptability judgement study with 17 participants,

<sup>7</sup>The original formulation of this constraint is: “Topics are sentence-initial. Failed by topics that are not sentence-initial, and by sentence-initial non-topics” (Costa, 1998, 4). An additional problem with this wording is that it categorically penalises topic-less sentences. My alternative proposal does not have this problem, since it is only violated by topics that are not at FinP’s left edge, but not by non-topical elements in this position. This is due to the fact that Generalised Alignment is directional; see its definition in section 8.1.2.

I asked native speakers of Czech to rate different sentences with respect to a given context, in order to control the topic of each item. For example, the phrase *posílat e-mail* ‘to send an e-mail’ was given in the context in 219, and participants were asked to judge sentences with only the object fronted (as in 219a), only the verb fronted (as in 219b), and verb and object fronted together (as in 219c). The same was done for VPs with a past participle, as in 220.

(219) Context: *Nezapomeň, že mu musíš posílat e-mail.*

‘Don’t forget that you must send him an e-mail.’

a. *E-mail mu budu posílat každý pátek.*  
e-mail him will.1SG send.INF every Friday

b. *Posílat mu budu e-mail každý pátek.*  
send.INF him will.1SG e-mail every Friday

c. *Posílat e-mail mu budu každý pátek.*  
send.INF e-mail him will.1SG every Friday

‘I will be sending him an e-mail every Friday.’

(220) Context: *Nezapomněl jsi, že mu musíš posílat e-mail?*

‘Did you remember that you must send him an e-mail?’

a. *E-mail jsem mu posílal každý pátek.*  
e-mail AUX.1SG him send.PTCP every Friday

b. *Posílal jsem mu e-mail každý pátek.*  
send.PTCP AUX.1SG him e-mail every Friday

c. *Posílal e-mail jsem mu každý pátek.*  
send.PTCP e-mail AUX.1SG him every Friday

‘I have been sending him an e-mail every Friday.’

The results displayed in figure 9.3 show that with VP topics, infinitives and past participles display different patterns: When infinitive + object is the topic (striped columns), only the object is fronted, or the entire phrase. When finite verb + object is the topic (dotted columns), the object or the verb is fronted, but not both together. Due to the small size of this pilot study, no z-scores were calculated. Appendix B contains a link to the OSF repository in which the materials and results can be found.

As predicted by the high ranking of UNICHAIN, verb and object can only be topicalised together when the verb is an infinitive. It is however surprising that with both verb types, topicalising just the object is always possible. The best explanation I have at this point is that these cases result from specifying only the object as the topic in the input; i. e. the grammatical alternatives with fronted objects result from a different candidate set than the ones where the whole VP is marked as the topic, and where either VP or V topicalisation wins. In the tableaux in 9.4 and 9.5, I illustrate the evaluation of these latter cases. What is marked as the topic in the input is given at the top left of the tableaux (since information-structural features are part

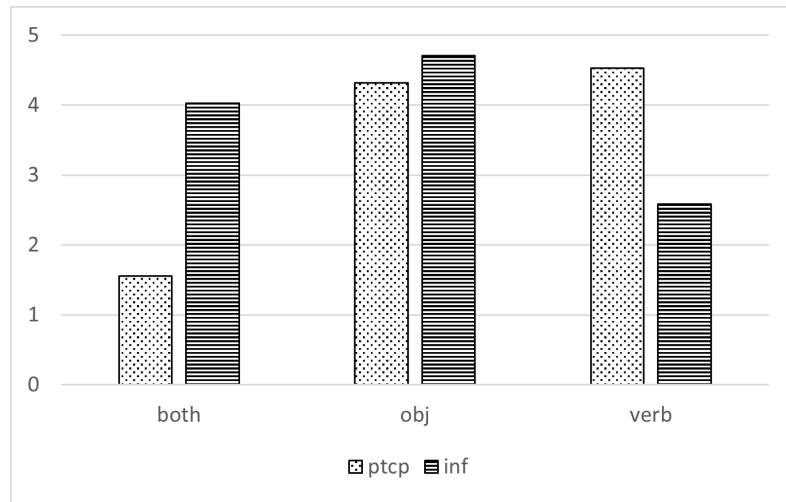


Figure 9.3: Acceptability judgements for (partial) fronting of VP topics (N=17; five-point-scale: 1=bad, 5=good)

of the input of an evaluation). For ease of exposition, in both tableaux, only candidates with optimally positioned clitics are considered.<sup>8</sup>

Topic: [VP V <sub>inf</sub> Obj ]	UNICHAIN	TOP-FIRST	H-CMPL
a. [FinP [VP V Obj ] [TP cl [vP VP ]]]			
b. [FinP V [TP cl [vP [VP V Obj ]]]]		*!	
c. [FinP Obj [TP cl [vP [VP V Obj ]]]]		*!	*

Table 9.4: OT tableau illustrating VP topicalisation with an infinitive verb

Topic: [VP V <sub>ptcp</sub> Obj ]	UNICh	TOP	H-CMPL
a. [FinP [VP V Obj ] [TP cl [vP V [VP V Obj ]]]]	*!		
b. [FinP V [TP cl [vP V [VP V Obj ]]]]		*	
c. [FinP Obj [TP cl [vP V [VP V Obj ]]]]		*	*!

Table 9.5: OT tableau illustrating VP topicalisation with a past participle

However, it is important to note that the judgement patterns in this pilot study with a given-VP context were very similar to those in the previous study without a context – compare the results for “verb” and “both” to those for “V” and “VX” for infinitives and past participles in main clauses in figure 9.2 above. Thus, participants in the follow-up arguably ignored the context – after all, the experimenter cannot force the participants to understand a certain element as the topic, especially with only one context sentence and the rather unusual VP topic. Alternatively, already in the earlier judgement task, participants could have imagined possible

<sup>8</sup>Note that when the clitic is a pronominal argument of the verb, it arguably has its base position within VP, thus VP fronting can be seen as a case of clitic lowering. Whilst this should clearly be restricted, it can also not be excluded entirely, since short subject-less sentences must also involve clitic lowering to satisfy higher-ranked constraints. For a discussion of this, see section 10.3.8 on VP fronting in BCMS, where lowering is apparently less tolerated.

contexts to accommodate the sentences. This means that the present section is best understood as a sketch for a potential analysis of VP topics and their linearisation that would benefit from further research.

### 9.2.5 A note on verb topicalisation

I have not yet addressed the issue of verb topicalisation in embedded clauses. For the topicalisation of verb + object, the results as given in figure 9.2 were similar in main and embedded clauses, which is fully in line with the present analysis: infinitives can be topicalised with their object, finite verbs/past participles cannot. But for topicalisation of only a verb (the “V” columns in figure 9.2) there is a clear difference between main clauses, where initial finite verbs and past participles are fully acceptable, and embedded clauses, where all constructions with only a verb preceding the clitic are degraded. The relevant examples are repeated in 221; the mean z-scores all lie between -0.4 and +0.05, which I indicate with a question mark here.

- (221) a. <sup>?</sup>*Myslím, že posílají mu dopisy každý týden.*  
 think.1SG COMP send.3PL him.DAT letters every week
- b. <sup>?</sup>*Myslím, že posílat mu budou dopisy každý týden.*  
 think.1SG COMP send.INF him.DAT will.3PL letters every week
- c. <sup>?</sup>*Myslím, že posílány mu byly dopisy každý týden.*  
 think.1SG COMP send.PASS him.DAT were letters every week
- d. <sup>?</sup>*Myslím, že posílali mu dopisy každý týden.*  
 think.1SG COMP send.PTCP him.DAT letters every week

All sentences in 221 are clear cases of verb topicalisation – if the verb was not above TP, the clitic would directly follow the complementiser. The restriction against verb fronting in embedded clauses, mentioned in section 5.1.2, largely holds, though possibly to slightly varying degrees for different verb types. It might thus turn out that verb topicalisation is generally dispreferred or even non-existent in Czech. According to Fehrmann and Junghanns (2012), verb-subject inversion (another clear case of a fronted verb) occurs when a subject is background (i. e. not focus), and comment (i. e. not topic); it then occupies its canonical position in Spec-TP. In this context, the verb moves out of the focused VP. Crucially, verb fronting does not indicate verb focus or the realisation of an abstract (event-time) topic (cf. Fehrmann & Junghanns, 2012, 76-7; see also section 3.5.2). In addition, topicalisation typically moves maximal projections, not heads.

For main clauses, the differences in fronting might relate to the crossing of non-clitic verbs as discussed in section 7.2.2: in the test sentences, the infinitive has to cross the future auxiliary (example 210a) and the passive participle has to cross the passive auxiliary (example 211a). Since this requires strong information-structural motivation, which was not provided in the experiment, these sentences might have fared worse than the ones where the finite verb or the *l*-participle were fronted, without crossing any non-clitic verb. Yet, for embedded clauses, this account does not suffice, since sentences 221b and 221c are similarly degraded as sentences 221a and 221d.

On the other hand, there is a discrepancy between the low judgement of sentence 221b and the much better judgements gathered for the corpus sentences in 202: In all these sentences, the infinitive is fronted to the position following the complementiser, and it crosses another verb. Given that the sentences in 202 are acceptable, these two restrictions cannot be absolute, but can be outranked by superior requirements.

We must also keep in mind that, given the clitic's flexible positioning, when a clitic occurs after a given element, this does not automatically entail that this element is above TP. For example, the winning candidate in tableau 9.5 may just as well have the past participle in *v*, with the clitic following it. An elegant way to block verb topicalisation would be to reformulate TOP-FIRST, so that topics must be in *Spec-FinP*. In consequence, every V that satisfies TOP-FIRST would at the same time violate the higher-ranked constraint UNICHAIN, since this would be movement of a head to a position for maximal projections.

At this point, it is important to observe that in the present account of clitic placement, many constructions that appear to involve fronted verbs actually only feature a clitic in a lower position. Separating these from real cases of verb fronting as the ones in 221 is thus an important achievement of constraint-based theories to clitic positioning, and also provides an important conceptual basis for further research.

### 9.2.6 Summary

The analysis of more complex data with respect to VP topics has illustrated how this approach to clitic placement can be incorporated into a wider OT-analysis of Czech constituent order. I have shown how the difference between infinitives and finite verbs with respect to topicalisation can be captured using a few very general, independently motivated constraints. I have also provided further arguments for classifying the *l*-participle as a finite verb, and consequently for *not* classifying the auxiliary clitics as finite verbs. The fact that the finite verb leaves VP prevents joint topicalisation of verb and object; this again shows that the first position is determined by general syntactic principles of the language, not by a clitic's requirements. We have also seen an additional argument against clitics in a fixed position.

## 9.3 Conclusions

This chapter has dealt with the interactions between clitic placement and argument topicalisation. We have seen that all phenomena described here can be captured with the constraints on clitic placement as summarised already in section 8.7 of the previous chapter. The extensions to the theory that we had to make all refer to other aspects of the grammar. They were captured using clearly defined constraints that are needed to model Czech syntax and its interaction with information structure also beyond the constructions that involve 2P clitics. The constraints are summarised in 222a-c, the ranking we have assumed for them in Czech is given in 222d.

- (222) a. UNICHAIN = A chain is uniform with regard to its phrase structure status S.  
Where:  $S \in \{\text{maximal, minimal, neither}\}$

- b. HEAD-COMPLEMENT = A head precedes all terminals originally dominated by its complement (cf. Broekhuis, 2008, 357).
- c. TOP-FIRST = Topics are sentence-initial. Failed by topics that are not sentence-initial, and by sentence-initial non-topics (cf. Costa, 1998, 4).
- d. UNICHAIN >> TOP-FIRST >> HEAD-COMPLEMENT



## Chapter 10

# Extending the Analysis to Further Slavic Languages

Having concluded the discussion of Czech clitic placement and its interaction with other syntactic principles of the language, the question is now whether this language-specific account can be extended to capture 2P systems in other languages. I will focus on three Slavic languages here: two further West Slavic languages, Slovak (section 10.1) and Upper Sorbian (section 10.2), and a South Slavic language, Bosnian-Croatian-Montenegrin-Serbian (BCMS; section 10.3). As will become clear, the more distant the relationship, the more different the clitic placement systems are in comparison to Czech. Yet, I will show that both minor and major differences can be captured with a restricted set of constraints and different constraint orderings.

### 10.1 Slovak: Czech's closest sibling

Slovak is a West Slavic language very closely related to Czech. It has roughly half as many native speakers as its sibling: whilst Czech has around 12 million, Slovak has slightly below 6 million. Since the language is thus relatively small, and because it shares many features with its larger sibling, it also receives less attention. This might be aggravated by the term “Czechoslovak” – yet despite sharing a state for the most part of the 20<sup>th</sup> century, Czech and Slovak have existed as distinct languages already before the creation of the Czechoslovak state and continue to do so today (cf. Sussex & Cubberley, 2006, 6-7).

The present section investigates the inventory and distribution of Slovak 2P clitics based on descriptions found in the literature, especially Franks and King (2000), as well as consultations with a native speaker, and acceptability judgement experiments. As for Czech, these studies were carried out using *L-Rex* and *Prolific*. Since the number of Slovak speakers registered as participants on *Prolific* is much smaller than that of Czech speakers, the Slovak tasks have a smaller sample size. However, as will become clear in the following subsections, the similarity of the patterns found for the two languages suggests that the Slovak data are reliable. On the other hand, some differences between Slovak and Czech occurred, and provide an interesting point of departure for future systematic research into the clitic system of Slovak and the differences between the two languages.

Section 10.1.1 describes the inventory of 2P clitics in Slovak. The basic placement patterns are presented in section 10.1.2, and the respective ordering of clitics in section 10.1.3. The question whether Slovak clitic placement is, as in Czech, purely syntactic, is discussed in section 10.1.4. Section 10.1.5 then turns to clitics in third position, including an investigation of VP topicalisation with different verb types. Clitic climbing is described and analysed in section 10.1.6. Section 10.1.7 provides the conclusions.

### 10.1.1 The clitic inventory

As for Czech, I will focus on these elements that are unambiguous, i. e. lexical, 2P clitics, thus setting aside elements that can optionally be placed in second position when unstressed.

According to Franks and King (2000), the lexical 2P clitics of Slovak come from two major groups: First, verbal auxiliary clitics based on *byt* ‘to be’ – here, as in Czech, the third person auxiliary is null. In contrast to Czech, the conditional marker *by* is invariant. And second, pronouns – again as in Czech, only some pronouns have both strong and weak forms (cf. Franks & King, 2000, 121-6). King (1996, 308) provides a list of Slovak clitics; but here, as for Czech, she does not clearly distinguish between lexical and optional 2P clitics. In combination with Franks and King’s (2000) full table of pronominal clitics, however, the unambiguous 2P clitics can be identified. Table 10.1 provides an overview that also includes optionally cliticising elements like the copula and plural pronouns. The lexical clitics are printed in bold.

	verbal clitics			pronominal clitics	
	COP	AUX	COND	DAT	ACC
1SG	som	<b>som</b>	<b>by</b>	<b>mi</b>	<b>ma</b>
2SG	si	<b>si</b>	<b>by</b>	<b>ti</b>	<b>t’a</b>
3SG	je	–	<b>by</b>	<b>mu/jej</b>	<b>ho/jej</b>
1PL	sme	<b>sme</b>	<b>by</b>	nám	nás
2PL	ste	<b>ste</b>	<b>by</b>	vám	vás
3PL	sú	–	<b>by</b>	im	ich
REFL				<b>si</b>	<b>sa</b>

Table 10.1: 2P clitics of Slovak; lexical clitics in bold

Thus, there are two differences to Czech: first, with respect to pronouns (cf. section 4.1.1), in the first person, the Slovak accusative pronoun is also an unambiguous 2P clitic; and second, with respect to verbs (cf. section 4.1.2), the conditional clitic is invariant. I will return to this latter point in section 10.1.3.

### 10.1.2 Clitic placement: the basic pattern

As in Czech, Slovak clitics appear after the first constituent of their clause. This can be a verb (223a), an adverbial (223b), or a complementiser (223c). Example 224 shows that the first constituent can also be a complex nominal phrase. In addition, sentence 224b provides some first evidence that clitics can appear after prosodic breaks.

- (223) a. *Chodili sme po meste.*  
 walk.PTCP AUX.1PL around town  
 ‘We walked around town.’
- b. *Naozaj si ma nevolal?*  
 really AUX.2SG me.ACC NEG:call.PTCP  
 ‘You really didn’t call me?’
- c. *D’akujem vám, slečna, že ste mne dali prednosť.*  
 thank you.DAT miss COMP AUX.2PL me give.PTCP preference  
 ‘Thank you, miss, for letting me go first.’

(cf. Franks & King, 2000, 121)

- (224) a. *Môj brat si kúpil psa.*  
 my brother REFL.DAT buy.PTCP dog.ACC  
 ‘My brother has bought himself a dog.’
- b. *Môj brat, ktorý žije v Japonsku, si kúpil psa.*  
 my brother REL live.3SG in Japan.LOC REFL.DAT buy.PTCP dog.ACC  
 ‘My brother, who lives in Japan, has bought himself a dog.’

According to Short (1993b), the reflexive clitic *sa* is most prone to movement and commonly appears directly after the verb (cf. Short, 1993b, 566-7). He provides sentence 225 as an example. It does not become clear whether he only means the accusative reflexive or also the dative *si*. Due to this ambiguity, and since my native speaker consultant expressed doubts about this claim, an acceptability judgement study was conducted in order to obtain comparable and also negative data.

- (225) *Z iniciatívy Jozefa Kohúta založil sa v Martine  
 from initiative.GEN Jozef.GEN Kohút.GEN found.PTCP REFL.ACC in Martine.LOC  
 hasičský zbor.  
 fire.ADJ brigade*

‘A fire brigade was founded in Martin on the initiative of Jozef Kohút.’

In the experiment, minimal pairs of sentences with clitic personal and reflexive pronouns were tested. The factor pronoun type was crossed with two further factors: position of the clitic (2P vs. 3P) and whether the clitic is adjacent to the verb (no adjacency, postverbal, preverbal). This yielded a total of twelve conditions, which were tested with six different test items. An example is given in 226 with a personal pronoun and in 227 with a reflexive pronoun. Table 10.2 gives a schematic overview of the experiment design and provides references to the respective examples. 24 native speakers of Slovak were recruited via the platform *Prolific*.

	2P		3P	
	pers	refl	pers	refl
no adjacency	226a	227a	226d	227d
postverbal	226b	227b	226e	227e
preverbal	226c	227c	226f	227f

Table 10.2: Design with example references for Slovak freedom of pronominal clitics

- (226) a. *Rada **Ťa** nové slovíčka učím.*  
 gladly.F you.ACC new words teach.1SG  
 ‘I like teaching you new words.’
- b. *Učím **Ťa** rada nové slovíčka.*  
 teach.1SG you.ACC gladly.F new words
- c. *Rada **Ťa** učím nové slovíčka.*  
 gladly.F you.ACC teach.1SG new words
- d. *Učím rada **Ťa** nové slovíčka.*  
 teach.1SG gladly.F you.ACC new words
- e. *Rada učím **Ťa** nové slovíčka.*  
 gladly.F teach.1SG you.ACC new words
- f. *Rada nové slovíčka **Ťa** učím.*  
 gladly.F new words you.ACC teach.1SG
- (227) a. *Rada **sa** nové slovíčka učím.*  
 gladly.F REFL.ACC new words teach.1SG  
 ‘I like learning (= teaching myself) new words.’
- b. *Učím **sa** rada nové slovíčka.*  
 teach.1SG REFL.ACC gladly.F new words
- c. *Rada **sa** učím nové slovíčka.*  
 gladly.F REFL.ACC teach.1SG new words
- d. *Učím rada **sa** nové slovíčka.*  
 teach.1SG gladly.F REFL.ACC new words
- e. *Rada učím **sa** nové slovíčka.*  
 gladly.F teach.1SG REFL.ACC new words
- f. *Rada nové slovíčka **sa** učím.*  
 gladly.F new words REFL.ACC teach.1SG

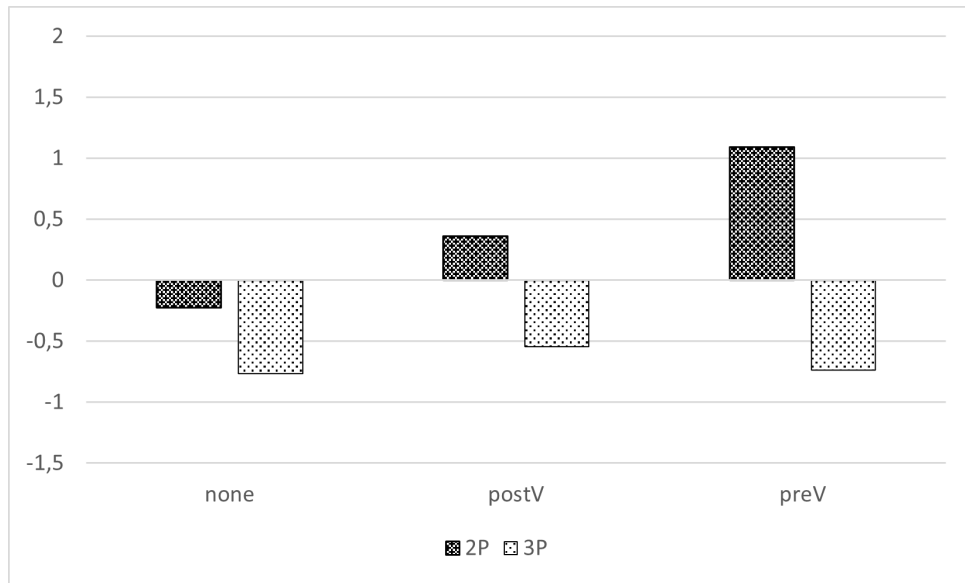


Figure 10.1: Acceptability judgements (z-scores) for 2P and 3P positions of personal pronouns, with no adjacency to the verb, right adjacency, and left adjacency (N=24)

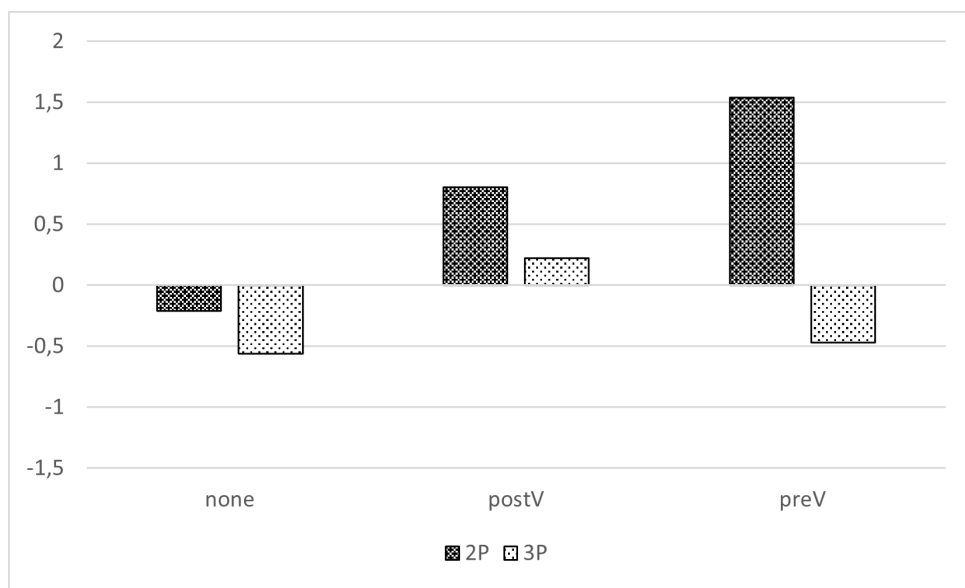


Figure 10.2: Acceptability judgements (z-scores) for 2P and 3P positions of reflexive pronouns, with no adjacency to the verb, right adjacency, and left adjacency (N=24)

The results show that there are only slight differences between personal and reflexive pronouns. In general, 2P placement combined with adjacency to the verb yielded positive z-scores. The fact that non-adjacency was never rated good might be related to the fact that in order to create this distance between verb and clitic, the verb had to be placed at the sentence edge, which yields a highly marked, context-dependent structure. The only other instance of positive z-scores was with postverbal 3P placement of reflexives. This is in line with Short's (1993b) observation – but the effect is small and clearly not categorical.

