

## Clitic Doubling in the Balkan Languages

## *Linguistik Aktuell/Linguistics Today (LA)*

*Linguistik Aktuell/Linguistics Today (LA)* provides a platform for original monograph studies into synchronic and diachronic linguistics. Studies in LA confront empirical and theoretical problems as these are currently discussed in syntax, semantics, morphology, phonology, and systematic pragmatics with the aim to establish robust empirical generalizations within a universalistic perspective.

### **General Editors**

Werner Abraham  
University of Vienna / Rijksuniversiteit  
Groningen

Elly van Gelderen  
Arizona State University

### **Advisory Editorial Board**

Cedric Boeckx  
Harvard University

Guglielmo Cinque  
University of Venice

Günther Grewendorf  
J.W. Goethe-University, Frankfurt

Liliane Haegeman  
University of Lille, France

Hubert Haider  
University of Salzburg

Christer Platzack  
University of Lund

Ian Roberts  
Cambridge University

Lisa deMena Travis  
McGill University

Sten Vikner  
University of Aarhus

C. Jan-Wouter Zwart  
University of Groningen

### **Volume 130**

Clitic Doubling in the Balkan Languages  
Edited by Dalina Kallulli and Liliane Tasmowski

# Clitic Doubling in the Balkan Languages

*Edited by*

Dalina Kallulli

University of Vienna

Liliane Tasmowski

University of Antwerp & KVAB

John Benjamins Publishing Company

Amsterdam / Philadelphia



The paper used in this publication meets the minimum requirements of American National Standard for Information Sciences – Permanence of Paper for Printed Library Materials, ANSI Z39.48-1984.

#### Library of Congress Cataloging-in-Publication Data

Clitic doubling in the Balkan languages / edited by Dalina Kallulli and Liliane Tasmowski.

p. cm. (Linguistik Aktuell/Linguistics Today, ISSN 0166-0829 ; v. 130)

Includes bibliographical references and index.

1. Balkan Peninsula--Languages--Clitics. I. Kallulli, Dalina. II. Tasmowski, Liliane.

P381.B3.C55 2008

415'.7--dc22

2008019877

ISBN 978 90 272 5513 6 (Hb; alk. paper)

© 2008 – John Benjamins B.V.

No part of this book may be reproduced in any form, by print, photoprint, microfilm, or any other means, without written permission from the publisher.

John Benjamins Publishing Co. · P.O. Box 36224 · 1020 ME Amsterdam · The Netherlands  
John Benjamins North America · P.O. Box 27519 · Philadelphia PA 19118-0519 · USA

# Table of contents

<b>List of contributors</b>	<b>VII</b>
<b>Preface</b>	<b>IX</b>
Introduction: Clitic doubling, core syntax and the interfaces <i>Dalina Kallulli and Liliane Tasmowski</i>	<b>1</b>
<b>PART I. Clitic doubling within the Balkan continuum: Rise and spread</b>	
1. Balkan object reduplication in areal and dialectological perspective <i>Victor A. Friedman</i>	<b>35</b>
2. Towards grammaticalization of clitic doubling: Clitic doubling in Macedonian and neighbouring languages <i>Olga Mišeska Tomić</i>	<b>65</b>
3. The genesis of clitic doubling from Ancient to Medieval Greek <i>Gunnar de Boel</i>	<b>89</b>
4. Clitic doubling and Old Bulgarian <i>Mila Dimitrova-Vulchanova and Valentin Vulchanov</i>	<b>105</b>
<b>PART II. Discourse functional properties of clitic doubling</b>	
5. Romanian Clitic Doubling: A view from pragmatics-semantics and diachrony <i>Virginia Hill and Liliane Tasmowski</i>	<b>135</b>
6. Clitic doubling from Ancient to Asia Minor Greek <i>Mark Janse</i>	<b>165</b>
7. Object clitic doubling constructions and topicality in Bulgarian <i>Zlatka Guentchéva</i>	<b>203</b>

