

©2010

Slavica Kochovska

ALL RIGHTS RESERVED

MACEDONIAN DIRECT OBJECTS, CLITICS AND THE LEFT PERIPHERY

by

SLAVICA KOCHOVSKA

A Dissertation submitted to the

Graduate School-New Brunswick

Rutgers, The State University of New Jersey

in partial fulfillment of the requirements

for the degree of

Doctor of Philosophy

Graduate Program in Linguistics

written under the direction of

Veneeta Dayal

and approved by

New Brunswick, New Jersey

May, 2010

ABSTRACT OF THE DISSERTATION

Macedonian Direct Objects, Clitics and the Left Periphery

By SLAVICA KOCHOVSKA

Dissertation Director:

Veneeta Dayal

This dissertation investigates the patterning of Macedonian direct object DPs with respect to clitic pronouns. The occurrence of clitics with direct objects in Macedonian varies along two dimensions, the type of DP and the position of DP. Two broad classes are considered, regular and *wh* DPs, and positions both within IP and at the left periphery.

The patterning of direct objects can be schematically represented as follows:

- (i) [_{IP} Subj (CL) V QP/*wh*]
- (ii) [_{IP} Subj QP/*wh* (CL) V]
- (iii) [_{XP}QP/*wh* [_{IP} Subj (CL) V]]

The clitic co-occurrence restrictions in Macedonian are argued to be sensitive to the feature specification for strength of the DPs. Within each class of DPs, I show that there are well defined distinctions which correlate with the DPs' distinctive semantic properties. In particular, the specification of the regular DPs is linked to their strength in the sense of Barwise & Cooper (1981); the specification of the *wh* DPs is linked to their

D-linking properties in the sense of Pesetsky (1987) and Comorovski (1996). The occurrence of the clitics is shown to be regulated by the Clitic Criterion (Sportiche 1998), with the clitic licensing a particular feature in the DP with which it co-occurs. The licensing is carried out through an agreement relation, in a Spec-Head configuration.

It is proposed that only [+strong] DPs satisfy the Clitic Criterion. As such, within IP only strong DPs trigger the appearance of the clitic pronoun. At the left periphery, however, weak DPs can optionally occur with clitics. This can be explained through the possibility of some weak DPs entering into a binding relation with *pro* in argument position, which satisfies the Clitic Criterion. This explanation depends on a characterization of the DPs as [+strong], [-strong] and unspecified.

I present arguments to establish that DPs at the left periphery that occur with clitics differ derivationally from those without clitics. I further establish on the basis of subject intervention effects, that wh-phrases with clitics occupy a higher position than wh-phrases without clitics, and suggest that the same holds for regular DPs.

Acknowledgements

Writing this dissertation has been an extraordinary experience in so many ways. It is now time to thank all the people who have made that experience enjoyable and finishing the dissertation possible.

I have been very fortunate to have Veneeta Dayal, Mark Baker, Roger Schwarzschild and Barbara Partee as my committee.

Veneeta Dayal has been my advisor, mentor and role model throughout my years at Rutgers. She was the Graduate Program Director who recruited me, my first semantics teacher and a constant source of support and encouragement throughout my studies. I would not have become the person and the linguist that I am without her. Her professionalism and integrity are unsurpassed and her breadth of knowledge remarkable. I am indebted to her for her wisdom, support, persistence, encouragement, and generosity with time and advice. I thank her for the joy and enthusiasm she brought to the work, for constantly asking the right questions and for guiding me through to the answers. Most of all, I thank her for believing in me and for believing in my work.

Mark Baker has been a constant source of inspiration throughout my studies at Rutgers, both as a teacher and as an advisor. I was very fortunate to have him on the committee of my second qualifying paper and the dissertation. His knowledge, helpful guidance and valuable insights greatly contributed to my understanding of the issues in the dissertation. I am very grateful to him for his unfailing support and interest in my work, for his willingness to listen, for his preciseness of thought, for his always insightful comments and helpful advice. Working with him has been an extraordinary experience.

Roger Schwarzschild has a unique ability to bring out the best in his students. I was very fortunate to have him on both the committee of my second qualifying paper and the dissertation. His understanding of theory, his genuine interest, his valuable comments and suggestions have helped me understand the significance of the work I was doing. I thank him for his generosity with time and attention, for his advice, kindness and support and for his unwavering enthusiasm. Working with him has been a truly rewarding experience.

I am very grateful to Barbara Partee for agreeing to be on my committee and I thank her for her wholehearted support and enthusiasm for the topic. Her detailed comments on every aspect of the analysis significantly contributed to the quality of the dissertation. I cannot imagine the dissertation being what it is without her input and valuable advice.

I thank my teachers and everyone at the linguistics department at Rutgers for their continuous support and encouragement and for helping me grow as a linguist. Alan Prince was my advisor on my first qualifying paper and a mentor ever since. I thank him for his unfailing interest in my work and steady encouragement. Special thanks for his help as a Graduate Program Director while finishing the dissertation. Viviane Déprez guided me through my second qualifying paper. I thank her for being always very supportive of my work, for her positive attitude and enthusiasm. Jane Grimshaw taught me how to be a teacher. I thank her for that wonderful gift, for her unwavering belief in me and for always having something helpful to say about my analyses. Ken Safir introduced me to the topics of this dissertation in my first year at Rutgers. I thank him for always being there to help, for his generosity with advice and for his enthusiasm for my work. I also thank Akin Akinlabi, Maria Bittner, Paul de Lacy, Bruce Tesar, José Camacho and Liliana Sánchez for their support, willingness to help and constant interest.

I thank Joanna Stoechr for always being there for me, for looking after me, for helping me with all things administrative and for making Rutgers feel like home. The department would not have been the same without her. I also thank Danielle Berlingieri for her assistance during the final stages of my studies and for her help in filing this dissertation.

I thank my friends and fellow students in and outside the department for contributing to my dissertation in numerous ways and for making my stay at Rutgers enjoyable. I thank Oluseye Adesola, Daniel Altsuler, Matthew Barrosa, Adrian Brasoveanu, Maia Duguine, Jose Elias-Ulloa, Carlos Fasola, Daphna Heller, Huynjoo Kim, Seunghun Lee, Xiao Li, John Manna, Vita Markman, Naz Merchant, Sarah Murray, Jessica Rett, Heather Robinson, Andrea Sims and Zuzana Tomková. Special thanks to my classmates Seunghun Lee, Xiao Li and Jessica Rett for being such wonderful friends and for always being there for me. Also, thanks to Heather Robinson, for being a great neighbor, an understanding friend and a patient listener.

Many other people have contributed to my understanding of the issues discussed in this dissertation and have helped me develop some of the ideas presented here, whether through passing comments or detailed discussions. They are Olga Arnaudova, Željko Bošković, Anna Cardinaletti, Steven Franks, Anastasia Giannakidou, Maria Guskova, Ileana Krapova, Meredith Landman, Jason Merchant, Jairo Nunes, Roumyana Pancheva, Liljana Progovac, Maribel Romero, Liliana Sánchez, Yael Sharvit and Ana Szabolcsi.

I became interested in linguistics during my undergraduate studies at the University of Auckland. I thank Sheila Dooley-Colberg, Harry Leder, Frank Lichtenberk and the late Scott W. Allan, for introducing me to generative linguistics, for being such wonderful teachers and for encouraging me to become a linguist. I have never regretted it. I am also

very grateful to Brian D. Joseph and Elizabeth Hume, at Ohio State University, for playing a significant role in that development.

I have been very fortunate to become part of the linguistic department at the University of Nova Gorica. I am indebted to Franc Lanko Marušič for making this possible. I also thank my colleagues Rok Žaucer, Penka Stateva, Arthur Stepanov, Amanda Saksida and Alja Ferme for making the department a great place to work. In particular, I thank Penka and Arthur for their help and numerous discussions on the ideas presented in this dissertation. I thank Lanko and Rok for always being willing to listen, for their suggestions and most of all for being so supportive and encouraging. The completion of the dissertation would have been impossible without their help. At the University of Nova Gorica, I also thank Martina Bergant for being a great friend.

I am indebted to Eduard Gulija, Liljana Kočovska, Gjorgi Kočovski, Igor Kochovski, Milka Kalajdziska, Panta Petrušev, Ivan Petrušev and Tatjana Rantaša for judgments on Macedonian and Mirela Dragomir for judgments on Romanian.

I thank my relatives in Macedonia for their love and affection and for always thinking of me. I especially thank my aunt Vera for being a constant source of strength and support. I also thank my friend Tanja for being the best friend possible.

Finally, my parents, Liljana and Gjorgi, and my brother Igor, have always been there for me. I thank Igor for the laughter and joy he brings, for being the most positive person ever and for showing me that life without art is impossible. And I thank my parents, for everything. Without the love of my family, their unlimited support and understanding I would not have been able to begin this journey. And it is only because of their faith in me that I have been able to complete it. This dissertation is dedicated to them, with love.

Dedication

*To my parents, Liljana and Gjorgi,
and my brother Igor*

Table of Contents

Abstract	ii
Acknowledgments	iv
Dedication	viii
Table of Contents	ix
 1. INTRODUCTION	 1
1.1 Clitic Doubling	2
1.2 Weak DPs with Clitics	4
1.3 Topicalization vs. Clitic Left Dislocation	5
1.4 Wh-Phrases and Clitics	6
1.5 Wh-Phrases and Subjects	8
 2. DIRECT OBJECTS AND CLITIC CO-OCCURRENCE RESTRICTIONS WITHIN THE IP	
DOMAIN	10
2.0 Introduction	10
2.1 Overview of the Basic Patterns of Clitic Doubling in Macedonian	11
2.1.1 Formal Representation of Clitic Doubling	22
2.2 Strong and Weak DPs in Macedonian	31
2.2.1 Barwise & Cooper (1981)	32
2.3 Preverbal Direct Objects	36

2.3.1	The Syntactic Relationship between the Clitic Phrase and Preverbal DPs	37
2.3.2	The Relationship between Pre- and Postverbal Positions	42
2.4	Conclusion	46
3.	CLITIC CO-OCCURRENCE RESTRICTIONS ON DIRECT OBJECTS IN THE LEFT PERIPHERY	48
3.0	Introduction	48
3.1	An Unexpected Division among Weak DPs	49
3.2	A Three-Way Distinction in Feature Specification	52
3.3	The Semantic Basis for the Distinction	60
3.4	Partitive Phrases	71
3.5	Conclusion	78
4.	DISTINCTIONS AT THE LEFT PERIPHERY	80
4.0	Introduction	80
4.1	Direct Objects and the Left Periphery	82
4.1.1	The Architecture of the Left Periphery	82
4.1.2	HTLD and CLLD	90
4.1.3	CLLD and Topicalization	95
4.2	Base-Generation vs. Movement	98
4.2.1	Arguments against Clitic Doubling as the Source for CLLD	98
4.2.2	Weak Crossover Effects	102

4.2.3	Island Effects	105
4.3	Semantic and Pragmatic Differences between CLLD and TOP	111
4.3.1	Scope Properties of CLLD and Topicalization	111
4.3.2	Available Readings of CLLD and Topicalization	115
4.3.3	Discourse Contexts for CLLD and Topicalization	122
4.4	Conclusion	126
5.	CLITIC CO-OCCURRENCE RESTRICTIONS IN WH-QUESTIONS	127
5.0	Introduction	127
5.1	Basic Facts about Wh-Questions in Macedonian	128
5.2	The Distribution of Clitics in Wh-Questions	130
5.3	D-Linking and Partitivity	133
5.4	Wh-Phrases and the Clitic Criterion.....	141
5.5	CLLDed Wh-Questions vs. Ordinary Wh-Questions	146
5.5.1	Weak Crossover Effects	146
5.5.2	Island Effects	149
5.6	A Loose End.....	153
5.7	Conclusion	158
6.	WH-EXPRESSIONS AT THE LEFT PERIPHERY	160
6.0	Introduction	160
6.1	Further Differences between D-Linked and Non-D-linked Wh-Phrases	161

6.1.1	Subject Intervention Effects in Wh-Questions	161
6.2	An Analysis	165
6.2.1	Adverbs in Wh-Questions	165
6.2.2	An Explanation for Subject Intervention Effects	173
6.3	Optional D-Linking and Subject Intervention Effects	184
6.4	Three Open Questions	188
6.4.1	The Position of CLLDed and Topicalized QPs	188
6.4.2	Topicalization of Strong QPs	193
6.4.3	CLLD and Topicalization in Embedded Contexts	194
6.5	Conclusion	198
REFERENCES		200

CHAPTER 1

Introduction

The goal of this dissertation is to explain the distribution of direct object DPs in Macedonian and their interaction with clitic pronouns.¹ In particular, it investigates direct objects in positions within IP and at the left periphery and the co-occurrence restrictions of clitic pronouns with the DPs in each of these environments.

Two types of expressions participate in these constructions, regular and *wh* DPs. I show, however, that within each type there are distinctions that affect the distribution of clitics. The distinctions are based on a proposed feature specification for strength. In the case of regular DPs it is related to their strength (in the sense of Barwise & Cooper (1981)) and in the case of the *wh* DPs it is related to their D-linking properties (in the sense of Pesetsky (1987) and Comorovski (1996)). Given this, the basic claim is that clitic co-occurrence with direct objects in Macedonian is regulated in a systematic way and is a result of broader principles of grammar.

The study also shows the close interplay between the occurrence of the clitics and the position of the DPs in the clause. I investigate the properties of the DPs both in the IP and CP domain and how this in turn informs their featural make-up. The question links to the more general issue of what positions are available to the DPs in both these domains.

¹ Macedonian is a South Slavic language spoken in the Republic of Macedonia, as well as in parts of eastern Albania, northern Greece (Aegean Macedonia), and south-western Bulgaria (Pirin Macedonia). Macedonian also belongs to the group of Balkan languages (which include both Slavic and non-Slavic languages, e.g. Bulgarian, Romanian, Albanian, etc.) and as such shares a number of characteristics with them (e.g. the co-occurrence of clitics with DPs).

The basic word order in Macedonian is SVO; permutations in word order, combined with differences in intonation, signal the topic and focus status of elements (Friedman 1993). Macedonian is a pro drop and a multiple fronting language. For a more in-depth overview of the properties of Macedonian, see Friedman (1993). The data in the dissertation comes from standard spoken Macedonian.

The dissertation does not address the behavior of subjects and indirect objects with respect to either of these two phenomena. I hope that the insights of the analysis will have some relevance for them as well, though I am aware that their behavior is not identical to that of direct objects and thus will require independent investigation and explanations.

To give an idea of the types of issues I will be concerned with, I give a brief overview of each of the chapters and the analyses offered there.

1.1 Clitic Doubling

Chapter 2 investigates the distribution of clitic pronouns with direct object DPs within IP. More specifically, it aims to explain the co-occurrence restrictions on clitics with direct objects in preverbal and postverbal position. The discussion thus concentrates on two instantiations of Clitic Doubling (CL-D), i.e. the doubling of DPs when they occur in their base position and when they occur in what I identify as a focus position. The two basic patterns are illustrated in (1a-b) and (2a-b).

(1) a. Petar *(ja) pročitā sekoja kniga.

Petar it read every book

‘Petar read every book.’

b. Petar (*gi) pročitā dve knigi.

Petar them read two books

‘Petar read two books.’

In (1a-b) the universal and numeral occupy a postverbal position. In this case, the clitic is obligatory with the universal, but unacceptable with the numeral. As (2a-b) show, the same requirements hold for DPs when they are in preverbal position.

(2) a. Petar sekoja kniga *(ja) pročitā.

Petar every book it read

‘It was every book that Petar read.’

b. Petar dve knigi (*gi) pročitā.

Petar two books them read

‘It was two books that Petar read.’

To explain the patterns in (1)-(2), I argue that the doubling of direct objects is driven by the checking of a [+strong] feature in the DPs, where [+strong] is based on the definition of strength in Barwise & Cooper (1981). Two types of DPs are posited, strong and weak; only strong DPs are predicted to trigger the presence of a clitic. The licensing of the clitic is regulated by the Clitic Criterion (following Sportiche 1998), which requires that the doubled constituent enter in an agreement relation with the clitic. In this context, I show that the distribution of clitics with preverbal DPs in (2a-b) is the same as that of postverbal DPs in (1a-b). The chapter also provides evidence for the existence of a separate FocP position above VP to which direct objects can move in Macedonian. This position is the one identified for the preverbal DPs.

1.2 Weak DPs with Clitics

Chapter 3 is a continuation of chapter 2 in the sense that it looks at the same set of DPs that are subject to clitic-doubling, i.e. strong vs. weak direct object DPs. The focus of this chapter, however, is on their behavior within the CP domain. While clitic co-occurrence in chapter 2 is argued to be a result of the characterization of the DPs as strong vs. weak, this chapter highlights a somewhat unexpected behavior in the patterning of the weak DPs in clause-initial position. Consider the data in (3a-b).

- (3) a. Sekoja kniga, Petar *(ja) pročitā.
 every book Petar it read
 ‘Every book, Petar read it.’
- b. Dve knigi, Petar (gi) pročitā.
 two books Petar them read
 ‘Two books, Petar read (them).’

In (3a-b), the examples are of a universal and a numeral in a clause-initial position. Both sentences are pronounced with a distinct ‘comma’ intonation, i.e. with a slight pause after the clause-initial element. As (3a) shows, the requirement that the universal obligatorily co-occurs with a clitic pronoun remains unchanged. Interestingly, (3b) shows that the numeral now optionally allows for a clitic.

The patterning of the weak DPs in the context of left-dislocation as in (3b) prompts a reanalysis of their properties. Based on their behavior with respect to the clitics, I show that the class of weak DPs in Macedonian are not homogenous (as suggested by their

behavior when in positions within IP), but that they split into two groups: [-strong] and unspecified. This gives rise to a three-way distinction within the DPs. Crucially, though, I show that the licensing of the clitic is still subject to the Clitic Criterion. Depending on the type of the DP, a left-dislocated DP will show up with a clitic either because of its [+strong] features or because of its binding relation to *pro* in object position, which is also [+strong]. The latter option is open not only to [+strong] DPs but also to unspecified DPs. Crucially, it is not open to [-strong] DPs whose features are incompatible with those of *pro*.

1.3 Topicalization vs. Clitic Left Dislocation

Chapter 4 argues that the clause-initial elements in Macedonian discussed in the previous chapter can be mapped onto two different structures. The DPs that co-occur with clitics are identified as instances of Clitic Left Dislocation (CLLD), following Cinque (1990); those that do not co-occur with clitics, are identified as Topicalization (TOP) constructions. The two constructions are schematically represented in (4) below.

- (4) a.
$$\begin{array}{c} [\text{XP DO}_i \quad [\text{IP S } pro_i \text{ CL}_i \text{ V } t_i \dots]] \\ \text{-----} \quad \wedge \end{array}$$
- b. i.
$$\begin{array}{c} [\text{XP DO}_i \quad [\text{IP S } t_i \quad \text{CL}_i \text{ V } t_i \dots]] \\ \wedge \quad | \end{array} \quad or$$
- ii.
$$\begin{array}{c} [\text{XP DO}_i \quad [\text{IP S V } t_i \dots]] \\ \wedge \quad | \end{array}$$

The structure in (4a) represents the derivation of left-dislocated strong or unspecified DPs. The structure in (4b-i) represent an alternative derivation for left-dislocated strong DPs; the structure in (4b-ii) is an instance of the same construction but with [-strong] or unspecified DP.

The chapter presents evidence that the two constructions, CLLD and Topicalization, differ derivationally. In particular, I propose that CLLDed elements are base-generated in their surface position while topicalized elements are generated within IP and move to their surface position. The base-generation vs. movement option entails two types of relations, i.e. the existence of binding and government chains, respectively. The binding type of relation is exemplified in (4a) above. The government type of relation is exemplified by the structures in (4b). The two types of constructions, CLLD and Topicalization, are shown to pattern differently with respect to weak crossover and syntactic islands.

In addition to their syntactic differences, I show that the two constructions can be distinguished in terms of their semantic and pragmatic properties. The chapter makes the case that clause-initial elements in Macedonian are interpreted at the left periphery.

1.4 Wh-Phrases and Clitics

Chapter 5 picks up on the concerns of chapters 2-4, relating the distribution of clitic pronouns to wh-phrases. The chapter offers further support for the claim advanced in chapter 3 that the distribution of the clitics correlates with the feature specification of the element it doubles, rather than its interpretation.

To illustrate the patterns of clitic co-occurrence with wh-phrases, consider the examples in (5a-c).

- (5) a. Što (*ja) pročitao Dino?
 what it read Dino
 ‘What did Dino read?’
- b. Koju knjigu Dino *(ja) pročitao?
 which book Dino it read
 ‘Which book did Dino read?’
- c. Koga Dino (ga) video?
 who Dino him saw
 ‘Who did Dino see?’

The data above shows that different types of wh-phrases behave differently with respect to the presence vs. absence of a clitic pronoun. In (5a) we see that the clitic is not allowed with mono-morphemic wh-phrases like *što* ‘what’; (5b) shows that the clitic is required with *koj N* ‘which N’ phrases. In (5c) we see that the clitic with *koj* ‘who’ is seemingly optional.

To account for the data in (5a-c), I argue that the wh-phrases separate into three classes: [+strong], [-strong] and unspecified for strength. I take the feature specification for strength in the case of wh-phrases to be based on their inherent possibilities for D-linking, as in Pesetsky (1987) and Comorovski (1996). Given their characterization, the distribution of the clitic follows from the Clitic Criterion. Drawing on evidence from

weak crossover and island effects, I also show that wh-questions in Macedonian can either be derived by movement or by base-generation.

1.5 Wh-Phrases and Subjects

Chapter 6 investigates the behavior of the three types of wh-phrases in (5a-c) with respect to other elements in the clause. In particular, it looks at the interaction between wh-phrases and subjects. The basic paradigm is given in (6)-(8).

(6) a. Što pročitao Dino?

what read Dino

‘What did Dino read?’

b. *Što Dino pročitao?

what Dino read

(7) a. Koja knjiga Dino ja pročitao?

which book Dino it read

‘Which book did Dino read?’

b. Koja knjiga Dino ja pročitao?

which book Dino it read

(8) a. *Koga Dino video?

who Dino saw

‘Who did Dino see?’

- b. Kogo Dino go vide?
 who Dino him saw

From the data, we see that in the case of mono-morphemic wh-phrases, like *što* ‘what’ in (6a), the subject (*Dino*) cannot intervene between the fronted wh-phrase and the verb, as indicated by the ungrammaticality of (6b). In the case of complex wh-phrases, like *koja kniga* ‘which book’ in (7a), we see that the subject (*Dino*) can intervene between the fronted wh-phrase and the verb, as illustrated by the grammaticality of (7b). The same pattern extends to *koj* ‘who’, as shown in (8a-b).

The explanation offered for the differences in (6)-(8) relies on the idea that [+strong] DPs like *koj N* ‘which N’ occupy different positions from [-strong] DPs like *što* ‘what’ at the left periphery in Macedonian. In particular, the former are placed in positions above CP, while the latter are in SpecCP. The proposal crucially relies on the positioning of the verb in C^0 in wh-questions, which is independently motivated by adverb placement tests.

The dissertation ends with a discussion of some issues that tie the analysis of wh-questions and clitics with the analysis of non wh quantifiers and clitics.

CHAPTER 2

Direct Objects and Clitic Co-Occurrence Restrictions within the IP

Domain

2.0 Introduction

This chapter focuses on the distribution of clitic pronouns in Macedonian and their co-occurrence with direct object DPs in post- and preverbal position. In particular, its goal is to explain what restrictions hold on the placement of clitics with DPs when the latter occur within the IP domain.

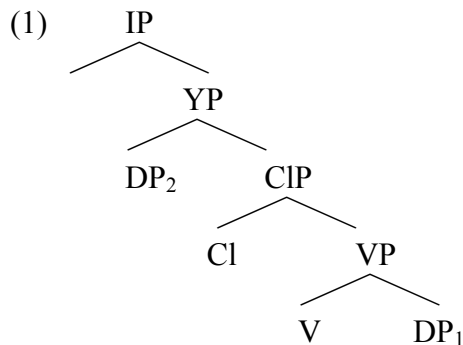
In trying to understand the properties of the doubling constructions in the language, I will concentrate on two instantiations of Clitic Doubling (CL-D), i.e. the doubling of direct object DPs when they occur in preverbal and postverbal position. My main goal here is to identify the conditions under which the doubling of DPs in Macedonian takes place. Specifically, I will argue that the doubling of direct objects is motivated by the checking of a [+strong] feature in the DPs. Thus, I will explain the presence of the clitic in relation to the semantic properties of the doubled DP. The implementation of the proposal relies on Sportiche's (1998) proposal that clitics are heads of their own maximal projections and that as such they license particular features on the doubled DP.

The chapter is organized as follows. In section 1, I give an overview of the basic data and generalizations regarding clitic doubling in Macedonian. I show that the doubling of the direct objects is not optional and I offer an analysis of the doubling constructions by showing that the determining factors for the co-occurrence of the clitics are the semantic properties of the DPs involved. The section also discussed the issue of the formal

representation of the doubling constructions. In section 2, I give a characterization of the Macedonian DPs in terms of their strong vs. weak properties and show how the two classes of DPs fit into the analysis of clitic doubling proposed in section 2. Section 3 discusses the properties of the preverbal direct objects and their relation to clitic-doubling. Section 4 contains the conclusion.

2.1 Overview of the Basic Patterns of Clitic Doubling in Macedonian

In this chapter, I concentrate on the distribution of DPs within IP and their co-occurrence restrictions with clitic pronouns. The part of the clause that is relevant to our discussion here is schematically represented in (1) below.



The structure in (1) represents one of the core claims of this chapter, i.e. that the clitic pronoun in Macedonian is a head of its own maximal projection (following Sportiche 1998). In section 2, I discuss the issues surrounding the occurrence and representation of the clitic in detail. For the moment, the structure in (1) is meant to represent the basic patterning of the two types of direct objects that I discuss here: direct objects in pre- and postverbal position. I thus begin by giving an overview of the basic patterns concerning the co-occurrence of such DPs with clitic pronouns.

Macedonian, like other languages, uses a set of so-called short pronominal forms in doubling constructions (Koneski 1987; Mišeska Tomić 2008). The pronominal clitics inflect for person and number, and, in the case of third person singular, for gender. They are given in (2) below.

(2) *Macedonian (full and short) pronominal clitics*

	ACC				DAT			
	SG		PL		SG		PL	
1 st	<i>mene</i>	<i>me</i>	<i>nas</i>	<i>ne</i>	<i>mene</i>	<i>mi</i>	<i>nam</i>	<i>ni</i>
2 nd	<i>tebe</i>	<i>te</i>	<i>vas</i>	<i>ve</i>	<i>tebe</i>	<i>ti</i>	<i>vam</i>	<i>vi</i>
3 rd (m./n.)	<i>nego</i>	<i>go</i>	<i>niv</i>	<i>gi</i>	<i>nemu</i>	<i>mu</i>	<i>nim</i>	<i>im</i>
3 rd (f.)	<i>nea</i>	<i>ja</i>	<i>niv</i>	<i>gi</i>	<i>nejze</i>	<i>ì</i>	<i>nim</i>	<i>im</i>

The distribution of the clitic pronouns with direct object DPs is as follows.¹

Definite DPs and demonstratives are obligatorily clitic-doubled, as shown in (3a).

Clitics also obligatorily co-occur with full pronouns, as shown in (3b).

¹ Indirect objects can also be clitic-doubled, in which case, the dative clitics are used. Subjects, on the other hand, are never clitic-doubled in Macedonian. In what follows, I concentrate on the distribution of clitic pronouns with direct objects. I have nothing to say about subjects or indirect objects and their clitic requirements. I hope that the insights will be applicable to the analysis of those two types of constituents. I leave the actual implementation aside for further research.

(3) a. Ivan *(ja) pročitá knigata/ovaa kniga.

Ivan it(f.sg.) read book-the/this book

‘Ivan read the book.’

b. Ana go vide nego.

Ana him saw him

‘Ana saw him.’

It is worth noting that the definite DP in (3a) can have both an attributive and a referential reading, in the sense of Donnellan (1966) (see Abbott 2004, among others). The two uses of the definite are also illustrated in (4).

(4) Ana *(go) bara profesorot.

Ana him look-for professor-the

a: ‘Ana is looking for the professor (i.e. she is looking for John Smith, who happens to be the professor).’

b: ‘Ana is looking for the professor (whoever he might be).’

On its (a) reading, the definite *profesorot* ‘the professor’ in (4) is used referentially, i.e. the description used holds of a particular individual that the speaker has in mind. In this case the definite is used to describe that particular individual. On its (b) reading, the definite is used attributively, i.e. the speaker says something about whoever happens to fit the description that is being used. Although the definite in (4) can receive both a

referential and an attributive interpretation, the clitic requirement does not change. In both cases, the definite must be clitic-doubled.

Just like the definites and pronouns, proper names are also obligatorily clitic-doubled, as shown in (5).

(5) Pero *(ja) pokani Ana na zabava.

Pero her invited Ana to party

‘Pero invited Ana to a party.’

At this point, we may note that the doubling of the direct object DPs in Macedonian is not subject to what is known as “Kayne’s generalization” (see Sportiche 1998, Anagnostopoulou 1999, 2006, etc.), which informally states that the doubling of a constituent is only possible if it is preceded by a special preposition (such as the object marker *pe* in Romanian, for example). The failure of languages to conform to “Kayne’s generalization” has already been established in previous studies, for example, Suñer (1988) for Spanish, Anagnostopoulou (1994) and Kallulli (1999) for Albanian and Greek, etc.. (For a general discussion on this issue, see Anagnostopoulou (2006)). That Macedonian, like other languages, violates this requirement is illustrated by the data in (3)-(5), where the direct objects are doubled, and yet not preceded by any prepositions.

Our next task is to determine how quantifiers behave in doubling constructions. When it comes to clitic doubling, the quantifiers in Macedonian split into two groups: those that obligatorily co-occur with a clitic and those that cannot co-occur with a clitic

pronoun.² The first group includes the strong quantifiers *sekoj* ‘every’, *site* ‘all’, and *povekjeto* ‘most’. The second group includes the weak quantifiers: bare numerals, *mnogu/malku* ‘many/few’, *nekolku* ‘several/a few’.³ Illustrative examples are given in (6) and (7), respectively.⁴

(6) Dino *(gi) pročita site/povekjeto knigi.

Dino them read every/most books

‘Dino read every/most books.’

(7) Dino (*gi) pročita dve/mnogu/malku/nekolku knigi.

Dino them read two/many/few/several books

‘Dino read two/many/few/several books.’

The behavior of the quantifiers with respect to clitic doubling is interesting in that it offers us a glimpse into the underlying factors that can potentially be responsible for the doubling of the direct objects. In particular, the fact that quantifiers in Macedonian can be

² Clitic doubling of quantifiers is also possible in other languages, e.g. Greek (Iatridou 1995, Anagnostopoulou 2006, Alexiadou 2006, Kallulli 1999, 2008), Albanian (Kallulli 1999, 2008), Spanish (Arregi 2003), etc. As it is the case with Macedonian, not all quantifiers allow for the co-occurrence of a clitic pronoun. Although none of these studies discusses the doubling of quantifiers in great detail, we can see from their examples that the restrictions on the clitics’ distribution are roughly the same as in Macedonian. It remains to be seen whether or not the exact parallels that I draw here between the clitics and the quantifiers hold for other languages as well.

³ The categorization of the quantifiers as weak/strong follows Barwise & Cooper (1981). In section 2.2, I return to this question and give a more detailed discussion of the classification of the Macedonian determiners with respect to this semantic property.

⁴ Interestingly, some quantifiers can be clitic-doubled, but the presence of the clitic in such cases is determined by the position the quantifier occupies in the clause. The group of quantifiers that participate in clitic constructions, but only when dislocated, includes: specific indefinites, bare numerals (on a specific reading), *several* (on a specific reading). In chapter 3, I return to the question of what motivates the co-occurrence of clitics with dislocated weak quantifiers and offer an explanation for their behavior.

clitic-doubled (and in the case of strong quantifiers the fact that they must be; see (6)) means that the referentiality of the DPs is not a determining factor for their doubling.⁵

It has been claimed that clitic doubling in Macedonian is driven by the specificity of the doubled object DPs (see Franks & King 2000).⁶ More generally, Sportiche (1998) also proposes that the property licensed by the clitic in a clitic doubling construction is specificity. This, however, cannot be the case for Macedonian.

⁵ The notion of referentiality I have in mind here is that of Cinque (1990). (Cinque's notion of referentiality is useful at this point in that it separates in some sense the class of quantifiers from that of the definites. In the discussion that follows, though, I show that the two types of DPs are grouped together as strong in the sense of Barwise & Cooper (1981).)

Cinque, following Rizzi (1990), takes referential phrases to be those that are used to refer to specific members of a preestablished set in the discourse. Referential phrases, according to Cinque, are the type of phrases that i) can establish coreference relations, and ii) do not show weak cross over effects. Quantifiers, according to these two tests, come out as nonreferential. They cannot corefer with a pronoun which they do not c-command and also give rise to weak crossover effects. Illustrative examples of both tests are given in (i) and (ii), respectively (from Cinque (1990:8-10)):

- (i) a. [Gli alunni che dovevano visitare *il museo*]*lo* hanno visitato in fretta.
the pupils who had to visit *the museum* visited *it* hurriedly
- b. *[Gli alunni che dovevano visitare *ogni museo*] hanno finito per visitar*lo* in fretta.
the pupils who had to visit *every museum* ended up visiting *it* hurriedly
- (ii) a. *Sua* madre ha presentato *Maria* ad un ragazzo.
her mother introduced *Maria* to a boy
- b. **Sua* madre ha presentato *ogni ragazza* ad un ragazzo.
her mother introduced *every girl* to a boy

The corresponding examples in Macedonian are given in (iii) and (iv), respectively.

- (iii) a. Učenicite koišto go_i posetija muzejot_i go_i posetija nabrzina.
pupils-the who it(m.sg.) visited museum-the it(m.sg.) visited hurriedly
- b. *Učenicite koišto go_i posetija sekoj muzej_i
pupils-the who it(m.sg.) visited every museum
go_i posetija nabrzina.
it(m.sg.) visited hurriedly
- (iv) a. Majka i_{i/j} ja_i pretstavi Marija_i na edno momče.
mother her.poss her.acc introduced Marija to one boy
'Her mother introduced Marija to a boy.'
- b. *Majka i_{*i/j} ja_i pretstavi sekoja devojka_i na edno momče.
mother her her.acc introduced every girl to one boy
'Her mother introduced every girl to a boy.'

Cinque goes on to establish a direct link between the referentiality of DPs and their co-occurrence with clitics. He argues that only referential phrases can be left-dislocated and coindexed with a clitic pronoun. The details of his proposal will be given in chapter 3 where I discuss the Clitic Left Dislocation constructions in Macedonian.

⁶ Franks & King (2000) claim that clitic doubling in Macedonian is obligatory for specific direct objects. Their list of specific DPs includes definites, proper names, and pronouns. They follow Enç (1991) in treating all definites as specific.

We have just seen that definites in Macedonian can receive two types of interpretation, referential and attributive. One could in this case think of the referential interpretation as specific interpretation and the attributive as nonspecific interpretation.⁷ However, even if we take the attributive to be nonspecific, the clitic requirement remains unchanged. In other words, the definite must always co-occur with a clitic pronoun, regardless of its interpretation. This was shown in (4) where a definite expression is interpreted both attributively and referentially, but is invariably doubled by the clitic.

A stronger argument against the claim that specificity plays a role in clitic doubling comes from the indefinites in data like (8). As (8) shows, an indefinite cannot be clitic-doubled in Macedonian. The indefinite can, however, receive a specific and a nonspecific interpretation, represented here as an (a) and (b) reading, respectively.⁸ Regardless of which interpretation it receives, the clitic is not allowed.

(8) Ivan (*ja) bara edna sekretarka.⁹

Ivan her look-for one secretary

a: 'Ivan is looking for a secretary (whom he spoke with yesterday).'

⁷ This parallel between referential/attribution and specific/nonspecific, respectively, has been suggested by Partee (1970, 2004). In particular, she suggests that the referential and attribution (i.e. nonreferential) interpretation of the definites (Donnellan 1966) be extended to the indefinites. A different parallel between referential/attribution and specific/nonspecific can be found in Kallulli (1999). Kallulli argues that the specific/nonspecific distinction does not extend to the attribution/referential since specific NPs can get both a referential and an attribution interpretation.

⁸ I take the specific reading of an indefinite to be same as the referential reading as identified by Fodor & Sag (1982). Thus, a specific indefinite is that which is used by the speaker (but not the hearer) to identify a particular referent in discourse. Semantically, I take specific indefinites to be singleton expressions, following Schwarzschild (2002). This means that the indefinites have an implicit restriction on their domain, i.e. the domain of the indefinite has a singleton extension. I return to the issue of specific indefinites in chapter 3.

⁹ When *eden* 'one' is interpreted as a numeral (in which case it is stressed), the clitic is not allowed.

(i) Ivan (*ja) bara edna sekretarka.
Ivan her look-for one secretary
'Ivan is looking for one secretary (not two, for example).'

b: ‘Ivan is looking for a secretary (whoever she may be)’

The (a) reading of the sentence in (8) means that Ivan is looking for a particular secretary, namely Elena. On the (b) reading, the indefinite receives a nonspecific interpretation, which means that Ivan is looking for a secretary regardless of who she turns out to be. Neither use of the indefinite in (8) licenses the use of the clitic. Thus, we can conclude that specificity is not a determining factor in the occurrence of the clitic.

We should note that the indefinite in (8) contains an overt determiner *eden* ‘one’. Bare singulars, as in (9) and (10), do not behave any differently with respect to clitic doubling. As (9) and (10) show, a bare singular which is interpreted as an indefinite can never be clitic-doubled.

(9) Ivan (*ja) bara sekretarka.¹⁰

Ivan her look-for secretary

‘Ivan is looking for a secretary (whoever she is).’

(10) Pero (*go) kupi kompjuter.

Pero it(m.sg.) buy computer

‘Pero bought a computer.’

Bare plurals in Macedonian cannot receive generic interpretation, but they can get an existential interpretation. In such cases, they cannot be clitic-doubled, as shown in (11a). Bare plurals cannot be used with a kind reading, as well. A definite must be used instead.

¹⁰ The indefinite in (9) can only receive a nonspecific interpretation.

When a definite is used with a kind reading (Carlson 1977), the clitic is obligatory, as shown in (11b).

(11)a. Ana (*gi) kupi jabolka.

Ana them bought apples

‘Ana bought apples.’

b. Ana *(gi/ja) saka knigite/knigata.

Ana them/it(f.sg.) like books-the/book-the

‘Ana likes books.’

To sum up the discussion so far, I have outlined the distribution of clitics with direct objects in Macedonian and in doing so, I have eliminated two potential factors that can be thought of as the driving factor behind the doubling constructions. First, we saw that the referentiality of the DP does not seem to play a role since quantifiers can (and in some cases must) be clitic-doubled in Macedonian. Second, we saw that the specificity of the doubled DP cannot be a determining factor. This stems from the fact that a) definite DPs are obligatorily clitic-doubled regardless of their interpretation (referential/specific vs. attributive/nonspecific), and b) specific indefinites cannot be clitic-doubled. Next, I show that the topichood of the DP in Macedonian is also not a driving factor in the doubling of the direct objects.

Kallulli (1999, 2008) makes a strong case that clitic doubling in Albanian and Greek correlates to topichood in that only nonfocused DPs can be clitic-doubled. That is to say, she argues that the clitic doubling of direct objects in these two languages is motivated by

the need to check the [-focus] feature of the DP in question. Given the fact that Albanian and Greek, like Macedonian, belong to the group of Balkan languages, it would be interesting to see if the same conditions holds for the doubling constructions in Macedonian. Below I show that this is not the case.¹¹

One of the environments that Kallulli (1999:31) identifies as incompatible with doubling of direct object DPs in Albanian and Greek is the out-of-the-blue sentences. In Macedonian, however, a doubling of the direct object in such cases is possible. The example in (12) illustrates this.¹²

(12) A: What happened here?

B: Ivan *(go) skrši stakloto.
 Ivan it(n.sg.) broke glass-the
 ‘[Ivan broke the glass]_F.’