Since half of the items featured accusative pronouns, and the other half dative pronouns, a retrospective analysis is possible in order to see whether this extended freedom applies only to the accusative reflexive *sa*, which is explicitly mentioned by Short (1993b), or also to the dative reflexive *si*. Figure 10.3 suggests that there are no differences between the two items. Yet, it must be noted that by splitting up the data in this way, the data set becomes quite small, with only three lexicalisations in each of the two groups. Yet, together with the general results shown in figures 10.1 and 10.2, it appears that if certain elements in Slovak have a tendency to appear postverbally, thereby sometimes violating 2P placement, it is the entire group of reflexives, not just one specific lexical element.

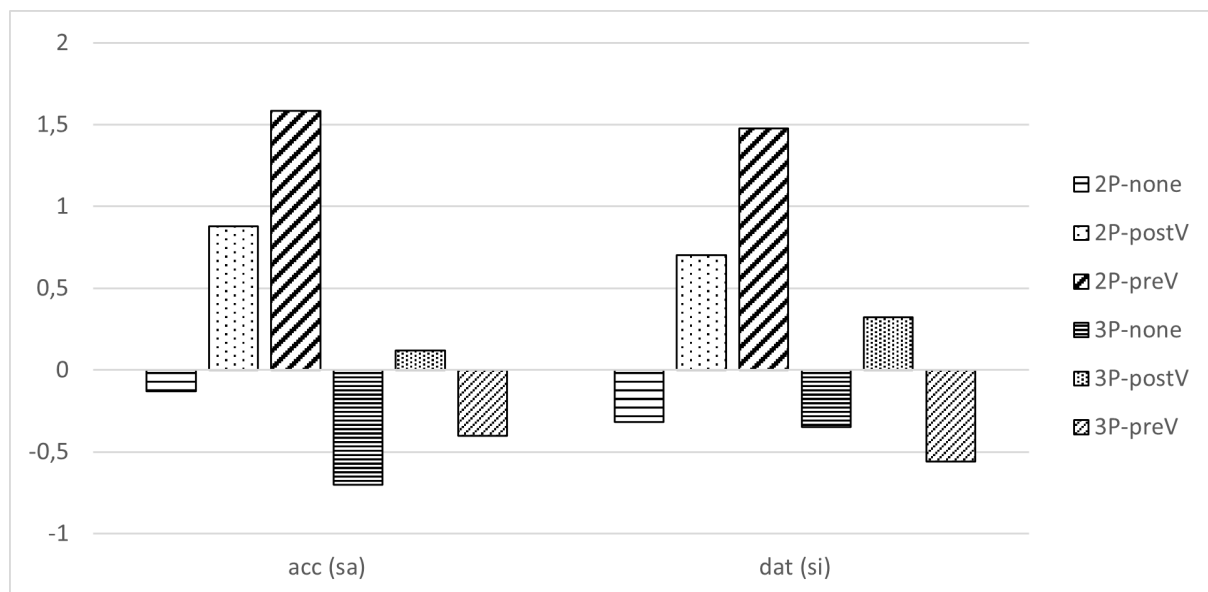


Figure 10.3: Acceptability judgements (z-scores) for different positions of the two reflexive pronouns *sa* and *si* (N=24)

We can thus conclude that Slovak clitics, like Czech ones, tend towards the second position, i. e. they appear after the first constituent of the clause, regardless of its length. Occurrences in lower positions are sometimes possible, especially with the reflexive pronominal clitics. Further experimental investigation of these patterns could help clarify the conditions under which such delayed placement is possible – note that sentence 225 is also an instance of postverbal 3P placement, as is sentence 227e, but the latter has a relatively low z-score of 0.2. This might be due to the brevity of the sentence, and to the lack of contextual motivation. A different case of lower clitic placement, 3P positioning in embedded clauses, will be discussed in section 10.1.5.

### 10.1.3 Ordering in the cluster

The main difference in the ordering of clitics between Czech and Slovak falls in the domain of clitic auxiliaries: in Slovak, the invariable conditional marker *by* always combines with a past auxiliary (and the *l*-participle) to form the conditional mood; this is shown in example 228. Thus, unlike in Czech, the two do not occupy one slot in the cluster, but have a clear ordering

relation. The full ordering template for Slovak clitics is given in 229 (cf. Franks & King, 2000, 126-9).

- (228) *Ja by som napísal list.*  
 I COND AUX.1SG write.PTCP letter  
 ‘I would write a letter.’

(cf. Franks & King, 2000, 126)

- (229) Clitic template for Slovak:  
 COND – AUX – REFL – DAT – ACC/GEN

Note that this behaviour of the conditional is fully in line with the analysis of the inflecting conditional clitic in Czech that I have presented in section 7.4: *by* only expresses the conditional, whereas the auxiliary expresses person and number. Since the elements used to form the past tense (auxiliary and participle) also occur in the present conditional, in both languages the past tense interpretation cannot arise compositionally. Also note that, because conditional is marked by a separate clitic in Slovak, the language does not have the bisyllabic clitics that are found in Czech. However, it is possible that this difference is merely orthographic, at least when considering the colloquial forms of the Czech conditional clitic.

For Czech, I have argued in section 4.1.3 that the conditional complementiser *-li* is not a 2P clitic, but a verbal clitic with the additional requirement that its host be clause-initial. Thus, it is not, as often assumed, in the first position of the clitic cluster. In Slovak, this issue does not arise, because something analogous to *-li* does not exist.

Based on the available data, I thus argue that we can analyse the order within the cluster based on the same principles as in Czech, as presented in section 8.5. One necessary adaptation is that the constraint referring to both the past and the conditional auxiliary should be split into two constraints in Slovak. The EDGEMOST constraint referring to the conditional is then ranked directly above the one referring to the past auxiliary. Whether or not PCC effects, which occur at least with some speakers in Czech, are also present in Slovak and display the same patterns, is subject to future comparative research.

#### 10.1.4 Prosodic enclitics or not?

Short (1993b, 566) describes Slovak clitics as enclitics, thereby implying that as in BCMS, the clitics require a host to their left. Given the close similarity between Czech and Slovak clitic placement we have observed so far, such a characterisation appears doubtful. And indeed, Franks and King (2000) point out Starke’s (1993) argument that Slovak clitic placement is syntactic, and not prosodic. They present several data points where the clitic follows an intonational boundary, and thus cannot be prosodically enclitic, yet potentially proclitic (cf. Franks & King, 2000, 133-4, and example 224 above).

When we translate the Czech sentences by Junghanns (2003), where a 2P clitic appears between two prosodic breaks and therefore cannot have a phonological host to either side (see

section 5.1.3, example 74), we find that Slovak clitics, like Czech ones, can in fact be phonologically free. My informant judged them all as good.

- (230) a. *Dôkazy, ktoré predložila, ho, ako je vidieť, nepresvedčujú.*  
 evidence REL present.PTCP him.ACC as is see.INF NEG:convince.3PL  
 ‘Obviously, the evidence that she has presented does not convince him.’
- b. *Knihy, ktoré tu vidíte, sa, ako tvrdia, dnes platia zlatom.*  
 books REL here see.2PL REFL.ACC as claim.3PL today pay.INF gold.INS  
 ‘They claim that the books you can see here are paid for with gold today.’
- c. *Ten televízor, ktorý sme u Vás našli, ste, ako vieme, nekúpili v obchode.*  
 this TV REL AUX.1PL at you find.PTCP AUX.2PL as know.1PL  
 NEG:buy.PTCP in shop  
 ‘We know that you didn’t buy the TV set that we found at your place at a shop.’
- d. *Ten človek, ktorý tu žije, by, ako tvrdia, mohol byť terorista.*  
 this man REL here live.3SG COND.3 as claim.3PL can.PTCP be.INF terrorist  
 ‘They claim that the man who lives here might be a terrorist.’

As in Czech, alternative positions of the clitic make the sentence unacceptable, as can be seen in 231 for sentence 230a above. The variant in 231b is not fully felicitous; my Slovak consultant judged it as “inelegant”, though probably not ungrammatical. Yet, the best version is sentence 230a, which shows that Slovak 2P clitics, like Czech ones, follow the first constituent independently of the prosodic environment that this position provides.

- (231) a. *\*Ho dôkazy, ktoré predložila, ako je vidieť, nepresvedčujú.*  
 him.ACC evidence REL present.PTCP as is see.INF NEG:convince.3PL
- b. *°Dôkazy, ktoré predložila, ako je vidieť, ho nepresvedčujú.*  
 evidence REL present.PTCP as is see.INF him.ACC NEG:convince.3PL
- c. *\*Dôkazy, ktoré predložila, ako je vidieť, nepresvedčujú ho.*  
 evidence REL present.PTCP as is see.INF NEG:convince.3PL him.ACC

### 10.1.5 Third-position placement

In embedded clauses, a single constituent can appear between the complementiser and the clitic, as shown in 232. Franks and King (2000) point out that this, too, is a syntactic restriction, and not dependent on prosodic factors such as the heaviness of the constituent. As in Czech, a verb cannot be placed in this pre-clitic position in embedded clauses, although, as shown before, the verb can be initial in main clauses. An example is given in 233 (cf. Franks & King, 2000, 135-6).



- (232) a. *Myslím, že **mu** Pavol dal to auto včera.*  
 think.1SG COMP him.DAT Pavol give.PTCP that car yesterday
- b. *Myslím, že Pavol **mu** dal to auto včera.*  
 think.1SG COMP Pavol him.DAT give.PTCP that car yesterday
- c. *Myslím, že to auto **mu** dal včera.*  
 think.1SG COMP that car him.DAT give.PTCP yesterday
- d. *Myslím, že včera **mu** dal to auto.*  
 think.1SG COMP yesterday him.DAT give.PTCP that car

‘I think that Pavol/he gave him that car yesterday.’

(cf. Franks & King, 2000, 135)

- (233) a. *Umyje **ho** zajtra ráno.*  
 wash.3SG him.ACC tomorrow morning
- b. *\*Myslím, že umyje **ho** zajtra ráno.*  
 think.1SG COMP wash.3SG him.ACC tomorrow morning

‘He will wash it/him tomorrow morning.’

‘I think that he will wash it/him tomorrow morning.’

(cf. Franks & King, 2000, 136)

Whereas sentences with the clitic(s) following the complementiser are perceived as neutral, 3P sentences are subject to discourse restrictions and therefore sometimes perceived as less felicitous by speakers (cf. Franks & King, 2000, 136-7). All this indicates that similarly to Czech, Slovak possesses a specifier position in the left periphery that information-structurally marked phrases, but not verbs, can move to. Whether the information structural conditions are exactly the same in the two languages however requires further investigation.

So far, the Slovak data can therefore be captured fully by the same constraints and ranking as proposed for Czech in section 8.2.1 for 2P placement and section 9.1.2 for 3P placement. The differences we have detected are all lexical in nature: divergences in the clitic inventory lead to divergences in the template governing the clitic cluster, as described in section 10.1.3 above.

In order to assess whether Slovak also displays similar patterns with respect to the topicalisation of verbal elements with and without their complements, the experiment reported in section 9.2.2 was translated to Slovak – for details of the design, see there. The materials and results can be found via the link provided in appendix B. The same 24 participants as in the judgement task regarding the freedom of pronominal clitics (cf. section 10.1.2) took part in the experiment. Figure 10.4 displays the results.

We find that in Slovak, finite verbs and *l*-participles are accepted in the pre-clitic position in main clauses, whilst infinitives and passive participles are less felicitous in this position. In embedded clauses, only infinitive+object receives a slightly positive z-score when placed between

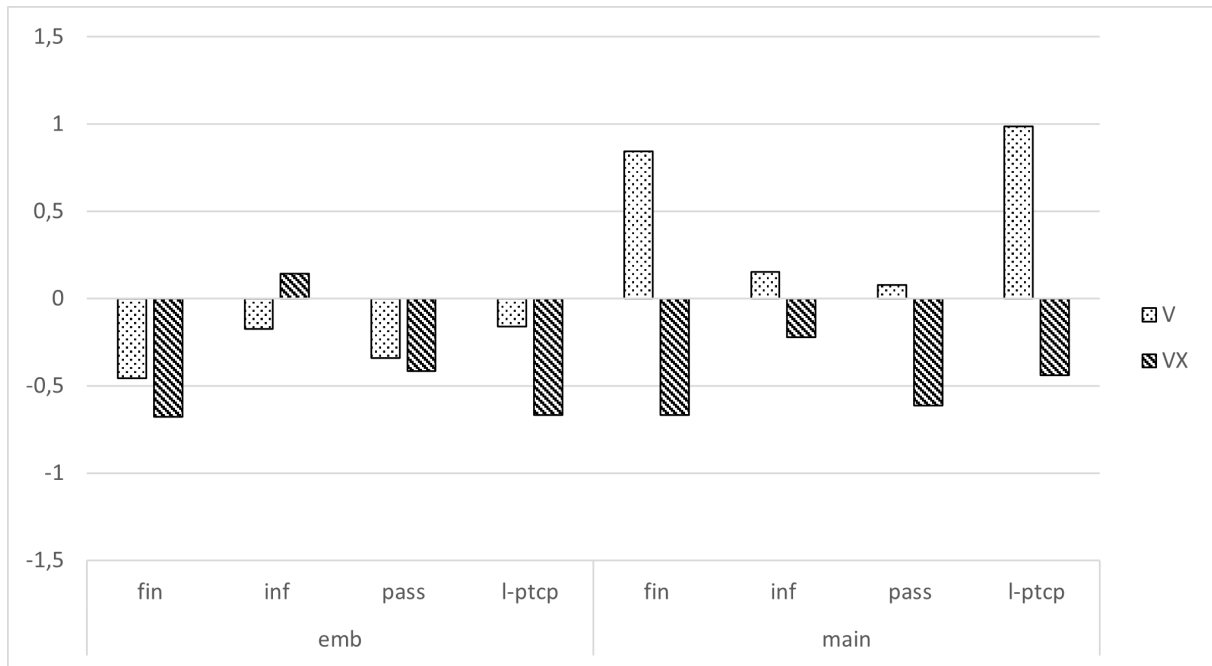


Figure 10.4: Acceptability judgements (z-scores) for fronting of different verb types with and without complements in embedded and main clauses (N=24)

complementiser and clitic. All other verb types are more strongly degraded for VP fronting. As in the Czech experiment, the passive poses the problem that here, the verb was fronted not with the object, but with the subject; and also the markedness of the morphological passive might play a role.

Regarding the three other verb types, the pattern found in Slovak is very similar to the results for Czech reported in section 9.2.2 (see figure 9.2 there). Yet, whilst in Czech, fronting of infinitive plus object yielded clearly positive z-scores, this is not so clear in Slovak. For Czech, I have argued that the infinitive remains within VP, and thus can be fronted together with it, whereas the finite verb and the past participle (which is also developing into a finite verb) move to *v*, and thus cannot be fronted inside the VP.

At this point, it is not possible to say whether the difference between Czech and Slovak is categorical, or whether it is simply due to a general tendency among the Slovak participants to give marked structures a lower rating. Yet again, the direct comparison of the two languages shows that whilst they are very similar in many respects, there are potential differences in rating patterns that could be explored in further comparative studies.

### 10.1.6 Clitic climbing

According to Franks and King (2000, 137-9), Slovak and Czech behave similarly with respect to clitic climbing, in the sense that both languages ban climbing out of finite clauses, as exemplified in 234. Furthermore, the authors state that in both languages, “clitics can only climb out of complements of modal verbs” (Franks & King, 2000, 137). That this description is too simplified for Czech has become clear in section 8.3.1. The acceptability judgement task reported there was replicated for Slovak. As in the Czech experiment, four verb groups were tested, with two

lexicalisations each: *začať* ‘begin’, *môcť* ‘can’, *musieť* ‘must’, and *chcieť* ‘want’. An example from this latter group in all three conditions is given in 235. The same 24 native speakers of Slovak as in the two previously reported experiments participated in this study.

- (234) a. *Ja si myslím, že mu to Marienka nepovedala.*  
 I REFL.DAT think.1SG COMP him.DAT it.ACC Marienka NEG:tell.PTCP  
 ‘I think that Mary did not tell him that.’
- b. \**Ja si mu myslím, že to Marienka nepovedala.*  
 I REFL.DAT him.DAT think.1SG COMP it.ACC Marienka NEG:tell.PTCP  
 (cf. Franks & King, 2000, 138)

- (235) a. *Lucia už dlho chce si kúpiť nový bicykel.*  
 Lucia already long.ADV want.3SG REFL.DAT buy.INF new bike
- b. *Lucia už dlho chce kúpiť si nový bicykel.*  
 Lucia already long.ADV want.3SG buy.INF REFL.DAT new bike
- c. *Lucia si už dlho chce kúpiť nový bicykel.*  
 Lucia REFL.DAT already long.ADV want.3SG buy.INF new bike  
 ‘Lucia has been wanting to buy herself a new bike for a long time now.’

In contrast to Franks and King’s (2000) claim, Slovak does not only permit climbing with modal verbs, but also with the phasic verb ‘begin’. All four verb groups behave the same way, as shown in figure 10.5: climbing to the second position of the main clause is judged good, but the two positions within the embedded clause are degraded. The results are thus very similar to those for Czech, and in both languages, the verb groups that are generally prone to exhibit climbing according to Spencer and Luís (2012) (namely modal verbs, verbs of desire, and phasic verbs) indeed all feature obligatory climbing.

Also similarly to the Czech data, two items, those with the control verb *odporučiť* ‘recommend’, had to be analysed separately because their behaviour was fundamentally different. In Czech, the results show that with such object control verbs, climbing is not possible, as discussed in section 8.3.3. Whereas the sentences in 236 pattern with the Czech results for this verb type (i. e. sentence 236c was rated low), the sentences in 237 received positive z-scores in all conditions: 237a (emb-1P) was rated 0.6, 237b (emb-2P) 1.5, and 237c (main-2P) 1.1. Note that the Slovak materials were exact translations of the Czech ones.

- (236) a. *Odporučila priateľom sa naučiť programovať.*  
 recommend.PTCP friend.DAT.PL REFL.ACC learn.INF program.INF
- b. *Odporučila priateľom naučiť sa programovať.*  
 recommend.PTCP friend.DAT.PL learn.INF REFL.ACC program.INF
- c. *Odporučila sa priateľom naučiť programovať.*  
 recommend.PTCP REFL.ACC friend.DAT.PL learn.INF program.INF  
 ‘She recommended her friends to learn how to programme.’

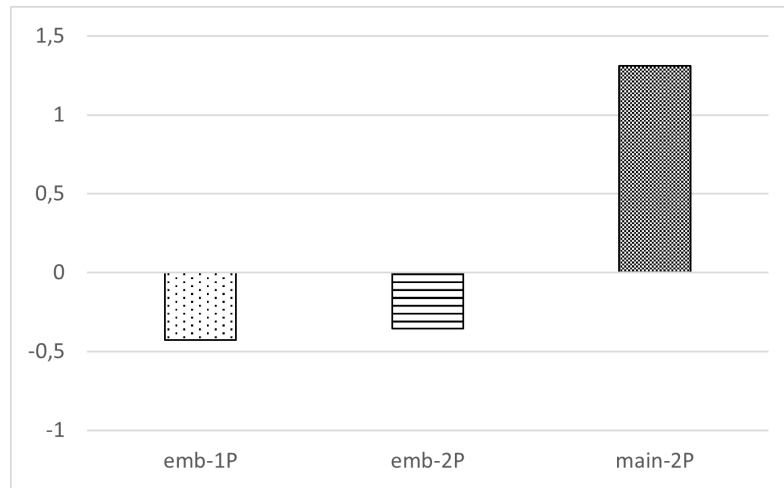


Figure 10.5: Acceptability judgements (z-scores) for different positions of clitics in a clause complex: climbing to main-2P or placement in emb-2P or emb-1P (N=24)

- (237) a. *Odporučil im sa umyt.*  
 recommend.PTCP them.DAT REFL.ACC wash.INF
- b. *Odporučil im umyt sa.*  
 recommend.PTCP them.DAT wash.INF REFL.ACC
- c. *Odporučil sa im umyt.*  
 recommend.PTCP REFL.ACC them.DAT wash.INF

‘He recommended them to wash themselves.’

Thus, Slovak clitic climbing is generally very similar to Czech, and can be analysed accordingly: An EDGEMOST constraint requires the clitics to climb to the highest TP within their CP. Due to NONINITIAL(Cl,CP) the clitics do not appear in first position. Climbing out of CPs follows from an interaction of this constraint with phase impenetrability: in order to leave CP, the clitic would have to go through the CP edge, from which it is banned. For the full analysis, see sections 8.3.1 and 8.3.2.

Yet, example 237 additionally shows the optionality of climbing mentioned by Junghanns (2002a, 82): here, we not only find the puzzling possibility of placing the clitic initially in the embedded clause (although with a lower rating than the other two options), but also optional climbing to the main clause. This difference to Czech and the surprising contrast between examples 236c ( $z=-0.9$ ) and 237c ( $z=1.1$ ) shows that despite the large overlap between the two languages, Slovak deserves a separate description and analysis.

### 10.1.7 Conclusions

We have seen that there are only minor differences in the clitic inventory between Czech and Slovak. The basic patterns of clitic placement in the clause are the same in the two languages: clitics tend to appear in second position after an initial phrase, verb, or complementiser. Possibly, reflexives enjoy a slightly greater degree of freedom than personal pronouns. Slovak also possesses

a position between the complementiser and the clitic that is targeted by information-structurally fronted phrases, possibly, but less clearly, including fronted VPs.

We have also seen that the Slovak clitics, like the Czech ones, defy the traditional phonological definition by Zwicky (1977) and others in that they do not require a prosodic host. Therefore, their 2P placement cannot be explained prosodically, but must be seen as a syntactic mechanism. Like Czech, we can capture Slovak clitic placement through conflicting constraints which require them to be relatively high in the clause, but not initial.