**PART III. Morpho-syntactic properties and modelling of clitic doubling**

- |     |  |     |
|-----|--|-----|
| 8.  | Clitic doubling, agreement and information structure:<br>The case of Albanian<br><i>Dalina Kallulli</i>  | 227 |
| 9.  | Clitic reduplication constructions in Bulgarian<br><i>Iliyana Krapova and Guglielmo Cinque</i>   | 257 |
| 10. | Clitic doubling, complex heads and interarboreal operations<br><i>Alexandra Cornilescu and Carmen Dobrovie-Sorin</i>   | 289 |
| 11. | Rethinking the Clitic Doubling parameter: The inverse<br>correlation between clitic doubling and participle agreement<br><i>Vina Tsakali and Elena Anagnostopoulou</i> | 321 |

**PART IV. Clitic doubling within the DP**

- |     |   |     |
|-----|---|-----|
| 12. | Romanian possessive clitics revisited<br><i>Larisa Avram and Martine Coene</i>                      | 361 |
| 13. | Possessive clitics in the DP: Doubling or dislocation?<br><i>Giuliana Giusti and Melita Stavrou</i> | 389 |

Name index	435
------------	-----

Language index	437
----------------	-----

Subject index	439
---------------	-----

## List of contributors

Elena Anagnostopoulou  
University of Crete

Larisa Avram  
University of Bucharest

Gunnar de Boel  
Ghent University

Guglielmo Cinque  
University of Venice

Martine Coene  
University of Antwerp and  
University of Leiden

Alexandra Cornilescu  
University of Bucharest

Mila Dimitrova-Vulchanova  
Norwegian University of Science and  
Technology

Carmen Dobrovie-Sorin  
Université Paris 7 & CNRS

Victor Friedman  
University of Chicago

Giuliana Giusti  
University of Venice

Zlatka Guentchéva  
LACITO – CNRS

Virginia Hill  
University of New Brunswick

Mark Janse  
Ghent University

Dalina Kallulli  
University of Vienna

Iliyana Krapova  
University of Venice

Melita Stavrou  
Aristotle University of Thessaloniki

Liliane Tasmowski  
University of Antwerp & KVAB

Olga Mišeska Tomić  
University of Skopje &  
University of Novi Sad

Vina Tsakali  
University of Crete &  
University of Ioannina

Valentin Vulchanov  
Norwegian University of Science and  
Technology



## Preface

The present volume has its origin in a workshop organized by the Royal Flemish Academy of Belgium for Sciences and Arts (KVAB) in Brussels in December 2004. However, the volume is independently structured and includes papers that were not presented at that event, such as those by Avram and Coene, Krapova and Cinque, and Tsakali and Anagnostopoulou.

The phenomenon of clitic doubling – the subject matter of this volume – is known to be especially prevalent in the Balkan Sprachbund. The various chapters of the volume deal with different aspects of this phenomenon, such as its morpho-syntactic, semantic, prosodic and pragmatic properties as attested not only synchronically but also from the perspective of their historical development both across the Balkan continuum and beyond, thus contributing to the understanding of the nature of the cross-linguistic variation, as well as the micro-variation observed with respect to clitic doubling.

The editors would like to thank the contributors for their co-operation, the external reviewers for their dedicated work, the series editors and Kees Vaes for their feedback and practical help, as well as the Royal Flemish Academy of Belgium for Sciences and Arts (KVAB) and the Flemish Foundation for Scientific Research (FWO) for financial support.

Dalina Kallulli, Liliane Tasmowski



## INTRODUCTION

# Clitic doubling, core syntax and the interfaces\*

Dalina Kallulli  
University of Vienna

Liliane Tasmowski  
University of Antwerp & KVAB

### 1. Clitic doubling and its theoretical significance

#### 1.1 Basic facts

Despite the controversies surrounding the proper constitution, genesis and development of the so-called “Balkan Sprachbund”, virtually all studies dealing with its core properties (see e.g., Tomić 2004; Tomić 2006 and references therein) list as a crucial characteristic of this group of languages what in the generative paradigm has since Jaeggli (1982) come to be known as the phenomenon of clitic doubling. That is, clitic doubling is a pervasive phenomenon in the languages of the Balkans, as already recognized in Miklosich (1862: 7–8); see Friedman (this volume).