The answer to the question in (12) provides new information about the subject, Ivan. If we take the focus of the sentence to be the most informative part of the sentence, following Kallulli (1999:25)¹³, the Macedonian example in (12) shows that the definite DP that is being used in such cases can, indeed must, be clitic-doubled.¹⁴

¹¹ Kallulli (1999) leaves open the question whether or not the same factor, i.e. topichood, plays a role in doubling constructions cross-linguistically. My goal here is not to provide an in-depth comparative study of Macedonian with respect to these (or any other) languages, but to investigate what role, if any, focus/topic play in clitic doubling in Macedonian. The motivation here is to understand the conditions under which clitic doubling takes place in Macedonian.

¹² Examples (12)-(15) are fashioned after Kallulli (1999:18, 31-32).

¹³ I will return to the question of focus in section 2.4. For the moment, it is sufficient for our purposes to treat focus as the most informative part of a sentence.

¹⁴ The Albanian and Greek examples in Kallulli (1999) also contain a definite object DP, but this seems to have no bearing on the co-occurrence of the clitic pronoun, which in this case is not allowed; see Kallulli (1999:31).

Along the same lines, it is perfectly possible in Macedonian to utter (13), as a response to a question like *Who did you see?*. As such, the direct object (in this case, *Petar*) is focused. (13) is also acceptable as an out-of-the-blue sentence. Albanian and Greek, on the other hand, do not allow this use (Kallulli 1999:18).

(13) *(Go) vidov Petar.

him saw Petar

‘I saw Petar.’

In Macedonian, the doubling of the direct object is also possible in contexts where the whole VP is focused, i.e. when the DP is part of the focused domain. This is shown in (14), where the focus domain extends to the whole VP.

(14) Ana ne *(ja) ispeče tortata, tuku *(gi) izede slivite.

Ana not it(f.sg.) baked cake-the but them ate plums-the

‘Ana didn’t [bake the cake]_F; she [ate the plums]_F.’

Finally, in constructions where the focus domain is delineated by the use of focus particles, such as *even* and *only*, Macedonian direct object DPs can be clitic-doubled. This is shown in (15).

- (15) Premierot go poseti duri i Ohrid (ne samo Skopje).
 prime minister-the it(m.sg.) visited even and Ohrid not only Skopje
 ‘The Prime Minister visited even [Ohrid]_F (not only Skopje).’

Based on the fact that the doubling of direct object DPs in environments that require focus interpretation (12)-(15) is possible in Macedonian, we can conclude that the clitic in Macedonian is not a licenser of topichood. Thus, we are left with the question of what exactly is the determining factor that motivates the doubling of direct object DPs. In the next section, I present my proposal and analysis of clitic doubling in Macedonian.

2.1.1 Formal Representation of Clitic Doubling

My main claim is that the doubling constructions in Macedonian are configurations that are sensitive to the semantics of the DPs.¹⁵ This section makes that relation explicit.

In order to understand what motivates the appearance of the clitic with direct objects, I take as a starting point their distribution with quantifiers. As we saw in the previous section, only strong quantifiers can be clitic-doubled in Macedonian. I take the strength/weakness of the quantifiers (defined as in Barwise & Cooper (1981); see also section 2.2) to be encoded as a feature specification. Thus, we can assume that strong quantifiers are characterized by a [+strong] feature. This, in turn, means that, structurally, the clitic doubling configuration arises as a result of a feature checking operation. In other words, I propose that the clitic is a licenser of a [+strong] DP, whereby the licensing is carried out through a spec-head relation (Chomsky 1995). My proposal that

¹⁵ In chapter 3, I show that this particular feature of the clitic doubling constructions sets them apart from Clitic Left Dislocation and Topicalization constructions, which I argue are driven by pragmatics.

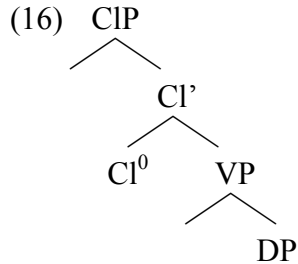
clitic doubling is sensitive to the [+strong] feature on DPs is, in some sense, a continuation of Comorovski's (1996) proposal that clitic doubling marks DPs that carry an existential presupposition.¹⁶

I assume with Sportiche (1998), Anagnostopoulou (1999, 2006) and others, that the doubled DP in clitic doubling constructions is the argument of the verb, and that the clitic is base-generated in its surface position. Following Sportiche (1998), I take clitics to be heads of their own functional projections, located in the IP domain.¹⁷ The clitic licenses a particular property or feature in the DP it doubles, which in this case is [+strong].¹⁸ (In this I depart from Sportiche who takes the view that clitics license specificity in the doubled DP.) The licensing of the feature must be carried out in an appropriate agreement relationship, which is achieved in a spec-head configuration. This means that the doubled DP must move (by LF) to the specifier position of the clitic phrase. The structure of the clitic phrase (and its relative location within the clause) is given in (16).

¹⁶ In this respect, the class of DPs that participate in clitic doubling constructions in Macedonian to a certain extent parallels the class of DPs that Comorovski identifies for Romanian. Comorovski (1996:73) lists the following set of DPs as characterized by clitic doubling: proper names, personal pronouns, [+animate] definites, specific indefinites, partitive NPs, and D-linked wh-phrases. As we have seen from examples like (3), clitic doubling of definites in Macedonian is not restricted to [+animate] DPs. As I show in chapter 3, the doubling of specific indefinites is also subject to further restrictions in Macedonian (i.e. the specific indefinite must be left-dislocated). I discuss the co-occurrence of clitics and partitives in Macedonian in chapter 3.

¹⁷ Sportiche (1998) distinguishes two types of clitics. The first type includes accusative clitics (e.g. French *en/le*). They are similar to [+wh] complementizers and [+neg] heads in that they license an operator-like feature in the DP they double. For Sportiche, the relevant feature in this case is specificity. The clitics in this case are heads of their own maximal projections (which Sportiche represents as Voices). The second type includes dative clitics (e.g. French *lui*). Sportiche treats these as agreement markers and analyzes them as heads of AgrPs.

¹⁸ See also Kallulli (1999, 2008), Alexiadou & Anagnostopoulou (1997), Anagnostopoulou (1999) who follow Sportiche in analyzing the clitic as the head of CIP. Kallulli argues that the clitic in Greek and Albanian licenses a [-focus] feature because focused elements cannot be clitic-doubled in these two languages. Alexiadou & Anagnostopoulou (1997) and Anagnostopoulou (1999) maintain that the clitic is as an agreement marker, following Suñer (1988).



The licensing of the feature by the clitic, i.e. the movement of the DP to the specifier position of the clitic, in Sportiche's terms, is regulated by the Clitic Criterion in (17).

(17) *Clitic Criterion* (Sportiche 1998:267)

- a. A clitic must be in a spec-head relationship with a [+F] XP at LF.
- b. A [+F] XP must be in a spec-head relationship with a clitic at LF.

What this means is that if a clitic is related to a particular DP, then the DP itself has to move to the specifier position of the CIP, in order to satisfy the Clitic Criterion.

The Clitic Criterion is part of the more encompassing Generalized Licensing Criterion, stated in (18).

(18) *Generalized Licensing Criterion* (Sportiche 1998:295)

- a. A [+F] head must be in a spec-head relationship with a [+F] XP at LF.
- b. A [+F] XP must be in a spec-head relationship with a [+F] head at LF.

Sportiche's proposal results in the following set of parameters:

(19) *Clitic Construction Parameters* (Sportiche 1998:268)

- a. Movement of the XP* to XP^ occurs overtly or covertly.¹⁹
- b. Head (H) is overt or covert.
- c. XP* is overt or covert.

The parameters enable Sportiche to account for the following cases of clitic-doubled constructions:²⁰

- (20) a. Covert movement of an overt XP*, combined with an overt H, gives rise to clitic doubling (as in Spanish, Romanian, etc.).
- b. Overt movement of an overt XP*, combined with a covert H, gives rise to scrambling (as in Dutch).
- c. Overt movement of an overt XP*, combined with an overt H, gives rise to Clitic Left Dislocation constructions (as in Italian).

Implementing Sportiche's proposal to Macedonian will mean the following. In all the cases of clitic doubling outlined in section 2.1, the direct object DP starts off in the argument position of the verb and then moves to the SpecCIP.²¹ From what we have seen

¹⁹ In Sportiche's terms, XP* is the doubled phrase, XP^ is the specifier position of the clitic phrase, and H is the clitic head.

²⁰ All clitic constructions, for Sportiche (1998:268) involve movement: the XP* must move to the specifier position of the clitic which licenses its feature.

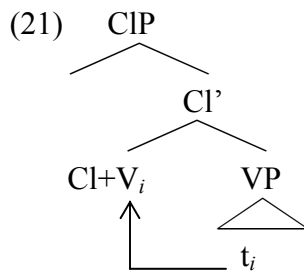
²¹ For completeness, the analysis of sentences like (i), which contain a clitic but no overt XP, in the framework developed here would be as follows:

- (i) Go vidov.
him saw
'I saw him.'

Following Sportiche (1998), I assume that the argument in (i) is a silent XP, which in this case is interpreted as a pronoun, i.e. *pro*. The XP moves to the spec position of the CIP, where it is identified by

so far, the movement of the doubled DP in Macedonian in such cases is covert. The movement of the DP into the specifier position of the CIP creates the necessary agreement relationship through which the clitic can license the [+strong] feature in the DP. In cases where the DP is not marked with a [+strong] feature, the feature-checking between the clitic and the DP cannot be established. Based on the principle of economy of representation (following Rizzi 1997, and others), I assume that the CIP will be projected only when needed. That means that the CIP will only be present in the structure when the object DP is [+strong].

In cases where the CIP is projected, I assume that it occupies a position above VP (following Sportiche 1998, among others). In order to account for the fact that clitics in Macedonian are proclitics and are phonologically dependent on the verb, I follow Rudin (1997) in that the verb raises to Cl and right-adjoins to it, forming a complex verb.



As mentioned earlier, the analysis of clitic doubling here explicitly assumes that the doubled DPs are arguments of the verb. Below, I briefly outline some of the arguments in

the coindexed clitic. *Pro* and the trace in argument position form a chain. The derivation of (i) is given in (ii):

(ii) $[[_{CIP} pro_i \text{ go}_i \text{ } [_{VP} \text{ vidov } t_i]]$

support of this view. More specifically, the following evidence shows that the direct object DP is not right-dislocated when it occurs in the direct object position.^{22, 23}

First, as is the case with dislocation in general, the dislocated element is always separated from the rest of the clause by a sharp intonation break. This holds for right-dislocation as well (Anagnostopoulou (2006) and references cited therein). However, in examples like (22), there is no sharp intonation break between the verb and the DP.²⁴ The ungrammaticality of (22a-b) relates to the particular intonation with which these sentences are pronounced, indicated here as a comma intonation. As we can see, the presence vs. absence of a clitic does not alter the acceptability of the examples.

(22) a. *Dino ja pročita, knjigata/sekoja kniga.

Dino it(f.sg.) read book-the/every book

‘Dino read it, the book/every book.’ *lit. translation*

b. *Dino pročita, dve knigi/mnogu knigi.

Dino read two books/many books

‘Dino read it, the book/every book.’ *lit. translation*

²² The question of whether the doubled DP is an argument or an adjunct is an ongoing issue in the literature on clitic doubling. For an overview see Anagnostopoulou (2006). Here, I do not present evidence to show that the DPs in clitic doubling are not adjuncts, as is the case for Mohawk, for example (see Baker 1996). For evidence against an adjunct analysis of the clitic-doubled DPs in Bulgarian, see Rudin (1997); see also Franks and King (2000).

²³ Having said this, arguing that clitic doubling is not an instance of right dislocation does not entail that the XP is the argument of the verb (rather than an adjunct to VP; see Philippaki-Warbuton et al. (2002)). The latter option is still available; the clitic in that case becomes the real argument. For analysis of this sort, see Philippaki-Warbuton et al. (2002), Sánchez (2002), and references cited in Anagnostopoulou (2006). For arguments against an adjunct analysis of clitic doubling in Bulgarian, see Rudin (1997).

²⁴ In addition to its intonational features, Anagnostopoulou (2006) gives the following arguments why clitic doubling should not be analyzed as right dislocation: i) Right dislocation is not subject to Kayne’s Generalization but clitic doubling is (however, that this is really the case is somewhat debatable; there have been arguments that clitic doubling is not, in fact, subject to Kayne’s Generalization; see Anagnostopoulou (2006) for Greek, among others.); ii) Some languages have right dislocation, but not clitic doubling (e.g. French).

Second, the direct object cannot be separated from the verb by an intervening PP or adverbial, as shown in (23b) and (24b), respectively.

(23)a. Ja staviv knjigata na masata.
 it(f.sg.) put book-the on table-the
 ‘I put the book on the table.’

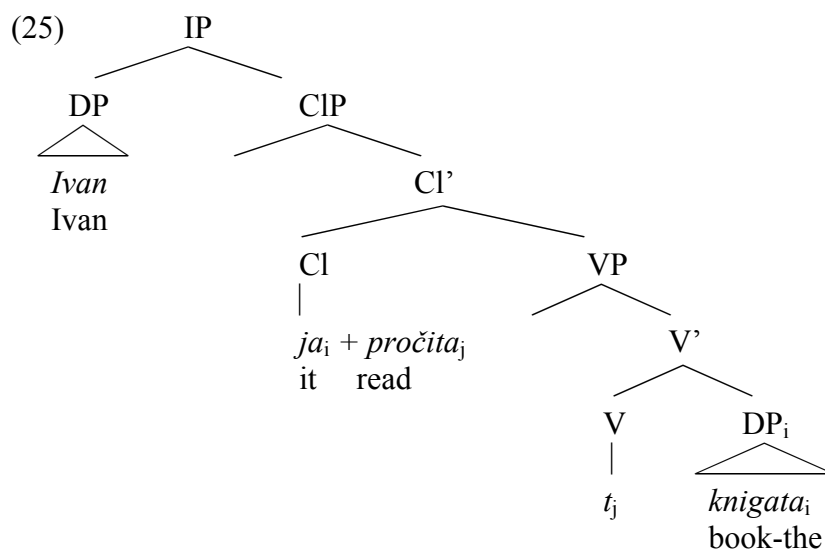
b. *Ja staviv na masata knjigata.
 it(f.sg.) put on table-the book-the

(24)a. Go znaeme Nikola mnogu ubavo.
 him know Nikola very well
 ‘We know Nikola very well.’

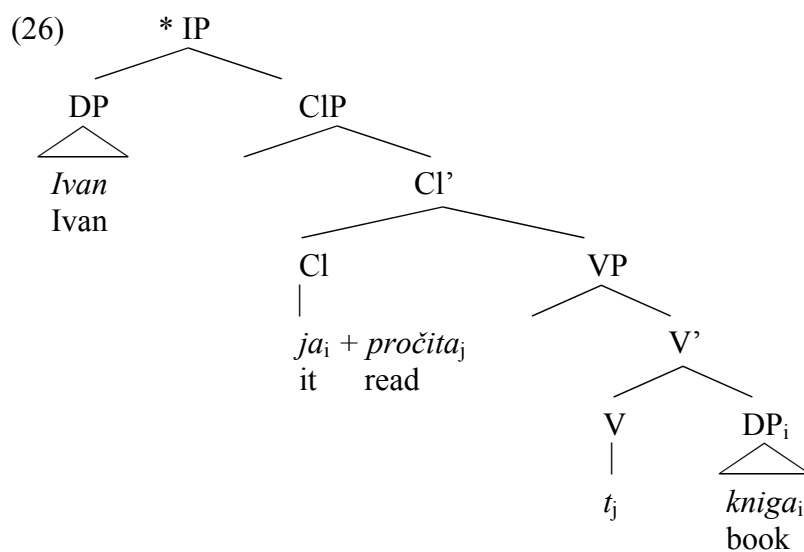
b. *Go znaeme mnogu ubavo Nikola.
 him know very well Nikola

I conclude this section with a brief summary of how the proposal made above accounts for the possible patterns of doubling constructions we find in Macedonian.

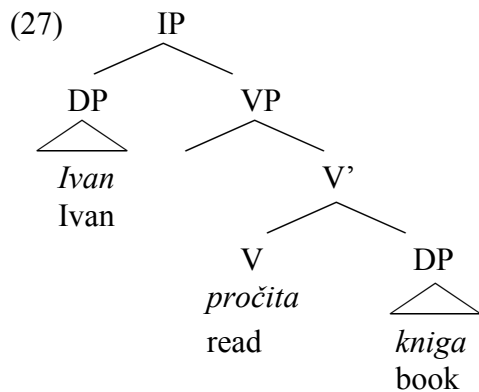
I claimed that the clitics in Macedonian license a particular feature in the DP they double. The doubling constructions arise in those cases where the direct object is [+strong], in which case the clitic is required to license this feature. An illustration of the derivation of a clitic-doubled construction is given in (25).



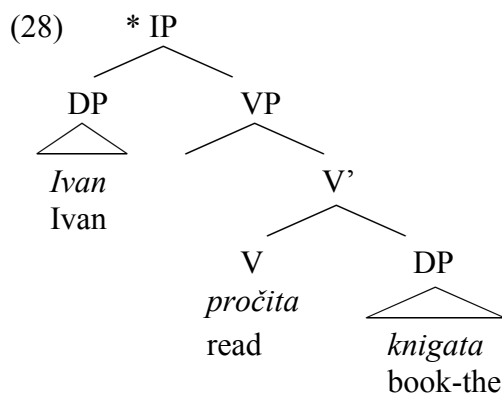
A weak DP in object position cannot license the clitic because the conditions for licensing the [+strong] feature do not obtain, i.e. the DP cannot enter into a spec-head relation with the clitic. This is what makes (26) ungrammatical.



A weak DP does not trigger the presence of a clitic as there are no features that need to be checked. In keeping with economy requirements, the CIP in this case is not projected, as shown in (27).



A construction with a [+strong] DP in object position without a clitic is ungrammatical because in this case the relevant feature of the DP cannot be checked. Hence, the derivation cannot converge, as illustrated in (28).



The analysis of clitic doubling outlined in this section crucially incorporates the idea that the derivation of such constructions is driven by the presence vs. absence of a [+strong]

feature. So far, I have maintained that the [+strong] feature is a representation of the semantic feature strong/weak in DPs. In the next section, I discuss the classification of the DPs in these terms and show how they fit into the analysis of clitic doubling I have outlined here.

2.2 Strong and Weak DPs in Macedonian

The first characterization of DPs in terms of the weak vs. strong distinction can be found in Milsark (1977). Milsark classified English determiners as weak vs. strong based on their ability or inability to occur in existential *there*-sentences, as in (29a-b). (From Milsark 1977:4)

- (29) a. There is a wolf at the door.
 b. *There is the wolf at the door.

Milsark noted that only weak determiners are allowed in existential *there*-sentences, as illustrated by (30a-b).

- (30) a. There are some/several/many/few/five wolves at the door.
 b. *There is/are every/most/all wolves at the door.

According to Milsark's diagnostic, English determiners separate into two classes: i) strong: the definites, universals, *most*, *both*, *(n)either*, and ii) weak: *a*, *sm*, the numerals, *many*, *few*, *several*, and the non-universal null determiner with plural and mass nouns.

There have been numerous attempts to formalize the strong/weak distinction among quantifiers: Barwise & Cooper (1981), Keenan (1987), Zucchi (1995), Keenan (2003), to name just a few. (See also Abbott 2004, Partee 1989, 2004, and references cited therein.) For the purposes of our analysis, below I offer a characterization of the strong vs. weak distinction as defined in Barwise & Cooper (1981).

2.2.1 Barwise & Cooper (1981)

Barwise & Cooper adopt Milsark's (1977) terminology of "strong" and "weak" quantifiers and set out to provide a formal characterization of the two types within the theory of generalized quantifiers where determiners express a relation between two sets: a set denoted by the common noun phrase (CNP) with which the determiner combines and a set denoted by the verb phrase. Barwise & Cooper's definition of strong/weak determiners is given in (31).

(31) DEFINITION. A determiner D is *positive strong* (or *negative strong*, resp.) if for every model $M = \langle E, || \rangle$ and every $A \subseteq E$, if the quantifier $||D||(A)$ is defined then $A \in ||D||(A)$. (Or $A \notin ||D||(A)$, resp.). If D is not (positive or negative) strong then D is *weak*.

Barwise & Cooper provide the test sentence in (32) which can be used to determine whether a determiner is to be classified as positive strong, negative strong, or weak.

(32) D N is a N / are Ns

Following the definition in (31), the definite article would come out as positive strong because of the presuppositional requirement that its first argument be a set that contains exactly one member. Thus, in a model in which the presupposition condition is satisfied, a sentence like *The unicorn is a unicorn* would be trivially true. In a model in which the presupposition requirement of the definite article is not satisfied, the sentence would be undefined. The characterization of the definite article as strong explains the exclusion of definites from *there*-sentences as in (29b), thus drawing a parallel between them and other strong determiners as in (30b). Proper names and pronouns also come out strong in Barwise & Cooper's model, since they, too, presuppose existence.

Following the definition in (31), English determiners are classified as follows:

- (34) a. positive strong: *all, every, most, the one, the two, both*
- b. negative strong: *neither*
- c. weak: *a, some, the numerals, many, few, a few, no*

Semantic definitions in set-theoretic terms for some of the determiners in (34a-c) are given in (35).²⁶

- (35) a. $\| \text{every} \| (A) (B) = 1$ iff $A \subseteq B$
- $\| \text{most} \| (A) (B) = 1$ iff $|A \cap B| > |A - B|$
- $\| \text{the two} \| (A) (B) = 1$ iff $|A|=2$ & $A \subseteq B$ where $|A|=2$
- $\| \text{both} \| (A) (B) = 1$ iff $|A|=2$ & $A \subseteq B$

²⁶ The definitions for *many* and *few* are from Partee (1989) and they reflect their ambiguity as cardinal vs. proportional. For arguments that *many* and *few* are ambiguous between the two readings, see Partee (1989).

b. $\| \text{neither} \| (A) (B) = 1$ iff $|A|=2$ & $A \cap B = \emptyset$

c. $\| \text{some} \| (A) (B) = 1$ iff $|A \cap B| \neq \emptyset$

$\| \text{no} \| (A) (B) = 1$ iff $|A \cap B| = \emptyset$

$\| \text{two} \| (A) (B) = 1$ iff $|A \cap B| = 2$ ‘exactly two’

$\| \text{two} \| (A) (B) = 1$ iff $|A \cap B| \geq 2$ ‘at least two’

$\| \text{many} \| (A) (B) = 1$ iff $|A \cap B| > n$ where n is contextually determined

$\| \text{many} \| (A) (B) = 1$ iff $\frac{|A \cap B|}{|A|} \geq k$ where k is a fraction or percentage

$\| \text{few} \| (A) (B) = 1$ iff $|A \cap B| \leq n$ where n is contextually determined

$\| \text{few} \| (A) (B) = 1$ iff $\frac{|A \cap B|}{|A|} \leq k$ where k is a fraction or percentage

Applying the definition in (31) and the test sentence in (32), the Macedonian determiners would be classified as in (36). I assume that the semantic definitions given in (35) apply to the determiners in (36).

(36)a. positive strong: *sekoj* ‘every’, *sie* ‘all’, *povekjeto* ‘most’, *obata* ‘both’,

edniot/dvata ‘the one/the two’ etc.

b. negative strong: *nitu eden od* Ns ‘neither’

c. weak: *nekoj* ‘some’ *eden/dva* ‘one/two’ etc., *mnogu* ‘many’, *nekolku*

‘several/a few’, *malku/nekolku* ‘few’, *nieden* ‘no’

In the previous section I argued that the strong vs. weak distinction is directly encoded in the grammar of Macedonian through the doubling of the direct objects. If we look at the

distribution of DPs and clitics in section 2.1, we can see that the presence vs. absence of clitic pronouns indeed correlates with the semantic properties of the doubled DPs that we have just identified. Thus, the semantic classification of the DPs in terms of strong vs. weak gives rise to a two-way distinction between the DPs in terms of their syntactic feature specification: those that are marked with a [+strong] feature (strong DPs) and those that lack such a feature altogether (weak DPs). The licensing of this feature underwrites the doubling constructions. That is to say, the clitic, being the licensor, must necessarily be present whenever a [+strong] DP enters the derivation.

The discussion so far has centered on direct object DPs that occur in their default, postverbal position. However, direct object DPs in Macedonian can also occur in other positions in the clause. In the next section, I look at the behavior of direct object DPs in preverbal position and their co-occurrence with clitics.

2.3 Preverbal Direct Objects

When we consider the clitic co-occurrence restrictions with preverbal direct objects, we notice that they are exactly identical to those of the postverbal ones. That is to say, clitics are obligatory with strong DPs and ruled out with weak DPs. Illustrative examples are given in (37)-(38), respectively.²⁷

(37) a. Pero vesnikot/ovoj vesnik *(go) pročita.

Pero newspaper-the/this newspaper it(m.sg.) read

‘It was a newspaper that Pero read.’

²⁷ Throughout this dissertation I translate all examples with direct object DPs in preverbal position with a cleft construction. This particular rendering of the Macedonian data into English presupposes their status as contrastively focused elements. The next section lays out arguments in favor of this view.

b. Ana site/povekjeto knigi *(gi) pročita.

Ana all/most books them read

‘It was all/most books that Ana read.’

(38)a. Petar kompjuter/jabolka (*go/gi) kupi.

Petar computer /apples it(m.sg.)/them bought

‘It was a computer/apples that Petar bought.’

b. Elena dve/mnogu/malku/nekolku knigi (*gi) pročita.

Elena two/many/few/several books them read

‘It was two/many/few/several books that Elena read.’

Based on this patterning, it seems plausible to consider the preverbal constructions to be derived from the postverbal constructions. In this section, I will try to elaborate on this relationship and in doing so, also shed light on the relationship between the preverbal position and the clitic phrase.

2.3.1 The Syntactic Relationship between the Clitic Phrase and Preverbal DPs

One of the implications of having the preverbal constructions being derived from the postverbal ones is that the doubling of preverbal DPs would have to rely on the same set of procedures that were invoked in the analysis of clitic-doubled postverbal DPs, as outlined in section 2.1.1. This entails that the presence vs. absence of clitics with preverbal DPs is tied to their featural makeup. A related question to this is that of the placement of the preverbal DPs, and in particular, the clitic-doubled ones.

Recall from our discussion that strong direct objects in postverbal positions move to SpecClP for feature checking purposes. The data in section 2.1 seemed to suggest that the movement of these DPs to SpecClP was covert. In this context, the clitic doubling of preverbal DPs is interesting because it now raises the issue of whether strong DPs can also move to the clitic phrase overtly. The assumption that the preverbal are derived from the postverbal constructions, leaves it open that, when fronted, the DPs would occupy the SpecClP position. Although this seems plausible, there are strong reasons to believe that doubled preverbal DPs in cases like (38a-b) move overtly, but to a position other than the clitic phrase.

Evidence that fronted strong DPs are not in SpecClP position comes from examples like (39a) where an adverb intervenes between the fronted universal and the clitic.

(39) a. Pero site knigi brzo *(gi) pročitā.

Pero all books quickly them read

‘It was all the books that Pero quickly read.’

b. Pero dve knigi brzo (*gi) pročitā.

Pero two books quickly them read

‘It was two books that Pero quickly read.’

If the strong DP in (39a) were in SpecClP position, the adverb would not have been able to intervene between it and the clitic pronoun. Thus, we can conclude that strong DPs, when fronted, move to a position different from that of SpecClP. Weak DPs, as in (38a-b) and (39b), presumably follow the same pattern. The association with the clitic between

the two types of constituents, though, remains unchanged: strong DPs require it because they need a licenser for their [+strong] feature; weak DP do not allow it because they lack the relevant features. Thus, the behavior of the preverbal phrases with respect to clitics is consistent to their behavior in postverbal position.

That the presence of the clitic with strong DPs is tied to their feature specification explains their obligatoriness, but it does not explain the DPs' placement in the clause. Examples like (39a) seem to suggest that strong DPs pass through the clitic phrase but do not stop there. Given this, it is important to explain what motivates the DPs' movement in the first place.

One option here is to say that the movement of the DPs is motivated by feature checking. But if this were the case, then the movement of the DP would stop at the point of the clitic phrase; it need not go any further. That is clearly not the case for Macedonian. Another option is to say that the movement is triggered by something other than feature checking. In the next section I show that interpretative differences between the two positions (pre- and postverbal) do exist and that the two positions are not simply variants of each other. Anticipating the discussion there, I claim that the preverbal position is different from the postverbal position in that it licenses contrastive focus.

Pursuing this line of analysis means that strong DPs move to a preverbal position for focus reasons. However, in the course of derivation, they necessarily go through ClP (where feature licensing takes place), but they do not remain there. If strong DP were to stay in SpecClP, then we should be able to get preverbal DPs without a focus interpretation. That is to say, if the movement of the strong DPs were due to feature checking, we should have the same interpretation of preverbal and postverbal DPs. This,

however, is not the case. The preverbal position in Macedonian is necessarily associated with contrastive focus (for arguments, see section 2.3.2). We can therefore conclude that movement of the strong DPs to a preverbal position is not motivated by the same factors that require their movement to the clitic phrase. (The question of whether the movement of weak DPs is motivated by feature checking does not even arise here. Given that they move to a preverbal position where they get focus interpretation (see discussion below), I assume that their movement to such a position is triggered by focus as well.)

I have already noted that the derivation of the preverbal doubling construction necessitates an overt movement of strong DPs to the clitic phrase, albeit on their way to the focus position. Positing such movement is in fact consistent with Sportiche's *Clitic Criterion* and his set of *Clitic Construction Parameters*. The implementation in this particular case would require overt movement of an overt phrase. An example of this type of movement in Sportiche's study is that of Dutch Scrambling. As we have just seen, another instantiation of this particular setting may be the preverbal constructions in Macedonian. We can thus conclude that for preverbal DPs the movement to CIP is presumably overt, though the movement is not triggered by the need to check their [+strong] feature.

The data in (37)-(38) show that both strong and weak DPs can be focused. This is strong evidence that there is an additional preverbal position, SpecFocP, to which both types of DPs can move, regardless of their semantic properties. That the FocP is not sensitive to the strength of the DP is further illustrated by examples such as (40) below.

- (40) a. Pero samo dve knigi (*gi) pročitā.
 Pero only two books them read
 ‘Pero read only two books_[F]’
- b. Eva samo Ljubljana *(ja) poseti.
 Eva only Ljubljana it(f.sg.) visited
 ‘Eva visited only Ljubljana_[F].’

In (40a), the focused element is the weak DP *dve knigi* ‘two books’; there is no clitic. In (40b) the focused element is a proper name; the clitic in this case is obligatory. As the examples show, the distribution of the clitic is independent from the discourse status of the elements themselves, in the sense that both DPs are foci, as evident from the use of the focus sensitive elements like *samo* ‘only’. Examples like these offer additional evidence that there are two preverbal positions, each of which is associated with different properties and as a result is able to accommodate different types of DPs. CIP continues to be associated with the intrinsic semantic properties of the DP, while FocP can accommodate DPs regardless of their strong/weak characterization.

The analysis of the preverbal doubling constructions put forth here relied on the idea that the preverbal and postverbal positions differ in their interpretation. As the translations of the examples in (38a-b) indicate, the interpretation of the preverbal objects is somewhat different to that of the postverbal ones. In the next section, I make this difference explicit.

2.3.2 The Relationship between Pre- and Postverbal Positions

The basic idea that I pursue here is that the difference in the interpretation of the pre- and postverbal DPs is a result of their occurring in positions that associate with different types of focus, the latter being informally defined as “the most informative part of the sentence” (Kallulli 1999:25). To bring out the difference between the pre- and postverbal positions, however, we need to look more closely into the type of focus they associate with. I take the main difference in this case to be that between information focus and contrastive focus. The definition (and, to some extent, terminology) comes from É. Kiss (1998).

É. Kiss (1998) argues for the existence of two types of foci: information and identificational. She takes information focus to be new, nonpresupposed information; identificational focus, on the other hand, to express exhaustive identification. The latter is defined as follows:

(41) The function of identificational focus: An identificational focus represents a subset of the set of contextually or situationally given elements for which the predicate phrase can potentially hold; it is identified as the exhaustive subset of this set for which the predicate phrase actually holds. (É. Kiss 1998:245)

Identificational focus, É. Kiss maintains, does not have to be present in every sentence, but when it is, it triggers movement. Information focus, on the other hand, is present in every sentence but does not trigger any syntactic reordering. É. Kiss points out that identificational focus is sometimes referred to as contrastive focus and realized as a cleft

construction. For this reason, I would refer to this type of focus as contrastive. Thus, the distinction we aim to capture between the pre- and postverbal position will be in terms of contrastive vs. information focus, respectively. In what follows, I will not go into the particulars of É. Kiss' analysis of the two types of foci. For our purposes it would be sufficient to remember that the main feature of the contrastive focus is its exhaustive identification "on a set of entities given in the context or situation" (É. Kiss 1998:248), and that the main feature of the information focus is its conveying new information, i.e. marking "the nonpresupposed nature of the information it carries" (É. Kiss 1998:248). Taking these features as a starting point, let us see how the two verbal positions fare in this respect.

Imagine a simple scenario in which Ana went shopping yesterday and then consider (42a-b) below.

- (42) a. Ana si kupi kapa.
 Ana refl. bought hat.
 'Ana bought a hat for herself.'
- b. Ana kapa si kupi.
 Ana hat refl. bought
 'It was a hat that Ana bought.'

By uttering (42a) we convey some new information about Ana, namely that she bought something. By uttering (42b), on the other hand, we claim that out of all the various pieces of clothing available in the given context, Ana chose a hat only, and she did not

choose anything else. Thus, the postverbal object in (42a) is information focus; the preverbal object in (42b) is a contrastive focus.²⁸

That new, nonpresupposed information is expressed by placing the object in postverbal position is further supported by the fact that (42a), but not (42b), can be uttered as an out-of-the-blue sentence. So, if we want to report on the fact the Ana bought something for her when she went shopping, we can say (42a), but not (42b).

That objects in postverbal position get nonexhaustive interpretation is illustrated in examples like (43) below.

(43) A: Where did you go on a holiday this summer?

B: a. Otidov vo Italija.

went in Italy

‘I went to Italy.’

b. Vo Italija otidov.

in Italy went

‘It was to Italy that I went.’

The answer in (43a) is understood to mean that the speaker went to Italy, among other places. The answer in (43b), on the other hand, is exhaustive, in that it means that the only place the speaker went to was Italy.

Szabolcsi (1981) provides the following test to bring out the difference in the interpretation between the two foci. The test involves a pair of sentences, the first of

²⁸ We can also interpret (42b) to mean that Ana chose a hat as opposed to something else, for example, a coat.

which contains two conjoined DPs. The second sentence contains only one of the two DPs. The focus expresses exhaustivity if the first sentence does not entail the second sentence.

(44) a. Ana si kupi kapa i šal.
 Ana refl. bought hat and scarf
 ‘Ana bought a hat and a scarf for herself.’

b. Ana si kupi kapa.
 Ana refl. bought hat
 ‘Ana bought a hat for herself.’

(45) a. Ana kapa i šal si kupi.
 Ana hat and scarf refl. bought
 ‘It was a hat and a scarf that Ana bought for herself.’

b. Ana kapa si kupi.
 Ana hat refl. bought
 ‘It was a hat that Ana bought for herself.’

The sentence in (44b) is entailed by (44a), hence the focused constituent in postverbal position does not express exhaustive information. (45b), on the other hand, is not entailed by (46a); the focus thus expresses exhaustive information.

It should become clear by now that the pre- and postverbal positions are not just interpretative variants, but rather do fulfill different functions. As a result of this, the DPs

that occupy these two positions are interpreted differently. The identical distribution of the clitics with post- and preverbal DPs we saw earlier suggests that the two constructions can easily be linked derivationally, in that the preverbal DPs get to their position by movement from their postverbal position. The important thing to remember, though, is that the movement of the DP from post- to preverbal position is not triggered by the need to check their strong features. Rather, the movement in such cases is triggered by contrastive focus. This analysis, of course, does not automatically preclude the possibility that the strong DPs stop in SpecCIP on the way to their focus position. Presumably they would be required to do so in order to check their strong features. The difference between this movement and that of the postverbal DPs is that the feature checking of the preverbal strong DPs can be overt if the DP is forced to move by focus.

To summarize, I have argued that the preverbal position in Macedonian is used as a contrastive focus position. I have not, however, probed into the question of whether or not the postverbal position can also be used for this purpose. Given the appropriate intonation (i.e. focal stress), I believe the postverbal position can also be used to express contrastive focus. The main point, however, is that the preverbal position can never be used to convey simple information focus, the way the postverbal position can. Even this, I think, gives us sufficient evidence to treat the two positions as different from one another.

2.4 Conclusion

In this chapter, I have shown that the clitic doubling of direct objects in Macedonian is regulated in a systematic way in that it licenses strong DPs. I have proposed that the clitic

is the formal licenser of the [+strong] feature of the object DPs generated within the VP. I have argued that direct object DPs split into two classes, strong vs. weak, and it is these properties that govern their interaction with the clitic pronouns.

The chapter also outlined the distribution of preverbal DPs and their co-occurrence restrictions with clitics. It was shown that their behavior follows from the analysis of clitic doubling of postverbal DPs. I have also proposed that the two positions accommodate different types of foci. Finally, I showed that there is a separate preverbal focus position, which can accommodate both strong and weak DPs. It was shown that this structural position is separate from that of the clitic phrase, which is only present when strong DPs are generated within the VP.

In the next chapter, I focus on the distribution of direct objects in the left periphery and their co-occurrence with clitics when the former occur in clause-initial position in Macedonian.

CHAPTER 3

Clitic Co-Occurrence Restrictions on Direct Objects in the Left

Periphery

3.0 Introduction

In this chapter, we turn our attention to direct objects in the left periphery, i.e. to direct objects that occur in clause-initial position in Macedonian. The chapter focuses on the distribution of clitic pronouns in the context of left dislocation.

The chapter aims to explain the conditions under which clitic doubling of dislocated constituents takes place. In this respect, it highlights a somewhat unexpected behavior in the patterning of weak DPs with regards to their possibility or impossibility to co-occur with clitic pronouns. We saw in chapter 2 that weak DPs in pre- and postverbal position uniformly disallow the presence of a clitic. Once dislocated, though, their behavior seems to change. The goal of this chapter is to explain the motivation behind these differences. The investigation will lead us to revise our original proposal with regards to their feature specification as put forth in chapter 2. In particular, I will argue that the weak DPs in Macedonian are not homogenous but that they split into two groups, each of which is characterized by a different set of features. The chapter thus is an exploration into the properties of weak DPs and the implications of it for the analysis of clitic doubling.

The chapter is organized as follows. Section 1 outlines the distribution of left-dislocated DPs, in particular weak ones, and their co-occurrence with clitic pronouns. Section 2 offers a novel characterization of the weak DPs which is then used to explain their different patterning with respect to clitics in the context of left dislocation. Section 3

outlines the semantic basis for the posited distinction within the weak DPs. Section 4 returns to the issue of the partitives and discusses how they fit in the overall analysis. Section 5 contains the conclusion.

3.1 An Unexpected Division among Weak DPs

We saw in chapter 2 a clear split between strong and weak DPs with respect to co-occurrence with clitics; the former obligatorily require them, while the latter uniformly disallow them. The distribution of the clitic was explained through the Clitic Criterion where the presence of the clitic was related to the presence of a particular feature on the doubled DP. We also saw that the pattern of clitic co-occurrence is preserved under focus movement. Thus, we were able to explain the parallel behavior of preverbal DPs to those of postverbal ones. Given the patterning of direct objects in these two contexts, we would expect a similar behavior of the DPs when they occur clause-initially, i.e. at the left periphery. Interestingly, though, this expectation is only partially met. While it is true that strong DPs occur with clitics obligatorily, as expected, the behavior of weak DPs becomes more nuanced. Weak DPs can occur without clitics when dislocated, but a subset of them can also appear with clitics.

For illustration, the basic patterns of left-dislocated strong and weak DPs is outlined in (1)-(4) below. The clause-initial position in Macedonian is characterized by a marked pause after the object, indicated here as a comma intonation.