The Slovak invariant conditional marker *by* represents an interesting analogy to our analysis of the Czech conditional auxiliary: whilst the latter is a combination of two clitics, the former has to combine with the past auxiliary, thereby yielding the same structure with the same implications as described for Czech in section 7.4. Further investigation is required to establish whether this difference is more than just an orthographical one, at least with respect to the Czech colloquial conditional clitics.

Finally, we have seen that Slovak also resembles Czech regarding clitic climbing, which permits an analysis where phase impenetrability and NONINITIAL(Cl,CP) interact in such a way that they prevent clitics from climbing out of CP. Slovak object control verbs also display variability in clitic positioning. Such variation is a challenge to many linguistic frameworks and deserves further investigation.

## 10.2 Upper Sorbian: 2P with variability

As is typical for the Slavic languages, Sorbian also displays strong differences between standard and colloquial languages. In order to limit the space of variation, I will focus exclusively on Upper Sorbian here. It has roughly twice as many speakers as Lower Sorbian, and there is more literature available (cf. Franks & King, 2000, 163).

The Upper Sorbian colloquial language is described for example by Scholze (2012). It is a variety between the standard and the catholic dialect, which is now only spoken by the elderly. The Upper Sorbian colloquial language is used in every-day communication in a number of villages in the district of Bautzen (sorbian *Budyšin*) in eastern Germany, not far from the border triangle with Poland and the Czech Republic. The Upper Sorbian standard language is based on the Bautzen dialect and more conservative and static than the colloquial language. All speakers are bilingual with German (cf. Scholze, 2012, 198-9).

Whilst the German language has had a continuous influence on Sorbian, normativisation in both Sorbian languages has been oriented towards Slavic, especially Czech (cf. Franks & King, 2000, 163). The consequence is that in many respects, the standard language is more similar to Slavic, for example regarding the lack of articles, the aspectual system, and subject drop. Colloquial Upper Sorbian, on the other hand, has more similarities with German: it has an article system, no subject drop, and adopted the German impersonal construction with *man* ‘one’. But interestingly, there are also features in which the colloquial language is closer to Czech than the standard, especially regarding the position of auxiliary and pronominal clitics (cf. Scholze, 2012, 199-210). This distinction will therefore be of crucial interest in the following subsections.

	COP/PRF	COND	FUT	PRET	PASS
1SG	sym	bych	budu	běch	buch
2SG	sy	by	budžeš	běše/bě	bu
3SG	je	by	budže	běše/bě	bu
1DU	smój	bychmoj	budžemoj	běchmoj	buchmoj
2DU	staj/stej	byštaj/-ej	budžetaj/-ej	běštaj/-ej	buštaj/-ej
3DU	staj/stej	byštaj/-ej	budžetaj/-ej	běštaj/-ej	buštaj/-ej
1PL	smy	bychmy	budžemy	běchmy	buchmy
2PL	sće	byšće	budžeće	běšće	bušće
3PL	su	bychu	budžeja/budu	běchu	buchu

Table 10.3: Upper Sorbian copular and auxiliary clitics

For the description of the clitic inventory, Franks and King (2000) are a valuable reference, since in much of the literature on Sorbian, the term *clitic* is used rather vaguely. However, many of the sources that Franks and King (2000) rely on were already several decades old when their handbook was published. I therefore additionally consulted newer literature, such as Kaiser and Scholze (2009) for Standard Upper Sorbian, Scholze (2007) for Colloquial Upper Sorbian, and Scholze (2012) for a comparison of the two.

I will first provide a description of the clitic inventory of Upper Sorbian in section 10.2.1. The basic placement patterns of clitics will be discussed in section 10.2.2, their ordering is addressed in section 10.2.3, and their placement in third position in section 10.2.4. Section 10.2.5 considers the role of prosodic boundaries for clitic placement, and section 10.2.6 provides the conclusions.

### 10.2.1 The clitic inventory

Like Czech and Slovak, Upper Sorbian possesses both auxiliary clitics and pronominal clitics. As in the other two languages, a crucial question is whether an element is a lexical clitic or whether it can just optionally appear in the second position. Table 10.3 presents the paradigm of the verb *być* ‘to be’ as provided by Franks and King (2000, 166). Interestingly, in contrast to Czech and Slovak, the entire paradigm appears to show the same behaviour, thus also the copular, future, and passive forms are candidates for lexical 2P clitics. Since Upper Sorbian, in contrast to its siblings, also has a preterite tense and a dual number, the verbal clitic inventory could potentially be much larger here. Also unlike in Czech, the perfect auxiliary has no gap in the third person, so it is displayed together with the copula in one column.

Table 10.3 also shows that Upper Sorbian, like Czech (and unlike Slovak), has bisyllabic forms, and additionally even trisyllabic ones. Whilst Franks and King (2000) view this as potentially problematic (cf. Franks & King, 2000, 167-8), I have argued in section 5.1.3 that even clitics that consist of several syllables could be prosodically dependent if they cannot project prosodic feet. Moreover, I have shown for Czech and Slovak that even monosyllabic clitics are in fact prosodically independent; this will be discussed for Upper Sorbian in section 10.2.5.

The examples in 238 show 2P placement of the perfect (238a), future (238b), and passive (238c) auxiliaries, and of the present tense copula (238d). In example 238b, the matrix clause additionally shows the copular use of the future auxiliary with 2P placement. However, sentences

with subject drop in the standard language also permit these elements to appear initially in the clause; two examples are given in 239. This phenomenon will be discussed in section 10.2.2.

- (238) a. *Ja sym z lékařom porěčala.*  
 1SG PRF.1SG with doctor.INS speak.PTCP  
 ‘I have spoken with the doctor.’
- b. *Potom budže jasne, što budu tam pokazować.*  
 then FUT.3SG clear what FUT.1SG there show.INF  
 ‘Then it will be clear what I will show there.’
- c. *Jeho mandželska bu z pjećomi džěćimi žohnowana.*  
 his wife PASS.2SG with five.INS.PL child.INS.PL bless.PTCP  
 ‘His wife was blessed with five children.’
- d. *Wokna su wubite.*  
 window.PL COP.3PL knocked.out  
 ‘The windows are knocked out.’

(cf. Franks & King, 2000, 164-5)

- (239) a. *Bychmy na nju sluchale, kaž hdy by bajki bala.*  
 COND.1PL on her.ACC listen.PTCP as when COND.3 story.ACC.PL tell.PTCP  
 ‘We would listen to her, as we used to do when she was telling fairy tales.’
- b. *Sym we wšelakich serbskich šulach našeje Lužicy  
 z wučerjom był.*  
 PRF.1SG in various.LOC.PL Sorbian.LOC.PL school.LOC.PL our.GEN Lusatia.GEN  
 with teacher.INS be.PTCP  
 ‘I have been a teacher in various Sorbian schools of our Lusatia.’

(cf. Franks & King, 2000, 164-5)

A potentially clitic verbal element that Franks and King (2000) do not mention is *měć* ‘have’. It is described as clitic by Kaiser and Scholze (2009); they provide the example in 240, where *mam* appears right after the first constituent of the embedded clause (cf. Kaiser & Scholze, 2009, 315). Also the first sentence of the article’s abstract contains a form of *měć* in second position (cf. Kaiser & Scholze, 2009, 305). Crucially, this verb appears after the first constituent, but before the lexical reflexive clitic *so*, making it plausible to assume that it is part of the clitic cluster. The distribution of modal verbs will be discussed in detail in section 10.2.2.

- (240) *...podžěl zamóženja, na kotryž mam prawo*  
 ...part inheritance.GEN on REL have.1SG right  
 ‘...part of the inheritance to which I have a right’

	ACC/GEN	DAT
2SG	<b>će</b>	<b>ći</b>
3SG.F		<b>(ji)</b>
REFL	<b>so</b>	<b>sej</b>

Table 10.4: Upper Sorbian pronominal lexical clitics

(cf. Kaiser &amp; Scholze, 2009, 315)

- (241) *W tutym přinošku ma so na zakładže bibliskeho teksta*  
 in this.LOC article.LOC have.3SG REFL.ACC on basis.LOC biblical.GEN text.GEN  
*městno werba w sadže [...] wopisać.*  
 place verb.GEN in sentence.LOC describe.INF

‘In this article, the position of the verb in the sentence [...] will be described on the basis of a biblical text.’

(cf. Kaiser &amp; Scholze, 2009, 305)

A further similarity to Czech and Slovak is that whereas the auxiliaries do not have strong, non-clitic counterparts, the pronominal system contains both strong and weak forms. Thus, a set of elements can be identified that consists only of these pronouns which have strong counterparts, and which are not homophonous with any strong or unpaired pronouns. The inventory of lexical pronominal clitics thus derived from Franks and King’s (2000) data is given in table 10.4. As in Czech and Slovak, the unstressed reflexive pronouns are clitic, but unlike in those languages, the set of lexically clitic pronouns is smaller: only the second person singular accusative, genitive and dative forms are clitic, and the literary and slightly archaic feminine singular dative pronoun. Other persons and cases, as well as all dual and plural pronouns, only have unpaired forms (cf. Franks & King, 2000, 170).

A problem also encountered in Czech appears to be more severe in Upper Sorbian: many elements are called “clitics”, and it is often not clear whether this means that an element can optionally appear in 2P, or whether it is a lexical clitic. For example, the strong pronominal forms *jeho* ‘him.ACC/GEN’ and *jemu* ‘him.DAT’ are labelled as clitics, but Kaiser and Scholze (2009) note that there is much agreement about these and other potentially clitic elements. For example, for *jemu*, the authors find both 242a and 242b in a modern translation of the bible; and for *jeho*, Franks and King (2000) provide the 1P example in 243. They also note that “any nonpaired pronoun” (Franks & King, 2000, 172) can be a *potential* clitic – but as the data in 242 and 243 show, the nonpaired masculine singular pronouns are not *lexical* clitics.

- (242) a. *Ale nan jemu rjekny:*  
 but father him.DAT say.3SG.PRET

‘But the father said to him.’



- b. *A syn rjekny jemu:*  
 and son say.3SG.PRET him.DAT  
 ‘And the son said to him.’

(cf. Kaiser & Scholze, 2009, 316)

- (243) *Jeho hišće widzał njejsym.*  
 him.ACC still see.PTCP NEG:AUX.1SG  
 ‘I still haven’t seen him.’

(cf. Franks & King, 2000, 173)

There are further elements that can potentially appear in clitic positions in Upper Sorbian, such as several short discourse particles like *pak* ‘but’ and *da* ‘then’. Like Czech, and unlike Slovak, the language also has a conditional complementiser *-li* ‘if’, which in all examples given by Franks and King (2000) appears after an initial verb. So it is plausible to assume that *-li* is distributed similarly as in Czech in that it imposes additional restrictions on its host. Since the status of both the particles and *-li* is not entirely clear, they will not feature in the following discussion. In contrast, Kaiser and Scholze (2009) describe the forms of the verb *być* ‘be’ as having “unambiguously clitic character” (Kaiser & Scholze, 2009, 317). I will thus focus on these elements, given in table 10.3, and the unambiguously clitic pronouns displayed in table 10.4, when describing the basic clitic placement patterns of Upper Sorbian in the following subsection.

### 10.2.2 Clitic placement: the basic pattern

As mentioned above, the standard and colloquial language differ in their treatment of the auxiliary and copula clitics: in standard Upper Sorbian, they can regularly appear in first position, as the examples in 244 illustrate. Scholze (2012) attributes this to the pro-drop property of the standard language: when the subject is dropped, the clitic ends up in first position. This is a striking contrast to Czech, where subject drop does not have this effect on clitic placement: instead, the clitic is realised lower under such circumstances (cf. sections 5.1.1 and 8.2.1). The colloquial language, in contrast, does not permit such initial placement; instead, the clitics appear in second position, as they would in Czech; see the examples in 245.

- (244) a. *Je we jstwě a spi.*  
 COP.3SG in room.LOC and sleep.3SG  
 ‘He is in the room and sleeping.’
- b. *Sym so mjeztym z wobydlerjemi rozmołwjala.*  
 PRF.1SG REFL.ACC meanwhile with inhabitant.INS.PL talk.PTCP  
 ‘In the meantime I have talked to the inhabitants.’

(cf. Scholze, 2012, 203-4)

- (245) a. *Ja sym dżensa moju famulaturu započinała.*  
 1SG PRF.1SG today my.ACC clinical.clerkship.ACC begin.PTCP  
 ‘I have started my clinical clerkship today.’
- b. *Por dny pozdžišo su woni po mne bóli.*  
 few days later PRF.3PL they at me.GEN be.PTCP  
 ‘A few days later they were with me.’

(cf. Scholze, 2012, 203-4)

A general problem encountered in the literature is the lack of negative data, since showing that an element frequently occurs in second position is not the same as showing that it *must* appear there. But indeed, according to Lenka Scholze (p. c.), placing the auxiliary in a lower position, such as in 246b and 246c, is not acceptable – the only position the clitic can appear in is second, as in 246a. And interestingly, with pro-drop in the standard language, the clitic *has* to appear in initial position, as in 247a. Lower clitic placement in 2P, as it would occur in Czech and Slovak, is not acceptable here. This is illustrated by 247b and 247c.

- (246) a. *Mój nan je jedne nowe awto kupił.*  
 my father PRF.3SG INDEF new car buy.PTCP  
 ‘My father has bought a new car.’
- b. *\*Mój nan jedne nowe awto je kupił.*  
 my father INDEF new car PRF.3SG buy.PTCP
- c. *\*Mój nan jedne nowe awto kupił je.*  
 my father INDEF new car buy.PTCP PRF.3SG
- (247) a. *Je nowe awto kupił.*  
 PRF.3SG new car buy.PTCP  
 ‘He has bought a new car.’
- b. *\*Kupił je nowe awto.*  
 buy.PTCP PRF.3SG new car
- c. *\*Nowe awto je kupił.*  
 new car PRF.3SG buy.PTCP

Also the positioning of modals verbs mentioned in the previous subsection can only be described fully with negative data. According to Lenka Scholze (p. c.), in both colloquial and standard language, the modal verbs *chcyć* ‘want’, *směć* ‘may’, *móc* ‘can’, *dyrbjeć* ‘must’, and *měć* ‘have’ all show the same distributional restrictions in the clause: they behave like 2P clitics in the sense that they must appear second in the clause as in 248a, and not in third position as in 248b. This applies to both standard and colloquial language. However, in a context where subject drop is licensed in the standard language, the modal can regularly appear initially, as

shown in 248c. Moreover, in both standard and colloquial language, polar questions are regularly formed with an initial modal verb, as can be seen in 248d. This means that even in the colloquial language, the 2P behaviour of the modals is overwritten by the rules of question formation.

- (248) a. *Dana chce/ smě/ může/ dyrbi/ ma sad kupić.*  
 Dana want.3SG may.3SG can.3SG must.3SG have.3SG fruit buy.INF  
 ‘Dana wants to/ may/ can/ must/ should buy fruit.’
- b. *\*Džensa Dana chce/ smě/ může/ dyrbi/ ma sad kupić.*  
 today Dana want.3SG may.3SG can.3SG must.3SG have.3SG fruit buy.INF  
 ‘Today Dana wants to/ may/ can/ must/ should buy fruit.’
- c. *Chce/ Smě/ Může/ Dyrbi/ Ma sad kupić.*  
 want.3SG may.3SG can.3SG must.3SG have.3SG fruit buy.INF  
 ‘She wants to/ may/ can/ must/ should buy fruit.’ (only standard language)
- d. *Chce/ Smě/ Může/ Dyrbi/ Ma Dana sad kupić?*  
 want.3SG may.3SG can.3SG must.3SG have.3SG Dana fruit buy.INF  
 ‘Does Dana want to buy fruit? / May/ Can/ Must/ Should Dana buy fruit?’

But crucially, also auxiliary verbs from the *być*-paradigm appear in initial position in polar questions, as shown in 249a. Franks and King (2000) argue that in this context, these elements must then be strong forms, because they can host the conditional clitic *-li*, as can be seen in 249b (although the use of *-li* in a matrix clause appears to be disputed; cf. Franks & King, 2000, 184, fn. 97). Additionally, as illustrated in 249c, they can appear as one-word answers. This means that, despite the fact that they have such a strong tendency towards the second position, we cannot simply label the elements given in table 10.3 as lexical clitics. It appears that other grammatical constraints can not only override their positional restrictions, but also their clitic status per se.

- (249) a. *Sy to hižo rozsudžil?*  
 PRF.2SG that already decide.PTCP  
 ‘Have you already decided that?’
- b. *Sy-li tu?*  
 COP.2SG-Q here  
 ‘Are you here?’
- c. *Sy to hižo rozsudžil? – Sym.*  
 PRF.2SG that already decide.PTCP PRF.1SG  
 ‘Have you already decided that? – Yes, I have.’

(cf. Franks & King, 2000, 183-4)

Regarding the placement of reflexive pronouns *so* and *sej*, Scholze (2012) observes that they are tied to the second position only in the colloquial language (cf. examples 250a and 250c),<sup>1</sup> whereas in the standard language, they may appear later (cf. the embedded clause in 250b) or initially (cf. 250d). Note that here, as in Czech, the conjunction *a* ‘and’ does not count as the initial constituent, arguably because it is extracausal.

(250) a. *Jej      jo      so      to   lubilo,      ka   jo      so      wón po*  
 her.DAT PRF.3SG REFL.ACC that please.PTCP how PRF.3SG REFL.ACC he in  
*jeje      měsće   pohíwał.*  
 her.GEN town.LOC move.PTCP

b. *Jej      so      lubješe,      kak wón so      po jeje      měsće*  
 her.DAT REFL.ACC please.PRET.3SG how he REFL.ACC in her.GEN town.LOC  
*pohibowaše.*  
 move.PRET.3SG

‘She liked how he moved around her town.’

c. *...a ja   sym      so      domóslil,      zo   leža      wone za*  
 and 1SG PRF.1SG REFL.ACC determine.PTCP COMP lie.PRS.3PL they after  
*přewrótom.*  
 turn.INS

d. *...a so      domyslich,      zo   wone leža      za   přewrótom.*  
 and REFL.ACC determine.PRET.1SG COMP they lie.PRS.3PL after turn.INS

‘...and I determined that they lie after the fall of the Berlin Wall.’

(cf. Scholze, 2012, 209)

I want to conclude this section with a comment on the parallels between Upper Sorbian and German word order. Since the auxiliaries behave like 2P clitics in many respects, and the non-finite main verb tends to appear in final position, there is strong resemblance between the Sorbian structure and the German “Klammerkonstruktion” (bracket construction). However, German sentence structure is not simply copied, because this bracketing also occurs in embedded clauses in Upper Sorbian, where it is clearly ungrammatical in German (cf. Kaiser & Scholze, 2009, 317). The sentences in 251 illustrate this.

(251) a. *...zo   bych      so      ze   swojimi      přećelemi   powjeselil*  
 COMP COND.1SG REFL.ACC with POSS.REFL.INS.PL friend.INS.PL rejoice.PTCP

<sup>1</sup>In Scholze (2012), only the standard language sentences are given in full, whereas for the colloquial language, just the different position of the reflexive is indicated. I would like to thank Lenka Scholze for providing me with the full colloquial equivalents.

- b. ...*dass* (\*würde) *ich* (\*würde) *mich* *mit* *meinen* *Freunden* *erfreuen*  
 COMP 1SG me.ACC with my.DAT.PL friend.INS.PL rejoice.INF  
würde  
 would.1SG

‘...that I would rejoice with my friends.’

(cf. Kaiser & Scholze, 2009, 317)

Note that the tendency towards 2P might be stronger in present-day Upper Sorbian than in older texts. In a bible translation from 1893, instances of verb-final placement can be found where also the auxiliary occurs at the right edge; either preceding the main verb, as in 252a, or following it, as it would be in German, cf. 252b. In addition, Scholze (2012) reports final positioning of the copula in the standard language; whilst it can also occasionally occur in colloquial speech, it is “exceedingly rare” there (Scholze, 2012, 205).

- (252) a. ...*kiž* *twoju* *žiwnosć* *z* *kurwami* **je** *póžrěl*  
 REL your.ACC farmstead.ACC with prostitute.INS.PL PRF.3SG devour.PTCP

‘...who has squandered your property with prostitutes’

- b. *Ale* *hdyž* *tón* *twój* *syn* *přišol* **je...**  
 but when that your son come.PTCP PRF.3SG

‘But when your son came...’

(cf. Kaiser & Scholze, 2009, 318-9)

### 10.2.3 Ordering in the cluster

Several Upper Sorbian sentences that I have discussed here already illustrate part of the ordering relations within the cluster: examples 244b, 250a, 250c, and 251a show that auxiliaries precede reflexives. In addition, example 241 shows that modal verbs also precede reflexives. Regarding personal pronouns, Franks and King (2000) show that they follow auxiliaries, as can be seen in 253. Their internal ordering appears to be dative before accusative, although this is more difficult to establish in Upper Sorbian than in Czech and Slovak due to the very small inventory of unambiguous pronominal 2P clitics (*je* ‘it/them’ is unpaired). The example also shows that when discourse particles integrate into the cluster, they appear initially; yet, they are not tied to the second position (cf. Franks & King, 2000, 177).

- (253) *Ja* *wšak* ***sym*** ***ći*** *je* *rjenje* *wumyl.*  
 1SG however PRF.1SG you.DAT them.ACC nicely wash.PTCP

‘But I washed it/them nicely for you.’

(cf. Franks & King, 2000, 177)

On the basis of Franks and King's (2000) analysis, I propose the Upper Sorbian template in 254. I slightly deviate from the template they propose by explicitly including copulas and modal verbs into the same slot as auxiliaries. The general pattern is the same as in Czech and Sorbian, as described in sections 4.3 and 10.1.3, respectively. The ordering within the cluster is thus clearly a Slavic property of Upper Sorbian, since it aligns with that of its sibling languages, and at the same time does not reflect the ordering of auxiliaries and pronouns found in German (cf. Franks & King, 2000, 185-6).

(254) Clitic template for Upper Sorbian:

(PARTICLES) – AUX/COP/MOD – REFL – DAT – ACC

#### 10.2.4 Third-position placement

I have shown in section 10.2.2 that despite the possibility for 1P placement, Upper Sorbian clitics are not placed freely in the clause: apart from sentences with subject-drop and polar questions, clitics appear in 2P, and not lower. What remains to be investigated is whether 3P constructions as in Czech and Slovak, with a fronted phrase following the complementiser and preceding the clitic, are available here as well. According to Lenka Scholze (p.c.), the answer appears to be yes: whilst the ideal position for the clitic is directly after the complementiser, as in 255a and 256a, a phrase can be fronted into the position between these elements, be it the subject (255b and 256c), the direct object (255c), or an adjunct phrase (256b). This applies to both colloquial and standard language. The fronting appears to be conditioned by information structure. For example, sentence 255c is best as a reaction to the statement *Peter has bought this old house*. – Then, *tutón stary dom* is a topic in 255c and *Jakub* is a contrastive focus: *No, I've heard that this old house was bought by JAKUB*.