As its name suggests, clitic doubling involves the doubling (or “reduplication”) by a clitic pronoun of a verbal argument (henceforth: the associate) inside the same propositional structure (that is, inside the same CP; see Adger 2003).<sup>1</sup> The associate is instantiated either by a full pronoun, a non-pronominal referring expression or, we will add, a CP.<sup>2</sup> When the associate is a pronoun or a DP,<sup>3</sup> the clitic bears the same

---

\*Research for this paper was funded by the Royal Flemish Academy of Belgium for Sciences and Arts (KVAB) and the Flemish Foundation for Scientific Research (FWO).

1. Though languages in which the subject is doubled by pronominal clitics exist (e.g., varieties of French, Italian, or Dutch), Balkan languages as a rule only exhibit object clitic doubling. Therefore, throughout this article clitic doubling is used with reference to internal (i.e., object as opposed to subject) verbal arguments.
2. The vast majority of generative studies on clitic doubling have focused on doubling of full pronouns and other DPs but not of CPs (but see Kallulli 2006, this volume).
3. Following Kallulli (2000), we take clitic doubled non-pronominal referring expressions to be DPs and not NPs.

phi-features and case as the pronoun or the DP, when the associate is a CP, the clitic is a singular (neuter) accusative. These patterns are illustrated through the Albanian examples in (1a), (1b) and (1c), respectively.<sup>4</sup>

- (1) a. *Ana mē<sub>i</sub> pa mua<sub>i</sub> në rrugë.* (Albanian)  
 Ana.the<sub>NOM</sub> me<sub>CL</sub> saw me<sub>FP</sub> in road  
 ‘Ana saw me on the road.’
- b. *Ana e<sub>i</sub> lexoi letrën<sub>i</sub> deri në fund.*  
 Ana.the<sub>NOM</sub> 3S,CL,ACC read letter.the<sub>ACC</sub> until in end  
 ‘Ana read the letter until the end.’
- c. *Ana e<sub>i</sub> dinte [CP që Eva kishte shkuar].*  
 Ana.the<sub>NOM</sub> 3S,CL,ACC knew that Eva had left  
 ‘Ana knew (it) that Eva had left.’

Clitic doubling was recognized as a Balkan phenomenon and was first studied as such in a thorough comparative perspective by Lopašov (1978). Generative studies on the other hand initially focused on its properties as exhibited in those Romance languages where it is attested. Thus, Jaeggli (1982), who first noticed the theoretical importance of clitic doubling, describes it for River Plate Spanish (spoken in Argentina, Uruguay and Paraguay), which exhibits patterns like those in (2a). Almost at the same time, similar patterns had been discussed for Romanian by Farkas (1978) and Steriade (1980); see (2b).<sup>5</sup>

- (2) a. *Lo<sub>i</sub> veo a Juan<sub>i</sub>.* (Spanish)  
 b. *Il<sub>i</sub> văd pe Ion<sub>i</sub>.* (Romanian)  
 him<sub>CL</sub> see-I on John  
 ‘I saw John.’

Clitic doubling in the examples in (2) might appear as an optional phenomenon since these sentences are also grammatical in the absence of the clitic. In other cases however, as in (3a) for River Plate Spanish and (3b) for Romanian, the clitic is not only possible but indeed obligatory.

4. As is obvious from the example in (1a), the distribution of clitic and full pronouns is different. In fact, their different distribution from full pronouns is perhaps the most striking property of what Zwicky (1977) refers to as ‘special’ clitics (i.e., word-like elements showing unique syntactic properties – pronominal clitics being one prototype) and indeed what fuelled the interest in them in generative grammar (Kayne 1975).

5. Jaeggli’s (1982) clitic doubling data drew immediate attention as they constituted a major challenge to Kayne’s (1975) analysis of French clitics, since in (standard varieties of) French clitic doubling is not possible.

- (3) a. \*(*Le*) *duele la cabeza a Mafalda.* (from Jaeggli 1982)  
       her<sub>CL</sub> hurts the head to Mafalda  
       ‘Mafalda has a headache.’
- b. \*(*Îi*) *îngheață lui Ion picioarele.* (from Reinheimer & Tasmowski 2005)  
       him<sub>CL</sub> freeze to J. feet-the  
       ‘John’s feet are freezing.’