In (1), we see that the presence of a clitic remains obligatory with strong DPs, such as definites and demonstratives (see (1a)), as well as strong quantifiers (see (1b)).

(1) a. Knigata/ovaakniga, Pero *(ja) pročita.

book-the/this book Pero it(f.sg.) read

‘The book/this book, Pero read it.’

b. Site/povekjeto knigi, Ana *(gi) pročita.¹

all/most books Ana them read

‘All/most books, Ana read them.’

On the other hand, weak DPs, such as *many/few* in (2), as well as bare plurals in (3a-b), disallow clitic pronouns in the IP.²

(2) Mnogu/malku knigi, Ana (*gi) pročita.

many/few books Ana them read

‘Many/few books, Elena read.’

(3) a. Slivi, Petar (*gi) kupi.

plums Petar them bought

‘Plums, Petar bought.’

¹ Restrictions on the co-occurrence of quantifiers (in particular weak ones) and clitic pronouns hold in other languages as well; see Iatridou (1995) and Alexiadou (2006) for examples from Greek. Iatridou shows that numerals cannot co-occur with clitics when in IP. Alexiadou shows that numerals can co-occur with clitics when CLLDed but not when within IP.

² Bare singulars in Macedonian cannot be left dislocated at all, as shown by the ungrammaticality of (i).

(i) *Kukja, Petar kupi.
house Petar bought
‘A house, Petar bought.’

In such cases, the presence of a clitic does not improve the grammaticality of the sentence, as shown in (ii).

(ii) *Kukja, Petar ja kupi.
house Petar it(f.sg.) bought
‘A house, Petar bought.’

I do not have an explanation for the restriction on bare (count) singulars with respect to left-dislocation. As pointed out to me by Liliana Sánchez (p.c.), the same restriction holds for Spanish. The reasons why this should be the case, however, are unclear to me. I leave this question aside for further research.

- b. Knigi, profesori (*gi) pišuvaat, statii, studenti
 books professors them write articles students
 ‘Books, professors write, articles, students do.’

As mentioned above, the generalization does not extend to all weak DPs. Some weak DPs seem to optionally allow clitics, as illustrated by *dve knigi* ‘two books’ and *nekolku knigi* ‘several books’ in (4).

- (4) Dve/nekolku knigi, Ana (gi) pročita.
 two/several books them read
 ‘Two/several books, Ana read (them).’

Given the fact that the dislocated weak DPs in (4) allow for the possibility of doubling, the question that immediately arises is why that should be the case. In other words, the issue here is to determine what causes the divergent behavior between the DPs in (2)-(3a-b) and those in (4). What I would like to claim is that, given their patterning, we are forced to revise our earlier assumptions that weak DPs in Macedonian function uniformly as a single class. In other words, the peculiar behavior of the weak DPs we witness here is indicative of a split within that class. The left-dislocation of such DPs seems to help bring out this difference.

In the next section, I argue that the divergent behavior of the weak DPs is a result of their different feature characterization. That is to say, the difference in the feature specification drives their different patterning with respect to the clitics. The challenge, of

course is to show that the modification we propose here does not alter the analysis of clitic doubling we proposed in chapter 2.

3.2 A Three-Way Distinction in Feature Specification

The analysis that I would like to propose relies on positing a three-way distinction in the feature specification of the DPs in Macedonian.³ Given the data in section 3.1, the dislocation of an element to the left periphery seems to bring out the differences between the various DPs more clearly, thus making their properties, so to speak, more prominent. One way of approaching the issue of their characterization is to look at the derivation of these constructions. The idea here is that the feature specification of the dislocated elements will turn out to play a determining role in the formation of syntactic chains.

To begin with, I assume that, in principle, direct objects can occur at the left periphery in one of two ways, by movement or by base-generation. At the moment, the data in section 3.1 is consistent with either one of these possibilities, so for the sake of argument, I will maintain this assumption throughout the discussion here. In chapter 4, I will look more closely into the arguments for deciding between the two options.⁴

The basic syntactic assumption then is that left-dislocated constructions in Macedonian can either be base-generated or derived by movement. The first option entails that the dislocated element is generated in its surface position and enters into a binding relation with its coindexed clitic. I adopt Cinque's (1990) view that such

³ I am indebted to Veneeta Dayal for the idea of a three-way distinction of the DPs in terms of strength features.

⁴ Deciding between the two options is not a trivial matter. However, maintaining an open view at the moment is legitimate as numerous analysis of left dislocation exist arguing either for movement (Sportiche 1998, for example) or base-generation (Cinque 1990, among others), or both (Aoun & Benmamoun 1998). Aoun & Benmamoun (1998), in particular, argue that clitic left dislocated elements in Lebanese Arabic can either move or be generated in positions at the left periphery. I return to this issue in chapter 4.

constructions are an instantiation of a binding chain and that constructions derived by movement are an instantiation of a government chain, which as Cinque argues, following Rizzi (1990), are defined in relation to the absence vs. presence of wh-movement. The conditions under which chains are formed in Cinque's system are outlined in Baker (1996:112), given here in (5).

(5) *The Chain Condition*

X and Y may constitute a chain only if:

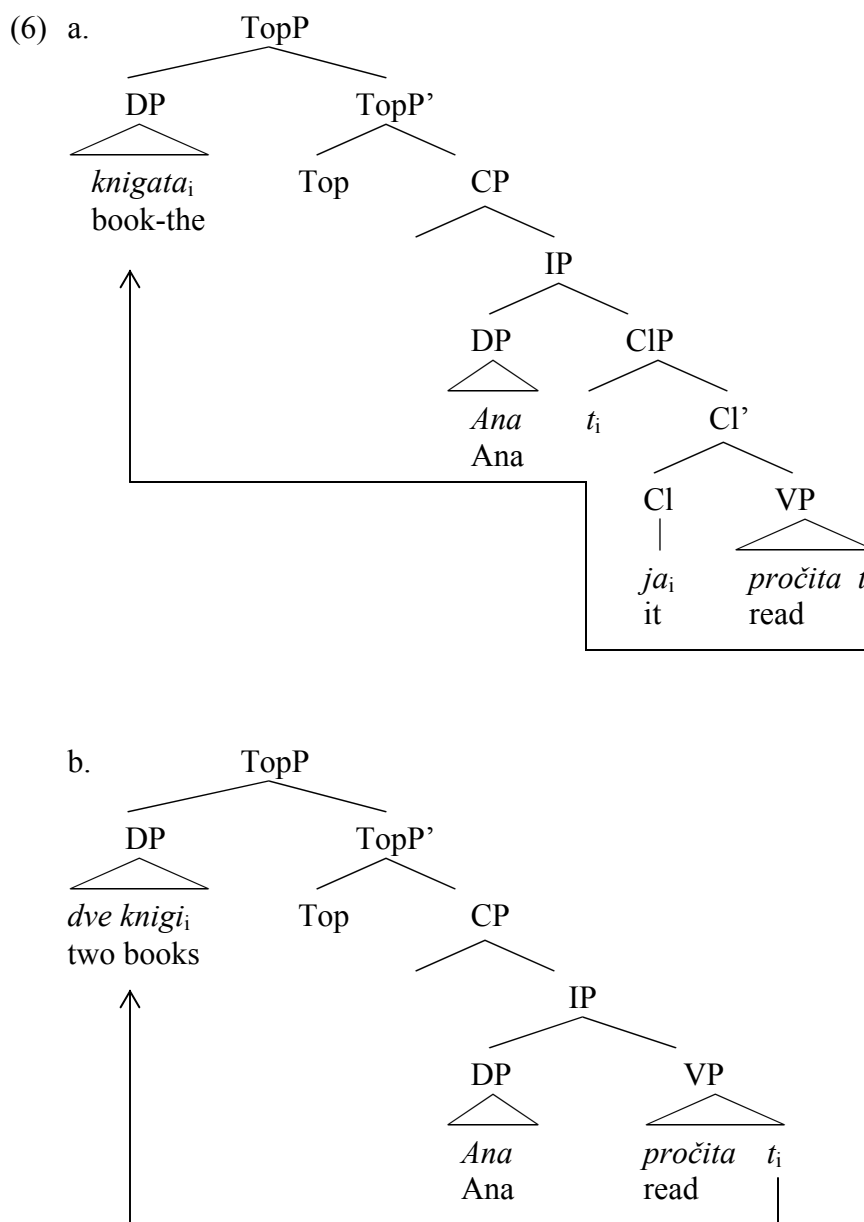
- (i) X c-commands Y.
- (ii) X and Y are coindexed.
- (iii) There is no barrier containing Y but not X.
- (iv) X and Y are nondistinct in morphosyntactic features (i.e. category, person, number gender, Case, etc.)

I also assume following Rizzi (1997) that dislocated elements in Macedonian occupy a TopP position within the CP field.⁵

Given the assumption about the mechanisms used in the derivation of the left-dislocated constructions, the empty category (EC) in argument position would either be a *pro* or a trace. One thing to note here is that the identity of the EC does not change the relation between the DP and the clitic, but it does affect the binding relation between the dislocated DP and the other elements of the chain.

⁵ Kallulli (1999) assumes that Clitic Left Dislocation in Albanian and Greek involves movement of the left-dislocated element to a topic position above CP. In particular, she assumes that CLLD in these two languages arises as movement of the doubled object to SpecCIP and then further up to SpecTopP above CP.

With these assumptions in place, let us consider the movement option first. When direct objects move to clause-initial position, they leave a trace in the argument position within IP. Derivations for strong and weak DPs are given in (6a) and (6b), respectively.



Under a movement analysis, the difference between the constructions in (6a) and (6b) comes out as a result of the Clitic Criterion. In the case of the strong DPs, a CIP is

projected, enabling the DP to check its [+strong] feature in the specifier position, on its way to the left periphery. In the case of the weak DPs, there is no [+strong] feature, so the CIP is absent, due to principles of economy. Thus, by evoking the same principles that explained clitic doubling of pre- and postverbal DPs in chapter 2, we are able to explain the doubling of strong DPs in the left periphery, as well as the absence of a doubling clitic with dislocated weak ones. This approach, however, explains only part of the data. What still remains unexplained is the presence of a clitic with left-dislocated weak DPs such as the numerals and *several*. By treating the weak DPs uniformly, we predict that they should behave the same with respect to the Clitic Criterion.

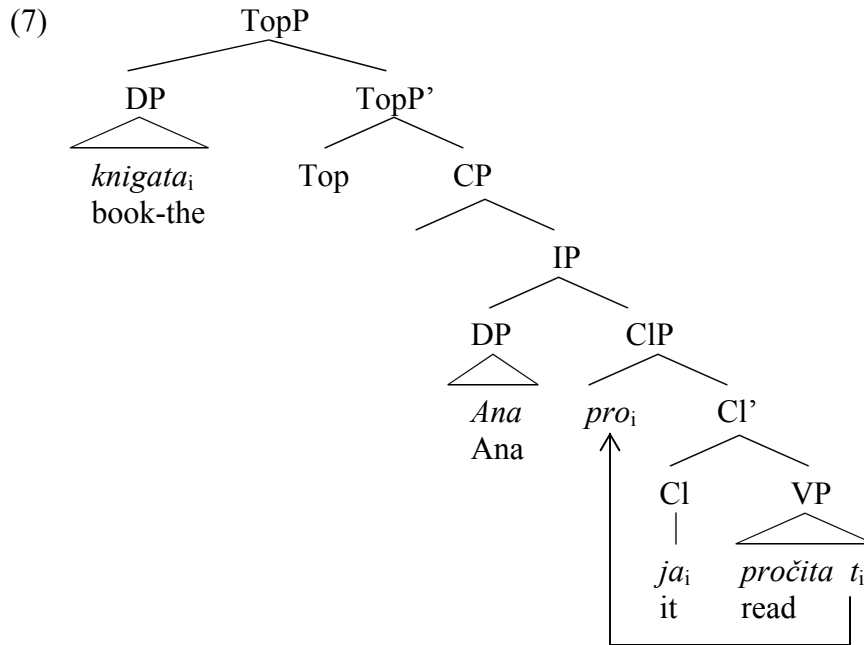
Before we tackle the question of why a clitic is possible with some left-dislocated weak DPs, I would briefly like to note that Sportiche (1998:277) also maintains a movement approach to constructions such as Clitic Left Dislocation (CLLD).⁶ He notes that CLLD may be an instance of overt movement of overt DP with overt clitic, the movement first being to SpecCIP then beyond. Such patterning, we should note, is predicted as part of his Clitic Parameters (see chapter 2). In other words, he notes that CLLD in languages such as Italian or Greek may be analyzed in terms of movement of the doubled object “to or beyond” SpecCIP (Sportiche 1998:297).⁷

I now turn to the second option, that of base-generating the direct objects at the left periphery. As already noted, the DP in the clause-initial position in this case, is connected to a *pro* in the argument position in IP. Given the conditions on chains outlined in (5)

⁶ Clitic Left Dislocation (CLLD) is a term that applies to cases of left-dislocated DPs coindexed with a clitic pronoun in the IP; see Cinque (1990) among many others.

⁷ Sportiche (1998:297), drawing a parallel between CLLD in Italian and Greek and clitic doubling of quantifiers in French when the latter occur in positions to the left of the clitic, notes that the movement should be obligatory because “true” clitic doubling is not allowed in these languages. This last remark is not quite right, though: it’s true that Italian does not have clitic doubling (this is one of Cinque’s (1990) arguments for base-generated CLLD constructions), but Greek does.

above, the dislocated DP enters into a chain relation with *pro*. The derivation of a construction involving a dislocated strong DP will be as in (7).



The presence of the clitic in (7) can be explained as follows. We know from earlier discussion that *pro* is [+strong] (see chapter 2 for details).⁸ Given that a [+strong] feature must be licensed for the derivation to converge, the clitic must be present. Thus, the explanation here is consistent with the analysis of clitic doubling constructions so far.

What happens when a weak DP is base-generated in the left periphery? In this case, too, the dislocated DP connects to *pro*, forming a binding chain. Suppose now that weak

⁸ In chapter 2, fn. 21, I followed Sportiche (1998) in analyzing examples like (i) below as doubling constructions involving a silent DP in the argument position of the verb, as in (ii). I repeat the original example below:

(i) Go vidov.
him saw
'I saw him.'

(ii) $[[_{CIP} pro_i \text{ go}_i \text{ } [_{VP} \text{ vidov } t_i]]$

The silent DP in this case is identified as a pronoun, i.e. *pro*. As (ii) shows, the analysis of such constructions is consistent with the analysis of clitic doubling outlined in chapter 2. At LF, *pro* moves to SpecCIP, leaving a trace in argument position, with which it forms a chain.

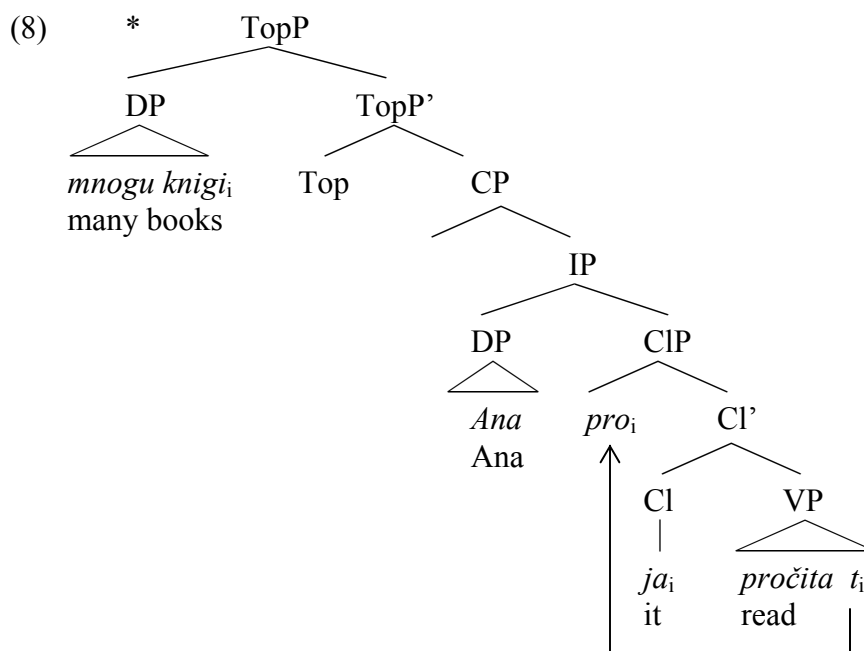
DPs are marked as [-strong].⁹ Given this, the configuration that involves a [-strong] DP, and a [+strong] *pro* (and clitic) would be ungrammatical. This is because the features of the weak DP clashes with the [+strong] feature of *pro* (as well as that of the clitic). This follows from the requirement that syntactic chains share the same feature specification. As pointed out by Suñer (1988:394) (and references cited therein), a chain is well-formed only when there is no clash in features between its elements. Suñer explains that a clitic may form a chain with a constituent only if it fulfills this requirement.¹⁰ In our case, we can extend this to include the dislocated DP as well. The ungrammaticality of constructions involving a left-dislocated weak DP and a clitic in the IP thus follows from general conditions on chains. This is also captured by the nondistinctiveness clause in the conditions on chains given in (5) above, as pointed out by Baker (1996:127).¹¹ The derivation of constructions involving a left-dislocated weak DPs is given in (8) below.

⁹ The assumption would in no way alter our explanation of their derivation by movement. The CIP would not be projected regardless of whether a DP is unspecified or [-strong]; its presence is only triggered by a [+strong] feature.

¹⁰ This condition is fulfilled in all cases of clitic doubling in Macedonian: both the clitic and the doubled element are marked as [+strong].

¹¹ Baker (1996:127) notes that the inability of indefinites to occur in CLLD constructions in the Romance languages may be a result of clash of features. The clash in this case, Baker explains, may be in definiteness features (the clitic pronoun is definite and as such can only license a DP that bares the same features). In other words, he notes that a chain can only be formed if both the doubled DP and the clitic pronoun are definite.

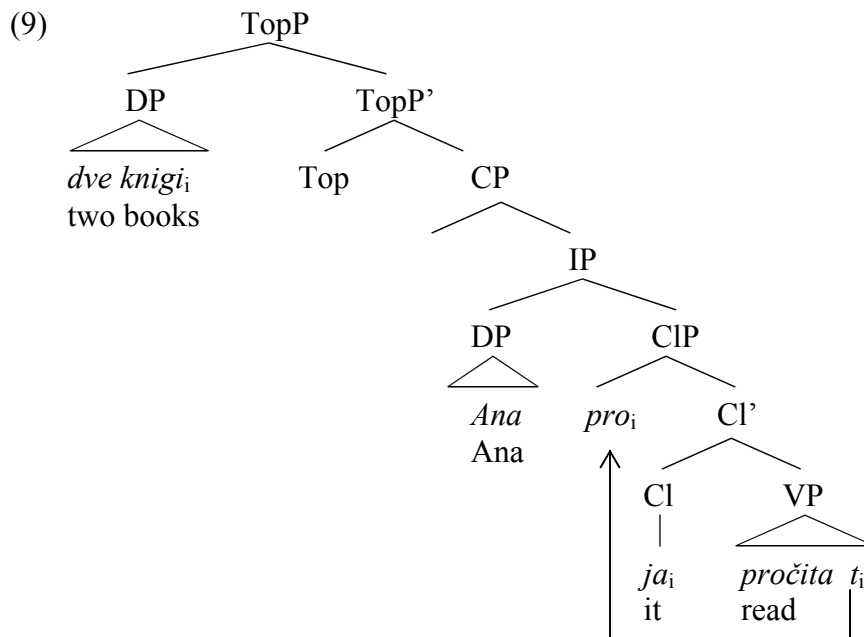
Baker points out that the clash of features can follow straightforwardly from clause (iv) of the Chain condition given in (5) above, if definiteness is considered to be a morphosyntactic feature. The same question would presumably arise with respect to [+strong]. I leave this open.



Thus, a strong DP is acceptable in a binding chain because its features are compatible to those of *pro* in the argument position. A weak DP in a binding chain yields an ungrammatical sentence because its features clash with the features of *pro*. We are now in a position to answer the question we started with, which in the context of the present discussion we can reformulate as follows: why is it possible for some weak DP to enter into a chain relation with *pro*, while others cannot. To answer this question, I follow a suggestion by Veneeta Dayal (p.c.) and propose that the weak DPs actually divide into two groups: those that are inherently [-strong] (this group of DPs would include *many/few* and the bare plurals in (2) and (3a-b)) and those that are unspecified for strength (this group of DPs would include the numerals and *several* in (4)). How does this help us explain the data?

To begin with, the clitic is generated because of *pro* (which is [+strong]) in accordance with the Clitic Criterion. A weak DP intrinsically marked as [-strong] is predicted not to be allowed to enter into a binding relation with *pro* because of clash of

features. Thus, a construction involving a clitic-doubled dislocated *many/few* or a bare plural is ungrammatical. A weak DP which is not specified for strength can enter into a binding chain because it does not clash with the features of *pro*. Since chains have to share features, elements participating in them will either have the features inherently (as in the case of strong DPs) or they will be able to acquire them from the chain (as in the case of the numerals, for example). Crucially, though, the numerals are not [-strong]. They remain unspecified or change to [+strong] through the binding relation with *pro*. The derivation of left-dislocated numerals is given in (9) below.



To summarize, by positing a three-way distinction among the DPs in terms of feature specification, we are able to account for the full set of data in section 3.1. By invoking the general condition that elements in a chain must not have clashing features, we were able to explain the divergent behavior of the weak DPs. Crucially, the analysis leaves it open that an unspecified weak DP can co-occur with a clitic when left-dislocated. In chapter 4,

I will try to account for both these options by arguing that the constructions without a clitic are derived by movement, while constructions with a clitic are base-generated.

Finally, it is important to note that the proposal I have just outlined does not in any way change the explanation for the effects of clitic doubling within IP. The doubling of strong DPs is accounted for by the Clitic Criterion. Weak DPs with a [-strong] feature also follow the Clitic Criterion: their features do not agree with that of the clitic, since the latter is a licenser for the [+strong] feature. Thus, these DPs are predicted not to occur in doubling constructions. As we know from chapter 2, this prediction holds. Weak DPs unspecified for strength would also not trigger the presence of the clitic. This follows from principles of economy. This explains the inability of weak DPs such as the numerals and *several* to co-occur with clitic pronouns.

3.3 The Semantic Basis for the Distinction

The explanation in the previous section depends crucially on the distinction among DPs with respect to the feature specification for strength. In this section, I will try to provide independent motivation for this claim. In section 3.2, I have maintained that weak DPs that allow for clitics when left-dislocated remain unspecified (or change to [+strong] through the binding relation with *pro*). Here, I provide evidence why such DPs are crucially not [-strong]. In order to do that, we need to look more closely at the semantic properties of these DPs, and in particular at their (in)ability to have a specific interpretation. In order to put this issue in context, we need to go back to Milsark's original proposal concerning the classification of DPs in terms of strength.

As we have noted in chapter 2, Milsark (1977) classified English determiners as weak vs. strong based on their ability or inability to occur in existential *there*-sentences. He noted, however, that the distribution of determiners with “property” predicates is also restricted, in that only strong determiners can co-occur with such predicates, as shown in (10a-b) for English.

- (10)a. Every student/the students are intelligent.
- b. *Sm students are intelligent.
- c. Some students are intelligent.

To account for the fact that weak determiners can also occur in such environments (cf. (10b) and (10c)), Milsark treated the weak determiners as ambiguous, in the sense that they could have strong correlates. On his analysis, the unstressed *sm* in (10b) is weak but the stressed *some* in (10c) is strong. Semantically, Milsark analyzed the strong determiners as quantificational expressions and the weak as nonquantificational, i.e. cardinality, expressions. He claimed that the treatment of weak determiners as cardinality expressions was warranted because “they do nothing more than express the size of the set of entities denoted by the nominal with which they are construed.” (Milsark 1977:23) The strong determiners, on the other hand, “can be regarded as being exactly the set of natural language quantifier words.” (*ibid.*)

Numerous other studies have also treated weak DPs as semantically ambiguous. Fodor & Sag (1982), for example, consider indefinites to be ambiguous between

quantificational and referential expressions.¹² Their main argument for distinguishing the two types stems from the observation that indefinites on their referential reading can get scope outside of syntactic islands; indefinites on their quantificational reading are unable to do that. In this context, Fodor & Sag argue that the exceptional wide scope reading of the indefinites stems from their referential reading. Kratzer (1998) also maintains that indefinites are ambiguous, but in this case the ambiguity is between a quantificational and specific interpretation. Kratzer argues that on their quantificational interpretation, indefinites are interpreted as generalized quantifiers; on their specific reading, they are interpreted via choice functions (following Reinhart 1995). Others have also taken the exceptional scope behavior of the indefinites to be a result of their non-quantificational status (on the view that true quantifiers are subject to syntactic constraints which limit their scope). Abusch (1993/4) and Diesing (1992), among others, adopt the view that indefinites are different from other quantifiers in that they do not have quantificational force of their own (following Kamp (1981), Heim (1982)).¹³

Schwarzschild (2002) takes a different view and argues that indefinites can uniformly be treated as existential quantifiers. He argues that their apparent exceptional scope is a result of a quantifier domain restriction, whereby the domain of the indefinite is restricted to a singleton set. He further shows that the resulting singleton indefinite

¹² Fodor & Sag (1982) define a referential indefinite to be that which is used by the speaker (but not the hearer) to identify a particular referent in the discourse. This is in essence the definition of a specific indefinite, which semantically I take to be interpreted as a singleton indefinite (following Schwarzschild 2002).

¹³ Abusch (1993/4) builds on the view that indefinites are different from other quantifiers (and thus have different scope properties) and derives their wide scope reading by allowing them to be bound by non-local operators via a special storage mechanism that “preserves the restrictions on [the] free variables corresponding to [the] indefinites” (p.108) until the indefinite is combined with its operator. Diesing (1992) distinguishes two types of indefinites: quantificational (in which case indefinites form restrictive clauses and are subject to QR) and non-quantificational (in which case indefinites introduce free variables that are bound by existential closure). Diesing argues that specific indefinites are presuppositional and as such undergo QR.

becomes similar to a definite, i.e. a true referring expression, and as such becomes essentially scopeless. In other words, he shows that the referential readings of indefinites argued for by Fodor & Sag (1982) arise from the domain restriction of the indefinite to a singleton set.

One of my main assumptions here is that the specific indefinites are not a distinct type of DP. That is, I explicitly assume that they are ordinary weak indefinites (existentials) which get a quantifier domain restriction, as proposed by Schwarzschild (2002). Crucially, I do not treat them as semantically ambiguous (following Schwarzschild (2002); contra Milsark (1977), Fodor & Sag (1982), Diesing (1992), and others). This in turn means that the analysis of clitic doubling constructions proposed in chapter 2 is not affected by the further distinction within the class of weak DP that I am proposing here.

As mentioned above, Fodor & Sag argue for ambiguity of indefinites, i.e. a separate referential interpretation in addition to their existential interpretation, and claim that the wide scope reading of an indefinite is a result of their referential interpretation. Schwarzschild removes the ambiguity condition: he argues that the indefinites are existential quantifiers and their observed referential interpretation is a result of a domain restriction (which can be explicit or implicit) containing a bound variable. This means that an indefinite is interpreted as an existential quantifier which ranges over a set that is contextually restricted to a singleton. The restriction of their extension to a singleton set enables the indefinites to refer to a particular referent, which in turn accounts for their referential properties. The singleton domain restriction essentially turns the indefinites

into a referential phrase; as such they become scopeless. This, in turn, explains their well-known ability to take wide scope even out of syntactic islands.

For concreteness, I will assume that the contextual delimitation is represented as in von Fintel (1994). The restriction on the quantifier domain is basically construed as a presupposition (Penka Stateva, p.c.). In other words, the restriction triggers a presupposition that there is single individual of which the assertion holds. This in essence captures the intuition that specific indefinites behave just like referential phrases, i.e. definites (as in the case of Schwarzschild (2002)) or indexicals (as in the case of Fodor & Sag (1982)). To illustrate how the context domain restriction would work on Fodor & Sag's original example, given here in (11), I follow von Fintel (1999, 2000:2):¹⁴

- (11) a. If a friend of mine from Texas had died in the fire, I would have inherited a fortune.
- b. if $\exists x (C(x) \ \& \ x \text{ friend of mine from Texas} \ \& \ x \text{ had died in the fire})$, I would have inherited a fortune

As von Fintel (2000) explains, the domain restriction in (11) is achieved by a free variable C , to which the context assigns a value. In this particular case, the value of the variable is such that it contains exactly one friend of mine from Texas.

¹⁴ Alternatively, the context domain restriction can be achieved through a subset selection function, i.e. a function from sets to subsets. In the case of a singular argument (as in (1a)), the subset would contain a single individual; in the case of a plural argument, the subset would contain a plural individual (von Fintel 2000:5).

(i) Call C a singleton subset selection function iff for any predicate P (that C is defined for if C is a partial function) $C(P)$ is true of exactly one individual.

For a discussion on how the idea of bound variable within the restrictor can be implemented, see also Stanley & Gendler-Szabó (2000), Schwarzschild (2002) and references therein.

There are two points of interest to us in the debate on indefinites, in general, and specific indefinites, in particular. First, there is the observation that some quantifiers allow for a referential interpretation, while others do not.¹⁵ Second, some quantifiers seem to have the ability to take scope outside of configurations that otherwise pose such restrictions; others do not have this ability. What studies have shown is that there is a correlation between the class of DPs that allow a referential interpretation and those that seem to have exceptional scope abilities.

My intention in bringing these issues up is to show that there is a correlation between the ability of DPs to have specific (i.e. referential) interpretation and their (subsequent) feature specification. More specifically, I will try to show that DPs that do not allow for a specific reading are inherently [-strong], while those that allow for a specific reading are unspecified for strength. In order to show the parallels, I will use the tests that bring out the referential reading and those that bring out the exceptional scope reading.

As previously mentioned, one of the key issues relating to the behavior of indefinites stems from the observation that they have a unique ability to take scope outside of

¹⁵ In their discussion on referential indefinites, Fodor & Sag (1982) give a rough classification of which DPs can and cannot have a referential interpretation. The first group includes: *a*, *some*, *several*, *many* and the numerals. The second group includes: *every*, *each*, *all*, *most*, *few*, *no*, and the null plural determiner. Fodor & Sag, however, note the difficulty in finding the determining factor that enables elements to have a referential reading. The classification of *many* and *few*, in particular, may not be as straightforward given their ambiguity between cardinal and proportional readings (see Partee (1989)), something that Fodor & Sag are also aware of.

Fodor & Sag (1982:393, fn. 2) speculate that the ability of determiners to get a referential interpretation may have something to do with their ability to get a ‘relative’ or ‘absolute’ interpretation. Determiners such as *every* or *most*, i.e. determiners that resist referential interpretation, they explain, are ‘relative’ in the sense that the number of individuals that they pick out depends on the number of individuals that make the predicate true in a given model. Thus, the interpretation of *every/most tree(s)* would crucially depend on how many trees there are in the model. Numerals, on the other hand, Fodor & Sag note, have an absolute interpretation; *many* and *several* are ambiguous between an absolute and relative interpretation. Based on this, Fodor & Sag suggest that “a referential interpretation is possible for a noun phrase only if its determiner is construed as absolute” (p. 393, fn. 2). They hasten to add that this may not be exactly true. They note the possibility that in phrases like *many friends of ours* the determiner can have a proportional reading. This would make it more compatible with a relative interpretation while at the same time allowing it to have a referential interrelation.

syntactic constructions that otherwise restrict the scopal behavior of quantified expressions. Thus, (12a) has a reading where the indefinite *a friend of mine* takes scope out of the *if* clause; this reading is not available to *each of the candidates* in (12b). Both examples are from Fodor & Sag (1982:369-370).

- (12)a. If a friend of mine from Texas had died in the fire, I would have inherited a fortune.
- b. If each friend of mine from Texas had died in the fire, I would have inherited a fortune.

The difference between (12a) and (12b) is that (12a) can be understood to be about a particular friend that the speaker has in mind. (12b), on the other hand, cannot be understood to mean that each friend (of the speaker) is such that if he/she had died the speaker would have inherited a fortune. This means that the indefinite in (12a) has the ability to take scope outside of a syntactic configuration that otherwise imposes restrictions on scope, as shown by the universal in (12b). This is why indefinites are said to take exceptional scope.

(12a) and (12b) illustrate the difference between indefinites and strong quantifiers like *each* in English. The test can be used to show that the same holds for Macedonian, keeping in mind that the goal here is to separate between the two types of weak DPs. The

examples in (13) and (14) are a variation of (12a-b), in that the relevant DP is in object position.¹⁶

(13)a. Ako požarot ubie eden rodnina, kje nasledam kukja.¹⁷

if fire-the kills one relative will inherit house

‘If the fire kills a relative (of mine), I will inherit a house.’

b. Ako požarot ubie dvajca rodnini, kje nasledam kukja.

if fire-the kills two relatives will inherit house

‘If the fire kills two relatives (of mine), I will inherit a house.’

c. Ako požarot ubie nekolku rodnini, kje nasledam kukja.

if fire-the kills several relatives will inherit house

‘If the fire kills several relatives (of mine), I will inherit a house.’

(14)a. Ako požarot ubie mnogu/malku rodnini, kje nasledam kukja.

if fire-the kills many/few relatives will inherit house

‘If the fire kills many relatives (of mine), I will inherit a house.’

b. Ako požarot ubie rodnini, kje nasledam kukja.

if fire-the kills relatives will inherit house

‘If the fire kills relatives (of mine), I will inherit a house.’

¹⁶ This is necessary because subject DPs in Macedonian cannot be clitic-doubled. What we are interested in is the correlation between a specific reading and the absence of a clitic when a weak DP occurs within IP. This can only be tested with direct object DPs.

¹⁷ Macedonian is a pro-drop language.

As was the case with the quantifier in (12a), the weak DPs in (13) are able to get a specific interpretation. Thus, (13a) can easily be understood to be about a particular relative of mine, such that if he dies, I would inherit a house. The numeral in (13b) can also be construed to as being about a specific set of two relatives. As Schwarzschild (2002:299) points out, the wide-scope existential in this case does not range over all my relatives, indiscriminately, but rather a specific set of two relatives. The same holds for *nekolku rodnini* ‘several relatives’ in (14c).

The weak DPs in (14a-b), on the other hand, do not lend themselves to such an interpretation. (14a) cannot be understood to mean that many (or few) relatives (of mine) are such that if they are killed by the fire, I would inherit a house. (14b) can only be understood to mean that if any relatives are killed by the fire, I would inherit a house. In other words, the weak DPs in (14a-b) cannot get a specific interpretation.

I should point out that, given Schwarzschild’s (2002) analysis, it follows that the weak DPs in (13a-c) allow for a specific reading because they can be interpreted as singleton expressions. (14a-b) then, would not allow for a specific reading because they cannot be interpreted as singleton expressions. For the numerals and *several* this means that they acquire a domain restriction such that they are interpreted as singleton expressions (on a par with indefinites). The set they quantify over is the set of relatives; the domain restriction is placed by the context. The only difference between the numerals/*several* and the indefinite in (13a) is that the denotation of the bare numeral and *several* would be a plural individual.¹⁸

¹⁸ A brief illustration of how the interpretation of a bare numeral as a plural individual works is given in (i) below.

- (i) a. Five ships are on dock.
- b. $\exists X [X \text{ is a group of ships} \ \& \ |X|=5 \ \& \ [\forall x \in X \rightarrow x \text{ is on dock}]]$

The analogy with the specific indefinites (and their treatment as singleton expressions) works for the numerals and *several* because they can also be interpreted as singleton expressions, i.e. they will denote a singleton set of a plural individual. In this case, the numeral will pick out the unique, maximal set of individuals. *Many/few*, on the other hand, are context-dependent, which means it is very difficult to fix the cardinality of the set denoted by the CNP (the set they denote).¹⁹ Moreover, *many/few* allow for a non-maximal interpretation, which means that the set of *n* number of individuals would never be construed in such a way so as to refer to a particular set of individuals. This means that it would be impossible (or in any case extremely difficult) to construe the referent of the quantifier as a singleton set (given the fact that a plural individual would have to pick out the unique maximal set of sum individuals). Thus, if we combine these properties of *many/few* with the idea that the specific interpretation of an indefinite/numeral is a result of its interpretation as a singleton, we can explain the nonreferential character of *many/few* to be a result of their inability (or extreme difficulty) to be interpreted as singleton expressions.²⁰

Five boats in (ia) is interpreted as a plural individual which means that out of the set of atomic and sum individuals in the lattice we pick out the maximal one. In this case, the implicit domain restriction, following von Stechow (1994) will be through a free variable *C* which will provide a contextually supplied set which will restrict the quantifier's domain. An illustration of this is given in (ii) below.

- (ii) a. $[[\text{five}_C \text{ ships}] \text{ are on dock}]$
 b. $\exists x [C(x) \ \& \ x \text{ is a ship} \ \& \ x \text{ is on dock}]$

In (ii), *five* is interpreted as a generalized quantifier of type $\langle\langle e, t \rangle, \langle\langle e, t \rangle, t \rangle\rangle$ and *C* denotes the contextually supplied set which contains exactly one plural individual which has the property of being on dock.

¹⁹ *Many* and *few* are by definition (and more so than any other quantifier) context dependent in that the set they denote is dependent on a contextually determined number. For example, how *many/few* *x*'s are *y*'s is dependent on how many *x*'s there are in the model.

²⁰ The lack of referentiality of *many/few* can be further explained within a general theory of presupposition and how these two quantifiers are understood. For example, in uttering the sentence *Five ships are on dock*, where *five ships* are understood referentially, i.e. the expression refers to a particular group of five ships. The actual grouping of the ships in this case would be contextually determined – just think of a context where there are five Australian ships, three New Zealand ones, and seven British in the harbor. In this context, the question *Are all Australian ships on dock?* can be answered by *Yes, there are five ships on dock.*, where *five ships* refers to the maximal set of Australian ships. Moreover, for this to happen, the

The importance of the examples in (13a-c) and (14a-b) for us is that the DPs that allow for a specific reading within IP, namely (13a-c), are exactly those that allow for clitics in left dislocation (see (4)). The DPs that do not allow for a specific reading within IP, namely (14a-b), are those that do not allow for clitics when left-dislocated (see (2)-(3a-b)). More importantly, this shows that the specific interpretation of a weak DP does not entail a [+strong] feature. We see this most clearly in case of the weak DPs in (13a-c). If such entailment relation existed, we would expect the numerals and *several* to be clitic-doubled even when they occur in IP internal positions.

I have thus shown that the DPs that have specific indefinite readings in Macedonian are those that are unspecified for strength. Crucially, we see that such DPs do not become [+strong]. If they did, they would be able to co-occur with clitics, as required by the Clitic Criterion. Given the fact that [-strong] DPs do not allow for a specific interpretation, we can conclude that weak DPs that do allow for such interpretation cannot be [-strong].

To conclude, I have provided semantic motivation for the syntactic features of Macedonian DPs with respect to strength. I have shown that numerals do not have inherent features for strength, while *many/few* are inherently [-strong]. The three-way distinction is completed with the inherently strong DPs, being specified as [+strong].

presupposition that there is a set of uniquely identified five ships is added to the common ground, before the assertion *are on dock* is being made. However, this is not the case with *many/few*. Imagine a context where there are sixty ships in the harbor, forty three of which are Italian, ten Spanish and seven Portuguese. In this case, *Many ships are on dock* cannot refer to a particular set of ships, because the presupposition that there are forty three Italian ships on dock is not added to the common ground before the assertion is made, and thus it is not shared information between the participants in the conversation. This, on the other hand, is possible with the numerals (and indefinites), which in turn enables their referential use.

3.4 Partitive Phrases

In this section, I would like to return to an issue that was briefly mentioned in chapter 2, fn. 16. The issue relates to the behavior of partitives and their co-occurrence restrictions with clitics. The partitives are relevant for the analysis of clitic doubling for at least two reasons. In the context of the discussion in chapter 2, the partitives pose an interesting problem in that they show a somewhat unexpected behavior with regards to their co-occurrence with clitics. In the context of the discussion in this chapter, the partitives offer evidence that strong readings of DPs do not necessarily entail [+strong] features. In what follows, I will try to show that the two issues are closely related and that a possible solution for the exceptional behavior of the partitives can be found if we maintain that clitic doubling in Macedonian is only licensed in contexts where a DP is unambiguously [+strong].