(255) a. (Ja) *sym slyšal, zo je Jakub tutón stary dom kupil.*  
 1SG PRF.1SG hear.PTCP COMP PRF.3SG Jakub that old house buy.PTCP

'I've heard that Jakub has bought that old house.'

b. (Ja) *sym slyšal, zo Jakub je tutón stary dom kupil.*  
 1SG PRF.1SG hear.PTCP COMP Jakub PRF.3SG that old house buy.PTCP

c. (Ja) *sym slyšal, zo tutón stary dom je Jakub kupil.*  
 1SG PRF.1SG hear.PTCP COMP that old house PRF.3SG Jakub buy.PTCP

(256) a. (Ja) *sym slyšal, zo je so při wohenju wohnjowy  
 1SG PRF.1SG hear.PTCP COMP PRF.3SG REFL.ACC during fire.LOC fire.ADJ  
 wobornik zranil.*  
 recruit wound.PTCP

'I've heard that a firefighter was wounded during the fire.'

- b. (Ja) *sym* *slyšal*, *zo* *při* *wohenju* *je* *so* *wohnjowy*  
 1SG PRF.1SG hear.PTCP COMP during fire.LOC PRF.3SG REFL.ACC fire.ADJ  
*wobornik zranil*.  
 recruit wound.PTCP
- c. (Ja) *sym* *slyšal*, *zo* *wohnjowy* *wobornik* *je* *so* *při*  
 1SG PRF.1SG hear.PTCP COMP fire.ADJ recruit PRF.3SG REFL.ACC during  
*wohenju zranil*.  
 fire.LOC wound.PTCP

Sentence 256c is perceived as less felicitous than sentence 256b. This might be due to the fact that *při wohenju* can easily be understood as a situation topic, i. e. the fire is already under discussion. In contrast, the sentence lends itself most easily to an interpretation where *wohnjowy wobornik* ‘firefighter’ is indefinite and the information focus, thereby favouring a position further to the right in the clause.

It thus appears that as in Czech and Slovak, we find 3P placement in Upper Sorbian when an element is fronted into the left periphery in an embedded clause. This is a crucial finding, because it suggests that as in the other two languages, clitics in Upper Sorbian do not have a fixed position in the C head. If they did, they would adjoin to the complementiser, and nothing would be able to intervene between these two elements.

Interestingly, the abstract to Kaiser and Scholze’s (2009) article also contains a construction which is not acceptable in the sibling languages: the sentence in 257 features a relative pronoun followed by a verb, then a potentially clitic particle, and only then comes the lexically clitic reflexive.

- (257) ...*we wšelakich* *pozicijach*, *kotrež* *dadža* *pak* *so* *jenož* *zdzěla*  
 in various.LOC.PL position.LOC.PL REL give.3PL however REFL.ACC only partly  
*na wliw němčiny wróco wjesć*.  
 on influence German.GEN back lead.INF

‘...in various positions, which can however only partly be traced back to the influence of German.’

(cf. Kaiser & Scholze, 2009, 305)

There are two possible explanations for this finding: One is that we are dealing with a case of lower placement of the reflexive pronoun as described for the standard language in section 10.2.2 (see example 250b there).<sup>2</sup> The other is that the verb *dać* ‘give’, which functions as a modal verb in this sentence (in the sense of ‘let’), is, like other modal verbs, also a 2P clitic, and the clitic cluster then consists of the entire sequence modal–particle–reflexive. Neither of these options can be excluded at the present point, and further research into the distribution of these elements is required to settle this issue.

<sup>2</sup>Regarding sentence 250b, one could argue that it represents a case of 3P as also found in Czech and Slovak: *kak* ‘how’ is an interrogative in the C-domain, and the subject pronoun *wón* ‘he’ is fronted into a position below C in the left periphery. Whilst I cannot completely exclude this interpretation, I deem it unlikely because Scholze (2012) describes this as a peculiarity of the standard language which is not permitted in colloquial Upper Sorbian. However, the 3P constructions described in the present subsection are independent of variety.

### 10.2.5 Prosodic enclitics or not?

We have seen that in both Czech and Slovak, despite their placement restrictions, clitics do not require a prosodic host. They are thus neither prosodically enclitic nor proclitic, at least not obligatorily. I have argued that due to this fact, Czech and Slovak clitic positioning should not be modelled through prosodic constraints. We have also seen that clitic placement in these languages can indeed be captured with reference only to syntactic constituency. The crucial sentences presented for Czech by Junghanns (2003) were discussed in section 5.1.3, their translation into Slovak in section 10.1.4.

It turns out that the Upper Sorbian language patterns with its two West Slavic siblings regarding the prosodic dependency of clitics: according to Lenka Scholze (p. c.), the sentences can be translated into the language<sup>3</sup> with the same clitic positions: following a pause after a relative clause and preceding a pause before a parenthetical. This applies to both the colloquial and the standard language. The placement of the clitic between two prosodic breaks indicates that Upper Sorbian clitic placement is not driven by the clitic's requirement for a prosodic host, but is instead a syntactic phenomenon.

- (258) a. *Dopokazy, kotrež je Marija předpoložila, će, kaž widžu,*  
 evidence REL PRF.3SG Marija present.PTCP you.ACC as see.1SG  
*njepřesvědča.*  
 NEG:convince.3PL

‘As I can see, the evidence that Marija has presented does not convince you.’

- b. *Knihy, kotrež tu widžiće, so, kaž so twjerdži, džensa ze*  
 books REL here see.2PL REFL.ACC as REFL.ACC claim.3SG today with  
*zlotom zaplaća.*  
 gold.INS pay.INF

‘They claim that the books you can see here are paid for with gold today.’

- c. *Tutón čłowjek, kiž tu bydli, by, kaž so twjerdži, móhl*  
 this man REL here live.3SG COND.3SG as REFL.ACC claim.3SG can.PTCP  
*terrorist być.*  
 terrorist be.INF

‘They claim that the man who lives here might be a terrorist.’

### 10.2.6 Conclusions

Whilst at first glance, the clitic system of Upper Sorbian seems very similar to that of its neighbour Czech, the preceding subsections have uncovered some relevant differences: First, the inventory of clitic verbs might be much larger in Upper Sorbian, including all forms of *być* ‘to

<sup>3</sup>The sentences were altered slightly in order to take into account the differences regarding pro-drop and the clitic inventory. Lenka Scholze (p. c.) also notes that sentences 258a and 258b would be realised with imperfective aspect in the standard language (*presvědčaja* and *plaća*, respectively).



be' in their copular and auxiliary uses, and also all modal verbs. The set of unambiguously clitic pronouns, in contrast, is slightly smaller.

Second, whilst the respective ordering of the clitics is in principle the same as in Czech (and Slovak), there are some relevant differences regarding the clitics' placement in the clause: In the Upper Sorbian standard language, there might be a connection between subject drop and the possibility of clitic-first placement (cf. Scholze, 2012, 204). And in both varieties, the verbal clitics appear in first position in polar questions. For a constraint-based analysis, I propose that in such cases, there is a constraint that requires the finite verb to appear in a certain position (below the subject or initially), and this constraint is ranked higher than NONINITIAL. We thereby capture the differences between languages simply by reranking of constraints.

There is also data that suggests that reflexive pronouns might display greater freedom of positioning than the other clitics. This is a tendency that we have also observed in Slovak (cf. section 10.1.2). For Slovak, a potential explanation might lie in the close ties the reflexive pronouns have with the verb: they tend to appear adjacent to it because, especially with inherently reflexive verbs, they are not arguments, but are part of the verb's lexical entry. In Upper Sorbian, however, in neither of the two sentences with low reflexive placement is the reflexive adjacent to the verb. Thus, despite the apparent similarity, the driving forces behind this variability might be different in the two languages.

Finally, the Upper Sorbian "Klammerkonstruktion" (bracket construction) as well as the possibility for initial placement of auxiliaries show resemblances with German sentence structure. One could conclude from this that Upper Sorbian clitics behave like the German finite verb: they move to the C head, and since in most clause types, another element also obligatorily occupies the C-domain, 2P placement is most frequent. Yet, polar questions and sentences with a dropped initial subject leave only the finite verb/clitic in C, and thus it appears initially. Above, I have already argued with Kaiser and Scholze (2009) that there are crucial differences between German V2 and Upper Sorbian 2P placement. Moreover, the 3P data presented in section 10.2.4 is crucial: as argued there, the intervention of the fronted phrase between complementiser and clitic shows that the clitic is *not* located in C. Also note that such constructions would be impossible in German. I conclude from this that whilst the 2P property of Upper Sorbian clitics can be overwritten by other constraints, it is at its core nevertheless the same as in Czech and Slovak: conflicting constraints lead to varying output positions of the clitics.

It has also become obvious that the general variability and possibilities of lower and initial placement make it much more difficult to determine whether an element is indeed obligatorily or only optionally clitic. In the end, the discussion is also one of terminology: Does Upper Sorbian have clitics that can appear in initial position, or are these elements that can appear in initial positions not clitics at all? What is clear is that, with their prosodic independence, the term "clitic" in Upper Sorbian, and also in Czech and Slovak, does not mean exactly the same as it does in languages where these elements clearly attach to a host. An example of such a language, namely BCMS, will be discussed in the following section.

Whilst within the limitations of this thesis, acceptability judgment experiments for Upper Sorbian could not be conducted, the preceding subsections have made it clear that there is a strong need for such data: the different placement options would have to be tested separately for

each individual lexical item in order to establish its status. This would also permit a description of the current usage of Upper Sorbian, instead of having to rely on partly decade-old data from a language whose clitic system appears to be changing. As in the other Slavic languages, a potential confounding factor is the difference between colloquial and standard language, since it is difficult to ensure that participants refer to only one of these systems when answering a survey.

### 10.3 BCMS: prosody and integrity

In this chapter, I explore how the analysis of Czech 2P placement through violable constraints can be extended to a Slavic language that is not part of the West Slavic group, namely BCMS. The main objective is to describe how clitic patterns in this language differ from what is found in Czech, and, based on that, to explore the roles of syntax and of phonology for clitic positioning. Following Zimmerling's (2012) claim that "it is essential to explain all existing language-specific options for the placement of (clusterizing) clitics [...] as instantiations of one and the same underlying linearization principles" (Zimmerling, 2012, 2), I will show how our constraint-based approach to clitic placement can be applied to patterns found in BCMS, how the variation can be explained, and I will identify issues that require future investigation. I will first present the basic clitic placement pattern of BCMS in section 10.3.1. Section 10.3.2 then lays out the approach proposed by Anderson (2005), and section 10.3.3 introduces some basic notions of the matching between syntactic and prosodic constituency, which are crucial for modelling the BCMS placement patterns. The three following sections then address the domains and definitions of the central constraints *EDGEMOST* (section 10.3.4), *NONINITIAL* (section 10.3.5), and *INTEGRITY* (section 10.3.6) in turn. Finally, a potential problem for the analysis, the identification of the *tP* in which the clitics are positioned, is investigated in section 10.3.9. Section 10.3.10 then provides a concluding summary.

#### 10.3.1 Clitic placement: the basic pattern

Similarly to Czech, the BCMS inventory of 2P clitics includes auxiliaries, pronominals, and the interrogative *li*. I will not go into the status of *li* any further here; yet, similarly to what we find in Czech and Upper Sorbian, this element also imposes additional restrictions on its host (cf. Franks & King, 2000, 26). The verbal auxiliary clitics comprise the past tense auxiliary and, unlike in Czech, the future tense auxiliary. Another difference to Czech is that the copula, which is isomorphic to the past auxiliary, is also a 2P clitic, as in Upper Sorbian. Pronominal clitics are, again as in Czech, found in the accusative, genitive, and dative cases; yet, the number of pronominal 2P clitics is larger than in Czech, since all persons and numbers display specific clitic forms (cf. Franks & King, 2000, 19-25).

We have seen in section 6.2, which has presented Radanović-Kocić's (1996) prosodic account of clitic placement, that in BCMS, clitics are placed with respect to prosodic constituency: they appear in second position of their intonation phrase (*tP*; cf. Radanović-Kocić, 1996, 441). This is a striking contrast to Czech, Slovak, and Upper Sorbian, where, as has become evident, only syntactic constituency is relevant. In (intonationally) simple sentences, the BCMS clitic

placement rules nevertheless yield a pattern superficially identical to the one found in Czech, as can be seen in example 259. In sentence 259a, the clitics follow the (pronominal) subject, in 259b they appear after the initial verb. Sentence 259c shows that also in BCMS, the complementiser constitutes a valid occupant of the first position.

- (259) a. *Mi smo mu je predstavili juče.*  
 we AUX.1PL him.DAT her.ACC introduce.PTCP yesterday
- b. *Predstavili smo mu je juče.*  
 introduce.PTCP AUX.1PL him.DAT her.ACC yesterday  
 ‘We introduced her to him yesterday.’
- c. *Ona tvrdi da smo mu je mi predstavili juče.*  
 she claims COMP AUX.1PL him.DAT her.ACC we introduce.PTCP yesterday  
 ‘She claims that we introduced her to him yesterday.’

(cf. Bošković, 2004, 38)

However, unlike in the West Slavic languages discussed so far, if a pause occurs after the first constituent, clitic placement is delayed, as example 260 (first introduced in section 6.2) with a restrictive relative clause and an appositive relative clause shows; the commas here indicate pauses. The delay occurs because the prosodic breaks which frame the appositive in 260b constitute *iP* boundaries. Compared to other Slavic languages, the behaviour of BCMS is exceptional: Zimmerling (2012) remarks that “in other Slavic languages parenthetical insertions do not change well-formedness conditions for the placement of clause-level clitics” (Zimmerling, 2012, 5).<sup>4</sup>

- (260) a. *Ona moja sestra koja je u Sarajevu vas se sjeća.*  
 that my sister REL is in Sarajevo you REFL remember.3SG  
 ‘My sister who is in Sarajevo remembers you.’
- b. *Moja sestra, koja je u Sarajevu, sjeća vas se.*  
 my sister REL is in Sarajevo remember.3SG you REFL  
 ‘My sister, who is in Sarajevo, remembers you.’
- c. *\*Moja sestra, koja je u Sarajevu, vas se sjeća.*  
 my sister REL is in Sarajevo you REFL remember.3SG

(cf. Radanović-Kocić, 1996, 437)

Other parentheticals also induce delayed clitic placement. Example 261 illustrates that their presence is crucial to the clitic’s position: placement after the subject occurs in 261a, when a

<sup>4</sup>Zimmerling (2012) also notes that the clitic distribution as described by Radanović-Kocić (1996) does not apply to all dialects of BCMS (cf. Zimmerling, 2012, 5). In contrast, Franks and King (2000) explain that “from the perspective of clitic phenomena, there is little variation among variants spoken in Serbia, Croatia, or Bosnia” (Franks & King, 2000, 18).

parenthetical intervenes between complementiser and subject; in 261b, however, we see that placement after the subject is not possible when this parenthetical is absent – then, the clitic must directly follow the complementiser, as in 261c. Example 262, repeated from section 6.2, shows that the same effect occurs with heavy initial constituents, which optionally form their own  $\iota$ Ps, i. e. which are optionally followed by a pause. When the sentence is uttered as one  $\iota$ P, the clitic appears after the heavy subject, but when a prosodic break occurs after *dima* ‘smoke’, the clitics appear one position further to the right.

- (261) a. *Znači da, kao što rekoh, oni će sutra doći.*  
 means COMP as said they will tomorrow arrive  
 ‘It means that, as I said, they will arrive tomorrow.’
- b. *\*Znači da oni će sutra doći.*  
 means COMP they will tomorrow arrive
- c. *Znači da će sutra doći.*  
 means COMP will tomorrow arrive  
 ‘It means that they will arrive tomorrow.’

(cf. Bošković, 2004, 39-41)

- (262) a. *Kolutovi plavičastog dima su se penjali.*  
 circles bluish smoke AUX.3PL REFL rise.PTCP
- b. *Kolutovi plavičastog dima penjali su se.*  
 circles bluish smoke rise.PTCP AUX.3PL REFL  
 ‘Bluish circles of smoke were rising.’

(cf. Radanović-Kocić, 1996, 435)

In order to develop an analysis of BCMS clitic placement, there are three central questions that we need to ask, and that will be addressed in the present chapter: First, what is the domain to whose left edge the clitics strive? More specifically, what is the domain of EDGEMOST in BCMS? Second, what prevents clitics in this language from appearing initially within that domain? And third, how can we define the first position, i. e. what is the nature of the entity that the clitics follow? Before I investigate these issues in more detail, I will present the analysis to BCMS clitic positioning by Anderson (2005) as well as some basic assumptions about the mapping of syntax and information structure to prosody.

### 10.3.2 Anderson’s (2005) analysis of BCMS

Anderson (2005) briefly takes up Richardson’s (1997) proposal for Czech and incorporates it into a cross-linguistic analysis. This includes a model of BCMS clitic placement using the constraints EDGEMOST and NONINITIAL – the same ones that he employs for Czech. What is crucial here is that the constraints refer to different domains in Czech and in BCMS (cf. Anderson, 2005,

150),<sup>5</sup> as can be seen in the definitions in 263. Whilst in Czech, NONINITIAL refers to CP, its domain in BCMS is  $\iota$ P. The domain of EDGEMOST, on the other hand, is TP in both languages.<sup>6</sup> For the OT architecture, this concurrence of phonology and syntax in clitic placement is, due to OT’s lack of distinct grammatical layers, entirely unproblematic.

- (263) a. NONINITIAL(Cl, $\iota$ P) = A clitic must not occur at the left edge of  $\iota$ P.  
 b. EDGEMOST(Cl,l,TP) = A clitic must occur as close to the left edge of TP as possible.  
 c. NONINITIAL(Cl, $\iota$ P) >> EDGEMOST(Cl,l,TP)

(cf. Anderson, 2005, 150)

The advantage of Anderson’s (2005) approach is that we can assume Czech and BCMS to function in entirely parallel ways with respect to clitics: they employ the same set of clitic-specific constraints – the only difference is the type of constituent, i. e. the domains, that these constraints refer to. This also means that we otherwise retain the formulation of the constraints, to the effect that in the analysis of BCMS, just as in Czech, the higher-ranked constraint is absolute, whilst the lower-ranked one is gradual, as was illustrated in section 8.2.2. This is necessary to derive the correct outcome in both languages; the effect is shown in the tableaux in 10.5 and 10.6, for simple main clauses (such as 259a) and for clauses with a heavy subject which induces a break (such as 262b), respectively.

	NONINI(Cl, $\iota$ P)	EDGEMOST(Cl,l,TP)
a. [TP [ $\iota$ P <b>cl</b> NP V ]]	*!	
☞ b. [TP [ $\iota$ P NP <b>cl</b> V ]]		*
c. [TP [ $\iota$ P NP V <b>cl</b> ]]		**!

Table 10.5: OT tableau illustrating clitic-second in simple clauses

	NONINI(Cl, $\iota$ P)	EDGEMO(Cl,l,TP)
a. [TP [ $\iota$ P <b>cl</b> heavy-NP ] [ $\iota$ P V ]]	*!	
b. [TP [ $\iota$ P heavy-NP ] [ $\iota$ P <b>cl</b> V ]]	*!	*
☞ c. [TP [ $\iota$ P heavy-NP ] [ $\iota$ P V <b>cl</b> ]]		**

Table 10.6: OT tableau illustrating clitic-third with a heavy subject

### 10.3.3 Constraints on prosodic constituency

Before we can undertake a more thorough review of the data, we must clarify what constitutes a  $\iota$ P. In the spirit of edge-based theories on the relation between syntax and prosody (for example Selkirk, 1986), I am assuming here that prosodic structure is built based on syntactic structure

<sup>5</sup>Note that although the domain specifications differ, the constraints themselves remain the same – they belong to the same constraint format, Generalised Alignment, which has language-specific arguments.

<sup>6</sup>Anderson (2005) labels EDGEMOST as LEFTMOST, which is of no consequence for the analysis. I will keep the constraint labels uniform for the Czech and BCMS analysis, for better readability.

on the basis of *MATCH* constraints. The relevant constraint for the building of *tP*s is given in 264; it requires that the edges of syntactic clauses coincide with the edges of intonational phrases. We have already seen the potential effects of this constraint for Czech in section 8.4.2.

- (264) *MATCH*<sub>CLAUSE</sub> = A clause in syntactic constituent structure must be matched by a corresponding *t*-phrase in prosodic representation.

(cf. Féry, 2017, 86)

But prosodic structure does not entirely mirror syntactic structure. First, embedding in prosody is more constrained and levels are subject to a stricter hierarchy. This is modelled through well-formedness constraints which capture the effects of the Strict Layer Hypothesis, i. e. the assumption that a constituent of level *n* may only directly dominate constituents of level *n*-1. Second, prosody is sensitive to rhythm, which is reflected in constraints such as *MINIMALBINARITY*, which requires a prosodic constituent to dominate at least two subconstituents (cf. Féry, 2017, 87-8). Since these two aspects will not play a decisive role for the data studied in the present chapter, I will not go into more detail here.

Another factor that has an influence on *tP* formation is length, measured in syllables or words: the longer a prosodic constituent is, the more likely a prosodic boundary is drawn (cf. Féry, 2017, 291-2). The effects of this are nicely visible in BCMS, where, as we have seen in 262, constituent length influences prosodic arrangements and through this, clitic placement. A final relevant factor is information structure: foci as well as topics have a strong tendency to be phrased independently, thus ideally they align with the edge of a *tP* (cf. Féry, 2017, 147-8). We will see in the following sections how such information-structurally grounded phrasing impacts clitic positioning in BCMS.

### 10.3.4 The effect of *EdgeMost*

As illustrated in section 10.3.2 above, Anderson (2005) proposes that the domain of *EDGEMOST* is the same in BCMS as it is in Czech. This also means that BCMS clitics refer in their placement to both syntactic and prosodic constituency. The present subsection investigates whether this view is justified, and then goes on to explore the effect that *EDGEMOST* has on clitic climbing in this language.

Remaining within a uniform analysis of Czech and BCMS, there is an alternative to Anderson's (2005) specification of the BCMS constraint domains: one might also assume that both *NONINITIAL* and *EDGEMOST* refer to *tP* in BCMS. Their ordering then prevents clitics from appearing at the left edge, and at the same time from appearing further to the right. In such an approach, one would thus reformulate *EDGEMOST* as in 265 – like this, clitic placement would be fully prosodic.

- (265) *EDGEMOST*(Cl,*l*,*tP*) = A clitic must occur as close to the left edge of *tP* as possible.