A comprehensive study of clitic doubling must then naturally address not only the question why this phenomenon appears at all (since it is missing in many languages), but also account for its (non-)obligatoriness, that is the question of what exactly triggers, forces or disallows clitic doubling. Grossly stated, traditional grammar has been mainly concerned with the latter aspect, while generative studies have focused on the former (see van Riemsdijk 1999 for an overview). However, in spite of the substantial body of literature on pronominal clitics, to date there is no consensus on either of these two fundamental issues.

The examination of the clitic doubling phenomenon breaks down into several topics for scrutiny, such as: its extension within and outside the Balkan Sprachbund and the observed variation; its realizational possibilities and the constraints on the argumental status of the associate DP (direct or indirect object), its categorial nature (pronominal or lexical), its semantics (definite, specific, presupposed, or neither) and pragmatics (topic or not, D-linked or not); its temporal and locational genesis, and whether or not influence under language contact can be detected; the relationship between the clitic and its associate, their respective positioning in relation to the governing verb, and the status of the doubled constituent when the latter is not in its canonical position.

In the rest of this section, we will dwell on past formal accounts of clitic doubling and the challenges that they have been confronted with.

## 1.2 Early formal accounts based on Romance

In the context of generative grammar, a central issue on the syntactic modelling of clitic pronouns has been whether or not their surface position, which in Romance and several Balkan languages is typically preverbal in simple declarative clauses, is a result of movement operations. Thus, according to Kayne (1975), clitics move to their (preverbal) surface position from the post-verbal argument position. In contrast, Strozer (1976), Rivas (1977), Jaeggli (1982) and Borer (1981) have argued that clitics are base-generated in their surface position.

Kayne’s influential movement analysis of clitics was based on several observations, one of which was the complementarity between clitics and DPs instantiating the same internal argument (i.e., lack of clitic doubling) in standard French (and Italian), as illustrated for French in (4c) versus (4a) and (4b). Thus, given the paradigm in (4),

base-generating the clitic in the post-verbal canonical position of the DP ensured that the DP and a corresponding (i.e., doubling) clitic would not co-occur.<sup>6</sup>

- (4) a. *Je vois Jean.* (French)  
 I see John
- b. *Je le vois.*  
 I him<sub>CL</sub> see
- c. \**Je le vois Jean*  
 I him<sub>CL</sub> see John

Kayne's analysis is also motivated by the locality conditions (reminiscent of A-movement) holding between the clitic and its purported movement base.

The proponents of the base-generation analysis, on the other hand, showed that one core empirical argument for the movement analysis, namely the complementarity between an argument DP and a corresponding clitic (i.e., the lack of clitic doubling), was not universal, as witnessed by the examples in (1) through (3) above.

Kayne's solution for this disturbing lack of complementarity between a DP and a doubling clitic, epitomized as *Kayne's Generalization* in Jaeggli (1982), was that a DP may in fact be doubled by a clitic, but only if this DP is endowed with an explicit (here: prepositional) case assigner. Simplifying somewhat, the idea was that the doubling clitic absorbs Case, so unless a preposition (or some other case-assigning device) could be inserted, the DP-argument would remain caseless and the Case Filter (Chomsky 1981) would cause the derivation to crash. As can be seen in (2) and (3), the Romance languages exhibiting clitic doubling (such as River Plate Spanish and Standard Romanian) seem to fit well into this picture, since they need a particular preposition-like element preceding the direct object in the presence of a doubling clitic. This element arguably acts as a case assigner (but see Tsakali & Anagnostopoulou, this volume, for a detailed overview of the problems arising from this solution).

However, as pointed out by Suñer (1988: 399–400), the examples in (5) from Portefño Spanish speak against the view that the prepositional element *a* in Spanish (a language where Kayne's Generalization was claimed to be operative, Jaeggli 1982) is a case assigning device.

---

6. Note however that Kayne's (1975) idea that clitics move from the post-verbal argument position principally allows for the possibility that the doubling clitic and its associate originate in the same position. In fact, this is the gist of the so-called "big DP" analysis of clitic doubling constructions (Uriagereka 1995; Torrego 1995, 1998; Cecchetto 2000, among others). That is, according to the big-DP analysis, the verb selects a complex DP with a clitic head taking a DP/PP complement.