Recall from chapter 2 that Macedonian makes a clear distinction between weak and strong DPs in that only the latter are allowed to co-occur with clitics pronouns. In this context, consider the partitives.

In terms of their semantic properties, partitives come out as strong on both tests introduced in chapter 2. On Milsark's test, they come out as strong because they are not allowed in *there*-sentences, as shown in (15).

- (15)a. *Ima edna od knigite na masata.
 have one of books-the on table-the
 'There is one of the books on the table.'

- b. *Ima povekjeto od knigite na masata.
 have most of books-the on table-the
 ‘There are most of the books on the table.’

On Barwise & Cooper’s (1981) test partitives also come out as strong. The sentences in (16a-b), for example, will come out true in every model due to the presupposition associated with the NP in the *of* phrase.

- (16)a. Dve od knigite se knigi.
 two of books-the are books
 ‘Two of the books are books.’
- b. Povekjeto od knigite se knigi.
 most of books-the are books
 ‘Most of the books are books.’

Given their characterization as strong, we would expect partitive DPs to freely co-occur with clitic pronouns. This, however, does not seem to be the case. Consider (17a-b) and (18).

- (17)a. (*Ja) pročitav edna od knigite.
 it(f.sg.) read one of books-the
 ‘I read one of the books.’

b. (*Gi) pročítav dve/nekolku od knigite.

them read two/several of books-the

‘I read two/several of the books.’

(18) *(Ja) pročítav sekoja edna od knigite.

it(f.sg.) read every one of books-the

‘I read every one of the books.’

In (17a-b), the partitives are headed by numerals or *several*; the clitic is not allowed. In (18), the partitive is headed by *sekoja* ‘every’; the clitic in this case becomes obligatory. Examples such as (17a-b)-(18) seem to show that the presence of a clitic is not determined by the features of the partitive DP as a whole, but rather by the properties of the head of the partitive. This is most evident in cases where the partitive is headed by a weak determiner, as in (17a-b).

The behavior of partitives in Macedonian as such poses an interesting problem for analyses like Comorovski’s (1996) which maintain that partitive NPs carry an existential presupposition. In Comorovski’s view, the presupposition associated with the partitive makes the DP strong. In Romanian this enables their clitic-doubling.²¹ Comorovski’s examples of partitives are in the context of *wh*-phrases, as in (19), but the obligatoriness of the clitic is clear. (From Comorovski 1996:73)

²¹ Comorovski (1996) takes this to be one of the properties they have in common with *which* NPs (i.e. D-linked *wh*-phrases).

- (19) Pe care (dintre ei)_i l_i-ai vazut e_i?
 prep which of them him you-have seen
 ‘Which one have you seen?’

We have just seen that partitive phrases in Macedonian follow a different pattern from that of their Romanian counterparts. Data such as (17a-b) and (18) show that the determining factor in the distribution of the clitic is the strength of the determiner, rather than the existential presupposition of the whole DP. This seems to suggest that for Macedonian, the features of the DP in the *of* phrase are inaccessible for feature checking.

The patterning of the partitives in (17)-(18) with respect to clitic-doubling is parallel to that of weak and strong DPs within IP. Once dislocated, partitives headed by numerals or *several* follow the unspecified weak DPs. This is shown in (20) where left-dislocation of the partitive DPs enables the presence of the clitic.

- (20) Dve/nekolku od knjige, Petar (gi) pročit_a.
 two/several of books-the Petar it(f.sg.) read
 ‘Two/several of the books, Petar read (it).’

A left-dislocated partitive headed by *many/few*, on the other hand, does not co-occur with a clitic. This is illustrated in (21) and (22) below.

(21)a. Mnogu od knjigite, Ana(*gi) ima pročitano.²²

many of books-the Ana them has read

‘Many of the books, Ana has read.’

b. Mnogu od studentite, Ana (*gi) zapozna lanska godina.

many of students-the Ana them met last year

‘Many of the students, Ana met last year.’

(22)a. Malku od knjigite, Ana(*gi) ima pročitano.

few of books-the Ana them has read

‘Few of the books, Ana has read.’

b. Malku od studentite, Ana(*gi) zapozna lanska godina

few of books-the Ana them met last year

‘Few of the students, Ana met last year.’

The partitives are important in that they shows that a strong reading of a DP does not entail a [+strong] feature. Although they are semantically strong, the inability of partitives to co-occur with clitics when they are in positions within IP, as in (17a-b), shows that these DPs remain unspecified for strength. If we assume that the head of the partitive is responsible for the feature specification of the phrase, then it follows that a partitive headed by an unspecified or a [-strong] weak determiners (numerals and *many/few*, respectively) would not be clitic-doubled; partitives headed with intrinsically [+strong] determiners would obligatorily require the presence of a clitic pronoun. Finally,

²² Some speakers seem to find the examples with a clitic marginally acceptable (at best). Even among those speakers, sentences such as (22) are judged ungrammatical.

this may explain the lack of clitics with weak DPs on their specific reading, when in positions within IP.

Before I conclude this section, I would like to briefly return to the partitives in Romanian. Given the similarities between Macedonian and Romanian with respect to clitic doubling, it would be interesting to see to what extent the proposal made above applies to them.

A cursory look at Romanian seems to show that partitives headed by weak determiners pattern like their Macedonian counterparts.²³ This is shown in (23) where the partitive headed by *two* is not clitic-doubled.

(23) Ana a citit două dintre aceste cărți.

Ana has read two of these books

‘Ana read two of the books.’

Given that partitives are strong, (23) is interesting in that it suggests that this property in Romanian also does not translate into a [+strong] feature. This is similar to what we saw in Macedonian. The parallel between the two languages, however, does not extend to all types of partitives. Consider (24):

(24) Ana a citit fiecare din aceste cărți.

Ana has read each of these books

‘Ana read each of these books.’

²³ I thank Mirela Dragomir for judgments on Romanian.

In (24), the partitive is headed by *each*, which is a strong determiner. Yet, the partitive is not clitic-doubled. This is in contrast to Macedonian.

Moreover, partitives in Romanian seem to optionally allow for a clitic when they are in positions in IP, regardless of whether or not they are headed by a weak or a strong quantifier; see (25a-b).

- (25) a. Ana le-a citit pe două dintre aceste cărți.
 Ana them-has read pe two of these books
 ‘Ana read two of the books.’
- b. Ana le-a citit pe majoritatea (acestor cărți).
 Ana them-has read pe most these books
 ‘Ana read most of these books.’

Their behavior remains unchanged in dislocated structures. In (26) we see examples of both types of partitives co-occurring with clitics; in (27) we see that the clitic may be absent.²⁴

- (26) a. Dintre aceste cărți, Ana le-a citit pe fiecare.
 of these books Ana them-has read pe each
 ‘Ana read each of these books.’

²⁴ Although a dislocated partitive without a clitic is acceptable, it is less so than a clitic-doubled one.

- b. Dintre aceste cărți, Ana le-a citit pe două.
 of these books Ana them-has read pe two
 ‘Ana read two of the books.’
- (27) a. Dintre aceste cărți, Ana a citit două.
 of these books Ana has read two
 ‘Ana read two of these books.’
- b. ?Dintre aceste cărți, Ana a citit pe fiecare/majoritatea.²⁵
 of these books Ana has read pe each/most
 ‘Ana read each/most of the books.’

The question of what exactly triggers the clitic with partitives in Romanian requires a more careful investigation and providing a definitive answer to such question is beyond the scope of this study. From the discussion above, we can conclude that for Macedonian, the strong interpretation of a constituent does not necessarily entail a [+strong] feature. This principle does not seem to be applicable to Romanian, which in turn may explain the differences in the clitic co-occurrence restrictions between the two languages.

3.5 Conclusion

In this chapter I have argued for a three-way distinction between the DPs in Macedonian. While inherently strong DPs are marked as [+strong], weak DPs split into two groups.

²⁵ Note, however, that the absence of a clitic with a left dislocated partitive headed by a strong quantifier is only possible if the special *pe* preposition is present. Leaving out the preposition results in an ungrammaticality.

Tests show that weak DPs that allow for specific interpretation are unspecified for strength, while weak DPs that do not allow for specific interpretation are specified as [-strong]. The fact that DPs that allow for specific interpretation, allow it indiscriminately, outside, and crucially within the IP, suggest that a strong interpretation does not entail a [+ strong] feature. Finally, the more nuanced characterization of the weak DPs enabled us to account for the optionality of the clitic when such DPs occur at the left periphery. The presence vs. absence of the clitic in all cases of left-dislocated DPs was argued to follow from general conditions on chains that disallow a clash of features between its elements.

CHAPTER 4

Distinctions at the Left Periphery

4.0 Introduction

In chapter 3, I explained the clitic co-occurrence patterns of clause-initial direct objects in Macedonian by making a three-way distinction among DPs in their feature specification for strength and by proposing that there are two distinct constructions involved. The two constructions are schematically represented in (1a-b). (Dotted lines represent a chain that is not formed by movement; solid lines indicate chains formed by movement.)

- (1) a. $[_{XP} DO_i \ [_{IP} S \ pro_i \ CL_i \ V \ t_i \dots]]$
 └──────────┘ └──────────┘
 └──────────┘ └──────────┘
- b. i. $[_{XP} DO_i \ [_{IP} S \ t_i \ CL_i \ V \ t_i \dots]]$ or
- └──────────┘ └──────────┘
- ii. $[_{XP} DO_i \ [_{IP} S \ V \ t_i \dots]]$
 └──────────┘ └──────────┘

(1a) represents the dislocation of a direct object that is [+strong] or unspecified for strength. (1b-i) represents the fronting of a [+strong] direct object; (1b-ii) represents the same construction with a [-strong] direct object or one unspecified for strength.

This chapter probes more deeply into the properties of the constructions containing such dislocated objects. In what follows, I will show that clause-initial elements in Macedonian can be distinguished derivationally and thus mapped onto two different constructions, which are identified as Clitic Left Dislocation (CLLD) and Topicalization

(TOP).¹ I show following Cinque (1990) and others that CLLDed DPs are base-generated in their surface position, and that topicalized DPs are derived by movement. The distinction arises most clearly in the case of weak DPs which allow for both options. Building on the derivational differences, I then show that the two constructions also differ in terms of their semantic and pragmatic properties.

Any discussion of dislocated elements inevitably touches on the issue of their positioning in the clause. For this reason, the chapter also addresses the question of where clause-initial elements fit into the left periphery in Macedonian. In this context, I also discuss a related construction called Hanging Topic Left Dislocation (HTLD) with a view of separating it from the two constructions under discussion, CLLD and Topicalization.

The chapter is organized as follows. I begin in section 1 with a discussion of the structure of the left periphery in Macedonian and the positions CLLDed and topicalized elements occupy there. Section 2 discusses the syntactic properties of the two constructions and presents arguments for their different derivation. Section 3 discusses the semantic and pragmatic properties of the two constructions. Section 4 contains the conclusion.

¹ A note on the terminology used: the term CLLD comes from Cinque (1990) and others; it identifies a construction where a DP in clause-initial position co-occurs with a pronoun in the IP. I use the term Topicalization to identify a construction with a fronted DP which does not co-occur with a clitic pronoun in the IP (see, for example, Ward & Birner (2004)). Since I assume that both constructions, CLLD and TOP, in Macedonian mark topics (see section 4.3), my use of the term Topicalization differs from that of Cinque (1990) and Rizzi (1997) who use this term for constructions that mark foci in Italian.

4.1 Direct Objects and the Left Periphery

4.1.1 The Architecture of the Left Periphery

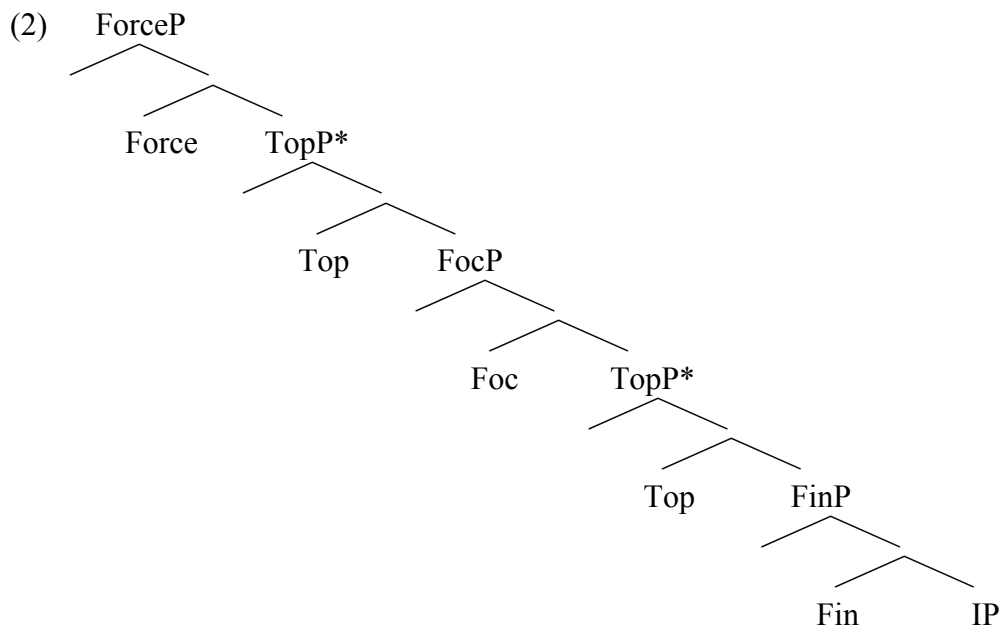
We saw in chapter 2 that direct object DPs in Macedonian can occur in two positions within IP, preverbal and postverbal. For a detailed discussion of their properties, see chapter 2. In this section, I focus on the structure of the left periphery as a way of orienting the discussion of how the clause-initial direct objects in Macedonian fit in this part of the clausal domain.

In what follows, I outline two possibilities about the placement of CLLDed and topicalized elements in the overall structure of the clause. In doing so, I build on my earlier assumptions about the representation of CLLD and Topicalization in Macedonian, as given in chapter 3. The structures of these two constructions there were drawn on the explicit assumption that both CLLDed and topicalized phrases occupy a position in the CP domain that hosts topics, which following Rizzi (1997), I identified as TopP (see chapter 3).² In order to understand better the motivation behind this particular analysis, I discuss below Rizzi's (1997) proposal about the left periphery, paying particular attention to the issue of the placement of topics within it.

Rizzi (1997) argues that a) the CP field has a more articulated structure in the sense that it no longer is defined in reference to a single C head but rather a host of functional projections and b) the CP field contains a set of fixed positions to which specific elements go. Rizzi's arguments come from ordering restrictions of topics and foci in wh-

² My analysis of CLLD as utilizing the Spec of TopP above CP falls in line with Kallulli's (1999) representation of CLLD in Albanian and Greek. It should be noted, however, that Kallulli assumes a movement analysis of the CLLD constructions in these two languages.

questions and relative clauses in Italian. The structure of the left periphery that he proposes is given in (2).³



As the schema in (2) shows, the CP field is defined by the presence of two functional heads, Force and Fin, responsible for clausal typing and encoding inflectional specifications (e.g. finiteness), respectively.⁴ More importantly, the CP field hosts other functional projections whose properties are defined not by selectional restrictions within the clause but by the discourse (see also Grohmann 2001). As such, the CP domain is responsible for encoding topic, focus, wh-expressions, etc..⁵

³ Topic projections are recursive, hence the asterisk (TopP*).

⁴ ForceP and FinP are included here for expository purposes. In the discussion that follows, I abstract away from Rizzi's original proposal concerning the existence of these two heads and explicitly assume a single C head as this is sufficient to delineate the CP domain.

⁵ One should immediately note that the existence of such projections in the left periphery does not preclude their existence elsewhere in the clause. For example, I have assumed that a focus position is available within the IP in Macedonian (see chapter 2). Languages differ in how they encode such discourse-related properties and whether or not they have specifically defined structural positions for this purpose.

As a way of illustration of how the schema in (2) applies in a particular language, consider the Italian examples below, from Rizzi (1997:288).

- (3) a. Credo che il tuo libro, loro lo apprezzerebbero molto.
 ‘I believe that your book, they would appreciate a lot.’
 b. *Credo, il tuo libro, che loro lo apprezzerebbero molto.
 ‘I believe, your book, that they would appreciate a lot.’
 c. *Credo di il tuo libro, apprezzarlo molto.
 ‘I believe ‘of’ your book to appreciate it a lot.’
 d. Credo, il tuo libro, di apprezzarlo molto.
 ‘I believe, your book, ‘of’ to appreciate it a lot.’

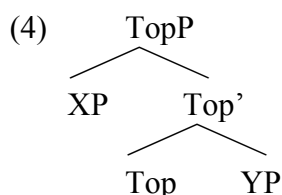
(3a-d) illustrates the distribution of the left-dislocated phrase *il tuo libro* ‘your book’ with respect to the finite complementizer *che* and the non-finite complementizer *di*. The distribution of left-dislocated phrases with respect to these two elements in Italian differs in that they always follow the complementizer *che* (see (3a-b)), but precede the complementizer *di* (see (3c-d)). Rizzi takes such distribution to be one of the arguments against the existence of a single C head and argues that the patterning can easily be explained by the schema in (2) where *che* and *di* are reanalyzed as heads of ForceP and FinP, respectively.⁶ This, coupled with the assumption that dislocated phrases like *il tuo libro* occupy topic positions within the CP field, explains the patterning of such phrases in (3a-d).

⁶ On a more general note, Rizzi (1997) maintains that ForceP hosts relative operators and FinP is activated by finiteness. As noted in fn. 4, I leave aside the details of the properties associated with these two heads as this is not directly relevant to the discussion here.

The positing of a more refined CP system has prompted a great number of studies and investigations into the mapping of various structures. Although undoubtedly of interest, I put aside these issues and focus on one aspect of Rizzi's proposal relevant for our purposes: the existence of topics within the CP domain.

Topics, for Rizzi, are preposed elements at the left periphery, characterized by a comma intonation. They express old information, i.e. information that is salient and available from previous discourse. Topicalized expressions enter into what is traditionally referred to as a topic-comment articulation.⁷ This means that a sentence that contains a topic expression can be separated into two parts: a topic, which is a constituent, and a comment, which is the part of the sentence that introduces new information about the topic. It is commonly assumed that the comment need not be a constituent (e.g. in cases where the topic is the direct object, the comment will consist of the subject and the verb).

Rizzi argues that topics are functional heads that project their own maximal projections within the left periphery (see (2) above). Their syntactic representation is as in (4) (with XP being the topic expression and YP being the comment):



⁷ In addition, Rizzi (1997) discusses a related construction called focus-presupposition, which also makes use of positions in the left periphery. Topic and focus differ in their interpretation, in that focus necessarily introduces new information. Rizzi notes that the focus-presupposition articulation in Italian contains a preposed element, which, unlike the topic, is not coindexed with any clitic pronouns. The focused element, he further notes, carries a special focal stress. An example of this is given in (i), from Rizzi (1997:286):

(i) IL TUO LIBRO ho letto (, non il suo).
 'Your book I read (, not his).'

Rizzi takes the topic-comment articulation in Italian to be represented through CLLD (as in Cinque 1990). That means that he explicitly associates the CLLD construction in Italian with the existence of a functional projection TopP in the CP domain. An example of a topic construction from Rizzi (1997:286) is given in (5).

(5) Il tuo libro, lo ho comprato.

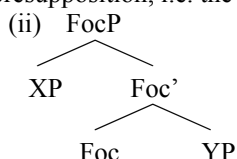
‘Your book, I bought it.’

Topics, Rizzi maintains, are recursive categories.⁸ This captures the fact that a sentence can contain more than one topic, as illustrated by (6) from Italian (from Rizzi 1997:290):

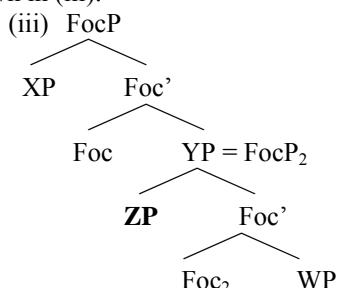
⁸ In contrast to this, focus phrases are unique. (i) shows that multiple focalization of elements in Italian is excluded (from Rizzi 1997:290):

- (i) *A GIANNI IL LIBRO darò (non a Piero, l’articolo).
 ‘TO JOHN THE BOOK I’ll give, not to Piero, the article.’

Rizzi speculates that the nonrecursiveness of the focus projection is tied to their interpretation. The structure of the focus construction is given in (ii). XP is the focal element, and YP is the part that contains the presupposition, i.e. the given information.



Rizzi notes that if FocP is recursive, then YP could also be realized as a FocP. In such cases, however, a clash in interpretation would result: the YP in a FocP only specifies given information, but the specifier position in a FocP introduces new information. In case of recursion, these two positions would overlap, as shown in (iii):



Such clash in TopP, Rizzi notes, does not arise because “nothing excludes a comment (the complement of the topic head) may be articulated in turn as a topic-comment structure, so the topic phrase can undergo free recursion” (Rizzi 1997:297). If this kind of reasoning is correct, Rizzi notes, it would be possible to retain the basic assumption that TopP and FocP are structurally identical, and that their differences with respect to recursion arise from their differences in interpretation.

(6) Il libro, a Gianni, domani, glielo darò senz'altro.⁹

‘The book, to John, tomorrow, I’ll give it to him for sure.’

Rizzi also assumes that topics are present in the structure only when needed. In other words, topics are subject to economy principles such as (7) below (from Rizzi (1997: 314).

(7) Avoid Structure

It should be noted that Rizzi assumes a non-movement analysis of topics in Italian (on a par with Cinque 1990).¹⁰ He does, however, note that movement of a constituent to the specifier position of TopP is possible. Movement in such cases would either be due to a satisfaction of a (topic) criterion or feature checking (as in Chomsky 1993). In case of the latter, an XP would have to have topic features which would be checked under Spec-Head agreement once the XP moves to SpecTopP. As I have argued in chapter 3, direct objects in Macedonian can either move or be base-generated in their clause-initial position. In cases where movement takes place, we can assume that such movement is motivated by the need for the DP to check its topic features.

Rizzi’s proposal regarding the positioning of topic phrases in Italian can be directly applied to the analysis of CLLD in Macedonian. As already indicated in chapter 3, the CLLDed element in such constructions would occupy the specifier position of a topic

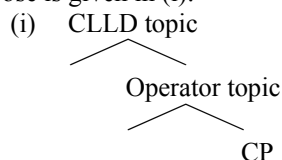
⁹ Both *il libro* and *a Gianni* are doubled by the “compound” pronoun *glie* ‘to him’ + *lo* ‘it’.

¹⁰ More specifically, topic constructions for Rizzi (1997) involve binding of a null constant by an anaphoric operator (in the sense of Lasnik & Stowell (1991)). In other words, the empty category in the argument position in examples like (5) is not a trace. Focus constructions as in (i) (fn. 7), are assumed to involve binding of a syntactic variable in argument position (following Cinque 1990 and others).

phrase. Moreover, since such topic positions can be made available to elements that move to the left periphery, the analysis can be extended to the cases of Topicalization in Macedonian as well. For this analysis to work, however, we would have to make one crucial assumption and that is that the two types of constituents occupy different topic positions. This in turn would mean that different topic projections would be associated with different properties (a possibility to which Alexiadou (2006:675) also alludes).¹¹ As I show in section 4.3, such interpretative differences between CLLD and Topicalization in Macedonian exist.

Having outlined the theory which supports an analysis of CLLD in terms of topic phrases within the CP domain, I would now like to note that an alternative analysis of such constructions is also available. Various studies have analyzed CLLD in terms of adjunction (either IP adjunction, as in Cinque 1990, Baker 1996, for example, or CP adjunction, as in Iatridou 1995).¹² The main argument that has been presented in support of this kind of an analysis comes from the fact that CLLD allows for multiple dislocation (see Cinque 1990, among many others). The evidence for this is robust and the property

¹¹ My view of CLLD and Topicalization as occupying different topic positions within the CP domain is in line with Krapova (2002) and Krapova & Cinque (2003), who argue for the existence of two different topic positions within the CP domain. It should be noted that their proposal is made in the context of *wh*-questions, and in particular with reference to D-linked and non-D-linked *wh*-phrases. D-linked *wh*'s in Bulgarian can optionally occur with a clitic pronoun. Krapova and Krapova & Cinque argue that D-linked *wh*'s which cooccur with a clitic occupy the position of a CLLD topic (on a par with CLLD constructions), while D-linked *wh*'s without a clitic occupy the position of Operator topic. (They also note that an adjunction analysis of CLLD constructions would work just as well. I return to this issue shortly.) The structure they propose is given in (i):



¹² For detailed implementation of this kind of analyses, see Iatridou (1995) (who analyzes CLLD in Greek as an adjunction to CP), Alexiadou & Anagnostopoulou (1997) (who analyze CLLD in Greek as an adjunction to IP), Baker (1996) (who takes CLLD to involve adjunction to IP), etc..

has been shown to hold of CLLD across languages. An analysis of the construction in terms of multiple adjunction naturally captures this fact.

The fact that CLLD constructions can be analyzed in terms of an adjunction opens up the possibility that the two constructions, CLLD and Topicalization, may have different structural representations, with the latter analyzed as movement to SpecCP, as Baker (1996) suggests, or movement to the specifier of some other functional projection.

As already noted, the main argument for the proponents of the adjunction analysis of CLLD is the fact that these constructions allow for multiple dislocation. Although this is a strong argument in favor of such an analysis, it is not a definitive argument against an analysis of CLLD in terms of TopP projections. Recall from our earlier discussion that Rizzi (1997) explicitly assumes that topics are recursive categories. Given this, multiple CLLDed elements would be analyzed in terms of multiple TopP projections. Topicalization, on the other hand, would require a single maximal projection.

To summarize, I have discussed two possibilities for the analysis of clause-initial elements in Macedonian. One option would be to analyze CLLDed and Topicalized elements as occupying the specifier positions of distinct topic projections in the CP domain, following Rizzi (1997). In light of the fact that CLLD allows for multiple dislocation, a second option would be to analyze such constructions as multiple adjunction (to CP or IP) and treat Topicalization as a process which targets the specifier of some functional projection. Given that both options work well in accounting for the fact that CLLD allows for multiple dislocation, I remain at this point neutral as to which of these structures applies to Macedonian. For consistency sake, I will continue to represent CLLD and Topicalization as involving TopP projections in the left periphery.

4.1.2 HTLD and CLLD

In the context of our discussion of the left periphery, I would next like to discuss another type of dislocated structure in Macedonian, Hanging Topic Left Dislocation (HTLD), with the intention of distinguishing it from CLLD. As we will see, HTLD also involves the placement of a constituent in clause-initial position, but has distinct properties from CLLD.

The properties of HTLD have been widely discussed in the literature. This type of dislocation is characterized by the presence of a constituent in clause initial position “which is connected with that clause through the intermediary of some anaphoric element referred to as the *resumptive* element” as in the English *John, I like him* (Alexiadou 2006:668). The very definition is reminiscent of the definition of CLLD. The two constructions, however, have different properties. (The list of features of HTLD against which I draw the comparison with CLLD is drawn from Cinque (1990), Anagnostopoulou (1997), Arnaudova (2002), Alexiadou (2006), Krapova & Cinque (2008), Vat (1997), and references cited therein.)

Examples of HTLD in Macedonian are given in (8a-b). The dislocated element *Petar* in (8a-b) is outside the CP and it is syntactically connected with that clause through the pronouns *toj* ‘he’ and *nego* ‘him’, respectively. Crucially though, the pronouns used in (8a-b) appear in their long form.¹³ This is especially evident in (8b), where the direct object *nego* ‘him’ is contrasted with the short form of the same pronoun, *go* ‘him’.¹⁴ The

¹³ See chapter 2 for the full paradigm of pronouns in Macedonian.

¹⁴ Only pronouns in accusative and dative distinguish between long and short forms: *nego*, *go* ‘he-ACC’, *nemu*, *mu* ‘he-DAT’. Thus, the pronoun *toj* ‘he’ does not have a corresponding short form. See chapter 2; see also Friedman (1993).

use of the long vs. short pronominal forms as resumptive elements is the first feature that crucially separates the HTLD from the CLLD constructions in Macedonian.

(8) a. Petar, toj e eden mnogu glupav čovek.

Petar he is one very stupid man

‘As for/speaking of Petar, he is a very stupid man.’

b. Petar, nego včera go vidovme na kino.

Petar him yesterday him saw at movies

‘As for/speaking of Petar, we saw him at the movies.’

In addition to the use of the long form of the pronouns, HTLD constructions are characterized with a sharper intonational break between the dislocated element and the rest of the clause than one finds in CLLD.

HTLD does not allow more than one dislocated element, as shown in (9a). Multiple CLLDed elements, on the other hand, are perfectly acceptable in Macedonian, as shown in (9b).

(9) a. *Knigata, na Petar, Ana nemu nea mu ja dade.

book-the to Petar Ana him(Dat) it(f.sg.) him it(f.sg.) gave

‘The book, to Petar, Ana gave it to him’

b. Knigata, na Petar, Ana mu ja dade.

book-the to Petar Ana him(Dat) it(f.sg.) gave

‘The book, to Petar, Ana gave it to him’

Also significant is the fact that the HTLD construction is not allowed in embedded contexts, as illustrated by the ungrammaticality of (10a). Again, we see that CLLD constructions are not restricted to the matrix clause, as shown in (10b).

- (10) a. *Znam deka Marija_i, Petar nea_j ne ja_j vikna na zabava.
 know-I that Marija she not her not her invite to party
 ‘I know that Marija, Petar didn’t invite her to the party.’
- b. Znam deka Marija_i, Petar ne ja_j vikna na zabava.
 know-I that Marija Petar not her invite to party
 ‘I know that Marija, Petar didn’t invite her to the party.’

It has been claimed that HTLD only allows for the dislocation of NPs (see Cinque 1990 and others). The facts from Macedonian, however, are slightly different here. Some cases of PP dislocation in Macedonian may be treated as instances of HTLD. Constructions like that in (11) are possible candidates, where the clause initial PP *vo London* ‘in London’ is associated with the pronoun *tamu* ‘there’.^{15, 16}

¹⁵ The PP in (11) is an adjunct. A left dislocation of an indirect object, on the other hand, is not allowed, as shown in (i).

(i) *Na Petar_i, nemu_i mu_i ja_j dadov knigata_j.
 to Petar him him it(f.sg) gave book-the
 ‘To Petar, I gave the book to him.’

Note that in English Left Dislocation, the dislocation of PPs is also not allowed, as shown in (ii) (from Alexiadou 2006:671):

(ii) *To John, I have already spoken to him.

¹⁶ As Mark Baker has pointed out to me, since the pronoun *tamu* ‘there’ in (11) is optional, the same example without the pronoun would be a case of Topicalization; see also section 4.1.3.

(11) [PP Vo London], (tamu) ušte ne sum bila.

in London there yet not have been

‘I haven’t been to London yet.’

It has been observed that quantifiers in general are marginal in HTLD (see Vat 1997).¹⁷

As (12a-c) show, the only type of quantifiers that is allowed in HTLD constructions in Macedonian are the numerals, as in (12a). HTLD with *many* as in (12b) and strong DPs as in (12c) are ruled out. This is in contrast to CLLD, which as we already know, allows for both strong DPs and numerals.

(12) a. ?Tri studentki, niv včera gi vidovme na kino.

three students(f.) them yesterday them saw at movies

‘As for/speaking of three students, we saw them at the movies.’

b. *Mnogu studenti, niv včera gi vidovme na kino.

many students them yesterday them saw at movies

‘As for/speaking of many students, we saw them at the movies.’

c. *Povekjeto studenti, niv včera gi vidovme na kino.

most students them yesterday them saw at movies

‘As for/speaking of most students, we saw them at the movies.’

It has been pointed out that HTLD is not subject to locality conditions, i.e. it can violate strong islands (Alexiadou 2006). This means that the HTLDED expression and the

¹⁷ Vat (1997) gives examples of HTLD in Dutch with indefinites, numerals, and universal. Of those, HTLD with the first two types if DPs are judged marginally acceptable; HTLD with a universal are ungrammatical.

pronoun in the IP do not enter into a chain relation: the hanging topic relates to some referent in the discourse and there is a loose relation between it and the pronoun with which it is coindexed. An example is given in (13), where the dislocated DP *Marija* is coindexed with the pronoun in the relative clause. As we will see in section 4.2.3, this is in contrast to CLLD where such coindexation between a CLLDed element and a pronoun is not possible.

(13) *Marija_i, go poznavam čovekot što nea_ija_i pobara za žena.*

Marija him know person-the that her her asked for wife

‘As for Marija, I know the man that asked to marry her.’

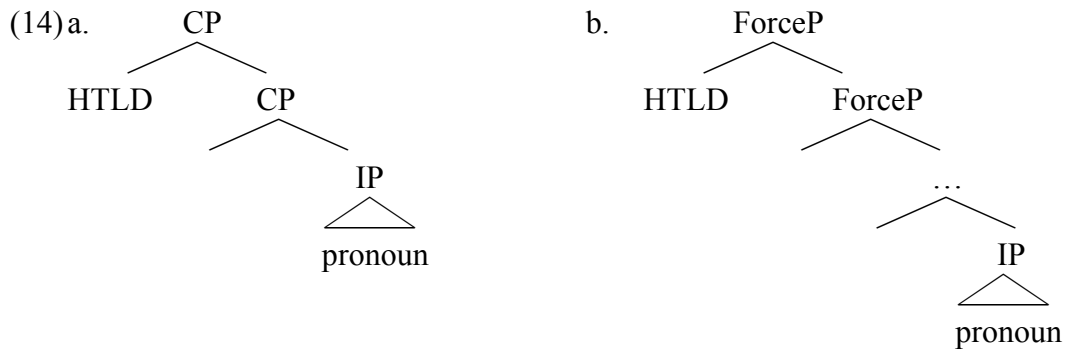
The behavior of HTLDed elements with respect to syntactic islands can perhaps explain the differences in the distribution of quantifiers in such constructions. Suppose that one of the conditions on HTLD is for the dislocated element to be referential (in the sense of Cinque 1990). If this speculation is correct, then only referential expressions and weak DPs that allow for a referential reading would be able to participate in such constructions. Weak DPs that do not allow for a referential reading, like *many* in (12b), would be ruled out, as would strong DPs, like *most* in (12c).

Based on the descriptive properties of HTLD, and especially the fact that they can only occur in matrix clauses, an argument can be made that HTLD must utilize a position that defines the left edge of the clause, i.e. left of ForceP (in Rizzi’s framework)¹⁸, while

¹⁸ FocP, the only nonrecursive category in the CP field in Rizzi (1997), must be ruled out as a candidate for obvious reasons: it is incompatible with HTLD because of its interpretation and it is also available in embedded contexts.

CLLD, given its ability to occur in embedded clauses, would utilize structures that are within the left periphery and are crucially recursive.

To capture the fact that HTLD is restricted to matrix contexts, we can follow Alexiadou (2006:674) and posit that HTLD in Macedonian has the structure in (14a) (or (14b), following Rizzi (1997)).¹⁹



To summarize, I have shown that HTLD in Macedonian is characterized by a distinct set of properties. This enables us to set the construction aside from here on in our discussion of the two constructions that are the focus of this dissertation, CLLD and Topicalization.

4.1.3 CLLD and Topicalization

Let us turn our attention to CLLD and Topicalization. In what follows I outline the similarities and differences between these two constructions.²⁰

¹⁹ There is an obvious drawback to this analysis in that adjunction, as we saw earlier, allows for multiple dislocation, which HTLD does not. At the same time, this kind of analysis is necessary given that HTLD must invoke a projection that is not available in embedded contexts. Given that our goal here is a modest one (to distinguish the construction from CLLD), I will adopt this analysis of HTLD, leaving these issues open for further research.

²⁰ The fact that the two constructions share certain features was observed by Cinque (1990); he also points out that topicalization involves the fronting of a single constituent, while CLLD can have multiple dislocated elements.

Aside from the fairly trivial fact that the dislocated element in CLLD and TOP are followed by a slight pause, both constructions are found in matrix and embedded contexts, as shown in (15a-b), respectively.

- (15) a. Znam deka sekoja kniga, Petar ja pročita.
 know(1sg.Pres) that every book Petar it(f.sg.) read
 ‘I know that every book, Petar read it.’
- b. Znam deka dve knigi, Petar gi pročita.
 know(1sg.Pres) that two books Petar them read
 ‘I know that two books, Petar read them.’
- c. Znam deka dve knigi, Petar pročita.
 know(1sg.Pres) that two books Petar read
 ‘I know that two books, Petar read.’

However, fronted elements in TOP (but not CLLD) can be of categories other than DP, as illustrated in (16) (based on examples in Cinque 1990).^{21, 22}

- (16) a. [_{AP} Ubava], taa nikogaš ne bila.
 beautiful she never not been
 ‘She’s never been beautiful.’

²¹ The only type of PPs that can be CLLDed are indirect objects. The PP *na Pero* ‘to Pero’ is coindexed with the clitic pronoun *mu* ‘him’.

(i) Na Pero_i, Marija odamna ne mu_i ima pišano.
 to Pero Marija long time not him has written
 ‘Marija has not written to Pero in a long time.’

²² As pointed out in fn. 16, the clause initial PP in (11) can also be taken as an instance of Topicalization when the PP is not coindexed with the pronoun *tamu* ‘there’ in the IP.

- b. [CP Deka pieš], site znaat.
 that drinks all know
 ‘Everyone knows that you drink.’

It should be noted that some languages, like Italian, allow for the dislocation of APs and CPs in CLLD (Cinque 1990:57-58); Macedonian seems to differ in this respect.²³

CLLD allows for multiple dislocation, as illustrated in (17a). In (17a) both the direct object *knigata* ‘the book’ and the indirect object *na Dino* ‘to Dino’ can be CLLDed.²⁴ TOP, on the other hand, does not allow for multiple dislocation, as shown by the ungrammaticality of (17b).²⁵

- (17) a. Knigata_i, na Dino_j, Eva mu_j ja_i dade.
 book-the to Dino Eva him it(f.sg.) gave
 ‘The book, to Dino, Eva gave it to him.’
- b. *Dve knigi, nekomu, Eva dade.
 two books someone(Dat) Eva gave
 ‘Two books, to someone, Eva gave.’

In light of their differences, and in particular, the ability of CLLD but not Topicalization to accommodate multiple dislocation, I suggest that the CLLDed and topicalized

²³ Cinque (1990:71) notes that the clitic in such cases is optional. He attributes the optionality to the syntactic category of the CLLDed element. Such elements, or rather their associated EC (empty category) in argument position in the IP “are not required to qualify as either PRO, pro, NP-trace, or variable (because they are not partitioned by the features [\pm pronominal, \pm anaphor]”. Direct objects, on the other hand, are necessarily coindexed with a *pro* in the argument position.

²⁴ As pointed out in fn. 21, indirect objects in Macedonian can be CLLDed.

²⁵ The same hold for CLLD and TOP in Italian; see Cinque (1990:63,73).

elements are allocated to different positions at the left periphery. Given the recursive properties of topic phrases, CLLD can easily be allocated to such functional projections. We can also assume that a dislocation in Topicalization triggers the projection of a single (and distinct) TopP in the CP domain. In Rizzi's framework they would both occur below ForceP.

4.2 Base-Generation vs. Movement

Having determined the positions which CLLDed and topicalized elements target in the left peripheral domain, let us turn to the relationship these two types of dislocated elements have to the positions within IP with which they are associated. The main goal here is to show that CLLD and Topicalization constructions are derivationally distinct. To that end, I will give three arguments for the claim that CLLD involves base-generation of an argument at the left periphery with a coindexed *pro* inside IP, while Topicalization is derived through movement of a direct object from IP to the left periphery.

4.2.1 Arguments against Clitic Doubling as the Source for CLLD

There is an obvious connection between clitic doubling (CL-D) and CLLD which is reflected in the fact that both constructions involve “doubling” of a constituent with a (clitic) pronoun. Therefore, it is very natural to assume that one construction is derived from the other, and in particular, that CLLD is derived from CL-D by movement of the doubled constituent to a clause initial position. On this view, the derivation of the CLLD construction in (18a) from the CL-D construction in (18b) would be as in (18c).