What kind of data do we need to decide between these two alternatives? The first type would be a TP that is divided into several *tP*s. If the clitics occur only within a lower *tP*, the constraint in 265 is at work; if the clitics appear in the highest *tP* within that TP, then we

should assume EDGEMOST(Cl,I,TP) as proposed by Anderson (2005). Consider in this respect the case of 267. Here, the clitics *su* and *se* occur in the higher  $\iota$ P, after the subject pronoun *oni* (which I assume to be above TP, given that BCMS is pro drop and an overt subject pronoun thus possibly focused), and not in the lower  $\iota$ P, in which case they would follow the main verb, *predstavili*. This is a first point in favour of Anderson's (2005) proposed constraint domain TP.<sup>7</sup>

- (267) [ $\iota$ P *Oni* [ $\text{TP}$  *su se*], [ $\iota$ P *kao što sam vam rekla*], [ $\iota$ P  
 they AUX.3PL REFL as AUX.1SG you.DAT say.PTCP  
*predstavili Petru*]].  
 introduce.PTCP Petar.DAT

'They have, as I told you, introduced themselves to Petar.'

(cf. Bošković, 2004, 53, fn. 15)

A second type of data that helps to differentiate between the two different approaches would be a  $\iota$ P that consists of several TPs. Such a case is instantiated by matrix clauses which are, due to their brevity, not followed by a prosodic break (arguably due to MINIMALBINARITY). This could plausibly be assumed for example 261c, repeated below as 268. If the domain of EDGEMOST is  $\iota$ P, we would expect the clitic to climb out of the embedded clause into the second position of the matrix clause, as in 268b. However, what we find is that the clitic remains within the second position of the embedded clause, immediately following the complementiser. Example 268a, if indeed constituting a single  $\iota$ P, could thus provide further evidence that clitics strive towards TP in BCMS – in other words, they are also sensitive to syntactic boundaries (cf. Bošković, 2001, 76-9). Even if a matrix clause is too short to form its own  $\iota$ P, the embedded clause's TP edge nevertheless plays a role for clitic placement.

- (268) a. *Znači da će sutra doći.*  
 means COMP will tomorrow arrive

'It means that they will arrive tomorrow.'

(cf. Bošković, 2004, 41)

- b. \**Znači će da sutra doći.*  
 means will COMP tomorrow arrive

<sup>7</sup>Bošković (2001) also discusses an alternative to 267, where the reflexive clitics remains within the lower  $\iota$ P. Interestingly, the sentence is tagged with a question mark in this monograph, but in Bošković (2004), this mark is lost:

- (266) <sup>(?)</sup>*Oni su, kao što sam vam rekla, predstavili se Petru.*  
 they AUX.3PL, as AUX.1SG you.DAT said, introduced REFL Petar.DAT

'They have, as I told you, introduced themselves to Petar.'

(cf. Bošković, 2004, 53)

The tableau in 10.7 shows how, if we assume that the entire clause complex is parsed into one  $\iota$ P, the  $\iota$ P-TP analysis correctly predicts 268a to be the winning candidate, whereas the  $\iota$ P- $\iota$ P analysis falsely predicts 268b to win, as shown in the tableau in 10.8. When the clitic directly follows the complementiser, it is ideally positioned at the edge of TP without being adjacent to a  $\iota$ P boundary. When it climbs higher, it is further from its own TP's edge without gaining anything with respect to NONINITIAL.

	NONINI(Cl, $\iota$ P)	EDGMO(Cl,l,TP)
$[\iota$ P [TP V [CP <b>cl</b> Comp [TP V ]]]]		* !
$[\iota$ P [TP V [CP Comp [TP <b>cl</b> V ]]]]		
$[\iota$ P [TP V [CP Comp [TP V <b>cl</b> ]]]]		* !

Table 10.7: OT tableau correctly predicting clitic placement in embedded clauses

	NOIN(Cl, $\iota$ P)	EDMO(Cl,l, $\iota$ P)
⊕ a. $[\iota$ P [TP V [CP <b>cl</b> Comp [TP V ]]]]		*
b. $[\iota$ P [TP V [CP Comp [TP <b>cl</b> V ]]]]		* * !
c. $[\iota$ P [TP V [CP Comp [TP V <b>cl</b> ]]]]		* * ! *

Table 10.8: OT tableau incorrectly predicting clitic placement in embedded clauses

Whilst part of the argumentation given here relies on assumptions about prosodic structure that are unfortunately not included in the data available to me, it still plausibly allows us to conclude that the assumption of TP as the domain for EDGEMOST is adequate. In contrast, the alternative view, namely that both NONINITIAL and EDGEMOST have prosodic domains, does not appear plausible for the discussed examples. Yet, in an OT architecture, the fact that clitic positioning refers to both syntactic and prosodic domains can be captured without cost: alignment constraints can refer to edges of either domain, and all constraints are evaluated in parallel.

Notice that the evaluation in tableau 10.7 implies that we must assume that BCMS clitics do not strive towards the *highest* TP, but only towards the TP that is an extended projection of the verb they belong to. This directly leads to the issue of clitic climbing. For Czech, I have argued in section 8.3 that the ban on clitic climbing out of CPs results from the interaction of NONINITIAL(Cl,CP) with CP-BARRIER – this counters the effect of EDGEMOST(Cl,l,MaxTP). In BCMS, the current constraint set predicts that not CPs, but already TPs, block climbing.

An interesting piece of evidence in this direction is clitic climbing out of indicative-like and subjunctive-like complements, as described by Progovac (1996). As example 269b shows, indicative-like verbs such as *kazati* ‘say’ do not permit clitic climbing; in contrast, climbing with subjunctive-like verbs such as *želeti* ‘wish’ is (to a certain extent) possible, as can be seen in 269d. Note that Progovac (1996) as well as Bošković (2001) point out that in all these sentences, the embedded clause forms a separate  $\iota$ P, thus the difference in the acceptability of climbing must have a syntactic cause. This is another strong clue that modelling BCMS clitic placement only with reference to prosodic constituents is not adequate.



- (269) a. *Milan kaže da ga vidi.*  
 Milan say.3SG COMP him see.3SG  
 ‘Milan says that he sees him.’
- b. \**Milan ga kaže da vidi.*  
 Milan him say.3SG COMP see.3SG
- c. *Milan želi da ga vidi.*  
 Milan want.3SG COMP him see.3SG  
 ‘Milan wants to see him.’
- d. <sup>?</sup>*Milan ga želi da vidi.*  
 Milan him want.3SG COMP see.3SG

(cf. Progovac, 1996, 423)

Now, crucially, Progovac (1996) shows that the second group, subjunctive-like verbs, only allows present-tense complements, i. e. that no independent expression of tense is possible on the embedded verb. In contrast, indicative-like verbs permit all types of temporal complements. This is illustrated in 270a-b for structures with subjunctive-like complements and in 270c for structures with indicative-like complements.

- (270) a. *Ne želim da idem.*  
 NEG wish.1SG COMP leave.PRS.1SG  
 ‘I don’t want to leave.’
- b. \**Ne želim da sam otišao / ću otići.*  
 NEG wish.1SG COMP AUX.1SG leave.PTCP will.1SG leave
- c. *Rekao sam da idem / sam otišao / ću otići.*  
 say.PTCP AUX.1SG COMP LEAVE.PRS.1SG AUX.1SG LEAVE.PTCP will.1SG LEAVE  
 ‘I said that I am going/have gone/will go.’

(cf. Progovac, 1996, 423)

The diverging acceptability of sentences 269b and 269d is compatible with an analysis in which subjunctive-like complements of verbs such as *želiti* ‘wish’ do not have a TP layer – for this reason, they allow no independent expression of tense on the verb (cf. Stjepanović, 2004, 189-90). In consequence, EDGEMOST(Cl,1,TP) favours clitic climbing out of these clauses into the second position of the matrix clause, thus avoiding the left edge of the first  $\iota$ P. The evaluation for sentence 269d is shown in the tableau in 10.9.<sup>8</sup>

<sup>8</sup>I am not considering potential intermediate landing sites of the clitics, which may be required for example for case assignment. Constraints such as CASE (cf. Grimshaw, 1997, 390-1) ensure that NPs appear in positions that allow them to be case-marked. However, since I assume that this is independent of the output realisation of clitics, I ignore this aspect here for the sake of simplicity.

	NI( <sub>i</sub> P)	EM(1,TP)	STAY
a. [ <sub>i</sub> P [TP NP V] [ <sub>i</sub> P [ <sub>v</sub> P da V <b>cl</b> ]]]		***	
b. [ <sub>i</sub> P [TP NP V] [ <sub>i</sub> P [ <sub>v</sub> P da <b>cl</b> V e <sub>l</sub> ]]]		**!	*
c. [ <sub>i</sub> P [TP NP V <b>cl</b> ] [ <sub>i</sub> P [ <sub>v</sub> P da V e <sub>l</sub> ]]]		**!	*
☞ d. [ <sub>i</sub> P [TP NP <b>cl</b> V] [ <sub>i</sub> P [ <sub>v</sub> P da V e <sub>l</sub> ]]]		*	*
e. [ <sub>i</sub> P [TP <b>cl</b> NP V] [ <sub>i</sub> P [ <sub>v</sub> P da V e <sub>l</sub> ]]]	*!		*

Table 10.9: OT tableau illustrating climbing out of subjunctive-like complements

As can be seen from the candidates in 10.9, I am assuming that subjunctive-like complement clauses not only lack a TP layer, but also a CP layer. This means that the element *da*, which is typically classified as a complementiser, is not a C head in these configurations. I hereby follow Stjepanović (2004), who argues that infinitival *da*-complements have no C-projection, based on multiple interrogative fronting data (cf. Stjepanović, 2004, 190). She concludes that “*da* in restructuring environments is not a C, but a counterpart of the English infinitive marker *to* or German *zu*” (Stjepanović, 2004, 204). Based on the placement of manner adverbs, she assumes that this variant of *da* is structurally located “in the vicinity of the verbal projections” (Stjepanović, 2004, 204). She does not specify its exact position, and neither will I, since this would require a deeper investigation of BCMS syntax, which exceeds the scope of the present dissertation. For my analysis, it suffices that *da* is neither a C- nor a T-element.

Generally, the fact that BCMS allows climbing out of clauses that have a complementiser, whereas the other Slavic languages only permit climbing out of complementiser-less clauses, becomes less irritating if *da* must not necessarily be a C-head. The exceptional status of BCMS is also noted by Franks and King (2000), who provide another example, this time with a causative complement introduced by *da*, as in 271. They write that there appears to be a greater “cohesion” between the matrix verb and its complement clause in such structures. We can conclude that this cohesion is the result of a reduced structural composition of the embedded clause.

- (271) *Marija ju je pustila da pliva.*  
 Marija it/her.ACC AUX.3SG let COMP swim  
 ‘Marija let it/her swim.’

(cf. Franks & King, 2000, 244)

A final point to be made with respect to climbing out of finite complements is its optionality. So far, I have presented an analysis of the climbing construction in 269d. Yet, for the alternative without climbing, sentence 269c, to arise, we must assume that the lack of the TP layer is optional. Thus, example 269c *does* have a TP projection in the embedded clause, and the clitic occupies its left edge. The same is true for the sentences with indicative-like complements in 269a-b – but here, due to the expression of tense on the verb, the TP layer is obligatory, and thus climbing is impossible.

I realise that this may not be a satisfying solution. Both in Czech and in BCMS, cases of optional climbing provide a significant challenge. Yet, this not only applies to the present

constraint-based approach: Stjepanović (2004) also points out that it is not clear why restructuring (i. e. the just discussed structurally reduced clausal complements) is a necessary condition for climbing, but not a sufficient one (cf. Stjepanović, 2004, 206-7). Clearly, more research is needed into the conditions that inhibit clitic climbing in contexts where it appears to be optional.

The (from a general Slavic perspective) more canonical cases of clitic climbing out of infinitival complement clauses without a complementiser, which we also find in Czech (cf. section 5.2.1), can be captured in the same way as climbing out of subjunctive-like complements in our analysis. An example is given in 272, where we see that clitic climbing to the second position of the matrix clause occurs obligatorily in these contexts. The evaluation of sentence 272b can be found in the tableau in 10.10. Yet, note that the analysis is contingent on the assumption that clauses which permit clitic climbing never have a TP layer in BCMS. In the data available to me through the literature, I have found no counter-evidence to this view.

- (272) a. *Želi ga vidjeti.*  
 want.3SG him see.INF  
 ‘She/he wants to see him.’
- b. *Jovan ga želi vidjeti.*  
 Jovan him want.3SG see.INF  
 ‘Jovan wants to see him.’
- c. *\*Želi vidjeti ga.*  
 want.3SG see.INF him

(cf. Bošković, 2001, 71)

	NoIN( <sub>i</sub> P)	EDMo( <sub>l</sub> ,TP)	STAY
a. [ <sub>i</sub> P [TP NP V <sub>fin</sub> [ <sub>v</sub> P V <sub>inf</sub> <b>cl</b> ]]]		**!*	
b. [ <sub>i</sub> P [TP NP V <sub>fin</sub> <b>cl</b> [ <sub>v</sub> P V <sub>inf</sub> <b>cl</b> ]]]		**!	*
☞ c. [ <sub>i</sub> P [TP NP <b>cl</b> V <sub>fin</sub> [ <sub>v</sub> P V <sub>inf</sub> <b>cl</b> ]]]		*	*
d. [ <sub>i</sub> P [TP <b>cl</b> NP V <sub>fin</sub> [ <sub>v</sub> P V <sub>inf</sub> <b>cl</b> ]]]	*!		*

Table 10.10: OT tableau illustrating clitic climbing out of infinitival complements

### 10.3.5 NonInitial vs. Suffix

Now that we have specified the domain to whose left edge clitics are drawn in BCMS – a syntactic domain, as in Czech – we need to explain why they do not occur in absolute initial position. According to both Radanović-Kocić (1996) and Anderson (2005), there is a ban against clitics in first position of the intonational phrase. Anderson (2005) captures this through the constraint NONINITIAL(Cl,<sub>i</sub>P). Section 6.2 has already provided evidence that it is plausible to assume that whereas Czech, Slovak, and Upper Sorbian clitics avoid a *syntactic* boundary, the left edge of CP, clitics in BCMS avoid following a *prosodic* boundary. It is thus clear that if we want to

maintain that the same constraint types are responsible for clitic placement in West Slavic and BCMS, we must assume different types of domain for NONINITIAL in these languages.

However, there is an alternative way of modelling BCMS. It involves replacing the NONINITIAL constraint with a constraint which requires the clitic to have a prosodic host to its left, for example through forming a prosodic word with left-adjacent material. Such a constraint has already been discussed in the present thesis: Billings's (2002) SUFFIX. In section 8.4.2, I have argued that it does not apply to Czech 2P clitics, which can appear after intonational boundaries without difficulty – yet, it is a good candidate for modelling the placement of BCMS clitics, which are clearly sensitive to prosodic constituency.

Assuming that SUFFIX applies to BCMS clitics would derive the fact that they must not be right-adjacent to a prosodic break from their need of a host to their left. On the other hand, one might argue that it prevents a parallel analysis of the two languages, since we still need to retain NONINITIAL for Czech, where clitics regularly show up after prosodic boundaries, i. e. without a host. Should we not assume a minimal number of clitic-specific constraints, which vary between languages only in the constituents that they refer to?

Whilst the intuitive answer to this might be yes, note that from a learner's perspective, it is irrelevant whether NONINITIAL applies to clitics in Czech – a child assumes certain constraints based on the input it receives in the language that it acquires. Thus, if all the data that the language learner receives includes clitics that are prosodically attached to a preceding element, a straightforward assumption is that these clitics are subject to a constraint SUFFIX (or, alternatively, ENCLITIC, cf. section 8.4.2, fn. 22).

According to Radanović-Kocić (1996), clitics in BCMS indeed “are attached as enclitics to the preceding word with which they form a single phonological word” (Radanović-Kocić, 1996, 430), which is evidenced by the fact that they form a single accentual unit, and that this unit is subject to phonological processes that typically apply within single words. I thus conclude that a constraint like SUFFIX is more adequate for BCMS than NONINITIAL(Cl,ιP) because it captures this prosodic incorporation as well as the clitics' inability to follow prosodic boundaries.

In fact, by assuming SUFFIX for BCMS, we are still maintaining an analysis that is parallel to Czech, because we can actually restate the constraint as formulated by Billings (2002) (“Morphemes marked as suffixes must follow some prosodic word”, cf. section 8.4.2) more precisely if we define it as an alignment constraint as given in 273. In both languages, thus, clitics appear in relatively high positions due to a positive alignment constraint, EDGEMOST(Cl,l,(Max)TP), but never initial in a certain domain due to a negative alignment constraint: NONINITIAL(Cl,CP) in Czech and SUFFIX in BCMS.

- (273) SUFFIX = \*ALIGN(Cl,l,ω,l)  
 = A clitic must not have a prosodic word (ω) boundary to its left (and thus forms an ω with the preceding element).

This specification as an alignment constraint also illuminates the relation between NONINITIAL(Cl,ιP) and SUFFIX: the latter includes the former, in the sense that it is violated by cases where the clitic follows a ιP boundary, plus all instances of the clitic following also minor prosodic boundaries, i. e. φ and ω. Finally, I should note that SUFFIX has nothing more to say about

the clitic's location than that it must not be placed initially or after a prosodic boundary – the precise position is regulated by the already discussed EDGEMOST constraint, as well as by INTEGRITY and other general syntactic constraints, which I am going to address in the following sections.

### 10.3.6 The nature of the first position in BCMS

The constraints employed so far drive clitics to a position as close to the left TP edge as possible, and they prevent them from following a prosodic boundary and from being initial in general. I will now address the question how the first position, i. e. the position that precedes the clitics, is defined in BCMS. I will first review some arguments for defining this initial constituent, the clitic host, as a syntactic entity. Then, I will investigate the issue of 2D and 2W placement in BCMS: is the INTEGRITY constraint ranked as it is in Czech, or is it lower, thus yielding a 2W language? This section will also show interactions of clitic placement with other more general syntactic constraints.

According to Radanović-Kocić (1996), BCMS clitics follow the first phonological phrase ( $\varphi$ ) of their intonation phrase ( $\iota$ P; cf. Radanović-Kocić, 1996, 441) – yet, this has been disputed, and I will review the arguments here. Whilst it is clear that prosodic factors play a role in BCMS clitic placement, the data I have discussed so far on the previous pages is fully consistent with a syntactic interpretation of the first position. In addition, sentence 260a, repeated here as 274, has as its first element an NP including a restrictive relative clause. This appears to be far better compatible with a syntactic interpretation than with a prosodic one – 14 syllables would be a surprisingly long  $\varphi$ , and given that  $\varphi$ 's are built based on syntactic phrases, one would expect that the constituent preceding the clitics is made up of two  $\varphi$ 's (cf. Féry, 2017, 86).

- (274) *Ona moja sestra koja je u Sarajevu vas se sjeća.*  
 that my sister REL is in Sarajevo you REFL remember.3SG  
 ‘My sister who is in Sarajevo remembers you.’

(cf. Radanović-Kocić, 1996, 437)

Bošković (2001) gives further evidence against the view that the first constituent is prosodic in nature: in configurations which permit the 2W placement of the clitic, the determiner would have to form its own  $\varphi$ , which, as Bošković (2001) remarks, is “far from obvious” (Bošković, 2001, 76). An example for such a construction is provided in 275b, first introduced in section 6.2. I will come back to the analysis of such cases later in the following section.

- (275) a. *Taj čovek je voleo Mariju.*  
 that man AUX.3SG love.PTCP Maria.ACC  
 b. *Taj je čovek voleo Mariju.*  
 that AUX.3SG man love.PTCP Maria.ACC  
 ‘That man loved Maria.’

(cf. Halpern, 1995, 16)

Another argument for assuming that the element preceding the 2P clitics is a syntactic constituent is instantiated by the sentences in 276b-c. Here, a topicalised VP contains a pronominal clitic and is followed by an auxiliary clitic.<sup>9</sup> Under the prosodic analysis, sentence 276b is especially problematic. The entire sentence forms a single *tP*, as can be seen from the fact that unlike in 276c, the auxiliary can appear directly after the fronted VP, thus there is no prosodic break (the break is marked by a comma in 276c). Now, the first  $\varphi$  in example 276b might be the initial verb, or the entire VP – in either case, we would expect both clitics to appear in the same position. It is much more plausible that clitic placement reflects the syntactic structure of the clause, with the pronominal following the verb in its VP, and the auxiliary following the entire VP.<sup>10</sup> Given all these arguments, I will thus follow Bošković (2001) in assuming that the first position in BCMS is a syntactic constituent.<sup>11</sup>

(276) a. *Ivan je ga želio sresti u Zagrebu.*  
Ivan AUX.3SG him.ACC want.PTCP meet.INF in Zagreb

‘Ivan wanted to meet him in Zagreb.’

b. *Sresti ga u Zagrebu je Ivan želio.*  
meet.INF him.ACC in Zagreb AUX.3SG Ivan want.PTCP

c. *Sresti ga u Zagrebu, Ivan je želio.*  
meet.INF him.ACC in Zagreb Ivan AUX.3SG want.PTCP

(cf. Franks &amp; King, 2000, 245)

### 10.3.7 Integrating integrity: 2W and 2D

As in our analysis of Czech, we can assume that the initial constituent is typically not broken up by the clitics because of constraints from the Integrity family, first introduced in section 8.2.2, and defined as in 277. These constraints penalise the placement of material inside certain phrase types, and in fact their general splitting, which accommodates the observation that structures which are inaccessible for clitics resist being discontinuous in general (cf. Anderson, 2005, 143-4).

(277) INTEGRITY(C) = A member of a category C may not be interrupted by phonological material that is not part of C.

Where:  $C \in \text{PCat} \cup \text{GCat}$

(cf. Anderson, 2000, 23)

<sup>9</sup>Note that typically, as in Czech, BCMS clitics appear adjacent to each other, following a templatic order. VP fronting including a clitic as we find it in the present example is possible in Croatian, but not in Serbian (cf. Bošković, 2004, 54, fn. 16). In Czech, according to my consultants, it is only very marginally accepted.

<sup>10</sup>Franks and King (2000) note that sentence 276a involves an unusual ordering of clitics, since *je* typically appears last in the cluster. Some speakers thus find the opposite ordering, *ga je*, more acceptable, but that does not affect how these constructions are generally judged (cf. Franks & King, 2000, 245, fn. 16). In section 10.3.8, I will come back to how the non-clustering of the clitics in examples 276b-c can be captured.

<sup>11</sup>For even more arguments that the entity preceding the clitics is a syntactic constituent, see Progovac (1996).

There are now two possible ways of analysing BCMS. The first alternative is that in BCMS, INTEGRITY is ranked *below* EDGEMOST – then, BCMS is a 2W language, as I have shown in section 8.6, which has presented the predictions of a factorial typology based on the major constraints of this analysis (see also Anderson, 2000, 24). In this case, occurrences of 2D placement must be explained by further principles. The second alternative is that in BCMS, as in Czech, INTEGRITY is ranked *above* EDGEMOST – then, BCMS is a 2D language, and cases of 2W placement require further explanation. I will investigate each of these alternatives in turn.