- (18)a. Filmot, Dino go gledaše.
 film-the Dino it(m.sg.) saw
 ‘Dino saw the film.’
- b. Dino go gledaše filmot.
 Dino it(m.sg.) saw film-the
 ‘Dino saw the film.’
- c. $[_{XP}[_{DP} \text{filmot}]_i [_{CP}[_{IP} \text{Dino } t_i \text{ go}_i \text{ gledaše } t_i]]]$

In (18c), the DP *filmot* ‘the film’ originates in the argument positions of the verb and subsequently is fronted to a position outside the clause. In accordance with the Clitic Criterion, the movement also leaves a trace in an intermediate position, in SpecCP.

Although some studies have argued for a movement analysis of CLLD from CL-D, others have argued against it.²⁶ Based on the facts from Macedonian, our proposal falls in line with the latter group. Below I present some of the arguments that have been put forth in support of this view. The discussion is based on Cinque (1990), Iatridou (1995), Arnaudova (2002), Anagnostopoulou (2006), and Alexiadou (2006).

One of the reasons that has been given against an analysis of CLLD in terms of CL-D is the fact that some languages are known to have CLLD, but not CL-D (Cinque 1990). Since CL-D is not available as a source, the argument can be made that CLLD is an independent construction. One such language is Italian, as the examples in (19a-b) show (from Cinque (1990:60-61)).

²⁶ For proponents of the movement analysis, see for example Sportiche (1998), Philippaki-Warbuton et al. (2002), among others. (The actual implementation and analyses differ from author to author.). For arguments against a movement analysis of CLLD, see Cinque (1990), Iatridou (1995), Anagnostopoulou (2006) and Alexiadou (2006), etc..

(19) a. *Lo conosciamo Gianni.

him-(we-)know Gianni

b. Gianni, lo conosciamo.

Gianni we know him

It is worth noting, however, that this particular asymmetry could be explained by assuming that CLLD in such languages is a direct result of an obligatory movement of a clitic-doubled DP. As Sportiche (1998) himself notes, CLLD might arise as a result of the inability of a language to have ‘true’ clitic doubling.

Another reason that has been given for an independent analysis of CLLD is the observation that CL-D and CLLD seem to target different types of maximal categories. As Anagnostopoulou (2006) notes, the cases of CL-D discussed in the literature seem to be restricted to DPs, while CLLD allows for the doubling of different types of phrases. Some examples of CLLD from Italian are given in (20a-b). (20a) contains a left-dislocated PP; (20b) contains a left-dislocated AP (from Alexiadou (2006:677)).²⁷

(20) a. A casa, non ci sono stato ancora.

to home not there am been yet

‘I haven’t been home yet.’

b. Bella, pare che non lo sia mai stata.

beautifulseems that not it is everbeen

‘It doesn’t seem that she has ever been beautiful.’

²⁷ In addition to PPs and APs, Italian allows for the CLLD of CPs; see Cinque (1990).

Finally, various studies have noted that in some languages weak quantifiers behave differently in CLLD and clitic doubling constructions in the sense that such elements can appear in CLLD constructions but they cannot be clitic-doubled. Examples from Greek are given in (21a-b) (from Iatridou 1995).²⁸

- (21) a. Tria provlimata mono o Kostas ta elise.
 three problems only the Kostas them solved
 ‘Only Kostas solved three problems.’
- b. *Mono o Kostas ta elise tria provlemata
 only the Kostas them solved three problems

Turning to Macedonian, we see that not all of the above arguments apply. To begin with, Macedonian has clitic doubling. This means that the construction can potentially be used as basis for a movement analysis of CLLD.²⁹

The second argument is also not applicable because CLLD of prepositional and adjectival phrases as in (20a-b) in Macedonian is not possible.

The third argument is applicable to Macedonian and provides the strongest evidence against conflating the clitic doubling and CLLD constructions. As we know already, the

²⁸ At the same time, strong DPs in such languages seem to be accepted in both types of constructions. That this holds for Greek is evident from examples like (i) and (ii), from Kallulli (1999:24, 57). (i)-(ii) are Kallulli’s (11) and (60), respectively. The clitic *ta* ‘them’ in (i) doubles the object *ola ta aghorja* ‘all the boys’; the clitic *ta* ‘them’ in (ii) doubles the clause-initial DP *ta luludha* ‘the floors’.

(i) I Anna ta misuse ola ta aghorja.
 the Ann them_{cl} HATED all the boys
 ‘Anna HATED all the boys.’

(ii) Ta luludhia_{acc} *(ta) efere o Jannis_{nom}.
 the flowers them_{cl, acc} brought John
 ‘As for the flowers, it was John who brought them.’

²⁹ It is important to note, however, that the very existence of CL-D in Macedonian is an argument against the view that takes CLLD to be a result of an *obligatory* movement of a clitic-doubled element, as in Sportiche (1998).

facts in (21) for Greek also hold for Macedonian: unspecified weak DPs cannot be clitic-doubled when they occur in IP internal positions, but they can co-occur with a clitic when they are in clause initial position (see chapters 2 and 3). The existence of true clitic doubling as distinct from CLLD, as well as the fact that weak DPs at the left periphery selectively occur with clitics, seem to be the strongest evidence against the view that CLLD in Macedonian is an instance of clitic doubling.

4.2.2 Weak Crossover Effects

The second type of evidence for a non-movement analysis of CLLD comes from weak crossover effects (WCO).

WCO violations arise in configurations where an operator binds a pronoun and a variable, neither of which c-commands the other. This is shown in (22a-b) for English and (23a-b) for Macedonian. Both cases are ruled out as a WCO violation because the variable t_i is coindexed with a pronoun *mu* ‘his’ to its left.³⁰


(22)a. *His_i mother loves everyone_i.

b. [everyone_i] [[his_i mother] loves t_i]

(23)a. *[Majka mu_i] go_i saka [sekoe dete]_i.


mother his him loves every child

‘His mother loves every child.’

b. [sekoe dete]_i [[majka mu_i] t_i go_i saka t_i]


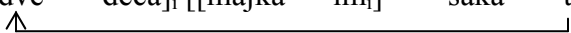
³⁰ The WCO effects in (22a-b) and (23a-b) follow from any proposal on WCO (for e.g., Koopman & Sportiche’s (1983) Bijection Principle, Lasnik & Stowell (1991), Safir (1996), etc.).

As has been noted by many (see Iatridou 1995, Cinque 1990, Rizzi 1997, Alexiadou 2006), CLLD constructions do not give rise to WCO effects.³¹ Macedonian is no exception to this. Thus, when we consider the universal in clause-initial position, as in (24a), the unacceptability seen in (23a) disappears.³² This is illustrated in (24a-b) below.

- (24)a. [Sekoe dete]_i, [majka mu]_i go_i saka.
 every child mother his him loves
 ‘Every child, his mother loves him.’
- b. [sekoe dete]_i [[majka mu]_i *pro*_i go_i saka t_i]


This is consistent with our proposal that CLLD is not derived by movement.

We can apply the same test to weak DPs. Weak DPs in clause-initial position without a clitic show WCO effects, as shown by the numeral in (25a-b). On the other hand, clause-initial weak DPs coindexed with a clitic, as the numeral in (26a-b), do not give rise to WCO effects.

- (25)a. *[Dve deca]_i, [majka im]_i saka.
 two children mother theirs loves
 ‘Two children, their mother loves them.’
- b. [dve deca]_i [[majka im]_i saka t_i]


³¹ Most of the literature bases this claim on the patterning of proper names. Iatridou (1995) gives examples of suspension of WCO effects with *each* in CLLD constructions in Greek.


³² On Koopman & Sportiche’s (1983) analysis, for example, the lack of WCO effects in (24) would follow from the fact that the trace and the pronoun are not bound by the same argument, as they would in (23a-b).

(26) a. [Dve deca]_i, [majka im]_i gi_i saka.

two children mother theirs them loves

‘Two children, their mother loves them.’

b. [dve deca]_i [[majka im]_i *pro*_i gi_i saka t_i]



As was the case with the universal in (24a-b), WCO effects do not arise in (26a-b) because the DP in clause-initial position does not bind a trace inside IP.

Our analysis of [-strong] DPs in clause-initial position, which treats them as instances of topicalization, predicts that they would give rise to WCO effects. The prediction seems to hold, as shown in (27).

(27) *[Mnogu deca]_i, [majka im]_i saka.

many children mother theirs loves

‘Many children, their mother loves them.’

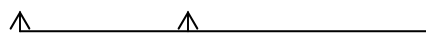
The correlation between movement and the presence of WCO violations is further replicated in the patterning of preverbal direct objects. Consider (28a-b) and (29a-b), for example.

(28) a. *[Majka mu]_i [sekoe dete]_i go_i saka.

mother his every child him loves

‘His mother loves every child.’

b. [[majka mu]_i [sekoe dete]_i t_i go_i saka t_i]



(29) a. *[Majka im_i] [dvedeca]_i saka.

Mother theirs two children loves

‘Two children, their mother loves.’

b. [[majka im_i] [dve deca]_i saka t_i]



(28a-b) and (29a-b) are ungrammatical under the bound variable construal due to the fact that the argument coindexed with the pronoun does not c-command the pronoun. In order for the DP to bind the pronoun, the DP would have to move covertly to a position c-commanding the pronoun to its left, leading to a WCO violation. Preverbal direct objects thus confirm that DPs that undergo movement consistently give rise to WCO effects.

To summarize, WCO tests show that CLLD has to be recognized as distinct from CL-D and furthermore that such construction is to be distinguished derivationally from Topicalization on the basis of base-generation vs. movement. Thus, we can conclude that CLLDed constituents originate in their clause-initial position, but topicalized elements move to the left periphery. This is shown most clearly in the case of unspecified DPs, where both options are available. It is also supported by strong DPs, which do not show WCO violations in clause initial position, showing that the possibility of base-generation is available to them.³³

4.2.3 Island Effects

The third type of evidence comes from the sensitivity of direct objects at the left periphery to syntactic islands. Under the proposed analysis, CLLD involves a binding

³³ We have no evidence to show whether a movement option is or is not available to them. In the absence of such evidence, I have assumed that it is.

chain while Topicalization involves a government chain, in terms of Cinque (1990).³⁴ One of the most striking properties of the two types of chains is the fact that they are both sensitive to strong islands, but that only government chains are sensitive to weak islands. For illustration, in (30)-(31), I give examples of both types of islands from English (from Cinque 1990:1).³⁵

(30) *Adjunct island*

- a. *To whom did you leave without speaking *t*?
- b. *How was he fired after behaving *t*?

(31) *Wh-Island*

- a. ??To whom didn't they know when to give their present *t*?
- b. *How did they ask you who behaved *t*?

The fact that binding chains are constrained by strong islands, but can violate weak ones, is evident in the examples in (32)-(33) from Alexiadou (2006:682) for Greek.³⁶ (32a-b)

³⁴ For Cinque (1990) the properties of the chains stem from the (non)referential properties of the elements that participate in them. Only referential phrases can form binding chains because they are able to connect to the gap in argument position through a binding relation (the latter is achieved through coindexation of the EC with their antecedent). Nonreferential phrases, on the other hand, Cinque maintains, identify the EC in object position as a variable, and connect to it only through a government chain. Referential phrases for Cinque (1990) are those which are D-linked (in the sense of Pesetsky 1987). As such, they refer to members of a set previously introduced in the discourse.

While Cinque's distinction between the two types of chains and their sensitivity to the two types of islands is applicable to Macedonia, his formulation of CLLD exclusively in terms of referentiality does not seem to hold. Strong DPs, such as the universals, are nonreferential on Cinque's tests (e.g. they have to c-command their pronoun in order to establish co-reference; see chapter 2), but they can be fronted to a clause initial position. This means that they can be part of a CLLD construction.

³⁵ *Whom* is an argument and as such can enter into a binding chain, while *how* is nonreferential and therefore can enter only in a government chain.

³⁶ The same effects are observed in Bulgarian, as argued by Arnaudova (2002), who shows that CLLD constructions show selective island sensitivity in that they violate weak islands, but obey strong ones. Other

are examples of strong islands created by relative and adjunct clauses; (33) is an example of a wh-island, which is a weak island.

- (32) a. *Ton Kosta sinandisa tin kopela pu ton ide.
 the Kostas met-1sg the girl that him saw-3
- b. *Tin efimerida apokimithike diavazondas tin.
 the paper fell-3sg asleep reading it

- (33) To forema den ksero pu na to valo.
 the dress neg know where subjt it put
 ‘As for the dress I do not know where to put it.’

As (32a-b) show, the fronting of the definites out of the relative and adjunct clause, respectively, results in ungrammaticality. At the same time, the fronting of the definite out of the wh-island is acceptable, as shown in (33).

We see the same effects in Macedonian. Consider the examples in (34a-b). The ungrammaticality of (34b) can be explained if we assume that topicalized phrases in Macedonian can only connect to the trace in object position via a chain of antecedent government.³⁷

languages go even further: CLLD constructions in Lebanese Arabic seem to violate strong islands (e.g. the Complex NP Constraint or the Adjunct island); see Aoun & Benmamoun (1998) for details.

³⁷ That the trace in such cases is a variable was proved by the presence of WCO effects in such constructions; see (25a).

(34) a. Sekoj student, se prašuvam kako kje go najdeš.³⁸

every students refl. wonder-I how will him find

‘Every student, I wonder how you will find him.’

b. *Nekogo, se prašuvam kako kje najdeš.

someone refl. wonder-I how will find

‘Someone, I wonder how you will find.’

Since the dislocation of the universal in (34a) does not result in a weak island violation, we can conclude that the relation between the clause initial element and *pro* in object position is one of a binding chain. (34b), on the other hand, results in a weak island violation, showing that a topicalized element and the trace in object position are part of an antecedent government chain. It should be noted that (34b), which has a DP unspecified for strength, can occur with a clitic in which case the grammaticality is repaired.

We see the same pattern emerging in (35a-b). A chain created by a clitic-doubled numeral as in (35a) does not create a weak island violation. On the other hand, the chain relation established by a fronted numeral not coindexed with a clitic, as in (35b), results in a weak island violation.

³⁸ That this is true of other strong DPs, like *most*, is evident from examples like (i) below where we see that such DPs freely violate weak islands.

(i) Povekjeto knigi, se prašuvam kako kje gi najdeš.
 most books refl. wonder-I how will them find
 ‘Most books, I wonder how you will find them.’

(35) a. Dve knigi, se prašuvam kako kje gi najdeš.

two books refl. wonder-I how will them find

‘Two books, I wonder how you will find them.’

b. ??/*Dve knigi, se prašuvam kako kje najdeš.

two books refl. wonder-I how will find

‘Two books, I wonder how you will find them.’

On the proposed analysis, direct objects that are [-strong] are predicted to form government chains only. The prediction is borne out as shown by the ungrammaticality of (36), where the fronting of the DP results in a weak island violation.

(36) *Mnogu knigi, se prašuvam kako kje najdeš.

many books refl. wonder-I how will find

‘Many books, I wonder how you will find them.’

We mentioned that both binding chains and government chains obey strong islands. That this is the case is shown in (37a-c)³⁹:

(37) a. *Petar, go ispiv kafeto pred da go viknam.

Petar it drank-I coffee-the before to him call

‘Petar, I drank the coffee before I called him.’

³⁹ The examples in (37a-c) and (38a-c) are fashioned after Baker (1996:104).

- b. *Site mački, plačev zatoa što Petar gi istepa.
 all cats cried-I because that Petar them beat
 ‘All cats, I cried because Petar beat them.’
- c. *Dve mački, plačev zatoa što Petar gi istepa.
 two cats cried-I because that Petar them beat
 ‘Two cats, I cried because Petar beat them.’

The ungrammaticality of (37a-c) is due to the fact that the dislocated phrases, whether strong as in (37a-b) or weak as in (37c), cannot enter into a chain with a pronoun inside an adjunct modifier clause which is a strong island. The same effects are present in relative clauses. Here, too, the binding relation between the dislocated phrase and a pronoun inside the relative clause is disrupted. This explains the ungrammaticality of both (38a-b) for strong DPs, and (38c) for weak DPs.

- (38)a. *Topkava, go izbrka kučeto što ja ukrade.
 ball-the it chased dog-the that it(f.sg.) stole
 ‘The ball, you chased the dog that stole it.’
- b. *Košnicava, go poznavam čovekot što ja isplete.
 basket-this him know-I man-the that it(f.sg.) wove
 ‘The basket, I know the man who wove it.’
- c. *Dve košnici, go poznavam čovekot što gi isplete.
 two baskets him know-I man-the that them wove
 ‘Two baskets, I know the man who wove them.’

The examples in (34a), (35a), (37)-(38), show is that the relation between a clitic-doubled phrase at the left periphery and a null *pro* in object position is sensitive to the presence of strong islands, but not weak ones. Aside from the fact that such examples show that binding and government chains pattern differently with respect to syntactic islands (cf. (34a), (35a) vs. (34b), (35b), (36)), such examples are in fact, as Baker (1996) points out, evidence that the dislocated phrase and *pro* (in CLLD constructions) are part of a chain to begin with.

To summarize, the behavior of clause initial DPs with respect to syntactic islands shows, as their behavior with respect to WCO did, that topicalization is derived by movement of a DP from an IP internal position to the left periphery, while CLLD instantiates a binding relation between a base-generated DP at the left periphery with a *pro* inside IP.

4.3 Semantic and Pragmatic Differences between CLLD and TOP

In section 2, I argued that the two constructions in which clause-initial direct objects participate differ with respect to their syntactic derivation. In this section, I present evidence from their semantic and pragmatic behavior that further supports the claim that CLLD and TOP are two distinct constructions in Macedonian.

4.3.1 Scope Properties of CLLD and Topicalization

It is a well-known fact that universals and bare numeral quantifiers in object positions in English can either have narrow or wide scope interpretation with respect to a subject

quantifier.⁴⁰ In (39), I give an example of scope interaction between a universal quantifier (*every book*) and a bare numeral (*two students*). The LFs corresponding to the two readings of (39a) are given in (39b) and (39c), respectively.

(39) a. Two students read every book.

b. $[_{IP} \text{two students}_i [_{IP} \text{every book}_j [_{IP} t_i \text{read } t_j]]]$

There are two students, namely John and Mary, such that they read every book.

c. $[_{IP} \text{every book}_j [_{IP} \text{two students}_i [_{IP} t_i \text{read } t_j]]]$

For every book, namely a, b, and c, there are two students such that they read that book. [The set of two students may be different for every book.]

When we look at object quantifiers within IP (i.e. DPs in post- and preverbal positions), we see that Macedonian patterns like English. In (40)-(41) we see that object DPs can either have wide or narrow scope with respect to a subject quantifier.

(40) a. Dvajca studenti ja pročitaa sekoja kniga. [S>O; O>S]

two students it(f.sg.) read every book

‘Two students read every book.’

b. Dvajca studenti sekoja kniga ja pročitaa. [S>O; O>S]

two students every book it(f.sg.) read

‘Two students read every book.’

⁴⁰ It is also well known that not all object quantifiers can have wide scope over a c-commanding quantifier; see for example, Beghelli (1993, 1995), Szabolcsi (1997), etc..

- (41) a. *Sekoj student pročitā dve knigi.* [S>O; O>S]
 every student read two books
 ‘Every student read two books.’
- b. *Sekoj student dve knigi pročitā.* [S>O; O>S]
 every student two books read
 ‘Every student read two books.’

When we turn to object quantifiers in clause initial position we find what appears to be an interesting difference in scope possibilities, as shown in (42a-b). In (42a) the subject *dvajca studenti* ‘two students’ cannot take wide scope over the object *sekoja kniga* ‘every book’; in (42b) the subject *sekoj student* ‘every student’ cannot take scope over the object *dve knigi* ‘two books’. The generalization that emerges from these data is that left-dislocated object DPs take wide scope with respect to the subject in Macedonian.

There is a puzzling exception to this generalization, though, as shown in (42c). In (42c), it is the subject quantifier *sekoj student* ‘every student’, rather than the object quantifier *dve knigi* ‘two books’ in clause initial position, that seems to get wide scope reading.

- (42) a. *Sekoja kniga dvajca studenti *(ja) pročitāa.* [*S>O; O>S]
 every book two students it(f.sg.) read
 ‘Two students read every book.’

- b. Dve knigi, sekoj student gi pročita. [*S>O; O>S]
 two books every student them read
 ‘Every student read two books.’
- c. Dve knigi, sekoj student pročita. [S>O; *O>S]
 two books every student read
 ‘Every student read two books.’

Given what we have seen in (42a-b) and (42c) we might be tempted to conclude that dislocated elements in CLLD only take wide scope in the sentence, while topicalized elements only take narrow scope in the sentence. If this were so, we would have to say that topicalized elements in Macedonian obligatorily reconstruct and as such are always interpreted in their base position. Aside from the problem of finding a principle that would force this, the implementation of this idea faces an obvious problem. We saw that direct objects in their base or focus positions (as in (40) and (41), respectively) allow for both wide and narrow scope readings. It would remain a bit mysterious that reconstructed topicalized DPs should not display the same possibilities. In addition, obligatory reconstruction would be problematic for the universals. Recall that nothing in our analysis rules out movement of the clitic-doubled universal in clause-initial position. If movement feeds reconstruction, then we would allow for two readings of the CLLDed universal in (45a). In other words, once the universal moves to its clause-initial position, it would be able to reconstruct to its argument position, and get a narrow scope reading. This is not what the data shows, however. In view of these considerations, an alternative explanation for the difference between CLLD and TOP is worth exploring.

4.3.2 Available Readings of CLLD and Topicalization

To see more clearly what is going on with the clause-initial DPs, consider the following scenario. The students in the English literature class were given a list of books to read over the holidays. At the beginning of the semester, the professor asks them who read what from the list of books and then tells her colleague of the results of the survey. It turns out that Ana and Elena have read *a* and *b*, Tanja and Petar have read *c*, *d*, and *e*, and Eli and Lidija have read *f* and *g*. The professor's report in this case may go something like this:

- (43) a. Dve knigi, Ana i Elena gi pročitaa.
 two books Ana and Elena them read
 'Two books, Ana and Elena read them.'
- b. Tri knigi, Tanja i Petar gi pročitaa.
 three books Tanja and Petar them read
 'Three books, Tanja and Petar read them.'
- c. Dve knigi, Eli i Lidija gi pročitaa.
 two books Eli and Lidija them read
 'Two books, Eli and Lidija read them.'

The statements in (43a-c) are all appropriate in the scenario given above. By using the CLLD constructions in (43a-c), the professor indicates that she knows who read what, without disclosing the identity of the books that each of the students read (although it should be noted that the professor could have opted to disclose the identity as well).

Given that the goal of the report (and the survey) is to know which books were read (and by whom), then a report that consists of statements with topicalized phrases, as in (44a-b), would not be appropriate. The use of such constructions seems odd or infelicitous. They seem to convey information about the number of books that each student read.⁴¹

- (44) a. Dve knigi, Ana i Elena pročitaa.
 two books Ana and Elena read
 ‘Two books, Ana and Elena read.’
- b. Tri knigi, Tanja i Petar pročitaa.
 three books Tanja and Petar read
 ‘Three books, Tanja and Petar read.’
- c. Dve knigi, Eli i Lidija pročitaa.
 two books Eli and Lidija read
 ‘Two books, Eli and Lidija read.’

The examples in (43)-(44) are informative in at least two respects. First, they clearly show that CLLD and Topicalization correspond to different readings of the dislocated numerals (to be made explicit very shortly). Second, they show that CLLD and Topicalization result in different pairings: CLLD seems to favor the pairing of books with the students who read them, while Topicalization seems to bring about a pairing between the students and the number of books that each of them read.

⁴¹ In effect, the inappropriateness of (44a-c) under these conditions can potentially be explained as flouting of Grice’s maxim of relevance.

The distinction is evident in the following situation as well. Consider a scenario where the top three students in the class, Ana, Petar, and Elena, were asked to solve a number of math problems. At the end of class, Ana had solved three problem sets (problem set no. 1, 2, and 3), while Petar and Elena had solved two (problem set no. 2 and 3). In this context, (45a) comes out true, while (45b) is false.

(45)a. Dve zadači, site studenti gi rešija.

two problems all students them solved

‘Two problems, all students solved.’

b. Dve zadači, site studenti rešija.

two problems all students solved

‘Two problems, all students solved.’

As (45a) shows, CLLD picks out the two problem sets such that every student solved. In this case, the numeral refers to problem sets no. 2 and 3. The numeral in (45b), on the other hand, seems to pick out, incorrectly, the number of problem sets such that every student solved.

Different judgments for (45a-b) are also obtained if we modify the above scenario as follows: Imagine that each of the three students solved two different problems, i.e. Ana had solved problem set no. 1 and 2, Petar had solved problem set no. 2 and 3, and Elena had solved problem set 1 and 3. In this case (45b), the construction with a topicalized DP, would come out as true, because each of the students has indeed solved two problem sets. At the same time, (45a), which has a CLLDed DP, would come out as false, because none

of the students have solved the same two problem sets.⁴² This confirms that the CLLDed and topicalized elements have different interpretation.

With this in mind, let us now return to our original examples in (43) and (44). The explanation that presents itself for the readings in (43)-(44) is based on the well-known distinction between the referential and cardinal readings of weak DPs (see Milsark (1977), Heycock (1995), among many others). To illustrate the basic distinction between the two readings consider (46).

(46) Pero vraboti dvajca studenti.

Pero hired two studenti

‘Pero hired two students.’

As a first step in separating the two readings of the numeral in (46), it helps to note that the sentence is a possible answer to either a question like *How many students did Pero hire?* in (47a) or a question like *Who did Pero hire?* in (47b).

(47) a. Kolku studenti vraboti Pero?

how many students hired Pero

‘How many students did Pero hire?’

⁴² Note that (i) and (ii), which contain direct objects in post- and preverbal position, respectively, come out true in both scenarios described here. The second scenario shows that a numeral in base or focus position can get a wide scope over the subject.

- (i) Site studenti rešija dve zadači.
all students solved two problems
‘All students solved two problems.’
- (ii) Site studenti dve zadači rešija.
all students two problems solved
‘It was two problems that all students solved.’

b. Kogo vraboti Pero?

who(Acc) hired Pero

‘Who did Pero hire?’

Under the referential reading, the answer in (46) signifies knowledge of a number with a commitment to identity (on the part of the speaker). Under the nonreferential (i.e. cardinal) reading, the same answer (46) signifies knowledge of a number without a commitment to identity.

While in the case of (46) both a referential or a cardinal interpretation is possible, there are other cases where one type of interpretation is favored. For example, “Seven people can fit into this car” is most likely an answer to “How many people can fit into this car?”, as it specifies an amount, and “I read *War and Peace*” is most likely an answer to a question like “What did you read?”, as it gives a specification of individuals.

We find similar preferences in the use of the two constructions, CLLD and Topicalization. Specifically, CLLD seems to force a referential reading of the initial element, while Topicalization seems to correlate with a cardinal reading. That CLLD is only possible with a referential reading is confirmed in contexts that force a cardinal reading. The prediction here is that Topicalization can be used in such contexts (because it allows for a cardinal reading), but CLLD cannot, (because it forces a referential reading). The prediction holds, as shown in (48a-b) below.

- (48) a. Šest časi, možes da staviš vo kutijava.
 six glasses can to put in box-this
 ‘Six glasses, you can put/fit in this box.’
- b. * Šest časi, možes da gi staviš vo kutijava.
 six glasses can to them put in box-this
 ‘Six glasses, you can put/fit in this box.’

Having established that the numerals in Macedonian have these two readings, we should immediately add that a cardinal reading of the numeral in (46) does not completely rule out a referential reading. We have just noted that certain preferences exist in terms of the correlation between questions and answers. It is important to note, however, that it is also possible to answer a question like “How many books did you read?” with “I read two, *War and Peace* and *Crime and Punishment*.”, whereby a specification of the amount does not rule out a subsequent specification of identity. Similarly, the statement in (46) can be modified in this manner. Thus, we can conclude that topicalized sentences in Macedonian are compatible with two kinds of follow-up statements – one that states explicitly that the speaker does not know the identity of the students (e.g. *Pero hired two students, but I don’t know who (they are).*) as well as one in which the individuals are specified (e.g. *Pero hired two students. They are Petar and Ana.*)

We have seen from the discussion here that there is a correlation between CLLDed and topicalized elements and referential vs. nonreferential reading, respectively. Based on

this insight, I would now like to propose an alternative solution to the scope problem posed by topicalized elements discussed in section 4.3.1.⁴³

Following suggestions by Veneeta Dayal and Barbara Partee (p.c.), I suggest that expressions at the left periphery are interpreted in that position.⁴⁴ Assuming that QR of expressions inside IP is an adjunction to IP, it is predicted that clause-initial direct objects will always have wide scope with respect to the subject. This is straightforward in the case of CLLDed expressions (cf. (42a-b), (43a-c) and (48b)). Given that we have identified topicalized numerals as having a cardinal interpretation, we can also explain the data in (42c) and (44a-c). The apparent low scope interpretation of the numeral is, in fact, due to a wide scope cardinal interpretation for it.^{45, 46}

To summarize, we have seen that the scope facts observed in CLLD and Topicalization can be explained by enforcing scope for these elements above the domain of QR and by maintaining a distinction between referential and cardinal readings of numerals.

⁴³ I am grateful to Veneeta Dayal (p.c.) for suggesting this particular approach.

⁴⁴ The claim that CLLDed and topicalized elements are interpreted at the left periphery will have repercussions for any analysis of binding facts in such constructions. I have not discussed the latter but I will have to assume that these can be explained with reference to the full chain that these elements form, which would enable binding theory to look at the base position of the elements. I am indebted to Mark Baker and Veneeta Dayal for bring my attention to the importance of this issue.

⁴⁵ As pointed out to me by Veneeta Dayal, we have to assume that the topicalized numeral does not have a referential reading, possibly due to the presence of CLLD, which is strictly referential.

⁴⁶ Roger Schwarzschild (p.c.) points out that the numerals on their referential (i.e. specific) reading are uninformative about their scope, because their interpretation as a singleton expression in essence neutralizes their scope with respect to other scope taking elements. Topicalized weak DPs and CLLDed strong DPs, on the other hand, are more informative in that they clearly show that they take the wide scope in the sentence. This means that left-dislocated quantifiers at LF are interpreted higher than other quantificational elements in the sentence (in our examples, the subject). This seems to be the simplest generalization that is compatible with the data.

4.3.3 Discourse Contexts for CLLD and Topicalization

In this section, I would like to address briefly the issue of the discourse status of the two dislocated structures, CLLD and Topicalization. The discussion that follows is by no means exhaustive and my intention in introducing this question is to highlight its importance and thus pave the way for future studies.

It is commonly assumed that clause-initial elements in both CLLD and Topicalization mark topics (see Ward & Birner (2004), Reinhart (1981), Rizzi (1997), Anagnostopoulou (1997), and references cited therein). This is confirmed in Macedonian by the fact that they are excluded as answers specifying the content of a *wh*-question.

(49) a. What did Ana read?

b. *Dve knigi, Ana pročitā.

two books Ana read

‘Two books, Ana read.’

c. *Dve knigi, Ana gi pročitā.

two books Ana them read

‘Two books, Ana read them.’

Both constructions are also ruled out in out-of-the-blue contexts, such as (50), giving further proof that the clause-initial DPs in CLLD and Topicalization are interpreted as topics.

(50)a. What happened at the races today?

b. *Dve nagradi, konjot na g. Vilson osvoi.

two prizes horse-the of Mr. Wilson won

‘Two prizes, Mr. Wilson’s horse won.’

c. *Dve nagradi, konjot na g. Vilson gi osvoi.

two prizes horse-the of Mr. Wilson them won

‘Two prizes, Mr. Wilson’s horse won them.’

Reinhart (1981:63) in her discussion of Left Dislocation (or HTLD) points out that such “sentences can be used appropriately in a given context only if the fronted NP can be understood as the topic, i.e. the sentence is used to assert something about its referent.” Since both CLLD and Topicalization fall into the class of constructions that generally mark topics, Reinhart’s remarks can be extended to include these two constructions as well. Numerous other studies have also noted this particular use of the two constructions (see, for example, Ward & Birner (2004) for claims that Topicalization (in their terms preposing) marks topichood, or Rizzi (1997) and Anagnostopoulou (1997) for claims that CLLD is used for the same purpose).

The claim that clause-initial elements in CLLD and Topicalization function as topics holds for Macedonian. Consider the question in (51a) and the answers in (51b-c).

(51)a. Koj gi pročita knigite?

who them read books-the

‘Who read the books?’

b. Dve knigi, Ana pročitā; tri knigi, Petar pročitā.

two books Ana read three books Petar read

‘Two books, Ana read; three books, Petar read.’

c. Dve knigi, Ana gi pročitā;

two books Ana them read

tri knigi, Petar gi pročitā.

three books Petar them read

‘Two books, Ana read them; three books, Petar read them.’

Although the question favors an individual answer where the answer specifies one or more individuals who read those books (e.g. *Petar read them* or *Petar and Ana read them*), it is also possible to answer the question with a topicalized or a CLLDed construction as in (51b-c), respectively. In a similar manner, (52a) can be answered as in (52b) or (52c).

(52)a. Što napravi so igračkite?

what did withtoys-the

‘What did you do with the toys?’

b. Dve igrački, i dadov na Ana.

two toys her gave to Ana

Dve igrački, i dadov na Tea.

two toys her gave to Tea

‘Two toys, I gave to Ana. Two toys, I gave to Tea.’

c. Dve igracki, i gi dadov na Ana(, topkata i kuklata).

two toys her them gave to Ana(, ball-the and doll-the)

Dve igracki, i gi dadov na Tea(, mečeto i žirafata).

two toys her them gave to Tea(, bear-the and giraffe-the)

‘Two toys, I gave them to Ana(, the ball and the doll). Two toys, I gave them to Tea(, the teddy bear and the giraffe).’

The use of a CLLD construction in (51)-(52) is also acceptable since the dislocated element is drawing from the set of toys, which has been already introduced in the discourse. The statements in (51) and (52) are interesting in that they show that the dislocated elements in both CLLD and Topicalization have to be given in the previous context and not only that but that they can introduce a further partitioning of the set introduced by the NPs in the question.⁴⁷ Such use of the topic comes close to what Büring (1995) identifies as partial topics, as they in some sense “narrow down” the topic set up in the discourse.⁴⁸

Given that both CLLD and Topicalization are similar to the extent that they unambiguously mark their clause-initial elements as topics, we can conclude that the only

⁴⁷ I am grateful to Veneeta Dayal for drawing my attention to this particular property of CLLD and Topicalization.

⁴⁸ ‘Partial topics’ are one of several types of topics that Büring (1999) identifies, in addition to ‘contrastive topics, as in (i), and ‘purely implicational topics, as in (ii). Büring goes on to argue that the different types of topics are different uses of what he calls S-Topics (sentence internal topics). From Büring (1999); the marking of topic and focus are Büring’s own.

(i) A: Do you think that Fritz would buy this suit?

B: Well, [I]_T certainly [WOULDN’T]_F.

Contrastive topics, Büring notes are used to “move the conversation away from an entity given in the previous discourse” (p.4). As he explains, by uttering B in (i), the speaker does not really answer A’s question but offers an alternative, though related statement. The ‘purely implicational topics’ as in (ii) indicate the speaker’s intention to discuss alternatives.

(ii) A: Did your wife kiss other men?

B: [My]_T wife [didn’t]_F kiss other men.

definitive difference that seems to emerge between the two is in terms of their referentiality. That is, as we noted earlier, CLLD seems to force a referential reading of the initial element, to the exclusion of a cardinal reading (see (48b)). Topicalization, on the other hand, correlates with a cardinal reading.

4.4 Conclusion

In this chapter, I offered evidence that the two constructions CLLD and Topicalization in Macedonian differ derivationally and in terms of their interpretation. Based on evidence from WCO effects and their patterning with respect to islands, CLLDed elements were shown to be base-generated in their clause-initial position, while topicalized elements were derived by movement of the DP from a position within IP. The two constructions were also shown to give rise to different readings of the dislocated DPs, with CLLD giving rise to a referential reading of the disclosed element. Finally, the fact that CLLDed objects take wide scope in the sentence was explained to arise from their being interpreted in the position in which they are generated. The observation that topicalized elements seemingly take narrow scope in the sentence in which they occur was shown to be a result of their wide scope cardinal interpretation, thus making it consistent with the view that dislocated elements are interpreted in the left periphery. The chapter also discussed the issue of the placement of CLLDed and topicalized elements in the left peripheral domain of the clause. It was shown that both types of elements can be analyzed as occupying separate topic positions within the CP domain.

CHAPTER 5

Clitic Co-Occurrence Restrictions in Wh-Questions

5.0 Introduction

One of the main insights from the previous chapters is that the distribution of the clitic pronouns in Macedonian correlates with the feature specification for strength in the DPs with which they co-occur. In this chapter, I would like to extend the discussion to wh-phrases and explore the consequences of the proposed analysis for these elements. Shifting the discussion to the domain of wh-questions is the next logical step in the exploration of the properties of clause-initial elements that I pursue in this study, as wh-phrases are the type of elements that in Macedonian require fronting for independent reasons. As such, they provide a useful testing ground for the analysis of left-dislocated elements we have been pursuing this far.

In this chapter I explore the clitic co-occurrence patterns and look at the factors determining their presence or absence with different types of wh-phrases. In particular, I investigate the extent to which the observed restrictions on the distribution of clitics as identified for the direct object DPs in Macedonian apply to such elements. I will show that the presence of clitics with wh-phrases is determined by their feature specification, rather than the interpretation they receive. The relevant feature in these cases would be D-linking, in the sense of Pesetsky (1987) and Comorovski (1996).

The chapter is organized as follows. Section 1 gives a brief overview of wh-questions in Macedonian, with an accent on its fronting requirements. Section 2 outlines the distribution of clitics with wh-phrases. Section 3 discusses the notion of D-linking

needed to explain the phenomenon. It introduces the role of partitivity as a crucial component in this. The section also gives a characterization of the different types of wh-phrases in Macedonian in terms of these two properties and it establishes a three-way distinction with regards to strength. Section 4 makes the argument that the co-occurrence of clitics with wh-phrases correlates with the D-linking/partitivity of these elements. Section 5 offers arguments for base-generation and movement analyses of wh-questions in Macedonian. Section 6 offers a tentative discussion of *koj* ‘who’ in Macedonian and its relation to *što* ‘what’. Section 7 contains the conclusion.

5.1 Basic Facts about Wh-Questions in Macedonian

I begin the discussion with a brief outline of the fronting requirements in Macedonian. Macedonian is a multiple fronting language, i.e. wh-phrases cannot stay in-situ (see Rudin 1988). The requirement holds for both mono-morphemic and complex wh-phrases. Illustrative examples for each type of wh-phrase are given in (1a-b) and (2a-b), respectively.

- (1) a. Koj što kupi?
 who what bought
 ‘Who bought what?’
 b. *Koj kupi što?
 who bought what

- (2) a. Koj student koja kniga ja kupi?
 which student which book it(f.sg) bought
 ‘Which student bought which book?’
- b. *Koj student ja kupi koja kniga?
 which student it(f.sg) bought which book

In (1b), the object wh-phrase *što* ‘what’ is left in-situ; in (2b), the object wh-phrase *koja kniga* ‘which book’ is left in-situ. In both cases, the resulting structures are ungrammatical.¹

Taking the main motivation for the obligatory fronting to be the Wh Criterion which requires that the wh-operator be in a Spec-Head relation with the head that carries the wh-feature (Rizzi 1991), we see that both mono-morphemic and complex wh-phrases in Macedonian are subject to it in the overt syntax. I will assume that C^0 carries the wh feature that triggers movement, as a result of which fronted wh-phrases are placed in SpecCP. I will refine this approach further in chapter 6, when two distinct positions for fronted wh-expressions will be motivated.

Since both types of wh-phrases must conform to the fronting requirements of the language, it is interesting to note that the two types of wh-phrases display differences when it comes to their interaction with clitic pronouns. Given what we have seen so far and anticipating the discussion that follows, the following emerge as likely structures for Macedonian wh-questions:

¹ This requirement seems to hold for echo questions as well. Thus, (1b) is ungrammatical even as an echo question. The same holds for (2b), though here an echo question interpretation is marginally acceptable. For the purposes of the present analysis, echo questions will be put aside. All of the examples discussed in this chapter are to be treated as regular multiple wh-questions.

- (3) a. $[_{CP} wh_i [_{IP} \dots pro_i Cl_i V t_i \dots]]$
 └──────────┘ └──────────┘
- b. i. $[_{CP} wh_i [_{IP} \dots t_i Cl_i V t_i \dots]]$
 └──────────┘ └──────────┘
- ii. $[_{CP} wh_i [_{IP} \dots V t_i \dots]]$
 └──────────┘

(3a) represents a wh-question with a base-generated wh-phrase, while (3b-i) and (3b-ii) represent wh-questions derived by movement of the wh-phrase to the left periphery.

The remainder of the chapter gives evidence for the existence of these structures in Macedonian. Given that the representations in (3a-b) differ minimally in the presence vs. absence of a clitic co-indexed with the wh-phrase, I begin the discussion with an outline of the basic generalizations pertaining to the co-occurrence restrictions that hold between these two types of elements.²

5.2 The Distribution of Clitics in Wh-Questions

Chapters 2 and 3 looked at this issue in the context of direct object DPs within IP and within the left periphery, respectively. It is now time to see how their distribution is regulated in the context of wh-questions.

In describing the clitic co-occurrence patterns in Macedonian wh-questions, it is useful to distinguish three broad classes of wh-expressions: *koj N* ‘which N’, *kakov N*

² To maintain continuity with the data considered so far, I concentrate on direct object wh-phrases. We can note that *koj* ‘who’ is never clitic-doubled, as is expected, given the fact that it is a subject wh-phrase, and subjects are never clitic-doubled in Macedonian. Indirect object wh-phrases (e.g. *komu* ‘to whom’) are obligatorily doubled. As noted previously, I leave the investigation on the restrictions on clitic co-occurrence with indirect objects, DPs and wh-phrases, for further research.

‘what N’, and *koj/što* ‘who/what’. The first two are complex wh-expressions; the latter are mono-morphemic.

Koj N ‘which N’ and *kakov N* ‘what N’ distinguish between singular and plural forms; when they are in the singular, they are inflected for gender.³ As Mišeska Tomić (2008) points out, *što* ‘what’ does not inflect for gender, number, or case; *koj* ‘who’ is [+human] and can inflect for case. When it does, it has the following forms: *koj* ‘who.NOM’, *kogo* ‘who.ACC’ and *komu* ‘who.DAT’.

The distribution of clitic pronouns with these wh-phrases is as follows.

Koj N ‘which N’ phrases obligatorily co-occur with a clitic pronoun, as illustrated in (4).⁴ *Kakov N* ‘what N’ phrases, on the other hand, can never co-occur with a clitic, as shown in (5).

- (4) *Koja* *kniga* **(ja)* *kupi* *Petar?*
 which book it(f.sg.) bought Petar
 ‘Which book did Petar buy?’

- (5) *Kakva* *kniga* **(ja)* *pročita* *Petar?*
 what book it(f.sg.) read Petar
 ‘What book did Petar read?’

³ The full paradigm of *koj* ‘which’ is: *koj* (m.sg.), *koja* (f.sg.), *koe* (n.sg.), *koi* (pl.). *Kakov* ‘what kind’ has the following set of forms: *kakov* (m.sg.), *kakva* (f.sg.), *kakvo* (n.sg.), *kakvi* (pl.)

⁴ The distribution of clitics with wh-expressions varies from languages to language. Bulgarian, for example, allows for the doubling of *which Ns*; see Krapova & Cinque (2003) and Krapova (2002) for details. They argue that the presence of a clitic is not really optional but that it correlates to the position the wh-phrase occupies in the clause (in this case a topic position above CP). Clitic doubling of *which Ns* in Romanian, is obligatory; see Dobrovie-Sorin (1994), Comorovski (1996) for details.

Mono-morphemic phrases *koj/što* ‘who/what’ present a mixed case. *Koj* ‘who’ is optionally possible with a clitic, as shown in (6a); *što* ‘what’ is not, as shown in (6b).

- (6) a. Kogo (go) vide?
 whom him saw
 ‘Who did you see?’
- b. Što (*go) kupi Petar?
 what it(m/n.sg.) bought Petar
 ‘What did Petar buy?’

The behavior of the wh-phrases is interesting when compared to the behavior of the DPs discussed in the previous three chapters. Recall that the patterning of the DPs inside IP supported a two-way distinction between DPs that required clitics and those that did not allow them. Their patterning outside of IP, however, gave rise to a three-way distinction between DPs that require clitics, DPs that disallowed them and DPs for which clitics were optional. Here we see that the wh-phrases behave like the DPs outside IP in that they fall into three groups: wh-phrases that consistently allow clitics (*koj N* ‘which N’ phrases), those that consistently disallow clitics (*što* ‘what’ and *kakov N* ‘what N’) and those wh-phrases that optionally allow for their presence (*koj* ‘who’). This poses an interesting set of questions about the application of the Clitic Criterion to wh-phrases and about the different types of wh-phrases in the language. In particular, the difference between *što* ‘what’ and *koj* ‘who’ is puzzling since they are both mono-morphemic argument phrases. The main question that I will attempt to answer in this chapter is how

closely the characterization of wh-phrases parallels that of the DPs and how this in turn determines their behavior with clitics.

5.3 D-Linking and Partitivity

The main question that arises is what governs the clitic co-occurrence patterns in (4)-(6) above. I will argue that the distribution of the clitics is governed by the feature specification of the wh-phrases, with the features being representative of their intrinsic semantic properties. As would be obvious, the wh-phrases that must occur with clitics are those that have been classified as D-linked by Pesetsky (1987). Since this is going to be the defining property for the Macedonian wh-phrases with respect to which I will explain the clitic co-occurrence restrictions, it is important at this point to make a brief clarification about what counts as a D-linked and what counts as a non-D-linked wh-phrase.

The distinction between D-linked (DL) and non-D-linked (NDL) wh-phrases comes from Pesetsky (1987). The difference can perhaps be best explained in reference to the answers that questions with such wh-phrases allow. Pesetsky notes that an answer to a question which contains a DL wh-phrase (*which N*) will have to pick out an object from a contextually defined set or a set that has already been mentioned in the discourse. Thus, a felicitous answer to the question *Which book did you read?*, he explains, would have to pick an object from the set of books that the speaker and hearer have in mind. Questions with NDL wh-phrases (*who/what*) pose no such requirements. This means that a question like *What did you read?* does not limit the range of felicitous answers to a particular set that the participants in the conversation have in mind.

Pesetsky uses the notion of D-linking to account for the different behavior of the two types of wh-phrases in examples like (7a-d) and (8a-d) below.

- (7) a. Who t_i read what?
 b. *What did who read t_j ?
 c. Which student t_i read which book?
 d. Which book did which student read t_j ?
- (8) a. Who did you persuade t_i to read what?
 b. *What did you persuade who to read t_j ?
 c. Which student did you persuade t_i to read which book?
 d. Which book did you persuade which student to read t_j ?

As (7a-b) and (8a-b) show, the fronting of mono-morphemic wh-phrases is subject to strict ordering restrictions. Here the fronting of the subject wh-phrase is preferred over the fronting of the object wh-phrase, giving rise to the well-known Superiority effects.⁵ In (7c-d) and (8c-d) we see that complex wh-phrases are not subject to such ordering restrictions, so the fronting of either the subject or the object is equally acceptable.

Pesetsky argues that the lack of Superiority effects with *which* Ns in (7c-d) and (8c-d) is due to their D-linking rather than their syntactic heaviness.⁶ His supporting evidence

⁵ Pesetsky (1987) explains the Superiority effects in questions in terms of the Nested Dependency Condition. He notes that wh-phrases which are discourse-linked are exempt from it and do not have to move at LF in order to get interpreted. For details see Pesetsky (1987). See also Dayal (2006).

⁶ Specifically, this means that the difference in Superiority effects between the wh-phrases does not derive from the fact that *which* is a specifier but *who/what* are heads.

comes from examples like (9). In (9) Superiority effects persist despite the fact that *how many* can also be considered syntactically heavy.

- (9) a. I need to know how many people_i t_i voted for whom.
 b. *I need to know who(m)_j how many people voted for t_j.

Having established that D-linking is a lexical property of *which* Ns, Pesetsky notes that *who* and *what* can also potentially be D-linked. When this happens, he notes, Superiority effects are suspended, as shown in (10) below (from Bolinger 1978):

- (10) I know what just about everybody was asked to do, but *what* did *who* (actually) do?

Thus, a crucial difference emerges between *which* Ns and *who/what* not only in the fact that they are D-linked and non-D-linked, respectively, but also in the fact that D-linking is an inherent property of the former, but not of the latter.⁷ This difference is also adopted by Comorovski (1996) in her analysis of Romanian multiple wh-questions.

Comorovski follows Pesetsky in distinguishing between the two types of wh-phrases, but offers a slightly different definition of D-linking. Comorovski argues that D-linked wh-phrases require a partitioning of the set that *which* takes as an argument. This, she explains, can be achieved in one of two ways: when the membership of the set denoted by the NP in the *which* phrase is known to both the hearer and the speaker or when the

⁷ The degree to which non-D-linked wh-phrases allow for optional D-linking varies. For example, *who* is said to be more flexible than *what* (Comorovski 1996). Wh-phrases like *why* and *how* are said to be even more resistant to D-linking, although they allow it (see Dayal 2006).

interlocutors share the same criterion according to which a partitioning of the set can take place. Thus, Comorovski emphasizes partitivity as intrinsically connected to D-linking.

Turning to Macedonian, it would be interesting to see how the D-linked characterization of the *wh*-phrases influences their patterning with respect to clitics. For example, given the D-linking of *koj N* ‘which N’, it is predicted that once *koj/što* ‘who/what’ are D-linked, they should pattern like *koj N* ‘which N’ and allow for the possibility of a clitic. We will see, however, that the change of the DL status of *who/what* in Macedonian does not seem to have the expected effects. In particular, we will see that it does not alter their behavior with respect to clitics. Before we look at the DL characterization of the *wh*-phrases in Macedonian, however, let us take a look at the second property that has been singled out as relevant for *wh*-phrases, and in particular *which N* and *what N* phrases.

Pesetsky in his discussion concentrates on *which Ns*, *who/what*, as well as *how many* and *where*.⁸ He does not, however, discuss *what N* phrases. A possible characterization of such phrases can be drawn in relation to *which Ns*.

One distinguishing property between *which Ns* and *what Ns* is partitivity. Heim (1987:33) argues that *which Ns* are inherently partitive and have an interpretation equivalent to *which of them (of the Ns)*. *What Ns*, on the other hand, are not partitive and can never get the interpretation **what of them (of the Ns)*.⁹ Comorovski (1996:11)

⁸ This in essence raises the question of the argument/adjunct distinction, which I will not address here. It should be noted, though, that adjunct *wh*-phrases like *where* and *when* in Macedonian can never co-occur with clitic pronouns.

⁹ A possible explanation for the inability of *what Ns* to receive such interpretation can be found in Heim (1987). Heim suggests that the inability of *what N* to get such a partitive interpretation may be due to the fact that *what* introduces a kind variable (in the sense of Carlson 1977). Since partitives in general are incompatible with kind-denoting heads, Heim explains, (for e.g., **Big ones of my turtles/This kind of my turtles withstand(s) freezing temperatures* (Heim 1987:33)), *what Ns* would not be able to get an interpretation of a partitive.

explains that the partitivity of *which* *Ns* comes from the fact that the argument selected by *which* refers to a set that has already been referred to in the previous discourse or is a set that is contextually salient. Note that partitivity as such is also a distinguishing factor between *which* *Ns* and mono-morphemic *who/what*, which in this respect pattern like *what* *N* phrases.

The distinction between *which* *N* and *what* *N* phrases shows up in Macedonian. If the ability to form an overt partitive is an indication of the partitivity of the *wh*-element, then Heim's and Comorovski's arguments hold for Macedonian *which* *Ns* and *what* *Ns* as shown below.

Which *Ns* can be used in overt partitive constructions; cf. (11a) and (11b). Note that, in both cases (11a) and (11b), the clitic *ja* 'it(f.sg.)' remains obligatory.

- (11) a. Koja kniga *(ja) pročita Petar?
 which book it(f.sg.) read Petar
 'Which book did Petar read?'
 b. Koja od knigite *(ja) pročita Petar?
 which of books-the it(f.sg.) read Petar
 'Which of the books did Petar read?'

What *Ns*, on the other hand, cannot be used in partitive constructions, hence the ungrammaticality of (12b). The issue of the clitic does not arise here.

(12) a. Kakva kniga (*ja) pročitā Petar?

what book it(f.sg.) read Petar

‘What book did Petar read?’

b. *Kakva od knjigite pročitā Petar?

what of books-the read Petar

‘What of the books did Petar read?’ *lit. translation*

To complete the paradigm, consider the behavior of *što* ‘what’ and *koj* ‘who’.¹⁰ Here we see an interesting pattern developing. As (13) shows, *koj* ‘who’ (when in the accusative) can be used in a partitive construction; the clitic remains optional in such cases (cf. (13) vs. (6a)). The same cannot be said of *što* ‘what’, since it cannot be used in a partitive construction, as shown in (14). As was the case with *what* Ns, the issue of the clitic in such cases does not arise.

(13) Kogo od prisutniti (go) intervjuiraše Ana?

who of attendees-the him interviewed Ana

‘Which of the attendees did Ana interview?’

(14) *Što od knjigite pročitā Petar?

what of books-the read Petar

‘What book did Petar read?’

¹⁰ I am grateful to Veneeta Dayal for pointing out that partitivity extends to these cases as well.

Given the data in (11)-(14), a pattern emerges: *koj N* ‘which N’ phrases can be used in a partitive and they obligatorily require the presence of a clitic. *Kakov N* ‘what N’ and *što* ‘what’ cannot be used in partitives and they do not allow for clitics. *Koj* ‘who’ can be used in a partitive and it can co-occur with a clitic. In sum, the data shows that when it comes to partitivity, *which Ns* and *what/what Ns* stand at opposite sides of the spectrum, with *who* being somewhere in the middle, seemingly patterning with *which Ns*.

The patterning of the partitives headed with wh-expressions is interesting for two reasons. On the one hand, we see that their behavior is somewhat parallel to that of the DPs in chapter 3. Recall that the cursory investigation of these constructions with DPs showed that the presence of the clitic with the partitive is determined by the properties of the head of the partitive rather than its interpretation (recall that partitive DPs come out as strong on both Barwise & Coopers (1981) and Milsark (1977) tests). Thus, a clitic was possible with unspecified weak DPs as in *Dve od knjige (gi) procitav* ‘Two of the books I read (them)’ but not with [-strong] DPs as in *Mnogu od knjige (*gi) procitav* ‘Many of the books I read.’ The same seems to hold for the wh-phrases, where the clitic correlates with the type of determiner that is being used. In this case, partitives headed with *which* obligatorily require a clitic, while those headed by *who* optionally allow it. Having said this, we also see that there is a difference between the two types of partitives. In the case of DPs, it was clear that both strong and weak determiners could form partitive phrases. In the case of wh-phrases, only D-linked ones can be used in such constructions. There is one notable exception to this, of course: partitives can be headed by *koj* ‘who’, as shown in (13).

I cannot explain why this should be the case. One possibility that presents itself would be to treat *koj* as D-linked but syntactically bare (in the sense that the head noun is missing, but can be understood from the previous context) (as suggested by Krapova (2002) for Bulgarian). This would imply that bare *koj* in Macedonian can either be D-linked or non-D-linked. Thus, when the inherently DL *koj* enters the derivation, it would trigger the clitic, as in (6a). The DL *koj* can also head a partitive. When an inherently NDL *koj* enters the derivation, the clitic cannot occur, which would account for the clitic-less version of the question in (6a). Although this might explain the presence vs. absence of clitics with *koj* in (6a), it still does not completely explain its behavior in the partitive in (13). If the partitive itself is dependent on the inherent DL properties of *koj*, once the partitive is formed, the clitic should become obligatory, as is the case with *which N* in (11b). This is not what happens in (13). The question, therefore, still remains open as to why *koj* ‘who’ behaves in this manner

For purposes of explaining the clitic co-occurrence patterns I will propose a three-way distinction in the feature specification for wh-phrases. If we classify *what N* and *what* as non-D-linked but *who* as neither D-linked nor non-D-linked, i.e. as unspecified for this feature, and we take non-D-linked wh-expressions to be unable to head partitives, the data discussed in this section follows. While I do not have an explanation for why wh partitives cannot be headed by non-D-linked expressions, I will show that making this three-way distinction explains the clitic co-occurrence patterns in Macedonian questions in a straightforward manner.

5.4 Wh-Phrases and the Clitic Criterion

We saw in the preceding section that wh-phrases in Macedonian split into three groups depending on whether they are D-linked (*which N*), non-D-linked (*what/what N*) and unspecified (*who*). We also know from previous discussion that wh-phrases belonging to the latter two groups can be D-linked contextually. Given this, the obvious question is whether the change of the DL status would change the distribution of the clitic. We will see shortly that this is not the case. I take that to be crucial evidence that the interpretation of the wh-phrases in Macedonian does not determine the presence vs. absence of the clitics.

We noted the fact that wh-phrases like *who* and *what* can optionally get a D-linked interpretation. To bring out this reading of the wh-phrases in Macedonian, imagine the following scenario: Eva, Petar and Ana went shopping for Lidija's birthday. The items they bought were a book, a perfume, and a scarf. In this scenario we can ask the question *What did Ana buy?* as in (15) below.

- (15) Što kupi Ana?
 what bought Ana
 ‘What did Ana buy?’

Što ‘what’ in (15) is D-linked: it asks to pick out an object from a contextually salient set (or a set of objects that has already been introduced in the discourse). Having established the D-linked status of *što* ‘what’ in (15), we can construct the same example as in (6) to see if the wh-phrase, under this interpretation, allows for a clitic.

(16) Što (*go) kupi Ana?

what it bought Ana

‘What did Ana buy?’

The ungrammaticality of (16) shows that the *wh*-phrase cannot co-occur with a clitic, despite the fact that the *wh*-phrase *što* ‘what’ receives a D-linked interpretation.

Kakva kniga in (5) can also be D-linked. Imagine a situation where Eva, Petar and Ana bought books on trees, shrubs and flowers. Now we want to know what type of book Petar bought. In the given context, it is possible to ask (17). Here, it is very clear that *what book* refers to an object in the contextually salient set of books. (17) is asking for the type of book that Petar bought.

(17) Kakva kniga (*ja) kupi Petar?

what book it(f.sg.) bought Petar

‘What book did Petar buy?’

As (17) shows, however, doubling of the now D-linked *what book* is still not allowed. Thus, we see from both (16) and (17) that changing the DL status of an otherwise NDL *wh*-phrase in Macedonian does not result in the presence of a clitic pronoun.

The behavior of the *wh*-phrases in this respect is reminiscent of the behavior of the IP internal weak DPs in chapter 3. Recall that in that case a specific interpretation did not result in the presence of a clitic either. If we assume that the defining semantic property of *wh*-phrases relevant for feature specification is D-linking, we can extend the proposal

made earlier that the presence of the clitic is tied to the intrinsic properties of the doubled element and that it is these properties that determine the clitic co-occurrence. In order to map out the feature specification of the wh-phrases, it would be useful to remind ourselves of the characterization of the DPs we have argued so far.

Recall that DPs fall into three categories: strong DP (e.g. *every*), [-strong] DPs (e.g. *many*) and unspecified DPs (e.g. the numerals and other indefinites). We have seen that wh-phrases differ with respect to their (in)ability to get a D-linked interpretation, with *koj N* ‘which N’ being inherently D-linked, *kakov N* ‘what N’ and *što* ‘what’ being inherently non-D-linked, and *koj* ‘who’ being unspecified. We also know that the latter can be optionally D-linked if an appropriate context is provided. On the assumption that D-linking, as an intrinsic semantic property, is the wh-correlate of the [+strong] feature we identified for the DPs, we can characterize the wh-phrases as follows: *koj N* ‘which N’ phrases are [+strong], because they are inherently D-linked; *kakov N* ‘what N’ and *što* ‘what’ are [-strong], because they can never participate in partitives; *koj* ‘who’ is unspecified for strength as shown by the possibility of partitive constructions. The features for strength then crucially reflect the partitivity of the wh-phrases. For ease of reference, I will mark strong wh-phrases as [+strong_{WH}].

This particular classification of the wh-phrases explains the presence vs. absence of a clitic in a straight forward manner. We know from previous discussion that the clitic is always triggered by a [+strong] DP, that it is never triggered by a [-strong] DP and that it may appear with DPs unspecified for strength (as in the case of CLLDed numerals). The same paradigm is evident with the wh-phrases: *koj N* ‘which N’ phrases obligatorily co-occur with clitics, *kakov N* ‘what N’ and *što* ‘what’ never allow for clitics, and *koj* ‘who’

optionally allows for them. Thus, as was the case with the DPs, the presence vs. absence of clitics with wh-phrases can be regulated by the Clitic Criterion: the clitic (i.e. the clitic phrase) will only be generated when it needs to license a $[+strong_{wh}]$ feature. In all other cases, the clitic phrase will not be projected due to economy principles.

We are now in a position to see what the derivations of our original examples in (4)-(6) are. These are given in (18)-(19) below. The structures reflect the two types of questions in Macedonian that are in principle possible: those derived by base-generation and those derived by movement. Evidence for the two types will be given in section 5.5. Anticipating the discussion there, the two types are represented in (18a-b) and (19a-c), respectively. As (18a-b) show, the argument position of base-generated wh-phrases would be taken by *pro*. In the case of a wh-phrases that moves to the left periphery, as in (19a-c), the trace in the argument position would be that of the wh-phrase itself.

- (18) a. $[_{CP} \text{koja kniga}_i \quad [_{IP} \dots \text{pro}_i \quad \text{ja}_i \text{ kupi}_j \quad [_{VP/IP} \text{Petar} \quad [_{VP} \text{t}_j \quad \text{t}_i]]]$
└──────────────────┘ \nwarrow ──┘
- b. $[_{CP} \text{kogo}_i \quad [_{IP} \dots \text{pro}_i \quad \text{go}_i \text{ vide}_j \quad [_{VP/IP} \text{Petar} \quad [_{VP} \text{t}_j \quad \text{t}_i]]]$
└──────────────────┘ \nwarrow ──┘

In (18a-b), the wh-phrase enters in a binding relation with *pro* in the argument position of the verb. As was the case with the direct object DPs, the clitic in this case is generated because of the strong features of *pro*.¹¹ Finally, the chains in (18a-b) are well formed

¹¹ Note that a movement analysis of *which* Ns cannot in principle be excluded. The important thing to remember though is that even in the case of movement the presence of the clitic would be predicted because of their being $[+strong_{wh}]$. In this, Macedonian *which* Ns parallel the behavior of $[+strong]$ DPs.

because there are no clashes in the feature specification of the elements involved as none of them are $[-\text{strong}_{\text{WH}}]$.¹²

- (19) a. $[_{\text{CP}} \text{kogo}_i \text{ C vide}_j \text{ } [_{\text{VP/IP}} \text{Petar } [_{\text{VP}} \text{t}_j \text{ } \text{t}_i]]]$
 \uparrow _____
- b. $[_{\text{CP}} \text{što}_i \text{ C kupi}_j \text{ } [_{\text{VP/IP}} \text{Petar } [_{\text{VP}} \text{t}_j \text{ } \text{t}_i]]]$
 \uparrow _____
- c. $[_{\text{CP}} \text{kakva kniga}_i \text{ C kupi}_j \text{ } [_{\text{VP/IP}} \text{Petar } [_{\text{VP}} \text{t}_j \text{ } \text{t}_i]]]$
 \uparrow _____

(19a) differs minimally from (18b) in that *kogo* ‘whom’ in this case is moved to the front of the clause from the object position. Given that it is unspecified for strength, *koj* cannot license the clitic pronoun. The same holds for *što* ‘what’ in (19b) and *kakva kniga* ‘what book’ in (19c), only in this case the lack of clitic is due to their $[-\text{strong}_{\text{WH}}]$ features. We can thus conclude that the distribution of the clitics with wh-phrases follows from the Clitic Criterion.

Before concluding this section, I would like to return to the question of how the three-way characterization of the wh-phrases we established here helps us explain the parallel between the wh partitives and their non-wh counterparts as noted in section 5.3.

In the case of the wh partitives, the features of the head DP percolate up to that of the partitive DP. This explains the presence of the clitic with partitives headed by *which* and

¹² We should add that *kogo* ‘whom’ differs from *što* ‘what’ in that it is $[+\text{human}]$. It is not clear to me if this might also be a factor in the licensing of the clitic. It has been noted that clitic doubling in some languages is sensitive to the feature $[+\text{animate}]$; see Suñer (1988), Comorovski (1996), etc.. We saw in chapter 2, however, that animacy or $[+\text{human}]$ is not a factor in the clitic doubling of DPs in Macedonian. The case of *kogo* ‘whom’ is the only case where $[+\text{human}]$ becomes more prominent. However, given its lack of WCO effects (see section 5.5.1), I will maintain that the clitic in the case of *kogo* ‘whom’ is linked to its unspecified features for strength. I leave the detailed investigation of the role the feature $[+\text{human}]$ may play for further research.

who. For reasons that we do not fully understand, [-strong] wh-phrases are not attested in the head position of this construction. However, once we accept that this restriction holds, the other aspects of the wh properties follows. Recall that partitives with non-wh DPs do not have this restriction, as a result of which all three types of DPs ([+strong], [-strong], and unspecified) can participate in this construction. The presence or absence of the clitic in those cases follows from the features of the head of the partitive.

To summarize, we have seen that the presence vs. absence of a clitic in wh-questions relates to the feature specification of the wh-phrases. Taking D-linking to be the relevant property that determines the strength of the wh-phrases, we distinguished between wh-phrases that are [+strong_{WH}] (*which N*), [-strong_{WH}] (*what/what N*) and those that are unspecified for strength (*who*). Given this, the distribution of the clitics in the context of wh-questions is regulated by the Clitic Criterion.

5.5 CLLDed Wh-Questions vs. Ordinary Wh-Questions

In this section, I present evidence from WCO and island effects that support the existence of two types of questions in Macedonian, those that are derived by movement and those that are base-generated. As I explain below, the differences in the derivation correlate with the properties of the wh-phrases themselves, as defined above.

5.5.1 Weak Crossover Effects


Wh-phrases in Macedonian differ with respect to WCO violations. Inherently DL wh-phrases do not give rise to WCO effects, but NDL wh-phrases do. (20a) shows that *koj N* ‘which N’ phrases do not create WCO violations.

(20) a. [Koe momče]_i [brat mu]_i go_i poseti?

which boy brother his him visit

‘Which boy did his brother visit?’

b. [koe momče]_i [[brat mu]_i *pro*_i go_i poseti t_i]




The lack of WCO effects in (20a) can be attributed to the fact that the *wh*-phrase does not bind a trace inside IP. *Kakov N* ‘what N’ phrases, on the other hand, give rise to WCO violations, as shown in (21a).¹³ The ungrammaticality of (21a) is due to the fact that the variable bound by the *wh*-phrase is coindexed with a pronoun to its left.

(21) a. *[Kakvo momče]_i [brat mu]_i poseti?

what boy brother his visit

‘What boy did his brother visit?’

b. [kakvo momče]_i [[brat mu]_i poseti t_i]



WCO effects also arise in questions with *što* ‘what’, as evident in the ungrammaticality of (22b), which is due to the coindexation of the variable with the pronoun.

¹³ A clitic cannot be used in (21a), as *what N* in Macedonian does not allow it (see also (5)).

(i) *[Kakvo momče]_i [brat mu]_i kje go_i poseti?
 what boy brother his will him visit
 ‘What boy will his brother visit?’

(22) a. Što_i [negoviot_i sopstvenik] prodade?¹⁴

what his owner sell

‘What did its owner sell?’

b. [što]_i [[negoviot_i sopstvenik] prodade t_i]

↑ _____ |

The test can be applied to questions with *koj* ‘who’. (23a) is a case where the wh-phrase does not co-occur with a clitic pronoun. The example shows that coindexation between the fronted wh-phrase and the pronoun leads to a WCO violation.

(23) a. Kogo_{j/*i} [brat mu_i] poseti?

whom brother his visit

‘Whom did his brother visit?’

b. kogo_{*i} [[brat mu_i] poseti t_i]

↑ _____ |

As expected, WCO violations become significantly weaker when *koj* ‘who’ co-occurs with a clitic in IP, as shown in (24a):

(24) a. ?Kogo_{i/j} [brat mu_i] go_i poseti?

whom brother his him visit

‘Whom will his brother visit?’

b. kogo_i [[brat mu_i] pro_i go_i poseti t_i]

----- ↑ _____ |

¹⁴ The use of the pronoun *negoviot* ‘his’ is more appropriate in this context. The short pronominal form *mu* ‘his’ is most naturally used with kingship terms.

The lack of WCO effects in (24a) supports the idea that wh-phrases coindexed with a clitic are base-generated in their surface position.

As in the case of DPs, we can take the lack/presence of WCO effects in wh-questions to be indicative of base-generation and movement, respectively. Since *koj N* ‘which N’ phrases and *koj* ‘who’ (when coindexed with a clitic) do not give rise to WCO effects (see (20a) and (24a)), we can conclude that they are base-generated in their surface position. The presence of the clitic in both cases would be due to strong *pro* in object position. Given that *kakov N* ‘what N’, *što* ‘what’ and *koj* ‘who’ (when not coindexed with a clitic) show WCO violations (see (21a), (22) and (23a)), we can conclude that they move to their clause-initial position. In such cases, the clitic is not licensed because the wh-phrases are [-strong_{WH}] or unspecified.

Finally, note that the characterization of *what* and *what N* as [-strong_{WH}] precludes the option of their being base-generated because in that case their features would clash with the features of *pro*. On the other hand, nothing stops *who* from being base-generated, since it is not specified for strength, i.e. crucially it is not [-strong_{WH}].

5.5.2 Island Effects

The proposed analysis distinguishes between wh-phrases that enter into a binding relation with their clitics and the null pronoun in the argument position and wh-phrases that enter into a government chain with the traces they leave after movement. Given this, we would expect questions with CLLDed wh-phrases to be able to violate weak islands. Wh-phrases that move, on the other hand, would be expected to pattern like topicalized DPs

and be sensitive to weak islands. These predictions are borne out. Consider the examples in (25a-c).

(25) a. *Koi studenti se misliš dali da gi isprašaš?*

which students refl think whether to them examine

‘Which students are you wondering whether to examine?’

b. **Kakvi studenti se misliš dali da isprašaš?*

what students refl think whether to examine

‘What students are you wondering whether to examine?’

c. **Što te prašaa dali pročita?*

what you asked whether read

‘What did they ask you whether you read?’

d. *Kogo se misliš dali da go isprašaš?*

who refl think whether to him examine

‘Who are you wondering whether to examine?’

e. **Kogo se misliš dali da isprašaš?*

who refl think whether to examine

‘Who are you wondering whether to examine?’

In (25a) *koi studenti* ‘which students’ escapes the weak island.¹⁵ In contrast, the fronting of *kakvi studenti* ‘what students’ and *što* ‘what’ in (25b-c), respectively, results in a weak island violation. This contrast is also present with *koj* ‘who’ in (25d-e).

¹⁵ I assume that *dali* ‘whether’ creates a weak island in Macedonian.

From this we can conclude that the two types of wh-phrases enter in different types of chain relations: a CLLDed wh-phrase forms a binding chain with *pro* in object position, while a moved wh-phrase and its trace are part of an antecedent government chain.

Given that binding chains, as well as those formed by movement, are constrained by strong islands, it is expected that both types of wh-phrases will obey these. In (26a), the CLLDed wh-phrase *koja prijateljka* ‘which friend’ cannot enter into a chain with the pronoun *ja* ‘her’ inside the adjunct clause. Fronted wh-phrases that enter into a government chain also obey strong islands, as shown in (26b-c). The patterning of CLLDed and fronted wh-phrases is also evident with *koj* ‘who’ in (26d-e).

- (26) a. **Koja prijateljka go ispi kafeto pred da ja vikneš?*
 which friend it drank coffee-the before to her call
 ‘Which friend did you drink the coffee before you called her?’
- b. **Kakva prijateljka go ispi kafeto pred da vikneš?*
 what friend it drank coffee-the before to call
 ‘What friend did you drink the coffee before you called?’
- c. **Što plačeše zatoa što Petar istepa?*
 what cried because that Petar beat
 ‘What did you cry because Petar beat?’
- d. **Kogo go ispi kafeto pred da go vikneš?*
 who it drank coffee-the before to him call
 ‘Who did you drink the coffee before you called him?’

- e. *Kogo go ispi kafeto pred da vikneš?
 who it drank coffee-the before to call
 ‘Who did you drink the coffee before you called?’

The same effects are replicated in strong islands created by relative clauses. In (27a) the binding relation between the CLLDed *koja topka* ‘which ball’ and the clitic *ja* ‘it(f.sg.)’ inside the relative clause is disrupted. This explains the ungrammaticality of (27a) for strong wh-phrases; it also explains the ungrammaticality of (27b-c) for weak wh-phrases. *Koj* ‘who’ follows the same pattern, as shown in (27d-e).

- (27) a. *Koja topka go izbrka kučeto što ja ukrade?
 which ball it chased dog-the that it(f.sg.) stole
 ‘Which ball did you chase the dog that stole it?’
- b. *Kakva topka go izbrka kučeto što ukrade?
 what ball it chased dog-the that stole
 ‘What ball did you chase the dog that stole?’
- c. *Što go izbrka kučeto što ukrade?
 what it chased dog-the that stole
 ‘What did you chase the dog that stole?’
- d. *Kogo go izbrka kučeto što go kasna?
 who it chased dog-the that him bit
 ‘Who did you chase the dog that bit him?’

e. *Kogo go izbrka kučeto što kasna?

who it chased dog-the that bit

‘Who did you chase the dog that bit?’

The behavior of wh-phrases with respect to syntactic islands in (25)-(27) shows that the relation between CLLDed wh-phrases and the null *pro* in the object position is that of binding. Their sensitivity to strong islands confirms that the relation between the wh-phrase and *pro* is that of a chain to begin with. Sensitivity of fronted wh-phrases to weak islands, on the other hand, showed these to be instances of government chains.

To summarize, the patterning of wh-phrases with respect to WCO and syntactic islands shows that wh-phrases coindexed with clitics are base-generated in their surface position while those that do not co-occur with clitics move to the left periphery.¹⁶

5.6 A Loose End

The proposal that the presence or absence of a clitic with unspecified wh-phrases correlates with their features for strength helps us explain the apparent optionality of the clitic with *koj* ‘who’ in (6a) and its unavailability with *što* ‘what’ in (6b). However, the proposal does not explain the overall divergent behavior of *koj* ‘who’ and *što* ‘what’, especially as seen in partitives. I have been unable to answer why that should be the case and can only speculate that the differences in their patterning with respect to the clitics is a reflex of deeper semantic differences between the two elements.

¹⁶ It should be noted that lack of WCO violations with CLLDed wh-phrases shows the possibility of base-generation of these elements. As with the strong DPs, given that the clitic with *koj* *N* ‘which N’ phrases is obligatory, we have no evidence to rule out movement as a possibility. In absence of evidence to the contrary, I would have to assume that it is. Note, though, that the presence of the clitic in such cases would also be a result of the Clitic Criterion.

It is worth knowing that difference between *who* and *what* have been noted to exist elsewhere.¹⁷ Dayal & Schwarzschild (2010) point out that the two wh-words pattern differently in English sluices: *what* is consistently disallowed as a wh correlate when the inner antecedent of a sluice contains a head noun; *who* is more flexible in this respect in that it is sometimes allowed as a wh correlate with a headed inner antecedent.¹⁸ The difference between the two wh-words is illustrated in (28a-b), respectively.¹⁹

- (28) a. *Joan was eating a doughnut. Fred didn't know what.
 b. At the LSA, Joan was talking to a philologist, but I don't know who (exactly).

Dayal & Schwarzschild suggest that the differences between *who* and *what* can be attributed to differences in their level of individuation (the latter referring to the manner of selecting or describing the elements in the domain; Dayal & Schwarzschild (2010:11)). They point out that questions with *who* allow for a greater level of individuation than questions with *what*. This, they show, is reflected in the fact that a question like *Who did you speak to?* can be answered by either *I spoke to a lawyer* or *I spoke to John*. The two answers, they explain, can pick out the same individual in a particular context; what they differ on is the level of individuation, with the latter being the more precise. A question like *What did you eat?*, they point out, can only be answered by *I ate a doughnut*, which means that a second level of individuation is not available here. Dayal & Schwarzschild show that the sensitivity to the level of individuation of *who* and *what* is reflected in their

¹⁷ I am grateful to Veneeta Dayal for bringing to my attention the differences between *who* and *what* in English sluices and in particular their different sensitivity with respect to individuation.

¹⁸ For the details of the analysis of the behavior of wh-words in sluices see Dayal & Schwarzschild (2010).

¹⁹ (28a-b) correspond to examples (22) and (35) in Dayal & Schwarzschild (2010), respectively.

interplay with *exactly*. Thus, the statement in (29a) can be followed by the question in (29b), but the statement in (30a) cannot be followed by the question in (30b). (From Dayal & Schwarzschild (2010: 11-12))

- (29)a. I spoke to a lawyer.
 b. Tell me exactly who you spoke to/Exactly who did you speak to?
 c. I spoke to Joe Smith.
- (30)a. I ate a doughnut.
 b. #Exactly what did you eat?

We may speculate that the differences between *koj* ‘who’ and *što* ‘what’ in Macedonian are possibly driven by these properties as well. The fact that *koj* allows for a greater level of individuation may explain the patterning between *točno* ‘exactly’ and *koj* ‘what’ in (31b-d).

- (31)a. Ana went to three conferences last year. She said she saw one of your students there.
- b. #Ti kaža kogo videla?
 you told whom saw
 ‘Did she tell who she saw?’

- c. Ti kaža kogo točno videla?
 you told whom exactly saw
 ‘Did she tell you exactly who she saw?’
- d. Ti kaža kogo (točno) go videla?
 you told whom exactly him saw
 ‘Did she tell you (exactly) who she saw?’

In (31) we see that the discourse in (31a) cannot be the trigger for the question in (31b) where *kog* occurs without a clitic. The questions in (31c-d) are, however, appropriate in this context, which seems to suggest that the CLLDed version of *kog* perhaps has the same contribution as *exactly* in that it requires a more individuation than the answer already available in the discourse.

A detailed investigation of the correlation between *exactly* and *who* in Macedonian and its subsequent role in fixing the degree to which *who* and *what* allow for individuation is beyond the scope of this study. My goal here was to point out that the two wh-phrases seem to differ in more respects than just their ability to co-occur with clitics. The extent to which their intrinsic properties determine their overall behavior in the language will have to await further studies.

In concluding this section, I would like to return to Comorovski (1996) who has pointed out that *who* in Romanian also shows divergent behavior with respect to clitics. Comorovski (1996:73) notes that the only wh-phrases that allow for clitics in the language are *which* Ns. However, she points out that *who* seems to optionally allow for a clitic when it is used in echo questions. (A non-echo *who*, she stresses, does not allow for

clitics; *ibid*). Comorovski goes on to explain that the clitic with echo-*who* in Romanian is only allowed if the echoed NP in the preceding discourse carries an existential presupposition, i.e. it is clitic-doubled itself.²⁰

As we have seen, Macedonian is different in this respect in that *koj* ‘who’ can co-occur with a clitic in regular questions. Thus, the occurrence of the clitic in this case cannot be restricted to echo questions. It is interesting to note, however, that the clitic with echo-*who* in Macedonian to a certain degree seems to correlate with the properties of the echoed NP, as Comorovski argues for Romanian. Consider (32a) and (33a) and the echo questions that follow in (32b) and (33b).²¹

(32)a. Go vidov profesorot po biologija.

him saw professor-the of biology

‘I saw the biology professor.’

b. Kogo (go) vide?

who him saw

‘Who did you see?’