I begin with the first alternative, the “2W theory”. In Halpern’s (1995) analysis, clitics p-adjoin to the  $\omega$  to their left if a phrase is positioned in front of them due to topicalisation. If topicalisation does not take place, there is no material to the left of the clitics and Prosodic Inversion occurs as a last-resort process (cf. Halpern, 1995, 5). This means that intact phrases are in the left periphery, whereas split phrases are lower. In the present approach, this would correspond to a ranking of INTEGRITY(XP) below EDGEMOST. Since, as I have already argued for Czech, topics are above TP, they will never be disturbed by the clitics. On the other hand, elements below TP trigger EDGEMOST violations if the clitics follow them, and will thus be broken up. The tableau in 10.11 shows how in this view, a sentence like 275b emerges when no phrase is topicalised, and tableau 10.12 illustrates how in the case of topicalisation, a sentence like 275a is selected as the winning candidate.

	SUFFIX	EM(Cl,l,TP)	ITGR(XP)
a. [TP [NP Taj čovek] V cl]		**!*	
b. [TP [NP Taj čovek] cl V]		**!	
☞ c. [TP [NP Taj cl čovek] V]		*	*

Table 10.11: OT tableau illustrating 2W placement in BCMS – “2W theory”

	SUFFIX	EM(Cl,l,TP)	ITGR(XP)
a. [FinP [NP Taj čovek] [TP V cl]]		*!	
☞ b. [FinP [NP Taj čovek] [TP cl V]]			
c. [FinP [NP Taj cl čovek] [TP V]]		*!	*

Table 10.12: OT tableau illustrating 2D placement in BCMS – “2W theory”

Whilst this approach works from a technical point of view, it is empirically problematic. Why should we characterise BCMS as a 2W language? In almost all the data that this chapter has presented so far, clitics appear after the entire initial constituent. If we assume that 2D occurs only in the context of topicalisation, we must conclude that fronted topics are ubiquitous in BCMS. This alone might not be a convincing argument against the 2W theory. Yet, there are other facts that point towards a secondary status of 2W constructions in BCMS.

First, Radanović-Kocić (1996) notes that cases of split constituents as in 275b are marginal and do not occur in all dialects.<sup>12</sup> Second, she also states that 2W constructions do not allow

<sup>12</sup>According to Anderson (2005), 2W only exists in Croatian (i. e. western dialects), but not in Serbian (i. e. eastern dialects). The latter linguistic group does not allow syntactic constituents to be broken up by clitics (cf. Anderson, 2005, 111).

multiple clitics, and gives the examples in 278 (cf. Radanović-Kocić, 1996, 436).<sup>13</sup> Whilst I do not have an explanation for this fact, it shows that 2W placement is subject to more restrictions than the – always available – placement in 2D. I conclude from this that it is not adequate to classify BCMS as a 2W language.

- (278) a. *Moj je brat došao.*  
 my AUX.3SG brother come.PTCP  
 ‘My brother came.’
- b. *⁑Moj ga se brat sjeća.*  
 my him REFL brother remembers  
 ‘My brother remembers him.’
- c. *\*Moj ti ga se brat sjeća.*  
 my you him REFL brother remembers

(cf. Radanović-Kocić, 1996, 435-6)

Let us therefore consider the second alternative, the “2D theory”. As in Czech, in order to counter the effects of  $\text{EDGEMOST}(\text{Cl}, \text{l}, \text{TP})$ , the constraint  $\text{INTEGRITY}(\text{XP})$  is ranked above it. In consequence, a candidate with a higher amount of  $\text{EDGEMOST}$  violations can nevertheless succeed because it maintains the initial constituent’s integrity. Thus, a sentence like 275a emerges as the winning candidate from the evaluation in the absence of any topicalisation. This is illustrated in the tableau in 10.13.

	SUFFIX	ITGR(XP)	EM(Cl,l,TP)
a. [TP [NP Taj čovek] V cl]			***!
b. [TP [NP Taj čovek] cl V]			**
c. [TP [NP Taj cl čovek] V]		*!	*

Table 10.13: OT tableau illustrating 2D placement in BCMS – “2D theory”

Now, the question arises how we can account for the alternative realisation of this sentence, with the clitic following the demonstrative *taj*, as in example 275b. Note that one cannot assume that sentence 275b emerges because a specific Integrity constraint for (demonstrative) NPs is ranked lower than  $\text{EDGEMOST}$ , because, as sentence 275a clearly shows, this type of constituent *can* be kept intact. The question that we must answer at this point is therefore: under which conditions do clitics split up constituents in BCMS?

Bošković (2004) provides an answer to this problem: he notes that contrastively focused elements cannot be broken up by the clitic, except if only one part of a phrase is contrastively focused. This is illustrated for complex geographical names in example 279. Sentences 279a-b show that contrastive focus on the entire NP prevents the clitic from entering it, whereas sentences 279c-d show that focus on just one part of the NP makes it possible for the clitic to intervene.

<sup>13</sup>Radanović-Kocić (1996) does not provide a translation for sentence 278c.



- (279) a. *U NOVI SAD su došli, ne ZRENJANIN.*  
in Novi Sad AUX.3PL arrive.PTCP not Zrenjanin  
‘In NOVI SAD they arrived, not in ZRENJANIN.’
- b. *\*U NOVI su SAD došli, ne ZRENJANIN.*  
in Novi AUX.3PL Sad arrive.PTCP not Zrenjanin
- c. *U GORNJI su Vakuf došli, ne DONJI.*  
in Gornji AUX.3PL Vakuf arrive.PTCP not Donji  
‘In GORNJI Vakuf they arrived, not in DONJI (Vakuf).’
- d. *U Bačku su TOPOLU došli, ne PALANKU.*  
in Bačka AUX.3PL Topola arrive.PTCP not Palanka  
‘In Bačka TOPOLA they arrived, not in (Bačka) PALANKA.’

(cf. Bošković, 2004, 47-8)

In addition to that, Bošković (2004) also remarks that not only clitics, but also other material may intervene between two parts of an NP if only one of them is focused, and gives 280 as an example, where the non-clitic subject pronoun *oni* ‘they’ also appears “inside” the NP. This strongly indicates that in 2W constructions, constituent splitting happens independently from clitic positioning.

- (280) *U Gornji su oni Vakuf došli.*  
in Gornji AUX.3PL they Vakuf arrive.PTCP  
‘In Gornji Vakuf they arrived.’

(cf. Bošković, 2004, 47, fn. 9)

Thus, we should assume that there is some other, focus-related constraint, which is responsible for the split positioning of the phrases in 279c-d. This constraint might require that focused phrases occupy certain focus positions (for example Spec-FinP in my analysis; Bošković (2004) concludes from the data that BCMS has two focus positions, a higher and a lower one; cf. Bošković, 2004, 48). I will not speculate any further about the nature of this constraint here. What is relevant, is that it is ranked *above* INTEGRITY(XP) and that it favours candidates with split constituents when only part of that constituent is marked for focus in the input. For the purpose of illustration, I am naming this constraint FOCUS; the tableau in 10.14 shows how its interaction with the other constraints yields an apparent 2W structure as in 279c.

Focus: X	SUF	FOC	ITGR(XP)	EM(l,TP)
a. [FinP X [TP Y V cl [PP X Y]]]			*	*!*
b. [FinP X [TP Y cl V [PP X Y]]]			(*)	*!
☞ c. [FinP X [TP cl Y V [PP X Y]]]			*	
d. [TP [PP X Y] cl V [PP X Y]]		*!		**
e. [FinP [PP X Y] [TP cl V [PP X Y]]]		*!		

Table 10.14: OT tableau illustrating 2W placement in BCMS – “2D theory”

At closer inspection, there is one problem with the tableau in 10.14: I have marked candidate b with an asterisk under INTEGRITY(XP). Yet, this constraint is by its definition violated only when there is *phonological* material inserted between the two parts of a constituent. Whilst candidate b clearly includes a kind of splitting of the XY-phrase, it is not in the sense of Integrity, for which reason I have set the asterisk in brackets; what separates X and Y is instead the TP boundary, and, above it, the empty Fin head. I therefore propose an additional constraint, CONTINUITY(XP), defined as in 281, which penalises the separating of parts of phrases like NPs and PPs in a structural sense. For the data discussed here, we can simply rank this constraint on a stratum with INTEGRITY(XP). But it is a generally interesting question whether a particular language is more sensitive to the structural or to the phonological disturbance of constituents.<sup>14</sup> With this constraint, I repeat the evaluation given in tableau 10.14 as shown in the tableau in 10.15 below.

- (281) CONTINUITY(XP) = An XP may not be interrupted by functional structure that is not part of XP.

Focus: X	SUF	FOC	CONT (XP)	ITGR (XP)	EM (L,TP)
a. [FinP X [TP Y V cl [PP X Y]]]				*	*!*
b. [FinP X [TP Y cl V [PP X Y]]]			*		*!
☞ c. [FinP X [TP cl Y V [PP X Y]]]				*	
d. [TP [PP X Y] cl V [PP X Y]]		*!			**
e. [FinP [PP X Y] [TP cl V [PP X Y]]]		*!			

Table 10.15: OT tableau illustrating 2W placement in BCMS – “2D theory” (including CONTINUITY)

This analysis is very similar to what I have proposed for Czech split topicalisation in section 9.1.3. This means that the differences between Czech and BCMS are not as categorical as is often claimed – they are both 2D languages, and they both permit the splitting of constituents for information-structural reasons. This latter phenomenon may simply be more widespread in at least some dialects of BCMS, and has therefore received more attention. Yet, it is not a consequence of clitic placement.

### 10.3.8 VP raising and clitic lowering

There is one case that still requires closer investigation: example 276b, repeated here as 282, involves a split of the clitic cluster – the pronominal clitic appears second in the topicalised VP, whereas the auxiliary clitic appears only after that VP. Such kinds of data have often been discussed as an argument against the view that clitics cluster together under the same syntactic node in the syntax (for example Bošković, 2001, 51). In the here-presented analysis, these constructions could simply result from the definition of Integrity: the object pronominal clitic is

<sup>14</sup>It makes sense to assume that like Integrity, Continuity also constitutes a constraint family, of which CONTINUITY(XP) is only one subconstraint.

part of VP, and thus appears inside it, whereas the auxiliary clitic is not, and therefore appears right after the fronted VP. In other words, due to how the Integrity family is defined, placement of the pronominal clitic inside VP does not violate the constraint.

- (282) *Sresti ga u Zagrebu je Ivan želio.*  
 meet.INF him.ACC in Zagreb AUX.3SG Ivan want.PTCP

‘Ivan wanted to meet him in Zagreb.’

(cf. Franks & King, 2000, 245)

Interestingly, the clitic not only *can* appear within VP, it *must* do so. The variant where both clitics follow the fronted VP, given in 283, is highly degraded. Yet, it is not plausible to explain this through EDGEMOST, because the VP is best analysed as being topicalised to the left periphery, and thus above TP. We would consequently expect the ideal position for the pronominal clitic *ga* to be after VP as well. However, VP is also the base position of this clitic, and in consequence, the positioning of *ga* in example 283 constitutes a case of clitic lowering. Since this is something that should not happen unrestrictedly anyway, we can formulate a constraint NOLOWERING as in 284.

- (283)  $?*[_{VP}$  *Sresti ga u Zagrebu] je ga Ivan želio  $-VP$ .  
 meet.INF in Zagreb AUX.3SG him.ACC Ivan want.PTCP*

Intended: ‘Ivan wanted to meet him in Zagreb.’

(cf. Franks & King, 2000, 245)

- (284) NOLOWERING = \*X ... X  
 = An element may not be realised lower than its copy.

Whereas NOLOWERING must be ranked above EDGEMOST, it must be below SUFFIX – otherwise, one would predict that in subject-less sentences with an auxiliary clitic, the clitic appears sentence-initially. Sentence 285 shows that this is not the case. Tableau 10.16 illustrates how with this ranking, NOLOWERING favours candidates a (example 282) and c over candidate b (example 283), and how the auxiliary clitic is prevented from entering VP by INTEGRITY(XP), thereby rejecting candidate c.<sup>15</sup> The fact that NOLOWERING is violable also means that it is compatible with accounts such as Bošković (2004) that assume a pronunciation of lower copies under certain conditions – even more, these accounts *require* a constrained availability of lowering.

- (285) *će Vidjet će Mariju.*  
 see.INF will.3 Marija

‘He/They will see Marija.’

<sup>15</sup>As mentioned in footnote 10, the order of *je* and pronoun in this example is untypical, but this does not affect the general judgements for these types of sentences. I assume that as in Czech (cf. section 8.5), clitic ordering in BCMS is regulated by subconstraints of EDGEMOST(CI,1,TP), which I ignore in tableau 10.16 for the sake of simplicity.

(cf. Franks &amp; King, 2000, 20)

	SUF FIX	NO LOW	ITGR (XP)	EDMO (l,TP)
☞ a. [FinP [VP V cl <sub>prn</sub> PP] [TP cl <sub>aux</sub> ...]]				**
b. [FinP [VP V cl <sub>prn</sub> PP] [TP cl <sub>aux</sub> cl <sub>prn</sub> ...]]		*!		*
c. [FinP [VP V cl <sub>aux</sub> cl <sub>prn</sub> PP] [TP ...]]			*!	****

Table 10.16: OT tableau illustrating pronominal and auxiliary clitic placement with a fronted VP

### 10.3.9 Which $\iota$ P is the clitics' domain?

In his constraint-based analysis of BCMS clitic placement, Anderson (2005) discusses a set of data that is not correctly predicted by the interaction of NONINITIAL and EDGEMOST. These are cases where the initial  $\iota$ P boundary coincides with the left TP boundary, and where the clitic appears only within a later  $\iota$ P. The examples he gives can be found in 286; sentence 286b was first introduced as 261 earlier in this chapter.

- (286) a. *Sa Petrom Petrovićem sreća se samo Milena.*  
 with Peter Petrović met REFL only Milena  
 ‘With Peter Petrović, only Milena met.’
- b. *Znači da, kao što rekoh, oni će sutra doći.*  
 means COMP as said they will tomorrow arrive  
 ‘It means that, as I said, they will arrive tomorrow.’
- c. *Ja, tvoja mama, obećala sam ti sladoled.*  
 I your mother promised AUX you ice.cream  
 ‘I, your mother, promised you an ice cream.’

(cf. Anderson, 2005, 150)

The delayed placement of the clitics in these examples is a clear indicator of a  $\iota$ P boundary before the constituent that the clitics follow – i. e. before *sreća* ‘met’ in 286a, before *oni* ‘they’ in 286b, and before *obećala* ‘promised’ in 286c. Anderson (2005) now argues that the effect of EDGEMOST(Cl,l,TP) would place the clitics second within the initial  $\iota$ P, because there, fewer violations of this constraint occur, whilst NONINITIAL(Cl, $\iota$ P) would still be respected.

In order to remedy this, Anderson (2005) makes the generalisation that the clitic is introduced into the  $\iota$ P that contains the clause head, i. e. the main verb (cf. Anderson, 2005, 151). He thereby substantiates earlier statements of other authors: Radanović-Kocić (1996) generalises that in BCMS, “clitics stay within *their own* intonational phrase” (Radanović-Kocić, 1996, 439; my emphasis), and Bošković (2004) reformulates this as “SC clitics occur in the second position of *their* intonational phrase” (Bošković, 2001, 39; my emphasis). These statements beg the

question what precisely “their” means – Anderson (2005) answers this by defining it as the  $\iota$ P that contains the main verb.

In the present OT analysis, we would thus require an additional constraint that demands that clitics appear within the  $\iota$ P that also includes the main verb of the clause. Whilst technically feasible, I will argue that for BCMS, such a constraint is first, not adequate, and second, not required.

With respect to the first point, i. e. that restricting clitics to the  $\iota$ P of the main verb is not adequate for BCMS, consider sentence 287, repeated from example 267 above, where both auxiliary and pronoun of the main clause are in the initial  $\iota$ P, whereas Anderson’s (2005) generalisation would predict them to occur within the  $\iota$ P of the main verb *predstavili* ‘introduced’.

- (287) [ $\iota$ P *Oni su se*], [ $\iota$ P *kao što sam vam rekla*], [ $\iota$ P  
 they AUX.3PL REFL as AUX.1SG you.DAT say.PTCP  
*predstavili Petru*].  
 introduce.PTCP Petar.DAT

‘They have, as I told you, introduced themselves to Petar.’

(cf. Bošković, 2004, 53, fn. 15)

Another argument against the proposal that clitics appear within their verb’s  $\iota$ P are copular clitics. Unlike in Czech, in BCMS, the copula, which is formally identical with the past tense auxiliary, is also a 2P clitic. An example is provided in 288 (cf. Franks & King, 2000, 19). In sentences with a present tense copula, the clitic is thus identical with the main verb of the sentence. The requirement would thus, trivially, always be met – clitic copulas could consequently show up in *any*  $\iota$ P. This is clearly undesirable.

- (288) *Ti si dobar student.*  
 you COP.2SG good student

‘You are a good student.’

(cf. Franks & King, 2000, 19)

I will now turn to the second argument against Anderson’s (2005) generalisation: we actually do not need it, because our model can already capture the allegedly problematic data. Let us consider the potential positions the clitic might occupy in the first  $\iota$ P of sentence 286a; they are listed in 289. The first two options, 289a-b, can immediately be excluded because, in the absence of any information-structurally induced phrase splitting, these sentences feature a fatal violation of INTEGRITY(XP). The final option, 289d, immediately fails at the high-ranked SUFFIX.

- (289) a. \* $[\iota$ P *Sa se Petrom Petrovićem*] [ $\iota$ P *srela samo Milena*].  
 with REFL Peter Petrović met only Milena  
 b. \* $[\iota$ P *Sa Petrom se Petrovićem*] [ $\iota$ P *srela samo Milena*].  
 with Peter REFL Petrović met only Milena

- c. \* $[_{iP}$  *Sa Petrom Petrovićem se*]  $[_{iP}$  *srela samo Milena*].  
with Peter Petrović REFL met only Milena
- d. \* $[_{iP}$  *Sa Petrom Petrovićem*]  $[_{iP}$  *se srela samo Milena*].  
with Peter Petrović REFL met only Milena

But what about candidate 289c? It is clear that we need a way to exclude such placement in our analysis, because if we do not,  $iP$  boundaries would no longer induce delayed clitic placement in BCMS, contrary to what we find in the data. So why does the clitic not simply attach prosodically to the PP, with the prosodic boundary occurring only after the clitic?

The answer lies in the matching of syntax and information structure to prosodic constituency, as discussed in section 10.3.3. I mentioned there that topics tend to be prosodically separated from the rest of the sentence. We have also already established that in this case, the relevant prosodic constituent is  $iP$ , because this is what drives delayed clitic placement. We can capture this with a simple alignment constraint, as given in 290: a constituent marked as a topic must have a  $iP$  boundary at its right edge. Thus, if the clitic attaches prosodically to the topicalised phrase, this boundary can no longer be at the topic's right edge, and the constraint is violated. The tableau in 10.17 shows how in this way, we can successfully exclude candidate 289c, and also how the other candidates in 289 do not succeed in the evaluation, leaving candidate e as the winner, which corresponds to sentence 286a.<sup>16</sup>

- (290) TOPICEDGE = ALIGN(Topic,r, $iP$ ,r)  
= There must be a  $iP$  boundary at the right edge of a constituent marked as a topic.

Topic: [PP P N N ]	SUF FIX	TOPIC EDGE	ITGR (XP)	EDMO (l,TP)
a. $[_{iP}$ [PP P <b>cl</b> N N ]] $[_{iP}$ V ...]			*!	***
b. $[_{iP}$ [PP P N <b>cl</b> N ]] $[_{iP}$ V ...]			*!	**
c. $[_{iP}$ [PP P N N ] <b>cl</b> ] $[_{iP}$ V ...]		*!		*
d. $[_{iP}$ [PP P N N ]] $[_{iP}$ <b>cl</b> V ...]	*!			
e. $[_{iP}$ [PP P N N ]] $[_{iP}$ V <b>cl</b> ...]				*

Table 10.17: OT tableau illustrating clitic placement in a lower  $iP$  in BCMS

The same reasoning applies to sentence 286c: the initial subject NP does not permit intervening material,<sup>17</sup> and presumably, appositives are also subject to a constraint that requires their edges to coincide with  $iP$  boundaries. Sentence 286b however is a different case, and requires closer investigation. Again, I am listing the alternative, ungrammatical clitic placement

<sup>16</sup>Anderson (2005) discusses an alternative solution, namely that a constraint NONFINAL(Cl, $iP$ ) prevents clitics from appearing last in a  $iP$ . Yet, he also observes that clitics in BCMS *do* occur in the final position of  $iP$ s (cf. Anderson, 2005, 150). TOPICEDGE has the advantage of being an independently required constraint that does not impose additional restrictions on clitic positioning, but traces their behaviour back to generally observed patterns of the language.

<sup>17</sup>Given the prosodic independence of the appositive clause, it might be surprising that such constructions are nevertheless subject to Integrity. Yet, as remarked by Georg Höhn (p.c.), in German, too, appositives are harder to separate from the noun than restrictives:

options for this example in 292 below. For the sake of conciseness, I am leaving out instances where the clitic is initial in  $\iota$ P, i. e. not enclitic, since we have sufficiently discussed the exclusion of such constructions.

- (292) a. \* $[_{\iota}P$  *Znači će da*],  $[_{\iota}P$  *kao što rekoh*],  $[_{\iota}P$  *oni sutra doći*].  
 means will COMP as said they tomorrow arrive
- b. \* $[_{\iota}P$  *Znači da će*],  $[_{\iota}P$  *kao što rekoh*],  $[_{\iota}P$  *oni sutra doći*].  
 means COMP will as said they tomorrow arrive
- c. \* $[_{\iota}P$  *Znači da*],  $[_{\iota}P$  *kao što će rekoh*],  $[_{\iota}P$  *oni sutra doći*].  
 means COMP as will said they tomorrow arrive
- d. \* $[_{\iota}P$  *Znači da*],  $[_{\iota}P$  *kao što rekoh će*],  $[_{\iota}P$  *oni sutra doći*].  
 means COMP as said will they tomorrow arrive

In fact, for all the sentences in 292, we can argue that their ungrammaticality is explained by the same principle that also explains why clitics do not climb out of finite clauses, as discussed in section 10.3.4: clitics do not climb out of clauses with a TP layer because this causes violation(s) of EDGEMOST(Cl,I,TP). It is plausible that this is the case in all the above examples, since the subject, *oni* ‘they’ is in the specifier of TP, or even higher in a focus position, and the adjunct clause *kao što rekoh* ‘as I said’ is attached to the entire embedded clause, thus to the lower CP.<sup>18</sup> In consequence, all of the clitic positions in 292 trigger unnecessary EDGEMOST violations. On top of that, sentence 292c can also be argued to feature a violation of an Integrity subconstraint, because the clitic breaks up the adjunct clause that it is not part of; and sentence 292d violates the alignment of an adjunct clause with a prosodic boundary to its right.