(33)a. Vidov germanski turist vo muzejot.

saw German tourist in museum-the

‘I saw a German tourist in the museum.’

²⁰ Comorovski (1996:73) defines clitic doubling in Romanian to be a property of presupposing NPs only.

²¹ These would be pronounced with a raising intonation on the *wh*-phrase. In presenting the echo questions in (32b) and (33b) my intention is only to highlight the difference between the CLLDed and fronted *koj* ‘who’ in contexts where the echoed NP does or does not carry an existential presupposition. A detailed investigation of the properties of echo questions in Macedonian is beyond the scope of this study.

- b. Kogo (*go) vide?
 who him saw
 ‘Who did you see?’

The discourse in (32a) can serve as a trigger for both the CLLDed and fronted *koj*, but the discourse in (33a) can only serve as a trigger for the fronted *koj*. This means that CLLDed *koj* ‘who’ in echo questions requires that the echoed NP is a presupposing one; *koj* ‘who’ without a clitic, on the other hand, does not seem to be subject to such restrictions. Putting details aside, Comorovski’s observation is relevant to the discussion here in that it highlights the fact that *who* does behave differently from all other non-D-linked wh-phrases in Romanian, which is exactly what we see in Macedonian.

Thus, we have seen from Dayal & Schwarzschild that *who* and *what* pattern differently in English sluices and we have seen from Comorovski that *who* patterns differently from all other non-D-linked wh-phrases when it comes to clitics in echo questions. We can conclude that although the patterning of Macedonian *who* and *what* with respect to clitics looks surprising, it appears to fit within the observed differences between these two elements cross-linguistically.

5.7 Conclusion

I have shown that Macedonian distinguishes between three types of wh-phrases depending on their inherent DL properties. On the view that the specification of their inherent DL properties is tied to partitivity, I argued that *which* *N*s are marked as [+strong_{WH}], *what* and *what N* are marked as [-strong_{WH}] and *who* is unspecified for

strength. The distribution of the clitics with *wh*-phrases then follows from the Clitic Criterion. The section also presented evidence from WCO effects and syntactic islands in support of the claim that *wh*-phrases in Macedonian can either be base-generated or moved to the clause-initial position.

In the next chapter, the differences between the two types of *wh*-phrases, those that can occur with clitics and those that cannot, will be shown to extend to their interaction with other elements in the clause.

CHAPTER 6

Wh-Expressions at the Left Periphery

6.0 Introduction

The previous chapter explored the clitic co-occurrence patterns in the context of wh-questions and it argued that the presence of clitics with wh-phrases is determined by the feature specification of the wh-element. In this chapter, I would like to look at another aspect of the wh-constructions, namely the position of the wh-phrases.

The analysis in the previous chapter assumed that the wh-phrases occupied the SpecCP position. Here I would like to look at this question a bit more closely. We already know from WCO and islands effects that Macedonian *who/what N/who* (without a clitic) behave differently from *which N/who* (with a clitic). Their different patterning was explained in terms of their derivation: the former move to the left periphery, the latter are base-generated there. In this chapter, I will show that the wh-phrases behave differently with respect to other elements in the clause, and in particular subjects. To explain the different patterning, I will argue that the wh-phrases occupy different positions at the left periphery.

As a way of concluding our discussion of the Macedonian direct objects, clitics and the left periphery, I will also explore to what extent our understanding of the wh-constructions can help us answer some of the open questions about the non-wh-constructions discussed in this study.

The chapter is organized as follows. Section 1 outlines the basic patterns of object wh-phrases with respect to the subjects in wh-questions in Macedonian. Section 2

addresses the question of the placement of verbs in wh-questions and it offers an explanation for the subject intervention effects by making the argument that the two types of wh-phrases occupy different positions in the clause. Section 3 looks at the question of whether optional D-linking influences the positioning of the wh-phrases. Section 4 brings the discussion back to the issues discussed earlier in the dissertation by exploring three open questions. Section 5 contains the conclusion.

6.1 Further Differences between D-Linked and Non-D-Linked Wh-Phrases

We have seen that in Macedonian *koj N* ‘which N’ phrases and CLLDed *koj* ‘who’ crucially differ from *što/kakov N* ‘what/what N’ and fronted *koj* ‘who’ phrases in that they can be base-generated in their surface position. That the two types of phrases pattern differently is further supported by their behavior with respect to other elements in the clause. In what follows, I outline the interaction of wh-phrases with respect to the subjects. I then explain the observed patterns by proposing that the two types of wh-phrases occupy different positions in the clause. The position they come to occupy will reflect their intrinsic D-linked properties.

6.1.1 Subject Intervention Effects in Wh-Questions

Evidence for the claim that the two types of wh-phrases occupy different positions in the clause comes from what I call Subject Intervention Effects (SIE). These are found in single constituent questions and are cases where the subject intervenes between a fronted wh-phrase and the verb. Let us see how [+strong], [-strong] and unspecified wh-phrases in Macedonian pattern in this respect.

In questions with [-strong] wh-phrases like *što* ‘what’, a subject can either follow the verb (1a), or precede the fronted wh-phrase (1b). Crucially, the subject cannot intervene between the fronted wh-phrase and the verb (1c).

- (1) a. *Što kupi Petar?*
 what bought Petar
 ‘What did Petar buy?’
- b. *Petar, što kupi?*
 Petar what bought
- c. **Što Petar kupi?*
 what Petar bought

The most natural way of asking a question in Macedonian is as in (1a), with a reversed order of the verb and the subject. (1b), where the subject precedes the fronted wh-phrase, is equivalent to the English ‘(As for) Petar, what did he buy?’ In effect, the fronting of the subject in (1b) is an example of topicalization, which is reflected in the fact that it is accompanied by a slight change in intonation, i.e. a comma intonation.¹ To reiterate, the clitic is absent in (1a-c).

¹ At the moment, I can only note that such questions have a different prosodic pattern, but I cannot offer any formal analysis of them. It should be noted that the intonational patterns of the wh-questions may very well prove to be of crucial importance in separating and identifying the relevant components of these constructions (e.g. deciding the topic or focus status of a particular element). I leave the role of intonation in wh-questions for further research.

In questions with [+strong] phrases, like *koja knjiga* ‘which book’ in (2a-c), the subject can follow the verb (2a) or it can precede the fronted wh-phrase (2b). Interestingly, the subject can also intervene between the wh-phrase and the verb (2c).²

- (2) a. *Koja knjiga ja kupi Petar?*
 which book it(f.sg.) bought Petar
 ‘Which book did Petar buy?’
- b. *Petar, koja knjiga ja kupi?*
 Petar which book it(f.sg.) bought
- c. *Koja knjiga Petar ja kupi?*
 which book Petar it(f.sg.) bought

As it was the case with (1b), I assume that (2b) is a construction which involves a topicalization of the subject. In this case, too, there is an intonation break between the subject and the fronted wh-phrase. In addition, the set of examples in (2a-c) displays the now familiar presence of the clitic *ja* ‘it’, co-indexed with the direct object *koja knjiga* ‘which book’.

The pattern we see in (1a-c) is also found with *kakov N* ‘what N’ phrases. Here, too, the subject cannot occur between the fronted wh-phrase and the verb. Examples are given in (3a-c).

² I should note that, for some speakers, (2c) may not be perfectly acceptable (at least when compared with (2a)). Even so, the main point here is that (2c) is definitely better than (1c). This is the intuition that the analysis tries to capture.

- (3) a. Kakva kniga kupi Petar?
 what book bought Petar
 ‘What book did Petar buy?’
- b. Petar, kakva kniga kupi?
 Petar what book bought
- c. *Kakva kniga Petar kupi?
 what book Petar bought

Finally, *koj* ‘who’ completes the picture outlined above in that it behaves differently when it co-occurs with a clitic and when it does not. A CLLDed *koj* ‘who’ patterns just like *koj N* ‘which N’ in that it allows for an intervening subject, as shown in (4c).

- (4) a. Kogo go vide Petar?
 who him saw Petar
 ‘Who did Petar see?’
- b. Petar, kogo go vide?
 Petar who him saw
- c. Kogo Petar go vide?
 who Petar him saw

A fronted *koj* ‘who’, on the other hand, patterns like *što* ‘what’ and *kakov N* ‘what N’ phrases and does not allow for the subject to intervene between it and the verb, as illustrated in (5c).

- (5) a. Kogo vide Petar?
 who saw Petar
 ‘Who did Petar see?’
- b. Petar, kogo vide?
 Petar who saw
- c. *Kogo Petar vide?
 who Petar saw

To summarize, the subject can intervene between a wh-phrase and the verb only when the wh-phrase co-occurs with a clitic, as in the case of [+strong] and unspecified wh-phrases in (2c) and (4c). The subject cannot intervene between a wh-phrase and the verb if the wh-phrase is fronted, as in the case of [-strong] wh-phrases in (1c) and (3c), as well as the unspecified wh-phrase in (5c). All wh-phrases can be preceded by the subject.

6.2 An Analysis

This section argues for an analysis of SIE which relies on the idea that the derivation of questions is sensitive to the DL feature of the wh-phrases. Given that the verb plays a role in the construction of the wh-constructions, I begin the discussion by addressing the question of its exact placement in the clause.

6.2.1 Adverbs in Wh-Questions

When considering the placement of the verb in wh-questions, there are two possibilities open, namely that the verb is in I or in C. In what follows, I will show that in

Macedonian, the verb in *wh*-questions raises to C^0 .³ Evidence for this comes from the placement and interpretation of adverbs. To bring out these differences, I will consider two types of adverbs: those that allow for subject-oriented and predicate readings, and those that only allow for subject-oriented reading.

It has been noted that some adverbs are ambiguous between a subject-oriented and a predicate reading (see McConnell-Ginet (1982), Izvorski (1995), Bošković (2004), Fanselow (2004), among others). An example of this is given in (6) for English.

(6) John correctly filled out the documents.

a reading: John was correct/right to fill out the documents.

b reading: John filled out the documents in a correct way/manner.

The same seems to hold for adverbs in Macedonian. In (7c), the adverb *pravilno* ‘correctly’ has both a subject-oriented and a predicate reading.

- | | | | | | |
|--------|-------|-----------|--------------------------------|----------------------|-------------------------------|
| (7) a. | Petar | gi | popolni dokumentite | pravilno. | b reading only |
| | Petar | them | filled documents-the | correctly | |
| b. | Petar | gi | popolni pravilno dokumentite. | | b reading only |
| | Petar | them | filled correctly documents-the | | |
| c. | Petar | pravilno | gi | popolni documentite | a and b reading |
| | Petar | correctly | them | filled documents-the | |

³ This in contrast to Bulgarian, for example, where it has been argued that the verb moves only up to I (see Izvorski 1995).

- d. Pravilno, Petar gi popolni dokumentite. **a** reading only
 correctly Petar them filled documents-the

The examples in (7) shows that the adverb has a predicate reading when it is within the VP (7a-b). When the adverb is outside of the VP, the subject-oriented reading also becomes available (7c-d). The availability of both readings in (7c) then can be a result of a structural ambiguity. The predicate (**b**) reading arises when the adverb is adjoined to VP; the subject-oriented (**a**) reading arises when the adverb is adjoined to some higher functional projection in the IP domain.⁴

- (8) a. [_{IP} Subj [_{VP} V Obj Adv]] **b** reading (7a)
 b. [_{IP} Subj [_{VP} V Adv Obj]] **b** reading (7b)
 c. [_{IP} Subj [_{AgrP/VP} Adv [_{VP} V Obj]] **a** and **b** reading (7c)
 d. [_{IP} Adv [_{IP} Subj [_{VP} V Obj]] **a** reading (7d)

Extending the discussion to *wh*-questions gives interesting results. As (9a) illustrates, when the adverb is within the VP, we only get the predicate reading. In (9b), though, the subject-oriented reading becomes available. This suggests that the verb is in C^0 , leaving the adverb and the subject in the IP (with the adverb being adjoined to IP). The predicate reading in (9b) would be a result of the adverb being adjoined to VP with the subject

⁴ In principle, there could be two possibilities here. One option is for the adverb to adjoin to IP, in which case the subject would be placed higher up in the structure (adjoined to IP or some higher TopP position), thus become topicalized. If this were the case, though, we would expect a difference in intonation. The second option is for the adverb to adjoin to some functional position in the IP (e.g. AgrP). This option enables us to place the adverb in a position higher than that of the VP, without having to topicalize the subject. Having said this, the adverb would have to be able to adjoin to IP (as in (7d) and (9b))) or even higher (as in (10b)), giving rise to subject-oriented reading readings.

being VP internal. Thus, I conclude that in wh-questions, the verb moves to a position higher than I^0 , i.e. it moves all the way to C^0 . The structural analysis of (9a-b) is given in (10a-b).

- (9) a. Što popolni Petar pravilno? **b** reading only
 what filled Petar correctly
 ‘What did Petar correctly filled out?’
- b. Što popolni pravilno Petar? **a** and **b** reading
 what filled correctly Petar

- (10)a. [_{CP} wh _C V_i [_{IP} Subj [_{VP} Adv [_{VP} ... t_i ...]]]
- b. [_{CP} wh _C V_i [_{IP} Adv [_{IP} Subj [_{VP} ... t_i ...]]]]

To complete the paradigm, we should note that the placement of the adverb between the fronted wh-phrase and the verb seems to result in ungrammaticality; see (11a). Its placement before the wh-phrase, however, is acceptable; see (11b). The structures for both (11a-b) are given in (12a-b), respectively.

- (11)a. *Što pravilno popolni Petar?
 what correctly filled Petar
 ‘What did Petar correctly filled out?’
- b. Pravilno, što popolni Petar? **a** reading only
 correctly what filled Petar

- (12) a. $*[_{CP} \text{ wh Adv } _C V_i \quad [_{IP} \text{ Subj } [_{VP} \dots t_i \dots]]]$
 b. $[_{XP} \text{ Adv } [_{CP} \text{ wh } _C V_i [_{IP} \text{ Subj } [_{VP} \dots t_i \dots]]]]$

The ungrammaticality of (11a) gives further support to the claim that the verb is in C. Its placement and that of the *wh*-phrase in SpecCP, makes it impossible for the adverb to occur between these two elements.

Compared to *što* ‘what’ in (10a-b), *koj tekst* ‘which text’ behaves differently in that it allows for the adverb to occur between the *wh*-phrase and the verb; compare (10a) with (13a).⁵ In both (13a-b), the adverb has the subject-oriented (**a**) reading.

- (13) a. Koj tekst pravilno go pročita Petar?
 which text correctly it(m.sg.) read Petar
 ‘Which text did Petar correctly read?’
 b. Pravilno, koj tekst go pročita Petar?
 correctly which text it(m.sg.) read Petar

This supports the initial observation that the subject-oriented reading is linked to the placement of the adverb in positions higher up in the IP domain. Given the view that *wh*-questions involve movement of the verb to C, the grammaticality of (13a-b) is even more

⁵ *What N* phrases, on the other hand, pattern like *what* in (11a-b):

- (i) a. *Kakov tekst pravilno pročita Petar?
 what text correctly read Petar
 ‘What text did Petar correctly read?’
 b. Pravilno, kakov tekst pročita Petar?
 correctly what text read Petar

important in that it shows that *which* *Ns* in Macedonian must occupy a different position from that of *what*.

Let us now look at adverbs that do not allow for the low, predicate reading. These adverbs offer additional evidence that the verb raises to C. Consider the placement of adverbs like *navodno* ‘allegedly’, which are placed at the top end of the hierarchy of adverbs in Cinque (1999), and subsequently the top end of the clause. That this is the case with *navoldno* ‘allegedly’ in Macedonian is illustrated in (14a-d).

- (14)a. *Petar gi popolni dokumentite navodno
 Petar them filled documents-the allegedly
 ‘Allegedly, Petar filled out the documents.’
- b. *Petar gi popolni navoldno dokumentite
 Petar them filled allegedly documents-the
- c. Petar navodno gi popolni dokumentite
 Petar allegedly them filled documents-the
- d. Navodno, Petar gi popolni dokumentite
 allegedly Petar them filled documents-the

The ungrammaticality of (14a-b) can be attributed to the fact that the adverb is in a position within the VP. Given that this position is associated with the predicate reading of adverbs, something that *allegedly* does not allow for, its placement in such a position is incompatible with its interpretation. Along the same lines, the grammaticality of (14c-d) can be attributed to the fact that the adverb occupies a position above VP within IP (in

(14c)) or outside of IP (in (14d)). From (14a-d) we can conclude that *navodno* ‘allegedly’ must occupy at least a position in IP (i.e. outside VP).⁶ *Allegedly*, thus, makes the same point as *correctly* except that it does not have the predicate reading.

The interaction of adverbs like *navodno* ‘allegedly’ with wh-phrases is not only relevant in that it provides supporting evidence for V-to-C movement, but it also shows that the adverbs are sensitive to the type of wh-phrases. Thus, when such adverbs are included in wh-questions, we see an interesting pattern developing. In questions with mono-morphemic *što* ‘what’, as in (15a), the adverb cannot occur between the wh-phrase and the verb. In questions with *koj N* ‘which N’ phrases, as in (15b), the adverb can occur between the wh-element and the verb.

- (15) a. *Što navodno popolni Petar?
 what allegedly filled Petar
 ‘What did Petar allegedly fill out?’
- b. Koi dokumenti navodno gi popolni Petar?
 which documents allegedly them filled Petar
 ‘Which documents did Petar allegedly fill out?’

The different patterning of the adverb in (15a-b) is further evidence for the placement of the verb in C. In (16a-b) we see that the adverb can follow the verb, which is consistent

⁶ We can assume that in (14c) the adverb is adjoined to AgrP or ClP, while in (14d) it is adjoined to IP or some other functional projection higher than IP, e.g. TopP.

with the analysis that such adverbs occupy high positions within IP, i.e. above VP; see (14c-d).⁷

- (16) a. Što popolni navodno Petar?
 what filled allegedly Petar
 ‘What did Petar allegedly fill out?’
- b. Koi dokumenti gi popolni navodno Petar?
 which documents them filled allegedly Petar
 ‘Which documents did Petar allegedly fill out?’

Having seen that adverbs like *navodno* ‘allegedly’ give additional support for the claim that the verb in wh-questions in Macedonian moves to C, let us go back to (15a-b) above. As noted earlier, the interesting fact about these examples is that the two types of wh-phrases *što* ‘what’ and *koi dokumenti* ‘which documents’ interact differently with the adverb in that the adverb cannot occur between *what* and the verb, but it can occur between *which N* and the verb. In the next section, I offer an explanation for this puzzling interaction between the adverbs and the wh-phrases.

To summarize, adverb placement tests offer proof that wh-questions in Macedonian involve raising of the verb to C. This fact will play a crucial role in determining the

⁷ That such adverbs indeed occupy high positions in the clause in Macedonian is also supported by (i) and (ii) below. The unacceptability of (i-ii) can be attributed to the placement of *navodno* ‘allegedly’ in the VP.

- (i) *Što popolni Petar navodno?
 what filled Petar allegedly
 ‘What did Petar allegedly fill out?’
- (ii) *Koi dokumenti gi popolni Petar navodno?
 which documents them filled Petar allegedly
 ‘Which documents did Petar allegedly fill out?’

positions of the wh-phrases, which in turn will help us explain the subject intervention effects observed in section 6.1.1 above.

6.2.2 An Explanation for Subject Intervention Effects

We saw that in single constituent questions the subject can intervene between a fronted wh-phrase and the verb if the wh-phrase is D-linked, as in (2c), but that the subject cannot intervene between a fronted wh-phrase and the verb if the wh-phrase is non-D-linked, as in (1c). As a reminder, the descriptive generalizations we aim to explain are given in (17a-b) below:

(17)a. DL wh Subj V

b. *NDL wh Subj V

One striking fact about the patterns in (17a-b) is that D-linked wh-phrases seem to behave exactly like CLLDed and topicalized DPs with respect to the subject. As it is evident from all the data that we have considered so far, CLLD and Topicalization allow for the subject to occur between the dislocated element and the verb. Added to this is the fact that the wh-phrases are also at the left periphery. Putting aside the distribution of the clitics with wh-phrases, which as we saw in chapter 5 is regulated by the Clitic Criterion (same as the DPs), the derivation of the structures in (17a-b) are subject to the Wh Criterion. We can suppose that its requirements could be satisfied by having the wh-phrases in SpecCP at least at some point in the derivation. With this in mind, one possible way of explaining the patterns in (17) above is to assume that V-to-C movement takes

place in (17b) but not in (17a). If this were the case, it would explain the word orders in both cases, and it would make (17a) compatible with CLLD and Topicalization which also lack V-to-C movement. Given the fact that the Wh Criterion is not an issue there, V-to-C movement is not expected to apply at all. However, there is an obvious problem with this proposal which is why D-linking should affect the application of V-to-C movement. I will instead propose a different solution, one which crucially relies on the availability of different positions for wh-phrases at the left periphery.

Recall our earlier discussion of the placement of adverbs in wh-questions. Adverb placement tests showed that the verb raises to C, both in questions with D-linked and non-D-linked wh-phrases. Given this, we can assume that in cases where the subject follows the verb, the subject occupies a SpecIP position.⁸ Added to this, we can further assume that the word order in (1b) and (2b), where the subject precedes the fronted wh-phrase, the subject is topicalized.

I would like to extend this analysis to the data in (1c) and (2c) and suggest that the subject can be topicalized even when it follows the wh-phrase.⁹ This, however, happens in highly restricted contexts, namely only in questions with *which N* phrases. The question then is why topicalization of the subject is possible in questions with DL wh-phrases (as in (2c)), but not allowed in questions with NDL wh-phrases (as in (1c))? To answer this question, we must look at the properties of the wh-phrases themselves. I would like to suggest that, given their D-linking properties, complex *which N* phrases

⁸ Another possibility would be to have the subject in the VP internal position. That this is an option is supported by examples like (9b) where a predicate reading of the adverb *correctly* is possible even when it appears before the subject.

⁹ This is similar to King's (1995) analysis of the word order patterns in Russian. A similar proposal is also made by Alexiadou & Anagnostopoulou (1998) who argue that preverbal subjects in Null Subject Languages are CLLDed.

behave like topics in the language, similarly to CLLDed elements (see also Krapova 2002, Krapova & Cinque 2003), so when they interact with other topics in the clause, intervention effects are suspended. In this case, I maintain that DL wh-phrases occupy a TopP position above CP.¹⁰ NDL wh-phrases, on the other hand, are not topics, in the sense that they do not refer to objects given in the context. Subsequently, they move to a different position than that of topics. Following standard analyses of wh-questions, I assume that NDL wh-phrases move to SpecCP. Subject intervention effects can then be explained in terms of the placement of the different wh-phrases and the position of the verb in C. The advantage of this analysis is that it also explains the patterning of adverbs and wh-phrases in (15a-b) above. Recall that there the adverb *navodno* ‘allegedly’ could not occur between a fronted *što* ‘what’ and the verb but it could occur between CLLDed *koi dokumenti* ‘which documents’ and the verb. The puzzling behavior can be explained if we assume that *which* Ns in Macedonian occupy positions higher than those of *what*.

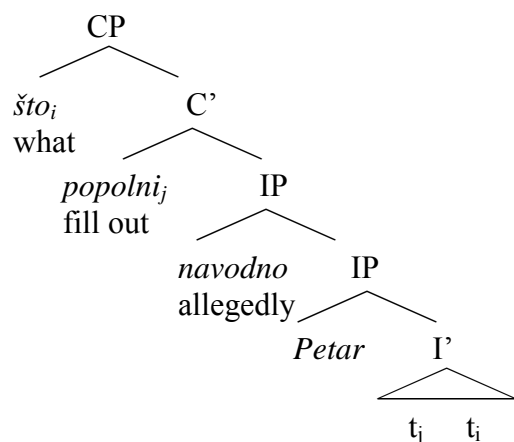
¹⁰ The proposal that DL wh-phrases move to a different position from NDL wh-phrases can be found in Richards (1997). Richards holds that in multiple fronting languages, NDL wh-phrases move to the SpecCP position (his analysis crucially relies on the existence of multiple specifiers). DL wh-phrases, he notes, may move to a topic position above CP. It should be noted that in his analysis movement of DL wh-phrases to topic positions is not obligatory. Their movement is to SpecCP, as is the case of NDL wh-phrases. Richards suggest that DL wh-phrases may have the ability to (subsequently) move to topic positions. It is this ability, he notes, that may be responsible for the different behavior of DL vs. NDL wh-phrases with respect to superiority effects.

Dayal (1996) also entertains the possibility of DL wh-phrases moving to a position above CP (in her analysis the position is marked as CP*). Dayal’s proposal is based on her findings of the parallels that hold between DL questions and echo questions. What is of particular interest to us is Dayal’s observation that “operations associated with CP* are triggered by discourse-sensitive elements such as echo or D-linked wh expressions” (p.130). The proposal is directly applicable to our analysis of wh-questions containing DL wh-phrases.

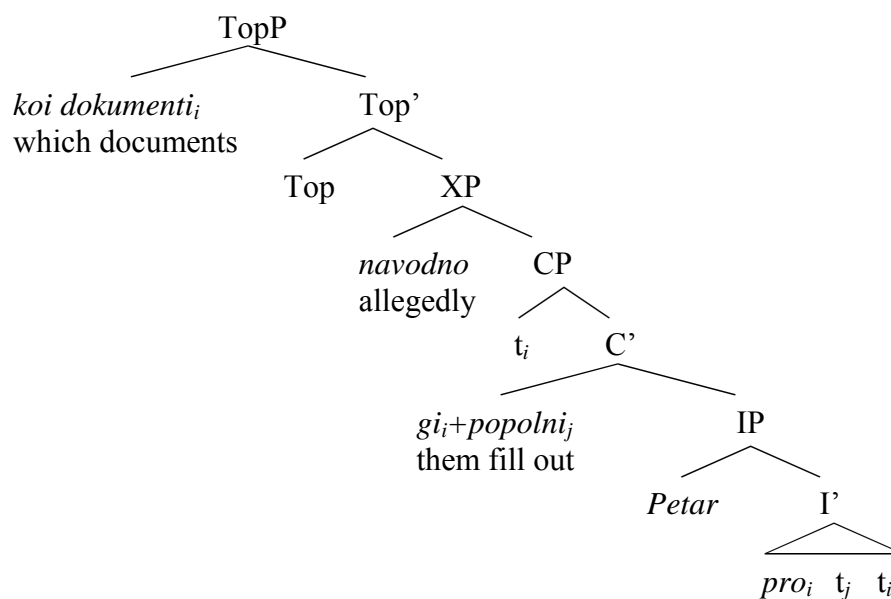
Dayal’s proposal is also relevant for our analysis because of the possibility that the CP* position may or may be accessed by movement (at LF). This is interesting for the proposal here because the TopP position is made available to both elements that move or are base-generated there. The question of what level these operations take place at (LF or in overt syntax) is perhaps the next logical step in the investigation of the phenomena that target the left periphery and I leave it open for further research.

The representations of the relevant parts of the structures in (16a) and (16b) are given in (18a-b), respectively.¹¹

(18)a.



b.



¹¹ The adverb in (18b) can either adjoin to CP or a TopP position. Another possibility would be a special function position, such as Mood_{evidential} (Rizzi 2004, following Cinque 1999). For our purposes what is relevant is that given the position of the *which N* phrase in (16b), the adverb can intervene between it and the verb. The question of how adverbs fit into the clausal structure in Macedonian (and especially the question of whether they adjoin or occupy specifier positions of adverbial phrases) is left open for further research.

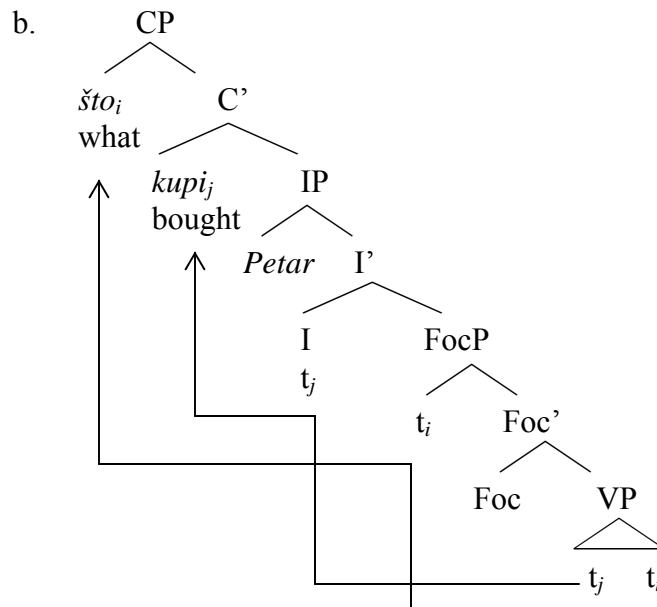
Having seen how the new proposal accounts for the interaction between *wh*-phrases and adverbs, let us see how it helps us explain the paradigms of subject intervention effects in (1a-c) and (2a-c). Although the discussion below concentrates on *what* and *which* *Ns*, the explanation is understood to hold for all fronted and CLLDed *wh*-phrases.

Consider (1a), repeated here as (19a). In (19a), the object *wh* moves to SpecFocP, to value its (intrinsic) focus features (following Bošković (2002), etc.). It then moves to SpecCP, to value the *wh* feature on C. The verb moves to C and the subject is in SpecIP.¹² The derivation of (19a) is given in (19b).

(19)a. Što kupi Petar?

what bought Petar

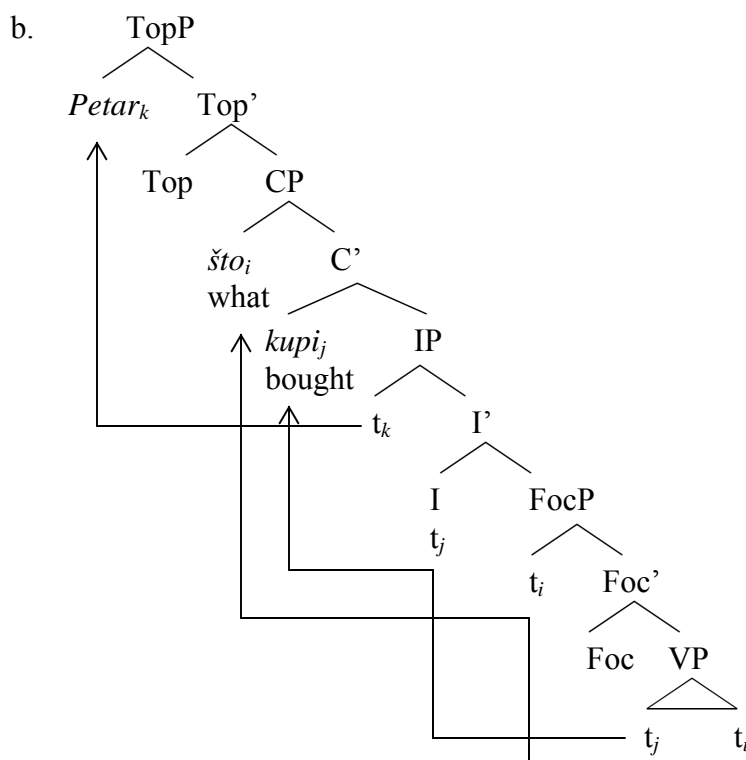
‘What did Petar buy?’



¹² I will assume that the movement of the subject from SpecvP position, for example, is motivated by an EPP feature on I.

We know from (1b), repeated here as (20a), that the subject can also precede the fronted wh-phrase. In such cases, I take the fronting of the subject to be a result of topicalization. The movement, I assume, is to a TopP position above CP (following Rizzi 1997). The movement of the NDL wh remains the same as in (20b) above. The derivation is given in (20b).

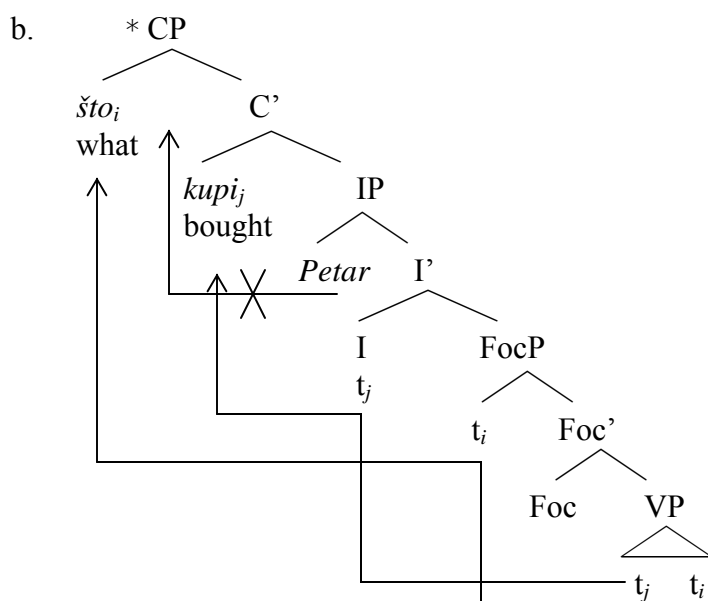
- (20) a. Petar, što kupi?
 Petar what bought
 ‘Petar, what did he buy?’



We are now in a position to explain the ungrammaticality of (1c). This would be a result of the configuration established by the fronted wh-phrase and the verb. The NDL wh-phrase must move to the front of the clause and it does so in the usual manner by moving

to SpecCP. The verb also moves to its C position. This creates a particular configuration that disallows the presence of an intervening element, which in this case is the subject. The derivation of the ungrammatical (1c), repeated here as (21a), is given in (21b).

- (21) a. *Što Petar kupi?
 what Petar bought
 ‘What did Petar buy?’



To summarize, we have seen that the subject intervention effects in questions with NDL wh-phrases arise due to the fact that the fronted wh-elements and the verbs occupy the SpecCP and C positions, respectively. There are only two options for the subject in such cases, and that is either to remain in SpecIP (which is the default case) or to topicalize to a position above CP.

We now turn to the data in (2a-c). The analysis of questions with *which* Ns is built on the observation that such elements do not give rise to WCO effects. Taking this to be

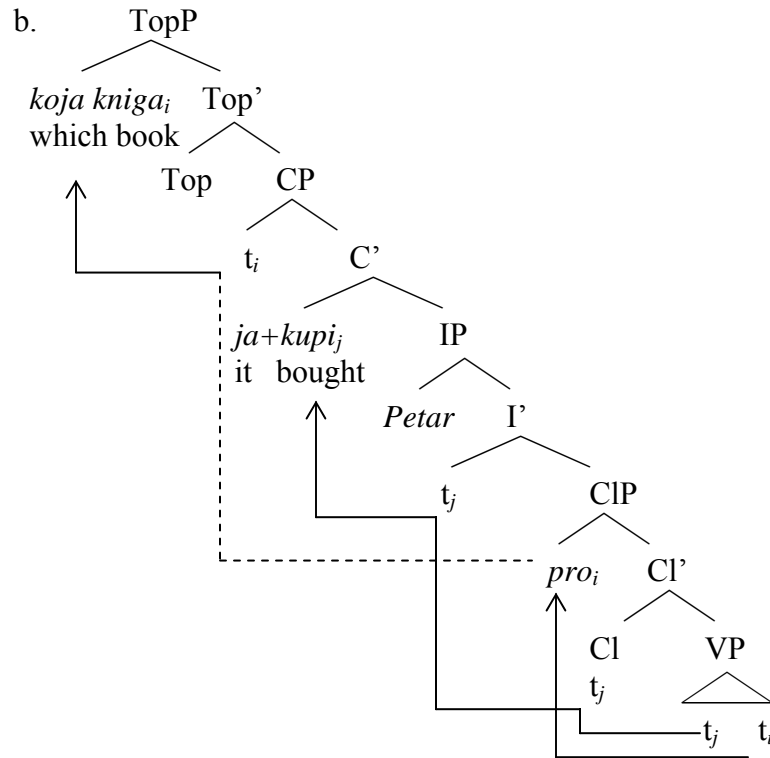
evidence for base-generation, it would mean that such questions would not involve a movement of the *wh*-phrase from the argument position within IP. Instead, the *wh*-phrase would be base-generated and would enter into a binding relation with a *pro* in the argument position.¹³ *Pro* in turn licenses the clitic pronoun (and a ClP). Finally, I take *which Ns* to be base-generated in SpecCP, from which they move to a higher position, TopP. This ensures that the Wh Criterion is satisfied.¹⁴

In (22a), the *wh*-phrase is base-generated in SpecCP and then moves to TopP; the verb undergoes movement to C. The *which N* phrase enters into a binding relation with *pro* in the IP. Given that *which N* phrases are inherently DL and therefore [+strong], the chain that they form with *pro* and the clitic satisfies the requirement that all its members share the same feature. The derivation of (22a) is given in (22b).

- (22) a. Koja kniga ja kupi Petar?
 which book it(f.sg) bought Petar
 ‘Which book did Petar buy?’

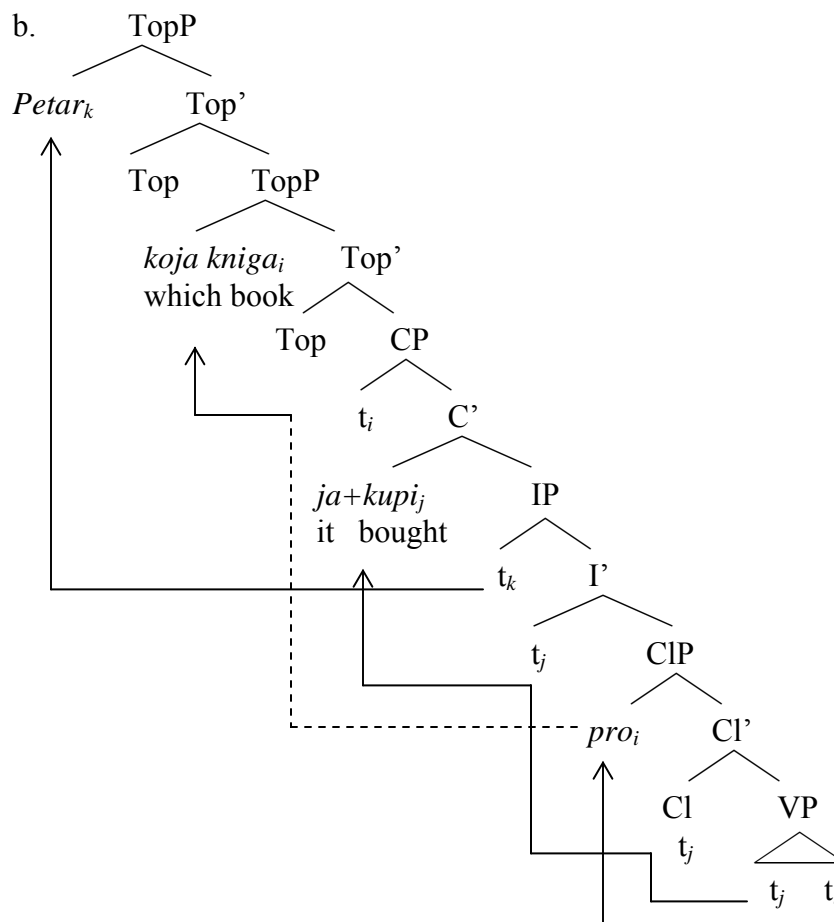
¹³ We should note that the subject intervention facts could also be explained even if *which Ns* undergo movement. In other words, the subject intervention facts are not dependent on the base-generation of the *wh*-phrase but rather on its being in a higher position than that of non-D-linked *wh*-phrases.

¹⁴ Alternatively, the Wh Criterion can be satisfied by the movement of *pro* from the SpecClP to SpecCP. I thank Veneeta Dayal for pointing this option.