I thus conclude that there is no empirical basis for a constraint that ties BCMS clitics to the  $\iota$ P that includes the main verb. Such a constraint would exclude structures that actually do have the clitic in a different  $\iota$ P than the head of its clause, as in example 287. It also has no impact on clitic copulas as in example 288, because these *are* the main verb of their sentence. Fortunately, the cases which lead Anderson (2005) to assume such a constraint (the sentences in 286) can all be captured with the general assumptions of the present analysis, using only two basic constraint types: alignment constraints (including the clitic-specific EDGEMOST and SUFFIX), and integrity constraints.

- 
- (291) a. <sup>?</sup>*Mein Vater hat angerufen, von dem ich dir schon viel erzählt habe.*  
 my father has call.PTCP of whom I you already much tell.PTCP have  
 ‘My father has called, who I’ve already told you a lot about.’
- b. *Der Typ hat angerufen, von dem ich dir schon viel erzählt habe.*  
 the guy has call.PTCP of whom I yo already much tell.PTCP have  
 ‘The guy has called, who I’ve already told you a lot about.’

<sup>18</sup>We obtain the same result if we assume that the adjunct clause is attached to TP; adjuncts that are even lower are also not accessible to the clitic due to the constraint NOLOWERING.

### 10.3.10 Conclusions

This chapter has shown that a constraint-based analysis of clitic placement can successfully be applied to the patterns found in BCMS. The fact that BCMS clitics, in contrast to West Slavic ones, are placed with reference to both syntactic and prosodic constituency does not pose a problem for the OT architecture, because it assumes no distinct layers of linguistic modules – all types of information are processed in parallel.

We have seen that in contrast to Czech, BCMS possesses a type of infinitival complements that lack both a TP and a CP layer (even when there is an apparent complementiser present), and thereby provide the context for clitic climbing. In contrast, the slightly different domain of EDGEMOST prevents climbing out of clauses that do have a TP layer: the clitics have nothing to gain from appearing higher in the clause in these contexts. The fact that clitics appear neither in clause-initial position nor after prosodic breaks is due to the fact that BCMS clitics, again in contrast to the Czech, Slovak, and Upper Sorbian ones, are true enclitics.

I have also argued against the view that BCMS is a 2W language. Instead, the data can be captured adequately by assuming that INTEGRITY(XP) is ranked higher than the clitic placement constraints, as it is in Czech, and that other information-structural requirements lead to the breaking up of phrases. I have also shown that we do not need an additional constraint that regulates the choice of tP in which the clitics appear.

In sum, I conclude that the basic principle behind clitic positioning in both languages is based on conflicting alignment constraints. Different domain specifications of these constraints lead to different patterns with respect to placement at prosodic boundaries and clitic climbing. Yet, in both languages, clitics do not interrupt otherwise intact phrases. BCMS has also provided us with deeper insights into the matching of syntactic and prosodic constituency, and how this can exclude prosodically dependent clitics from certain positions.

Some of the arguments in this chapter had to be made based on provisional assumptions about the data's prosody. This is, of course, partly due to the fact that the focus of this thesis is Czech, and that I did not have access to native speakers. However, it also reflects a general lack of information about prosody in the literature. In languages such as BCMS, where there is a clear impact of phonology on clitic placement, prosodically annotated data, ideally gathered systematically in experiments, is thus a clear desideratum for future research.



## Chapter 11

# Conclusions and Outlook

This dissertation has provided a fully optimality-theoretic analysis of clitic placement, in which clitics occupy varying overt positions due to the interaction of two conflicting constraints of the Generalised Alignment family. The analysis captures the placement of clitics in second and third position in simple clauses, as well as clitic climbing and its absence in complex clauses. I have also proposed an account of clitic ordering including person-ordering (PCC) effects in Czech. Another important aspect has been the interaction of clitic placement with several types of syntactic operations, including constituent and non-constituent fronting. For BCMS, the analysis also encompasses the prosodic effects of such operations. This final chapter provides some concluding remarks and observations on some of the major issues addressed in this thesis.

### 11.1 Modelling clitic placement

Both in Czech and in BCMS, the requirement for clitics to not appear initially in a certain domain was traced back to more general grammatical principles: in Czech, information-structurally non-salient elements cannot occupy the left CP-edge; in BCMS, clitics are suffixal in the sense that they require a prosodic host to their left. This means that the non-initiality requirement on clitics is based on their information-structural or prosodic properties – since a deficiency in these areas is one of the core properties of typical clitics, this could be a factor that accounts for the cross-linguistic frequency of 2P clitic placement.

With respect to the relatively high position of clitics, the responsible constraint must remain a clitic-specific one. Whilst we thus still do not have a full explanation why clitics prefer to be placed in second position, and not lower in the clause, I believe that the present thesis has nevertheless achieved a coherent model that accurately predicts their peculiar behaviour.

### 11.2 The nature of (West Slavic) clitics

In this dissertation, I have established that Czech, Slovak, and Upper Sorbian clitics deviate from the prototypical clitic in an important aspect: they are not prosodically dependent, since they can appear without any prosodic host. I have therefore assumed that they do not form an  $\omega$  with a neighbouring word. However, they are not prevented from doing so – the crucial observation is that their placement within the clause is independent from potential en- or proclitic prosodic

attachment. The fact that clitics do not appear alone plausibly results from their inability to be information-structurally prominent: since in a single-word utterance, the only overtly realised element must necessarily be the focus, clitics are illicit in such configurations. This applies to Czech, Slovak, and BCMS clitics. In Upper Sorbian, the clitic status of auxiliaries, copulas, and modal verbs can be overwritten in certain contexts, and they can then appear initially and in one-word answers.

I have argued that Czech auxiliary clitics should not be viewed as typical verbs, but rather as reduced elements which lack a verbal stem. However, this conclusion still begs the question whether clitics are phrasal affixes that do not have a position in the syntax, as argued by Anderson (2000/2005), or whether they are a mere spell-out of features and do occupy syntactic positions, as proposed for example by Franks and King (2000). As can be seen from the assumptions made in the further course of this thesis, I have adopted the second view: clitics occupy syntactic positions which allow them, for example, to receive theta roles and case. They are also subject to general syntactic constraints. Yet, note that in an OT account, syntax is not a separate linguistic module, so the question whether clitics are “syntactic” or “morphological” is of lesser consequence.

The crucial property that clitics must have for the present analysis to work from a technical viewpoint is a feature [+clitic], whereby they can be identified by the constraints that apply to them. It is therefore this feature that is the basis for their unusual placement properties, and possibly also what prevents them from fulfilling certain salient prosodic and information-structural roles. That this feature is idiosyncratic and can probably not be traced back to a more principled underlying feature becomes evident through the differences in the clitic inventories of the four investigated languages.

### 11.3 The interaction of syntax and phonology

Within an OT framework, syntax, phonology, and other parts of grammar can interact freely. In this respect, the present thesis has shown that alignment constraints are a particularly adequate tool for modelling clitic placement. They can capture the fact that what is summarised as 2P placement can be purely syntactic in one language, whilst referring to both syntax and prosody in another. Given that prosodic and syntactic boundaries coincide in many cases, this is not surprising: confronted with a great amount of inconclusive evidence, different learners might form deviating hypotheses about the constraints that regulate clitic placement, and this over time leads to language change and therefore also to variation within the same language family.

The present approach also allows us to capture the interactions between syntax and phonology that are mediated by information structure. We have seen that the heavy influence information structure has on word order in Czech is also evident in clitic placement. First, NONINITIAL(Cl,CP) is actually derived from a more general constraint that prevents elements that cannot be contrastive topics or foci from appearing initially in CP. Second, the elements that can precede the clitics are often, though by no means necessarily, fronted due to their information-structural status. Since such processes also have an impact on prosody, they directly interact with clitic positioning in BCMS, where clitics are sensitive to phonological structure. All this shows that the different linguistic modules are highly entangled, and this entanglement

can be represented in a flexible, yet at the same time precisely defined manner, with a purely optimality-theoretic architecture.

On a more general note, we have seen that for the analysis to work, we must assume that both syntactic and prosodic projections are labelled, and that constraints can refer to these labels. Both these assumptions are in opposition to what has been postulated by Grimshaw (1997) – yet, the fact that we find elements that are actually sensitive to such labels is a clear indicator that we should not abandon them merely for the sake of simplicity.

## 11.4 Transferability to interface theories

Whilst I have based a central part of my analysis of clitic placement on constraints as proposed by Anderson (2000) and related works, I have also drawn on the ideas developed in what I have called syntax-and-PF approaches. In consequence, many of the assumptions I have made could also successfully be transferred into such an approach, where clitic positioning is also flexible due to conflicting requirements.

One major difference is that many of these approaches work with the assumption that, if necessary, a lower copy of the clitics is pronounced. In contrast, in my approach, clitics do not rise to a high position in the first place, but are directly realised in their optimal position. Yet, this is a rather technical point that could most probably be resolved without great cost.

Another aspect is far more critical: In Czech, Slovak, and Upper Sorbian the constraints that drive the placement of clitics refer to syntactic constituents – not only the one that drives them upwards,  $EDGE_{MOST}(Cl, I, MaxTP)$ , but crucially also the one that prevents their initial placement,  $NONINITIAL(Cl, CP)$ . However, in interface approaches, the constraints that intervene with copy realisation are *PF* constraints, and these should not have syntactic domain specifications. Thus, whereas BCMS can be adequately captured within such a framework, I argue that the West Slavic languages I have discussed here cannot.

## 11.5 On clitics as a diagnostic

If 2P clitics occupy a fixed position in the syntactic structure, they can be used to identify the edge of the preceding constituent – in the case of 2P clitics, this is typically assumed to be the left periphery. However, the present thesis has shown that clitics do not demarcate any specific point in the syntactic structure. Their placement is flexible, always striving to optimally satisfy conflicting constraints.

On the other hand, the fact that these constraints refer to specific syntactic and prosodic constituents means that clitics can serve to identify their presence within a given structure. For example, they can aid in clarifying the status of embedded clauses: in Czech and Slovak, clauses that permit clitic climbing cannot have a CP layer, and clauses that block it must presumably be CPs; and in BCMS, it appears that clitic climbing is sensitive to the presence of a TP layer.

## 11.6 Objectives for further research

There are some areas in which the present dissertation has dealt with a lack of sufficiently fine-grained data. One desideratum for future research is therefore a systematic collection of clitic placement data through both acceptability judgement studies and corpus research, in order to enhance both the amount and the quality of the data. Especially in the case of BCMS, but also in Czech, there is an additional need for data that includes prosodic information.

One area that would strongly benefit from such investigations is the optionality that can be found in some cases of clitic climbing. Ideally, we will find that, as with 3P placement, optionality is only apparent, and conditions such as information-structural factors instead determine the clitics' position. If on the other hand we are dealing with true optionality in one or several of these cases, this would be a completely different type of challenge, for constraint-based as well as derivational approaches.

This thesis has also shown that comparable data is required for understanding the commonalities and differences between the Czech and Slovak clitic patterns: whilst the two languages are certainly similar in many respects, especially Slovak risks being neglected if potential differences are not investigated systematically. I have only provided a few small comparative studies here, but I believe that parallel investigations of the two languages should feature in future research, too. An even more striking case is Upper Sorbian, whose clitic system in fact offers some surprising features: the language appears to have enlarged the set of clitic verbs significantly, but at the same time, these 2P elements can be placed initially and even in isolation under the right conditions. In order to establish the status of different potentially clitic elements in the language, more in-depth systematic research is required.

Another promising field of investigation are the VP and verb fronting constructions for which I have sketched an analysis. The precise information-structural contexts that condition them need more research in Czech and Slovak. In addition, a comparison of such research results to BCMS could provide further insights into how information-structural displacements and integrity requirements interact with the demands posed by clitics.

Finally, the status of the Czech past participle and the clitic past auxiliary deserve further investigation, especially from a diachronic perspective: How has the status of these elements shifted in the standard as well as the colloquial language? And into which direction do the current developments in these two grammatical systems point? How does the prosodic independence of clitics fit to the fact that they are partly eroding phonologically and giving up their status as proper verbs and pronouns?

It is an intriguing fact that behind a statement as simple as “clitics are placed in second position”, we actually find quite different and surprisingly complex patterns that evade a straightforward explanation in any approach to the functioning of grammar. The Slavic languages provide a particularly abundant field of research in this respect, and clearly their investigation and analysis is far from complete.

# Appendix A

## List of constraints

ALIGN(Clause,l,tP,l)

= A clause's leading edge must coincide with the leading edge of a tP (Billings, 2002).

CLITICORDER

= If the foot of the chain of  $\alpha$  precedes the foot of the chain of  $\beta$ , then the head of the chain of  $\alpha$  precedes the head of the chain of  $\beta$ . Where:  $\alpha, \beta \in \{x \mid x \text{ is a lexical clitic}\}$ .

CONTINUITY(XP)

= An XP may not be interrupted by functional structure that is not part of XP.

CP-BARRIER

= \*...  $X_i$  ... [CP ... [YP  $X_i$  ... ] ... ]

= An element which has a copy within CP must also have a copy at the left edge of that CP.

EDGEMOST(Cl,l,TP)

= ALIGN(Cl,l,TP,l)

= A clitic must occur as close to the left edge of TP as possible (Richardson, 1997).

EDGEMOST(Cl,l,MaxTP)

= ALIGN(Cl,l,TP<sub>i</sub> such that ALIGN(TP<sub>i</sub>,l,CP,l), l)

= A clitic must occur as close to the left edge of the highest TP in its clause complex as possible.

For subconstraints of EDGEMOST responsible for the cluster-internal ordering of clitics, see section 8.5.

HEAD-COMPLEMENT

= A head precedes all terminals originally dominated by its complement (Broekhuis, 2008).

INTEGRITY(C)

= A member of a category C may not be interrupted by phonological material that is not part of C. Where:  $C \in \text{PCat} \cup \text{GCat}$  (Anderson, 2000).

## INTEGRITY(XP)

= An XP may not be interrupted by phonological material that is not part of XP (Anderson, 2000).

## MATCHCLAUSE

= A clause in syntactic constituent structure must be matched by a corresponding  $\iota$ -phrase in prosodic representation (Féry, 2017).

## NOLOWERING

= \* $X$  ...  $X$  = An element may not be realised lower than its copy.

## NONINITIAL(Cl,CP)

= \*ALIGN(Cl,l,CP,l)

= A clitic must not occur at the left edge of CP (Richardson, 1997).

## STAY

= Trace is not allowed (Grimshaw, 1997).

## SUFFIX

= \*ALIGN(Cl,l, $\omega$ ,l)

= A clitic must not have a prosodic word ( $\omega$ ) boundary to its left (and thus forms an  $\omega$  with the preceding element).

## THETA

= An argument must be theta-marked (Grimshaw, 1997).

## TOP-FIRST

= ALIGN(Topic,l,FinP,l)

= An element marked as a topic must appear at the left edge of FinP.

## TOPICEDGE

= ALIGN(Topic,r, $\iota$ P,r)

= There must be a  $\iota$ P boundary at the right edge of a constituent marked as a topic.

## UNICHAIN

= A chain is uniform with regard to its phrase structure status S. Where:  $S \in \{\text{maximal, minimal, neither}\}$ .

## Appendix B

# Acceptability judgement tasks

The table below lists all experiments discussed in this thesis in alphabetical order of their short names (ID's), and gives their sample size, the respective section number, and the number of the chart displaying the results. The materials and results for each experiment can be found in a repository at the Open Science Framework (OSF). It also contains all charts as well as the instructions in Czech and Slovak. For ease of reference, all files relating to a specific experiment are named according to its ID. This is the link to the repository: <https://osf.io/b4y6c/>.

ID	Name	N	Section	Chart
CZ-climb	Czech clitic climbing	63	8.3	8.2, 8.3
CZ-front	Czech verbal fronting	61	7.2.1, 9.2.2	7.1, 9.2
CZ-front2	Czech (partial) VP fronting in context	17	9.2.4	9.3
CZ-li	Czech verbs with -li	8	7.2.3	7.2
CZ-pass	Czech passive participle fronting	63	9.2.3	see OSF
CZ-pos	Czech clitics in 1st and 4th position	16	8.2.3	see OSF
CZ-s	Czech contracted auxiliary	8	7.1.2	see OSF
CZ-split	Czech split topicalisation	8	9.1.3	9.1
SK-climb	Slovak clitic climbing	24	10.1.6	10.5
SK-free	Slovak freedom of pronominal clitics	24	10.1.2	10.1, 10.2, 10.3
SK-front	Slovak verbal fronting	24	10.1.5	10.4

# Appendix C

## German Summary and affidavit

### Deutschsprachige Zusammenfassung

Die vorliegende Dissertation befasst sich mit der Positionierung von Klitika im Tschechischen. Diese umfassen Teilmengen der Pronomina und Hilfsverben und treten typischerweise in der zweiten Position des Satzes auf, weswegen sie auch als 2P-Klitika bezeichnet werden. Das zentrale Ziel dieser Arbeit ist es, diese Elemente zu charakterisieren sowie ihre Distribution zu modellieren.

In Kapitel 1 werden Ziele und Grundannahmen der Arbeit präsentiert und erläutert, aus welchen Quellen die verwendeten Sprachdaten stammen. Hier sind zwei Arten von Daten zu unterscheiden: jene, die in der existierenden Literatur zu tschechischen Klitika diskutiert werden und dort von den Autor\*innen bezüglich ihrer Grammatikalität markiert werden; und jene, die ich mithilfe von Akzeptabilitätsfragebögen selbst erhoben habe und die einen Einblick auch in die feineren Abstufungen von Sprecher\*innenurteilen erlauben. Einige Daten wurden zudem auch dem Tschechischen Nationalkorpus entnommen.

Kapitel 2 liefert wichtige Hintergrundannahmen zu Klitika im Allgemeinen: ihre Definition, verschiedene Klassifikationssysteme, ihre Rolle im Sprachwandel sowie kurze typologische Betrachtungen. Bei der Definition von Klitika ist besonders zu beachten, dass sie häufig im Vergleich zu sowohl vollen Wörtern als auch Affixen definiert werden und hierbei mit beiden Gruppen Eigenschaften teilen.

Weitere Hintergrundannahmen werden in Kapitel 3 besprochen: Hier geht es um die Grundlagen der tschechischen Grammatik. Es wird klargestellt, dass das zentrale Untersuchungsobjekt das Standardtschechische ist, was sich teils sehr stark von der Umgangssprache unterscheidet. Besonderes Augenmerk liegt in diesem Kapitel auf der Syntax und Informationsstruktur des Tschechischen, da diese direkt relevant sind für die Platzierung von Klitika.

Kapitel 4 stellt im Anschluss das klitische Inventar der tschechischen Sprache im Detail vor. Eindeutig an die zweite Position gebundene klitische Elemente finden sich unter den Pronomen und den Hilfsverben. Auch der Konditionalkomplementierer und die verbale Negation können als Klitika analysiert werden, stehen allerdings nicht in 2P. Elemente, die nur in dieser Position auftreten können, lassen sich dadurch identifizieren, dass sie nicht satzinitial stehen, keinen anderen Klitika als *host* dienen und nicht fokussiert oder isoliert stehen können. Die Anordnung von mehreren 2P-Klitika in einem Satz ist fest geregelt.



In Kapitel 5 wird genauer auf die Positionierung der Klitika eingegangen. In einfachen Sätzen stehen 2P-Klitika nach der ersten Konstituente. Es gibt allerdings auch Kontexte, die eine Verschiebung in die dritte Position (3P) bewirken: Dies tritt auf, wenn zwischen Klitikon und Komplementierer oder Fragewort ein topikalisiertes Element steht. Es wird gezeigt, dass die 2P-Beschränkung auf Klitika keinesfalls darauf basieren kann, dass diese Elemente enklitisch sind, also einen *host* zu ihrer Linken benötigen, da sie prosodisch auch völlig isoliert stehen können. In bestimmten Satzgefügen tritt im Tschechischen *clitic climbing* auf, womit das Klitikon nicht in dem Satz steht, zu dem es gehört, sondern im höheren Matrixsatz.

Kapitel 6 bietet einen Überblick über verschiedene Theorien zur Positionierung von Klitika. Rein syntaktische Ansätze weisen eine Reihe konkreter Probleme bei der Modellierung der Daten auf. Prosodische Ansätze liefern gute Ergebnisse für das Serbokroatische, sind allerdings aufgrund der oben erwähnten Eigenständigkeit von Klitika nicht für das Tschechische geeignet. Vielversprechender sind hier sogenannten Syntax-und-PF-Theorien sowie morphologisch-optimalitätstheoretische Ansätze, da beide ermöglichen, dass neben der Syntax auch die Phonologie bzw. Morphologie in die Platzierung von Klitika eingreift. Das Kapitel stellt zudem kurz die Annahmen verschiedener Theorierichtungen über die Anordnung der Klitika untereinander dar.

Den ersten Teil meiner eigenen Analyse tschechischer Klitika bildet das Kapitel 7. Es wird aufgezeigt, dass der morphologisch-lexikalische Unterschied zwischen den klitischen Perfektauxiliaren und den auf den ersten Blick formgleichen nicht-klitischen Passivauxiliaren und Kopula eklatant ist. Zudem weist das l-Partizip, mit dem die klitischen Auxiliare auftreten, ein mit finiten Verben identisches syntaktisches Verhalten auf. Auf Basis der Daten lässt sich folgern, dass Auxiliarklitika keinen verbalen Stamm besitzen, während das l-Partizip die Funktion des finiten Verbs im Satz einnimmt.

Im Anschluss hieran untersucht Kapitel 8 die Positionierung von 2P-Klitika aus einer optimalitätstheoretischen Perspektive. Zunächst werden die Grundlagen der Optimalitätstheorie (OT) dargelegt: Verletzbare Constraints auf sprachliche Strukturen sind in verschiedenen Sprachen unterschiedlich zueinander hierarchisch sortiert. In einem einschrittigen Evaluationsprozess wird so aus einer Menge von Kandidaten der optimale ausgewählt, das heißt jener, der die Constraints unter Berücksichtigung ihrer Sortierung am besten erfüllt. Mithilfe zweier für Klitika spezifischer Constraints,  $\text{NONINITIAL}(\text{Cl}, \text{CP})$  und  $\text{EDGEMOST}(\text{Cl}, \text{L}, \text{MaxTP})$  können sowohl 2P in einfachen Sätzen als auch in Satzgefügen, also *clitic climbing*, modelliert werden. Es wird außerdem gezeigt, dass sich ersterer Constraint auf eine allgemeinere Beschränkung zurückführen lässt, nämlich dass informations-strukturell "schwache" Elemente nicht am linken Rand der CP stehen dürfen.

In Kapitel 9 wird nun die Analyse auf 3P-Kontexte ausgeweitet, was eine genauere Betrachtung von Topikalisierung im Allgemeinen, und verbaler Topikalisierung im Besonderen, beinhaltet. Hierfür werden weitere Constraints definiert und diskutiert, die allgemeinere Beschränkungen auf syntaktische Strukturen darstellen. Somit wird die Erfassung der Positionierung von Klitika in einen weiteren, OT-syntaktischen Kontext eingebettet.

Kapitel 10 erweitert den Ansatz auf drei weitere slawische Sprachen. Das eng mit dem Tschechischen verwandte Slowakische weist nur leichte Abweichungen auf und lässt sich im Grunde analog analysieren. Die ebenfalls westslawische Sprache Obersorbisch hingegen ist vor

allem in Hinblick auf klitische Verben interessant, deren Inventar in der Sprache stark ausgeweitet ist. Zugleich können diese Elemente aber auch initial stehen. Dies lässt sich durch Constraints, die NONINITIAL(CI,CP) übergeordnet sind, erfassen. Schließlich wird noch die südslawische Sprache Serbokroatisch analysiert, wo neben syntaktischer Konstituenz auch prosodische Grenzen eine Rolle für die Positionierung von 2P-Klitika spielen. Die einschrittige, nicht geschichtete Architektur von OT lässt eine freie Interaktion dieser beiden Ebenen ohne weiteres zu. Es wird gezeigt, dass der Typus des Alignment-Constraints, zu dem NONINITIAL und EDGEMOST zählen, zur Modellierung klitischer Phänomene in beiden Sprachen sehr gut geeignet ist. Außerdem werden Gemeinsamkeiten und Unterschiede zwischen dem Tschechischen und Serbokroatischen diskutiert.

Kapitel 11 liefert schließlich die Schlussfolgerungen der Arbeit und bietet einen Ausblick auf weitere Forschungsarbeit zu diesem Thema. Zentral ist hier, dass sich das Verhalten tschechischer Klitika mithilfe von Alignment-Constraints modellieren lässt, die sich ausschließlich auf syntaktische Konstituenz beziehen. Während Prosodie keine direkte Rolle hierbei spielt, scheint im Gegenteil die Rolle der Informationsstruktur bisher unterschätzt worden zu sein. Dieser muss sowohl in OT- als auch in Syntax-und-PF-Ansätzen mehr Aufmerksamkeit zukommen, wenn das Verhalten von Klitika adäquat erfasst werden soll.

## Versicherung

Hiermit versichere ich an Eides statt, dass ich die schriftliche wissenschaftliche Abhandlung (Dissertation) selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe und dass anderweitig keine entsprechende Promotion beantragt wurde und hierbei die eingereichte Dissertation oder Teile daraus vorgelegt worden sind.

Nina Adam, am 23. Januar 2024

# Online resources

L-REX – Starschenko, A. & Wierzba, A. (2022). L-Rex: Linguistic rating experiments. Version 1.0. GNU General Public License v3.0.

Online access: <https://github.com/2e2a/l-rex/>

PROLIFIC – Prolific (2023). London, UK. First released in 2014.

Online access: <https://www.prolific.com/>

SYN2015 – Křen, M., Cvrček, V., Čapka, T., Čermáková, A., Hnátková, M., Chlumská, L., Jelínek, T., Kovářiková, D., Petkevič, V., Procházka, P., Skoumalová, H., Škrabal, M., Truneček, P., Vondříčka, P. & Zasina, A. (2015). *SYN2015: reprezentativní korpus psané češtiny*. Ústav Českého národního korpusu, Filozofická fakulta Univerzity Karlovy, Praha.

Online access: <https://www.korpus.cz/>

# References

- Anderson, S. R. (1992). *A-Morphous Morphology*. Cambridge: Cambridge University Press.
- Anderson, S. R. (1996). How to put your clitics in their place, or why the best account of second-position phenomena may be something like the optimal one. *The Linguistic Review*, 13, 165–91.
- Anderson, S. R. (2000). Towards an optimal account of second position phenomena. In J. Dekkers, F. van der Leeuw & J. van der Weijer (Eds.), *Optimality Theory: Syntax, phonology and acquisition* (pp. 302–333). Oxford: Oxford University Press.
- Anderson, S. R. (2005). *Aspects of the theory of clitics*. Oxford: Oxford University Press.
- Avgustinova, T. & Oliva, K. (1995). *The position of sentential clitics in the Czech clause*. Saarbrücken: Universität des Saarlandes.
- Benveniste, E. (1946). Structure des relations de personne dans le verbe. *Bulletin de la Société de Linguistique de Paris*, 43(126), 1–12.
- Billings, L. A. (2002). Phrasal clitics. *Journal of Slavic Linguistics*(10), 53–104.
- Bošković, Ž. (2001). *On the nature of the syntax-phonology interface: Cliticization and related phenomena*. Amsterdam: North Holland.
- Bošković, Ž. (2004). Clitic placement in South Slavic. *Journal of Slavic Linguistics*, 12, 37–90.
- Broekhuis, H. (2008). *Derivations and Evaluations: Object shift in the Germanic languages*. Berlin: Mouton de Gruyter.
- Ćavar, D. & Wilder, C. (1994). Long head movement? Verb movement and cliticization in Croatian. *Lingua*, 93, 1–58.
- Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. N. Li (Ed.), *Subject and topic* (pp. 25–55). New York: Academic Press.
- Chomsky, N. (1994). *Bare phrase structure*. Cambridge: MIT Working Papers in Linguistics.
- Chomsky, N. (1995). *The minimalist program*. Cambridge: MIT Press.
- Chomsky, N. (1999). *Derivation by phase*. Cambridge: MIT Press.
- Chomsky, N. (2000). Minimalist inquiries: the framework. In R. Martin, D. Michaels & J. Uriagereka (Eds.), *Step by step* (pp. 89–155). Cambridge: MIT Press.
- Corbett, G. (1983). *Hierarchies, targets and controllers: agreement patterns in Slavic*. London: Croom Helm.
- Costa, J. (1998). Word order typology in Optimality Theory. In G. L. Jane Grimshaw & S. Vikner (Eds.), *Optimality theoretic syntax*. Cambridge: MIT Press. (Ms., Universidade de Lisboa and University of Leiden)
- Daneš, F. (1957). *Intonace a věta ve spisovné češtině*. Praha: Nakladatelství Československé Akademie Věd.

- Dotlačil, J. (2007). Why clitics cannot climb out of CP: A discourse approach. In M. G. Ulyana Savchenko & R. Compton (Eds.), *Formal Approaches to Slavic Linguistics (FASL): The Toronto meeting 2006* (pp. 76–93). Ann Arbor: Michigan Slavic Publications.
- Dryer, M. S. (2013a). Determining dominant word order. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Leipzig: MPI for Evolutionary Anthropology.
- Dryer, M. S. (2013b). Order of subject, object and verb. In M. S. Dryer & M. Haspelmath (Eds.), *The World Atlas of Language Structures Online*. Zenodo.
- Dvořák, V. (2010). On the syntax of ditransitive verbs in Czech. In *Proceedings of FASL 18 – Formal Approaches to Slavic Linguistics: The second Cornell meeting*.
- Engels, E. & Vikner, S. (2006). An optimality-theoretic analysis of Scandinavian object shift and remnant VP-topicalisation. In H. Broekhuis & R. Vogel (Eds.), *Optimality Theory and Minimalism: A possible convergence?* (pp. 195–231). Universitätsverlag Potsdam.
- Fanselow, G. (2020). Is the OV-VO distinction due to a macroparameter? In M. Tanaka, T. Tsutsui & M. Hashimoto (Eds.), *Linguistic research as an interdisciplinary science* (pp. 1–26). Tokyo: Hitsuji Publishers.
- Fanselow, G. & Lenertová, D. (2011). Left peripheral focus: mismatches between syntax and information structure. *Natural Language & Linguistic Theory*, 29(1), 169–209.
- Fehrmann, D. & Junghanns, U. (2012). “Missing” and “misplaced” subjects. In P. Bański, B. Łukaszewicz, M. Opalińska & J. Zaleska (Eds.), *Generative investigations: Syntax, morphology, and phonology* (pp. 63–92). Newcastle upon Tyne: Cambridge Scholars Publishing.
- Franks, S. (1998). *Clitics in Slavic*. Position paper presented at the Comparative Slavic Morphosyntax Workshop, Indiana University, Bloomington.
- Franks, S. (2016). Clitics are/become minimal(ist). In F. Marušič & R. Žaucer (Eds.), *Formal studies in Slovenian syntax*. Amsterdam: John Benjamins.
- Franks, S. (2017). *Syntax and spell-out in Slavic*. Bloomington: Slavica Publishers.
- Franks, S. & Bošković, Ž. (2001). An argument for multiple spell-out. *Linguistic Inquiry*, 32(1), 174–83.
- Franks, S. & King, T. (2000). *A handbook of Slavic clitics*. Oxford: Oxford University Press.
- Fried, M. (1994). Second-position clitics in Czech. *Lingua*, 94, 155–75.
- Féry, C. (2017). *Intonation and prosodic structure*. Cambridge: Cambridge University Press.
- Grimshaw, J. (1997). Projection, heads, and optimality. *Linguistic Inquiry*, 28(3), 373–422.
- Haider, H. & Szucsich, L. (2022). Slavic languages – “SVO” languages without SVO qualities? *Theoretical Linguistics*.
- Halpern, A. (1995). *On the placement and morphology of clitics*. Stanford: Center for the Study of Language and Information.
- Huntley, D. (1993). Old Church Slavonic. In B. Comrie & G. G. Corbett (Eds.), *The Slavonic languages* (pp. 125–187). London: Routledge.
- Junghanns, U. (2002a). Clitic Climbing im Tschechischen. In U. Junghanns (Ed.), *Untersuchungen zur Syntax und Informationsstruktur slavischer Deklarativsätze* (pp. 57–90). Leipzig: Institut für Linguistik der Universität Leipzig.

- Junghanns, U. (2002b). Klitische Elemente im Tschechischen: Eine kritische Bestandsaufnahme. In T. Daiber (Ed.), *Linguistische Beiträge zur Slavistik IX* (pp. 117–49). München: Otto Sagner.
- Junghanns, U. (2002c). Zur Verbstellung im Deklarativsatz des Tschechischen. In U. Junghanns (Ed.), *Untersuchungen zur Syntax und Informationsstruktur slavischer Deklarativsätze* (pp. 7–56). Leipzig: Institut für Linguistik der Universität Leipzig.
- Junghanns, U. (2003). *Non-cliticizing clitics?* Paper presented at the Fifth Conference on Formal Description of Slavic Languages (FDSL-5), University of Leipzig, 26–28 November.
- Junghanns, U. (2021). Grammatische Regel oder Möglichkeit? In B. Brehmer, A. Gattnar & T. Perevozchikova (Eds.), *Von A wie Aspekt bis Z wie zdvořilost. ein Kaleidoskop der Slavistik für Tilman Berger zum 65. Geburtstag* (pp. 171–184). Tübingen: Tübingen University Press.
- Junghanns, U. & Lenertová, D. (2007). Basic processes of information structuring in Czech. In U. Junghanns (Ed.), *Linguistische Beiträge zur Slavistik XIII* (pp. 233–58). München: Otto Sagner.
- Junghanns, U. & Zybatow, G. (1997). Syntax and information structure of Russian clauses. In E. W. B. et al. (Ed.), *Annual workshop on formal approaches to slavic linguistics. the cornell meeting 1995* (pp. 289–319). Ann Arbor: Michigan Slavic Publications.
- Kaiser, G. A. & Scholze, L. (2009). Verbstellung im Sprachkontakt – das Obersorbische und Bündnerromanische im Kontakt mit dem Deutschen. In L. Scholze & B. Wiemer (Eds.), *Von Zuständen, Dynamik und Veränderung bei Pygmäen und Giganten. Festschrift für Walter Breu zu seinem 60. Geburtstag* (pp. 305–30). Bochum: Universitätsverlag Brockmeyer.
- Kayne, R. (1994). *The antisymmetry of syntax*. Cambridge: MIT Press.
- Keydana, G. (2011). Wackernagel in the language of the Rigveda. A reassessment. In *Historische Sprachforschung* 124 (pp. 106–33). Göttingen: Vandenhoeck & Ruprecht.
- King, T. H. (1996). Slavic clitics, long head movement, and prosodic inversion. *Journal of Slavic Linguistics*, 4(2), 274–311.
- Kiparsky, V. (1967). *Russische historische Grammatik, Band II: Die Entwicklung des Formensystems*. Heidelberg: Carl Winter Universitätsverlag.
- Kučerová, I. (2012). Grammatical marking of givenness. *Natural Language Semantics*, 20, 1–30.
- Kulik, I. (2019). Clitic Climbing im Tschechischen, Slovenischen und Polnischen: ein Vergleich. In *Linguistische Beiträge zur Slavistik. XXVI. Jungslavistinnen-Treffen in Bamberg, 6.-8. September 2017*. München: Otto Sagner. (Preprint)
- Lazzeroni, R. (1994). Rileggendo Benveniste: Le relazioni di persona nel verbo. *Rivista di Linguistica*, 6(2), 267–74.
- Legendre, G., Raymond, W. & Smolensky, P. (2006). Optimality in syntax I: Case and grammatical voice typology. In P. Smolensky & G. Legendre (Eds.), *The harmonic mind: From neural computation to optimality-theoretic grammar* (pp. 161–81). Cambridge: MIT Press.
- Lenertová, D. (2001). On clitic placement, topicalization and CP-structure in Czech. In G. Zybatow, U. Junghanns, G. Mehlhorn & L. Szucsich (Eds.), *Current issues in formal Slavic linguistics* (pp. 294–305). Frankfurt: Peter Lang.

- Lenertová, D. (2004). Czech pronominal clitics. *Journal of Slavic Linguistics*, 12(1/2), 135–71.
- Lenertová, D. & Junghanns, U. (2007). Wide focus interpretation with fronted focus exponents in Czech. In K. Schwabe & S. Winkler (Eds.), *On information structure, meaning and form: Generalizations across languages* (pp. 347–363). Amsterdam/Philadelphia: Benjamins.
- Marantz, A. (1988). Clitics, morphological merger, and the mapping to phonological structure. In M. Hammond & M. Noonan (Eds.), *Theoretical morphology: Approaches in modern linguistics* (pp. 253–70). London: Academic Press.
- McCarthy, J. J. & Prince, A. (1993). Generalized alignment. In G. Booij & J. V. Marle (Eds.), *Yearbook of morphology*. Dordrecht: Springer.
- Meyer, R. (2004). *Syntax der Ergänzungsfrage: Empirische Untersuchungen am Russischen, Polnischen und Tschechischen* (P. Rehder, Ed.). München: Otto Sagner.
- Munn, A. B. (1993). *Topics in the syntax and semantics of coordinate structures* (Unpublished doctoral dissertation). University of Maryland.
- Naughton, J. (2005). *Czech. An essential grammar*. Abingdon: Routledge.
- Nespor, M. & Vogel, I. (2007). Prosodic domains of external sandhi rules. *Annali Online di Ferrara*, 1, 14–53. (Originally published in 1982.)
- Pesetsky, D. M. (1982). *Paths and categories* (Unpublished doctoral dissertation). MIT, Department of Linguistics and Philosophy.
- Prince, A. & Smolensky, P. (1993). *Optimality Theory: Constraint interaction in generative grammar*. MS, Rutgers University, New Brunswick, and University of Colorado, Boulder.
- Prince, A. & Smolensky, P. (2006). Optimality: From neural networks to Universal Grammar. In P. Smolensky & G. Legendre (Eds.), *The harmonic mind: From neural computation to optimality-theoretic grammar* (Vol. 1, pp. 123–43). Cambridge: MIT Press. (Originally published 1997 in *Science*, Vol. 275, pp. 1604–10.)
- Progovac, L. (1996). Clitics in Serbian/Croatian: Comp as the Second Position. In A. Halpern & A. Zwicky (Eds.), *Approaching second. Second position clitics and related phenomena* (pp. 411–28). Stanford: Center for the Study of Language and Information.
- Radanović-Kocić, V. (1996). The placement of Serbo-Croatian clitics: A prosodic approach. In A. Halpern & A. Zwicky (Eds.), *Approaching second. Second position clitics and related phenomena* (pp. 429–46). Stanford: Center for the Study of Language and Information.
- Richardson, M. (1997). Czech clitics: 2P or not 2P, that is the question. In L. Kaiser (Ed.), *Studies in the morphosyntax of clitics* (pp. 131–50). New Haven: Yale University.
- Rivero, M. L. (1991). Long head movement and negation. *Linguistic Review*, 8, 319–51.
- Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (Ed.), *Elements of grammar: Handbook in generative syntax* (pp. 281–337). Dordrecht: Kluwer.
- Rudin, C. (1988). On multiple questions and multiple wh fronting. *Natural Language & Linguistic Theory*, 6, 445–501.
- Scholze, L. (2007). *Das grammatische System der obersorbischen Umgangssprache unter besonderer Berücksichtigung des Sprachkontakts* (Unpublished doctoral dissertation). Universität Konstanz, Fachbereich Sprachwissenschaft.
- Scholze, L. (2012). Deutscher Einfluss auf die obersorbische Wortfolge – Vergleich von Schrift- und Umgangssprache. In S. Wölke & H. Bartels (Eds.), *Einflüsse des Deutschen auf*

- die grammatische Struktur slawischer Sprachen. Internationale Konferenz des Sorbischen Instituts 2011* (pp. 198–213). Bautzen: Domowina-Verlag.
- Schütze, C. T. & Sprouse, J. (2013). Judgment data. In R. J. Podesva & D. Sharma (Eds.), *Research methods in linguistics*. Cambridge: CUP.
- Selkirk, E. (1986). On derived domains in sentence phonology. *Phonology Yearbook*, 3, 371–405.
- Short, D. (1993a). Czech. In B. Comrie & G. G. Corbett (Eds.), *The Slavonic languages* (pp. 533–92). London: Routledge.
- Short, D. (1993b). Slovak. In B. Comrie & G. G. Corbett (Eds.), *The Slavonic languages* (pp. 455–532). London: Routledge.
- Siewierska, A. & Uhliřová, L. (1998). An overview of word order in Slavic languages. In A. Siewierska (Ed.), *Constituent order in the languages of Europe* (pp. 105–150). Berlin, New York: De Gruyter Mouton.
- Šimík, R. & Jasinskaja, K. (2022). There is no single Slavic word order type. *Theoretical Linguistics*.
- Šimík, R. & Wierzba, M. (2017). Expression of information structure in West Slavic: Modeling the impact of prosodic and word-order factors. *Language*, 93(3), 671–709.
- Šimík, R., Wierzba, M. & Kamali, B. (2014). Givenness and the position of the direct object in the Czech clause. In *Proceedings of FASL 22* (pp. 418–34).
- Skopeteas, S. (2010). Syntax-phonology interface and clitic placement in Mayan languages. In V. Torrens, L. Escobar, A. Gavarro & J. Gutierrez (Eds.), *Movement and clitics: Adult and child grammar* (pp. 306–332). Cambridge Scholars Publishing.
- Smolensky, P. & Legendre, G. (2006). Harmony optimization and the computational architecture of the mind/brain. In P. Smolensky & G. Legendre (Eds.), *The harmonic mind: From neural computation to optimality-theoretic grammar* (Vol. 1, pp. 3–61). Cambridge: MIT Press.
- Smolensky, P., Legendre, G. & Tesar, B. (2006). Optimality Theory: The structure, use, and acquisition of grammatical knowledge. In P. Smolensky & G. Legendre (Eds.), *The harmonic mind: From neural computation to optimality-theoretic grammar* (Vol. 1, pp. 453–544). Cambridge: MIT Press.
- Song, J. J. (2012). *Word order*. Cambridge: Cambridge University Press.
- Spencer, A. & Luís, A. (2012). *Clitics: An introduction*. Cambridge: Cambridge University Press.
- Stjepanović, S. (2004). Clitic climbing and restructuring with “finite clause” and infinitive complements. *Journal of Slavic Linguistics*, 12(1/2), 173–212.
- Sturgeon, A., Harizanov, B., Polinsky, M., Kravtchenko, E., Gallo, C. G., Medová, L. & Koula, V. (2012). Revisiting the person case constraint in Czech. In *Proceedings of the 28th meeting of Formal Approaches to Slavic Linguistics (FASL)*. Michigan Slavic Publications.
- Sussex, R. & Cubberley, P. (2006). *The Slavic languages*. Cambridge: Cambridge University Press.
- Szucsich, L. & Haider, H. (2015). Freie Wortstellung in slavischen Sprachen und die VO/OV-Unterscheidung. In E. Dieser (Ed.), *Linguistische Beiträge zur Slavistik: XX. JungslavistInnen-Treffen in Würzburg, 22.-24. September 2011 (Specimina Philologiae*



- Slavicae*) (pp. 94–124). München: Otto Sagner.
- Toman, J. (1980). Weak and strong: notes on *be* in Czech. In G. Brettschneider & C. Lehmann (Eds.), *Wege zur Universalienforschung. Sprachwissenschaftliche Beiträge zum 60. Geburtstag von Hansjakob Seiler* (pp. 305–310). Tübingen: Gunter Narr.
- Toman, J. (1986). Cliticisation from NPs in Czech and comparable phenomena in French and Italian. In H. Borer (Ed.), *The syntax of pronominal clitics* (pp. 123–47). New York: Academic Press.
- Toman, J. (1996). A note on clitics and prosody. In A. Halpern & A. Zwicky (Eds.), *Approaching second. Second position clitics and related phenomena*. Stanford: Center for the Study of Language and Information.
- Toman, J. (1999). On clitic (dis)placement. In M. Dimitrova-Vulchanova & L. Hellan (Eds.), *Topics in South Slavic syntax and semantics* (pp. 205–28). Amsterdam/Philadelphia: John Benjamins.
- van Riemsdijk, H. C. (1978). *A case study in syntactic markedness: The binding nature of prepositional phrases*. Lisse: The Peter de Ridder Press.
- Veselovská, L. (1995). *Phrasal movement and X<sup>0</sup>-morphology* (Unpublished doctoral dissertation). Palacký University Olomouc.
- Veselovská, L. (2008). The extended verbal projection in Czech: Three variants of the verb *be*. In G. Zybatow, L. Szucsich, U. Junghanns & R. Meyer (Eds.), *Formal Description of Slavic Languages: The fifth conference, Leipzig 2003* (pp. 555–69). Frankfurt: Peter Lang.
- Vintr, J. (1991). Das Tschechische. In P. Rehder (Ed.), *Einführung in die slavischen Sprachen* (pp. 74–87). Darmstadt: Wissenschaftliche Buchgesellschaft.
- Wackernagel, J. (1892). Über ein Gesetz der indogermanischen Wortstellung. *Indogermanische Forschungen*(1), 333–436.
- Wanner, D. (1996). Second position clitics in medieval Romance. In A. Halpern & A. Zwicky (Eds.), *Approaching second. Second position clitics and related phenomena* (pp. 537–578). Stanford: Center for the Study of Language and Information.
- Zimmerling, A. (2012, August). *Sentential clitic hosts in Slavic languages*. (Handout of a talk given at SLS 7)
- Zwart, J.-W. (2017). An argument against the syntactic nature of verb movement. In L. R. Bailey & M. Sheehan (Eds.), *Order and structure in syntax I: Word order and syntactic structure* (pp. 29–47). Berlin: Language Science Press.
- Zwicky, A. (1977). *On clitics*. Bloomington: Indiana University Linguistics Club.
- Zwicky, A. & Pullum, G. (1983). Cliticization vs. inflection: English n't. *Language*, 59(3), 502–13.