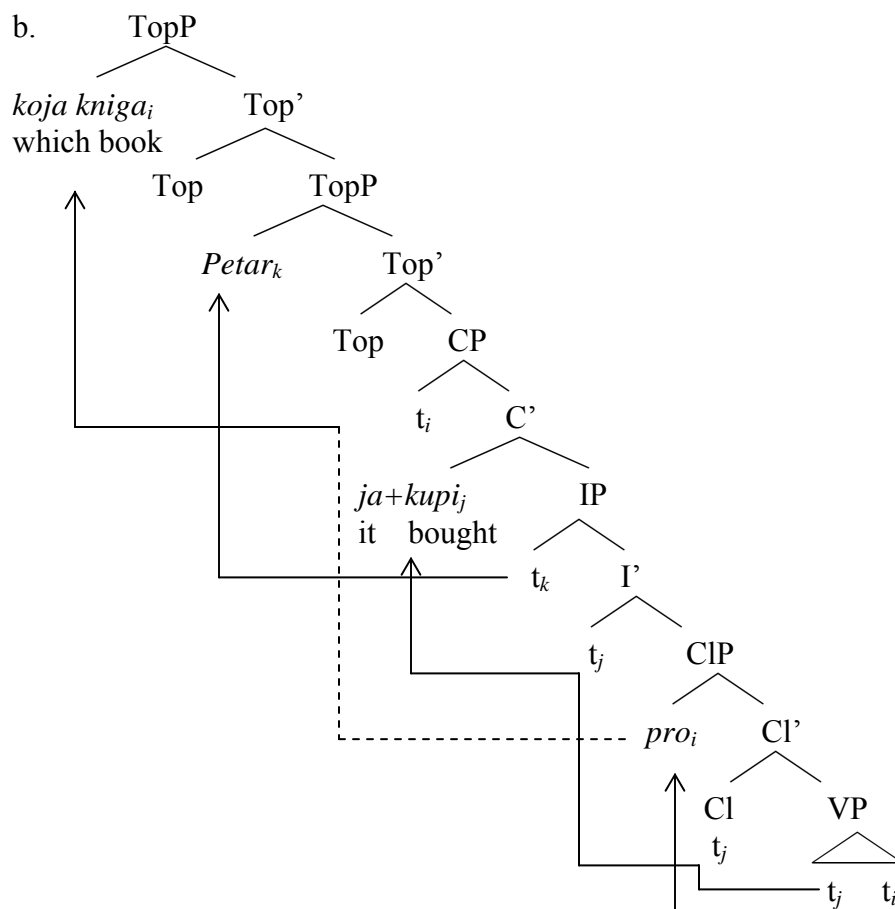
In the case where the subject precedes the wh-phrase (2b), I assume that the subject has moved to a TopP position above that of the wh-phrase. In this case, the derivation of (23a) would be as in (23b).

- (23) a. Petar koja kniga ja kupi?
 Petar which book it(f.sg) bought
 'Which book did Petar buy?'



Finally, the pattern in (2c) can be explained as follows: given that *which Ns* occupy TopP positions above CP, they can either be preceded or followed by expressions occupying other TopP positions. In this case, the subject can not only precede the wh-phrases (as in (23a)), but it can also follow it. This explanation of the subject intervention in questions with *which N* phrases crucially relies on the assumption that the subject in such cases is indeed topicalized. The derivation of the (24a) is given in (24b).

- (24) a. Koja kniga Petar ja kupi?
 which book Petar it(f.sg) bought
 ‘Which book did Peter buy?’



To summarize, we have seen that the lack of subject intervention effects in questions with [+strong] *wh*-phrases is due to the fact that the fronted *wh*-elements and the subject both occupy positions above *CP*. This makes it possible for the *wh*-phrase to be separated from the verb by an intervening subject.

6.3 Optional D-Linking and Subject Intervention Effects

We have already noted the fact that the D-linking of otherwise non-D-linked wh-phrases does not trigger the presence of a clitic (see chapter 5). This was taken as evidence that the distribution of the clitic corresponds to the features for strength of the wh-elements rather than their interpretation. In this section, I would like to show that the change in interpretation (i.e. optional D-linking) does not have an impact on SIE as well.

We have already noted that NDL wh-phrases like *what/what N* can optionally be D-linked. It is interesting to note, however, that their D-linking does not alter their behavior with respect to their interaction with the subject. In other words, the subject intervention effects we saw displayed in questions like (1a-c) are not alleviated in situations where the same set of wh-phrases receive a DL interpretation.

For illustration consider the question in (1a), repeated here in (25a). We noted that the wh-phrase in such case gets a D-linked interpretation (see section 5.4, chapter 5). As (25b) shows, however, the change in the interpretation does not seem to have an effect on the positioning of the subject in the clause. (25c) shows that the subject can precede the fronted wh-phrase regardless of the interpretation it receives.

(25) a. Što kupi Petar?

what bought Petar

‘What did Petar buy?’

b. *Što Petar kupi?

what Petar bought

- c. Petar, što kupi?
 Petar what bought

The same effects are present with *what N* phrases, as illustrated in (26). The positioning of the subject between the fronted wh-phrase and the verb is disallowed under a non-D-linked and a DL interpretation of the wh-element (for example, in the scenario outlined in section 5.4, chapter 5).

- (26) Kakva knjiga kupi Petar?
 what book bought Petar
 ‘What book did Petar buy?’
- b. Petar, kakva knjiga kupi?
 Petar what book bought
- c. *Kakva knjiga Petar kupi?
 what book Petar bought

On the one hand, the behavior of the wh-phrases in (25)-(26) is consistent with their behavior with respect to clitics, because in both cases we see that the patterning is determined by the intrinsic properties of the wh-elements, so their acquired interpretation does not alter their syntactic behavior. Given that the positioning of the subject was used as evidence that *which Ns* and *what* occupy different positions in the clause, we would have to conclude that the optionally DL phrases in (25)-(26) occupy the same position as

their NDL counterparts.¹⁵ Below, I will explore two possible alternatives of how to capture this behavior, but would leave the definitive answer for further research.

One alternative is to maintain that NDL wh-phrases move to SpecCP. I have argued that inherently DL wh-phrases can be base-generated in that position as well. We can speculate the base-generation at the left periphery is compatible with their intrinsic DL properties. After being generated in SpecCP, *which N* phrases move to a position above CP. NDL phrases, however, once they move to SpecCP remain there. How does this solve our original problem? When it comes to the clitics, their distribution would still be a result of the intrinsic properties of the wh-phrases: in the case of the *which Ns*, the clitic would be generated because of strong *pro*; in the case of [-strong] wh-phrases they would not be generated at all. In terms of the positioning of the wh-phrases with respect to the subject, those that remain in SpecCP (the NDL and optionally DL phrases) would force a verb-subject inversion, as usual. Those that move to a higher position (the inherently DL phrases), would optionally allow for the subject to intervene. The ordering of fronting operations would give us the different word order: in cases where the subject topicalizes first and is followed by the *which N* phrase the result would be wh-Subj-V; in cases where the *which N* moves out of SpecCP first and is followed by a topicalization of the subject the result would be Subj-wh-V.

This sort of analysis would also explain the patterning of *koj* ‘who’. Recall that *koj* ‘who’ can either occur with a clitic or without it. In the first case, it patterns like *which N* in that it does not give rise to WCO effects and can violate weak islands. In the second case, it patterns like *što/kakov N* ‘what/what N’ and gives rise to WCO violations and

¹⁵ Note also that the optionally DL wh-phrases differ from topicalized DPs as well in that the latter consistently allow for the subject to intervene between the fronted element and the verb. This is again supported by all the data of Topicalization presented here.

obeys weak islands. In the first instance, *who* would be generated in SpecCP. In the second instance, *who* would move to this position. We can suppose that the base-generation at the left periphery is compatible with the DL interpretation that it has. Moreover, in the case of *koj* ‘who’ this kind of analysis might force us to conclude that there are indeed two *who* wh-phrases in Macedonian, one D-linked like *which N*, one non-D-linked like *what/what N*. The clitic would correlate with the intrinsic properties of each: it will be triggered by the first one because of its [+strong] features, but will be absent in the latter case because of the wh’s [-strong] features.

Another alternative would be to maintain the original proposal that the DL and NDL wh-phrases occupy different positions in the clause, TopP and SpecCP, respectively.¹⁶ In this case, the Wh Criterion would be satisfied by the movement of *pro* from CIP to SpecCP, so base-generation of *which N* would not be ruled out because of it. Additionally, wh-phrases that receive a DL interpretation could come to occupy a yet different position from either of the two, for example a position adjoined to CP. We could speculate that such positioning would be compatible with (or due to) their DL interpretation (cf. Dayal 1996).¹⁷ At the same time, their being adjoined to CP would force the topicalization of the subject to be to a position above CP (TopP), thus crucially blocking the subject from intervening between the wh-phrase and the verb. One might even suppose that an optionally D-linked wh-phrase in such cases moves to some other functional projection (e.g. FocP; as in Rizzi 1997). (The positing of subsequent

¹⁶ Iatridou (1995) analyzes CLLDed elements in Greek as base-generated, in a position adjoined to CP. She calls this position the “DL position” and restricts it to elements that have been mentioned in previous discourse.

¹⁷ Krapova (2002) and Krapova & Cinque (2003) argue that DL and NDL wh-phrases occupy different positions (the latter are in SpecCP). Furthermore, DL wh-phrases that are clitic doubled occupy a position different to that occupied by non-clitic-doubled DL wh-phrases. The former occupy a CLLDTopP position, the latter an Operator TopP position. Note that the DL wh-phrases in both cases are inherently DL. Macedonian is different in that it obligatorily doubles the *which N* phrases.

movement on inherently and optionally D-linked wh-phrases would mean that the DL interpretation crucially is associated with positions above the CP.) We can further speculate that in such case the subject inversion is still forced by the Wh Criterion, in that the verb would have to move to Foc, where the wh feature would be checked under a Spec-Head agreement.

All of this, however, remains, speculative. It seems clear though that a solution for the problem would have to take into consideration the nuanced interplay between the interpretation of the wh-phrases and the positions which they ultimately come to occupy.

6.4 Three Open Questions

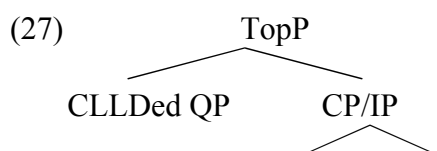
In this section I would like to look back at some of the issues discussed in the first part of the dissertation and see how what we have learned about wh DPs may inform our understanding of the QPs. There are three open questions that can potentially be answered. The first question goes back to the discussion of topicalized and CLLDed QPs, specifically to the issue of whether or not they occupy the same position at the left periphery. The second questions relates to the issue of whether strong QPs undergo topicalization, i.e. whether they also move to the left periphery. The third question relates to the behavior of QPs in embedded contexts.

6.4.1 The Position of CLLDed and Topicalized QPs

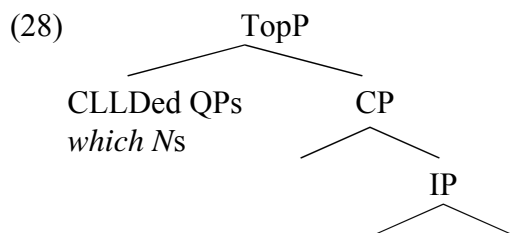
In this chapter we have seen evidence that [+strong] wh-phrases like *which N* and [-strong] wh-phrases like *what/what N* occupy different positions in the clause in Macedonian. We did not have clear evidence that this was the case for CLLDed and

topicalized QPs (see chapter 4). We have seen evidence though that *which Ns* pattern like CLLDed QPs with respect to WCO and island sensitivity. Given these parallels and the fact that *which Ns* and *what/what N* occupy different positions, we can speculate that the same holds for their non wh counterparts.

The analysis of CLLD non-wh-constructions relied on the existence of topic positions within the CP domain which the QPs can occupy.



In section 6.2, I extended this analysis to wh-questions and *which Ns* in particular. Based on the similarities between *which Ns* and CLLDed elements in Macedonian, such as the ability to co-occur with clitics as well as similarities in their interpretation (in the sense that CLLDed elements are necessarily interpreted as specific while *which Ns* are necessarily interpreted as D-linked), it was natural to assume that *which Ns* occupy the position of CLLDed elements above CP.¹⁸



¹⁸ This has also been proposed by Krapova (2002) and Krapova & Cinque (2003) for Bulgarian and Kallulli (1999) for Albanian.

That this view of CLLDed DPs is correct is supported by data like (29), which shows the interaction between CLLDed QPs and wh-phrases.

- (29) a. Site/dve knigi koj student gi pročitā?
 all/two books which student them read
 ‘All/two books, which student read them?’
- b. Koj student site/dve knigi gi pročitā?
 which student all/two books them read
- c. Site/dve knigi koj gi pročitā?
 all/two books whothem read
 ‘All/two books who read them?’
- d. *Koj site/dve knigi gi pročitā?
 who all/two books them read

As (29a-b) show, CLLDed QPs can either precede or follow [+strong] wh-phrases. The difference in grammaticality between (29c) and (29d) shows that CLLDed QP can only precede non-strong wh-phrases, but cannot follow them.¹⁹ The descriptive generalizations that correspond to the data in (29a-d) are given in (30)-(31) below.

¹⁹ I use non-strong here as a neutral term here, to cover both [-strong] and unspecified wh-phrases. More detailed investigations are needed to show whether differences between these two types of wh-phrases exist in such contexts. The subject wh-phrase in (29c-d) is *koj* ‘who’ which in our analysis is unspecified; it is assumed that CLLDed QPs would behave similarly with [-strong] wh-phrases as well. To determine whether this would be the case, we would have to look at questions with CLLDed QPs as subjects and wh-phrases as objects. Added to this, we would have to determine the status of subjects with respect to CLLD and Topicalization. Only then we would be able to establish a more definitive correspondence between CLLDed and topicalized DPs of both kinds, regular and wh ones. Although of great relevance, I have to leave the investigation of these issues for the future.

- (30) a. CLLDed QP + [+strong] wh
 b. [+strong] wh + CLLDed QP

- (31) a. CLLDed QP + non-strong wh
 b. *non-strong wh + CLLDed QP

We can explain the data in (29a-d) as follows. Assuming that fronted wh-phrases occupy the SpecCP position, CLLDed QPs must be higher than SpecCP. The fact that the order of CLLDed QPs and [+strong] wh-phrases is interchangeable, suggest that they both occupy the same position. Supposing that that position is TopP as in Rizzi (1997), we can account for the multiple instances of CLLDed constituents (29a-b)/(30a-b) and the strict ordering of CLLDed QPs and wh-phrases in (29c-d)/(31a-b). Note also that the proposal here is compatible with our earlier claim that DPs can intervene between wh-phrases and the verb only if the wh-phrase is [+strong]. This explains the ungrammaticality of (29d)/(31b).

In light of the fact that CLLDed and fronted wh-phrases occupy different positions in the clause (see section 6.2.2), the question is do we have sufficient evidence to conclude that the same applies to their non wh counterparts. We already saw that CLLDed QPs and [+strong] wh-phrases patterns alike. Topicalized QPs seems to behave like their CLLDed counterparts, as can be seen in (32a-b).

(32)a. Dve knigi koj pročitā?

two books who read

‘Two books who read?’

b. *Koj dve knigi pročitā?

who two books read

The data in (32a-b) show that topicalized QPs can precede a non-strong wh-phrase but cannot follow it. This suggests that topicalized DPs occupy at least positions above that of these wh-expressions (i.e. SpecCP). The generalizations corresponding to (32a-b) are given in (33a-b), respectively.

(33)a. Topicalized QP + non-strong wh

b. *non-strong wh + Topicalized QP

We are considering the option that CLLDed DPs (wh and non wh) may occur higher than non-strong DPs. If this were the case, we would predict that it would be possible for a [+strong] wh to precede but not follow a topicalized QP. It is not clear that this is the case though, as shown in (34a-b), schematically represented in (35a-b).

(34)a. ?Dve knigi koj student pročitā?

two books which student read

‘Two books, which student read?’

b. ?Koj student dve knigi pročitā?

?which student two books read

(35)a. ?Topicalized QP + [+strong] wh

b. ?[+strong] wh + Topicalized QP

Following Rizzi's (1997) framework, we can assume that topicalized DPs can occur either in SpecCP or FocP. (Both positions would capture the fact that multiple topicalization is not allowed.²⁰) Topics in this framework are dispersed throughout the CP domain. In Italian, for example, Rizzi notes that they can either precede or follow FocP. On the assumption that [+strong] wh-phrases, like CLLDed QPs, occupy topic positions, the occurrence of topicalized QPs in Foc may explain the patterns in (35a-b) above.

Although the behavior of CLLDed and topicalized regular DPs seems more nuanced than that of wh-phrases, given the patterning of CLLDed and moved wh DPs, it is likely that their non wh counterparts also occupy different positions, with CLLDed DPs possibly occurring higher than moved DPs.

6.4.2 Topicalization of Strong QPs

In this section I would briefly like to address the question of the topicalization of [+strong] DPs. In chapter 4 we saw evidence from WCO effects and island sensitivity which shows that base-generation of [+strong] DPs in Macedonian has to be allowed. However, we saw no evidence that would rule out movement of such DPs, so there is no

²⁰ For SpecCP this follows from the fact that it is a specifier position and therefore unique (see Baker 1996). For FocP this follows from its non-recursiveness (see Rizzi 1997).

empirical argument for having only base-generation as a possible means of deriving clause-initial [+strong] DPs. Compared to this, weak QPs separate into those that are derived by movement and those that are base-generated at the left periphery. The former give rise to WCO and island effects while the latter show lack of such effects.

As suggested by Veneeta Dayal (p.c.), given the option of base-generation, it might be possible to rule out a movement account of [+strong] DPs on grounds of economy, perhaps in the spirit of Fox (1995, 2000). The motivation for this could be found in the lack of interpretative differences between CLLDed and topicalized strong DPs. Again, this is different from unspecified DPs where such difference exists. We saw in chapter 4 that CLLDed numerals receive a specific interpretation, while topicalized ones do not. Such differences were not observed with the strong DPs. Given this, and the possibility of base-generation, there is no a priori reason why movement should also be an option for the strong QPs. Thus, we can conclude that strong QPs do not undergo movement to the left periphery, unlike weak DPs which allow for both movement and base-generation.

6.4.3 CLLD and Topicalization in Embedded Contexts

We saw in chapter 4 that CLLDed and Topicalized QPs can occur both in matrix and embedded contexts. I repeat some of the examples of embedded constructions in (36) below.

- (36) a. Znam deka sekoja kniga, Petar ja pročita.
 know-I that every book Petar it(f.sg.) read
 ‘I know that every book, Petar read it.’

b. Znam deka dve knigi, Petar gi pročita.

know-I that two books Petar them read

‘I know that two books, Petar read them.’

c. Znam deka dve knigi, Petar pročita.

know-I that two books Petar read

‘I know that two books, Petar read.’

The data in (36a-c) seem to suggest that both types of constituents occur in positions between CP and IP. In Rizzi’s (1997) framework, this would translate into their occupying positions below ForceP.

Compared to this, we know that HTLD constructions are restricted to matrix clauses only. Examples from chapter 4 are repeated below in (37).

(37)a. *Znam deka Marija_i, Petar nea_j ne ja_j vikna na zabava.

know-I that Marija she not her not her invite to party

‘I know that Marija, Petar didn’t invite her to the party.’

b. Znam deka Marija_i, Petar ne ja_j vikna na zabava.

know-I that Marija Petar not her invite to party

‘I know that Marija, Petar didn’t invite her to the party.’

In chapter 4, we took evidence like this to mean that HTLD targets positions above CP, or in Rizzi’s framework, positions above ForceP. Thus, given (36) and (37), a clear distinction between the dislocated constructions emerges, with CLLD and TOP targeting

positions other than HTLD. The facts are more complicated than this, though. As (38a-c) show, CLLDed and topicalized QPs can also precede the complementizer *deka* ‘that’.

(38)a. Znam sekoja kniga deka Petar ja pročita.

know-I every book that Petar it(f.sg.) read

‘I know that every book, Petar read it.’

b. Znam dve knigi deka Petar gi pročita.

know-I two books that Petar them read

‘I know that two books, Petar read them.’

c. Znam dve knigi deka Petar pročita.

know-I two books that Petar read

‘I know that two books, Petar read.’

Given (38) we would have to allow for the possibility that CLLDed and topicalized constituents occupy positions above CP.

Before we conclude this section, consider the interaction of CLLDed and topicalized QPs with *wh*-phrases in embedded clauses. In (39a-b), a CLLDed QP can precede but cannot follow a moved *wh*-phrase. In (39c-d), a CLLDed QP can either precede or follow a CLLDed *wh*-phrase.

(39)a. Se prašuvam dve knigi koj gi pročita.

refl. wonder-I two books whothem read

‘I wonder two books who read them.’

- b. *Se prašuvam koj dve knigi gi pročita.
 refl. wonder-I whotwo books them read
- c. Se prašuvam dve knigi koj student gi pročita.
 refl. wonder-I two books which student them read
 ‘I wonder two books which student read them.’
- d. Se prašuvam koj student dve knigi gi pročita.
 refl. wonder-I which student two books them read

The behavior of CLLDed DPs, both regular and *wh*, is consistent with their behavior in matrix clauses and supports the view that they occupy the same type of positions.

Similarly, topicalized QPs can occur before fronted *wh*-phrases, but cannot follow them. This is shown in (40a-b). As in matrix contexts, the pairing of a topicalized QP and a CLLDed *wh* is less acceptable, as shown in (40c-d).

- (40)a. Se prašuvam dve knigi koj pročita.
 refl. wonder-I two books whoread
 ‘I wonder two books who read.’
- b. *Se prašuvam koj dve knigi pročita.
 refl. wonder-I whotwo books read
- c. ??Se prašuvam dve knigi koj student pročita.
 refl. wonder-I two books which student read
 ‘I wonder two books which student read them.’

d. ??Se prašuvam koj student dve knigi pročita.

refl. wonder-I which student two books read

The data in this section showed that there are clear differences between HTLD constructions, on the one hand, and CLLD and Topicalization constructions, on the other, which is reflected in their patterning in embedded contexts. The behavior of CLLDed and fronted DPs is more complex. We can conclude with some certainty that such constituents occupy positions above CP, when the CP is present in the structure. However, their exact positioning with respect to each other remains open.

A detailed analysis of these issues would have to involve a closer look at the types of embedded clauses. This is especially relevant in the context of Rizzi's theory where the left periphery is defined in reference to multiple C heads. It is likely that CLLD and Topicalization constructions in Macedonian patterns differently with different types of complementizers. However, probing into the properties of embedded clauses, though clearly of great relevance, I must leave for the future.

6.5 Conclusion

I have shown in this study that the distribution of direct objects and clitics in Macedonian can be explained in a principled manner in that the co-occurrence restriction that hold between the DPs and clitics correlate with the intrinsic properties of the DPs themselves. The investigation targeted two classes of DPs, regular quantifiers and wh-expressions. On closer examination, however, well-defined distinctions within each class were shown to exist. The distribution of the clitics was argued to be a result of the feature specification

of the DPs for strength, which in the case of the regular DP was linked to their strength (as in Barwise & Cooper (1981)) and in the case of wh DPs was linked to their D-linking properties (as in Pesetsky (1987) and Comorovski (1996)). I argued that the clitic pronoun in Macedonian is licensed by a [+strong] feature in the DP and as such follows from the Clitic Criterion (Sportiche 1998).

The study looked at the behavior of DPs in two syntactic domains, within IP and at the left periphery. By doing so, it showed that the feature specification for strength has different consequences for the DPs that occur in such positions. While strong DPs were the only ones to satisfy the Clitic Criterion within IP, a subset of the weak DPs was shown to be able to co-occur with clitics when left-dislocated. To account for such behavior of the weak DPs, it was argued that they enter into a binding relation with strong *pro* in argument position. More generally, it was shown that DPs in clause-initial position enter into two types of relations, binding and government. The resulting structures were proven to be derivationally distinct. Finally, it was proposed that CLLDed and moved wh-phrases occupy different positions at the left periphery, with CLLDed DPs occupying a higher position to that of moved DPs. It was suggested that the same might apply to their non wh counterparts.

References

- Abbott, B. 2004. Definiteness and Indefiniteness. In *The Handbook of Pragmatics*, L. R. Horn and G. Ward (eds), 122-149. Oxford: Blackwell.
- Abusch, D. 1993/4. The Scope of Indefinites. *Natural Language Semantics* 2: 83-135.
- Alexiadou, A. 2006. Left Dislocation (including CLLD). In *The Blackwell Companion to Syntax*. M. Everaert and H. van Riemsdijk (eds.), 668-699. Oxford: Blackwell.
- Alexiadou, A. & E. Anagnostopoulou. 1997. Toward a Uniform Account of Scrambling and Clitic Doubling. In *German: Syntactic Problems – Problematic Syntax*. W. Abraham and E. van Gelderen (eds.), 142-161. Tübingen: Max Niemeyer.
- Alexiadou, A. & E. Anagnostopoulou. 1998. Parametrizing AGR: Word Order, V-Movement and EPP-Checking. *Natural Language and Linguistic Theory* 16: 491-539.
- Anagnostopoulou, E. 1994. *Clitic Dependencies in Modern Greek*. PhD dissertation, Salzburg University.
- Anagnostopoulou, E. 1997. Clitic Left Dislocation and Contrastive Left Dislocation. In *Materials on Left Dislocation*, E. Anagnostopoulou, H. C. van Riemsdijk, and F. Zwarts (eds.), 151-192. Amsterdam: John Benjamins.
- Anagnostopoulou, E. 1999. Conditions on Clitic Doubling in Greek. *Clitics in the Languages in Europe* 3: 762-798.
- Anagnostopoulou, E. 2006. Clitic Doubling. In *The Blackwell Companion to Syntax*. M. Everaert and H. van Riemsdijk (eds.), 519-581. Oxford: Blackwell.
- Aoun, J. & E. Benmamoun. 1998. Minimality, Reconstruction, and PF Movement. *Linguistic Inquiry* 29: 569-597.
- Arnaudova, O. 2002. Clitic Left Dislocation and Argument Structure in Bulgarian. In *Annual Workshop on Formal Approaches to Slavic Linguistics* 10, J. Toman (ed.), 19-38. Ann Arbor: Michigan Slavic Publications.
- Arregi, K. 2003. Clitic Left Dislocation is Contrastive Topicalization. In *Proceedings of the 26th Penn Linguistics Colloquium, Penn Working Papers in Linguistics* 9.1, E. Kaiser and S. Arunachalam (eds.), 31-44. Philadelphia: Penn Linguistics Club.
- Baker, M. 1996. *The Polysynthesis Parameter*. Oxford/New York: Oxford University Press.

- Barwise, J. & R. Cooper. 1981. Generalized Quantifiers in Natural Languages. *Linguistics and Philosophy* 4: 159-219.
- Beghelli, F. 1993. A Minimalist Approach to Quantifier Scope. In *Proceedings of NELS 23*. University of Massachusetts, Amherst: GLSA.
- Beghelli, F. 1995. *The Phrase Structure of Quantifier Scope*. PhD dissertation, University of California, Los Angeles.
- Bolinger, D. 1978. Asking More than One Thing at a Time. In *Questions*. H. Hiz (ed.), 107-150. Dordrecht: Reidel.
- Bošković, Ž. 2002. On Multiple Wh-Fronting. *Linguistic Inquiry* 33: 351-383.
- Bošković, Ž. 2004. PF Merger in Stylistic Fronting and Object Shift. In *Minimality Effects in Syntax*. A. Stepanov, G. Fanselow and R. Vogel (eds.), 37-72. Berlin/New York: Mouton de Gruyter.
- Büring, D. 1999. Topic. In *Focus –Linguistic, Cognitive, and Computational Perspectives*, P. Bosch and R. van der Sandt (eds.), 142-165. Cambridge University Press.
- Carlson, G. 1977. *Reference to Kinds in English*. PhD dissertation, University of Massachusetts, Amherst.
- Chomsky, N. 1993. A Minimalist Program for Linguistic Theory. In *The View from Building 20: Essays in Honor of Sylvain Bromberger*, K. Hale and S. J. Keyser (eds.), 1-52. Cambridge: MIT Press.
- Chomsky, N. 1995. *The Minimalist Program*. Cambridge: MIT Press.
- Cinque, G. 1990. *Types of \bar{A} -Dependencies*. Cambridge: MIT Press.
- Cinque, G. 1999. *Adverbs and Functional Heads*. Oxford/New York: Oxford University Press.
- Comorovski, I. 1996. *Interrogative Phrases and the Syntax-Semantics Interface*. Dordrecht: Kluwer.
- Dayal, V. 1996. *Locality in Wh Quantification*. Dordrecht: Kluwer.
- Dayal, V. 2006. Multiple-Wh-Questions. In *The Blackwell Companion to Syntax*. M. Everaert and H. van Riemsdijk (eds.), 275-326. Oxford: Blackwell.
- Dayal, V. & R. Schwarzschild. (2010). Definite Inner Antecedents and Wh-Correlates in Sluices. Rutgers University, ms..

- Diesing, M. 1992. *Indefinites*. Cambridge: MIT Press.
- Dobrovie-Sorin, C. 1994. *The Syntax of Romanian: Comparative Studies in Romance*. Berlin/New York: Mouton de Gruyter.
- Donnellan, K. 1966. Reference and Definite Descriptions. *The Philosophical Review* 75: 281-304.
- É. Kiss, K. 1998. Identificational Focus versus Information Focus. *Language* 74: 245-273.
- Enç, M. 1991. The Semantics of Specificity. *Linguistic Inquiry* 22: 1-25.
- Fanselow, G. 2004. The MLC and Derivational Economy. In *Minimality Effects in Syntax*. A. Stepanov, G. Fanselow and R. Vogel (eds.), 73-124. Berlin/New York: Mouton de Gruyter.
- von Fintel, K. 1994. *Restrictions on Quantifier Domains*. PhD dissertation, University of Massachusetts, Amherst.
- von Fintel, K. 1999. Quantifier Domains and Pseudo-Scope. MIT, ms.
- von Fintel, K. 2000. Singleton Indefinites (re. Schwarzschild 2000). MIT, ms.
- Fodor, J. & I. Sag. 1982. Referential and Quantificational Indefinites. *Linguistics and Philosophy* 5: 355-398.
- Fox, D. 1995. Economy and Scope. *Natural Language Semantics* 3: 283-341.
- Fox, D. 2000. *Economy and Semantic Interpretation*. Cambridge: MIT Press.
- Franks, S. & T. H. King. 2000. *Clitics in Slavic*. Oxford: Oxford University Press.
- Friedman, V. 1993. Macedonian. In: *The Slavonic Languages*, B. Comrie and G. Corbett (eds.), 249-305. London: Routledge.
- Grohmann, K. 2001. Clausal Tripartition, Anti-Locality and Preliminary Considerations of a Formal Approach to Clause Types. In *ZAS Papers in Linguistics* 24, K. von Heusinger and K. Schwabe (eds.). Berlin: ZAS.
- Heim, I. 1982. *The Semantics of Definite and Indefinite Noun Phrases*. PhD dissertation, University of Massachusetts, Amherst.
- Heim, I. 1987. Where does the Definiteness Restriction Apply? Evidence from the Definiteness of Variables. In *The Representation of (In)definiteness*, E. Reuland and A. ter Meulen (eds.), 21-42. Cambridge: MIT Press.

- Heycock, C. 1995. Asymmetries in Reconstruction. *Linguistic Inquiry* 26: 547-570.
- Iatridou, S. 1995. Clitics and Island Effects. In *University of Pennsylvania Working Papers in Linguistics* 2, R. Izvorski and V. Tredinnick (eds.), 11-30. Philadelphia: University of Pennsylvania Linguistics Club.
- Izvorski, R. 1995. Wh-Movement and Focus-Movement in Bulgarian. In *ConSole II Proceedings*, R. Eckardt and V. van Geenhoven (eds.), 54-67. The Hague: Holland Academic Graphics.
- Kallulli, D. 1999. *The Comparative Syntax of Albanian: On the Contribution of Syntactic Types to Propositional Interpretation*. PhD dissertation, University of Durham.
- Kallulli, D. 2008. Clitic Doubling, Agreement and Information Structure: The Case of Albanian. In *Clitic Doubling in the Balkan Languages*, D. Kallulli and L. Tasmowski (eds.), 227-256. Amsterdam: John Benjamins.
- Kamp, H. 1981. A Theory of Truth and Semantic Representation. In *Formal Methods in the Study of Language*, J. Groenendijk, T. Janssen and M. Stokhof (eds.), 277-322. Amsterdam: Mathematisch Centrum.
- Keenan, E. 1987. A Semantic Definition of "Indefinite NP". In *The Representation of (In)definiteness*, E. Reuland and A. ter Meulen (eds.), 286-317. Cambridge: MIT Press.
- Keenan, E. 2003. The Definiteness Effect: Semantics or Pragmatics?. *Natural Language Semantics* 11: 187-216.
- King, T. H. 1995. *Configuring Topic and Focus in Russian*. Stanford: CSLI.
- Koneski, B. 1987. *Gramatika na makedonskiot literaturni jazik*. Skopje: Kultura.
- Koopman, H. & D. Sportiche. 1983. Variables and the Bijection Principle. *The Linguistic Review* 2:139-160.
- Krapova, I. 2002. On D-Linking and Multiple Wh-Fronting. Paper presented at *FASSL 4*, Sofia 2002.
- Krapova, I. & G. Cinque. 2003. On the Order of Wh-Phrases in Bulgarian Multiple Wh-Fronting. In *Proceedings of Formal Description of Slavic Languages 5*, G. Zybatow, L. Szucsich, U. Junghanns, and R. Meyer (eds.), 318-336. Frankfurt: Peter Lang.
- Krapova, I. & G. Cinque. 2008. Clitic Reduplication Constructions in Bulgarian. In *Clitic Doubling in the Balkan Languages*, D. Kallulli and L. Tasmowski (eds.), 257-288. Amsterdam: John Benjamins.

- Kratzer, A. 1998. Scope or Pseudo-Scope? Are there Wide-Scope Indefinites. In *Events and Grammar*, S. Rothstein (ed.), 163-196. Dordrecht: Kluwer.
- Lasnik, H. & T. Stowell. 1991. Weakest Crossover. *Linguistic Inquiry* 22: 687-720.
- McConnell-Ginet, S. 1982. Adverbs and Logical Form: A Linguistically Realistic Theory. *Language* 58: 144-184.
- Milsark, G. 1977. Toward an Explanation of Certain Peculiarities of the Existential Construction in English. *Linguistic Analysis* 3: 1-31.
- Mišeska Tomić, O. 2008. Towards Grammaticalization of Clitic Doubling: Clitic Doubling in Macedonian and Neighbouring Languages. In *Clitic Doubling in the Balkan Languages*, D. Kallulli and L. Tasmowski (eds.), 65-88. Amsterdam: John Benjamins.
- Partee, B. 1970. Opacity, Coreference, and Pronouns. *Synthese* 21: 359-385. [Reprinted in B. Partee. 2004 *Compositionality in Formal Semantics: Selected Papers by Barbara Partee*, 26-49. Oxford: Blackwell.]
- Partee, B. 1989. Many Quantifiers. In *Proceedings of the 5th Eastern States Conference on Linguistics*, J. Powers and K. de Jong (eds.), 383-402. Columbus, OH: Ohio State University. [Reprinted in B. Partee. 2004 *Compositionality in Formal Semantics: Selected Papers by Barbara Partee*, 241-258. Oxford: Blackwell.]
- Partee, B. 2004. *Compositionality in Formal Semantics: Selected Papers by Barbara Partee*. Oxford: Blackwell.
- Pesetsky, D. 1987. Wh-in-Situ: Movement and Unselective Binding. In *The Representation of (In)definiteness*, E. Reuland and A. ter Meulen (eds.), 98-129. Cambridge: MIT Press.
- Phillipaki-Warbuton, I. S. Varlokosta, M. Georgiadjis, & G. Kotzoglou. 2002. On the Status of 'Clitics' and their 'Doubles' in Greek. *Reading Working Papers in Linguistics* 6: 57-84.
- Reinhart, T. 1981. Pragmatics and Linguistics: An Analysis of Sentence Topics. *Philosophica* 27: 53-94.
- Reinhart, T. 1995. *Interface Strategies*. OTS, University of Utrecht.
- Richards, N. 1997. *What Moves Where In Which Language?*. PhD dissertation, MIT.
- Rizzi, L. 1990. *Relativized Minimality*. Cambridge: MIT Press.

- Rizzi, L. 1991. Residual Verb Second and the Wh Criterion. *Technical Reports in Formal and Computational Linguistics* 2, University of Geneva.
- Rizzi, L. 1997. The Fine Structure of the Left Periphery. In *Elements of Grammar: Handbook in Generative Syntax*, L. Haegeman (ed.), 281-337. Dordrecht: Kluwer.
- Rizzi, L. 2004. Locality and Left Periphery. In *Structures and Beyond. The Cartography of Syntactic Structures*, A. Belletti (ed.), 223-252. Oxford: Oxford University Press.
- Rudin, C. 1988. On Multiple Questions and Multiple Wh-Fronting. *Natural Language and Linguistic Theory* 6: 445-501.
- Rudin, C. 1997. Agr-O and Bulgarian Pronominal Clitics. In: *Formal Approaches to Slavic Linguistics* 5, M. Lindseth and S. Franks (eds.), 224-252. Ann Arbor: Michigan Slavic Publications.
- Safir, K. 1996. Derivation, Representation, and Resumption: The Domain of Weak Crossover. *Linguistic Inquiry* 27: 313-339.
- Sánchez, L. 2002. Discourse Topic Constraint on Left Dislocated Subjects and CLLD Structures. In *Features and Interfaces*, K. Zagana, E. Mallen, and J. Herschensohn (eds.), 255-266. Amsterdam: John Benjamins.
- Schwarzschild, R. 2002. Singleton Indefinites. *Journal of Semantics* 19:289-314.
- Sportiche, D. 1998. *Partitions and Atoms of Clause Structure: Subjects, Agreement, Case and Clitics*. London/New York: Routledge.
- Stanley, J. & Z. Gendler Szabó. 2000. On Quantifier Domain Restriction. *Mind and Language* 15: 219-261.
- Suñer, M. 1988. The Role of Agreement in Clitic-Doubled Constructions. *Natural Language and Linguistic Theory* 6: 391-434.
- Szabolcsi, A. 1981. The Semantics of Topic-Focus Articulation. In *Formal Methods in the Study of Language*, J. Groenendijk, T. Janssen and M. Stokhof (eds.), 513-541. Amsterdam: Mathematisch Centrum.
- Szabolcsi, A. 1997. Strategies for Scope Taking. In *Ways of Scope Taking*, A. Szabolcsi (ed.), 109-154. Dordrecht: Kluwer.
- Vat, J. 1997. Left Dislocation, Connectedness and Reconstruction. In *Materials on Left Dislocation*, E. Anagnostopoulou, H. C. van Riemsdijk, and F. Zwarts (eds.), 67-92. Amsterdam: John Benjamins.

- Ward, G. & B. Birner. 2004. Information Structure and Non-Canonical Syntax. In *The Handbook of Pragmatics*, L. R. Horn and G. Ward (eds), 153-174. Oxford: Blackwell.
- Zucchi, A. 1995. The Ingredients of Definiteness and the Definiteness Effect. *Natural Language Semantics* 3: 33-78.

Curriculum Vita

Slavica Kochovska

Education

- 2002-2010 Ph.D. in Linguistics, Rutgers University
- 1998- 2000 M.A. in Linguistics (First Class Honours), The University of Auckland
- 1995-1998 B.A. in Russian and Linguistics (Highest Honours), The University of Auckland

Publications

- 2007 Quantifier Dislocation in Macedonian. To appear in *Proceedings of the 2nd Congress of the Slavic Linguistics Society*. ZAS. Berlin.
- 2006 (Non) D-Linking at the Macedonian Left Periphery. In R.Compton, M. Golezinowska and U. Savchenko (eds.) *Proceedings of Formal Approaches to Slavic Linguistics 15: The Toronto Meeting 2006*. Ann Arbor, Michigan Slavic Publications. pp. 141-160.
- 2006 Intervention Effects in Macedonian *Wh*-Questions. In J. Bunting, S. Desai, R. Peachey, C. Straughn and Z. Tomková (eds.) *Proceedings from the 42nd Annual Meeting of the Chicago Linguistic Society*. The Chicago Linguistic Society, Chicago. pp. 149-162.
- 2004 Palatalization in Macedonian in a Feature Geometry Model. In Brian D. Joseph and M.A. Johnson (eds.) *Macedonian Studies: Papers from the Fifth International Macedonian-North American Conference on Macedonian Studies, 1-4 May 2003*. Columbus, Ohio. pp. 167-194.