- (5) a. *Yo lo voy a comprar el diario justo antes de subir.* (Porteño Spanish)  
 ‘I am going to buy it-the newspaper just before coming up.’
- b. *Yo la tenía prevista esta muerte.*  
 ‘I had foreseen it-this death’
- c. *Ahora tiene que seguir usándolo el apellido.*  
 ‘Now she has to go on using it-the surname.’

Suñer (1988) argues instead that *a* is an animacy marker, which is why it is missing in the examples in (5) even though the direct object DPs here are clitic doubled. Her argument is strengthened by the fact that the South-Danubian Aromanian and Megleno-Romanian dialects double the direct object in the same contexts as in (5) while lacking a preposition-like element even before animates (see the contributions by Friedman and Tomić, this volume, and references therein):

- (6) a. *nu-l<sup>u</sup> ávdu fičórlu* (Aromanian; Caragiu-Marioțeanu 1975: 237)  
 not-him<sub>CL</sub> hear boy-the  
 ‘I don’t hear the boy’
- b. *jeł nu lă vreà țela fițșór*  
 he not him<sub>CL</sub> wanted that child  
 ‘He did not like that child’ (Megleno-Romanian; Caragiu-Marioțeanu *ib.*: 277)

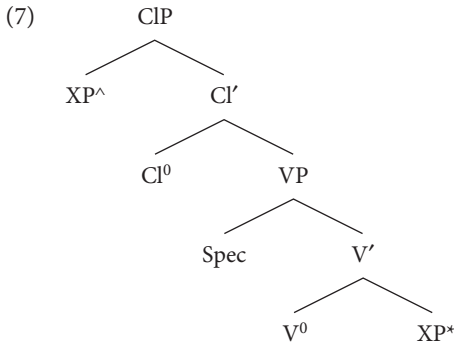
Another objection to Kayne’s Generalization brought forth in Dobrovie-Sorin (1994), namely that indirect object clitic doubling in (Standard) Daco-Romanian involves no such prepositional element either, presented a further challenge, since Kayne’s Generalization does not obviously hold in these cases. Dobrovie-Sorin solved this problem by appealing to the fact that in Standard Romanian indirect objects bear morphological dative case, which in pre-minimalist Case theory was considered to be inherent, or lexical, and as such, different from structural case that had to be assigned by a governing category. Further research soon revealed that this phenomenon is quite extent in the Balkan languages: Albanian and Modern Greek (henceforth: MG) exhibit clitic doubling not only of inherently case marked indirect objects, but also of structurally governed direct objects bearing morphological accusative case but that nonetheless do not co-occur with a prepositional element (see the Albanian examples in (1), and Anagnostopoulou (1994) and references therein for MG).<sup>7</sup> Moreover, Macedonian, Bulgarian, (and among the Romance languages) Aromanian and Megleno-Romanian can double a purely structurally governed direct object that bears no morphological case, as in (6) above.

7. In fact, PPs cannot be doubled in these languages (see Kallulli 1995 for Albanian and Anagnostopoulou 1994 for MG).

### 1.3 Combining base-generation and movement analyses: Sportiche (1996)<sup>8</sup>

Starting from the basic idea that both base-generation and movement analyses to clitic constructions must be correct, as indicated by strong arguments (mentioned in section 1.2) in favour of either view, Sportiche (1996) provides a structural analysis that manages to combine both approaches.<sup>9</sup>

Specifically, Sportiche claims that all clitics are always base-generated in pre-existing slots, namely as  $X^0$ -elements heading their own projections in whose specifier position they license a particular property/feature  $F$  (the nature of which is discussed below). For the derivation to converge, this feature has to be saturated or checked off (Chomsky 1995). Since features may only be checked off in Spec-Head configurations, the (doubled argument)  $XP^*$  in (7) must by LF move to the  $XP^\wedge$  position for the (licensing) Spec-Head configuration to obtain.<sup>10</sup>



In Sportiche's terms, movement of  $XP^*$  to the  $XP^\wedge$  position is motivated by the Clitic Criterion given in (8). This is an analogue of Rizzi's (1991) Wh-Criterion, and yet another instantiation of the so-called Generalized Licensing Criterion in (9), according to which feature-licensing may only obtain in Spec-Head configurations.

8. Sportiche's influential paper was circulated widely as a manuscript since 1992.

9. Sportiche's structural analysis is also discussed in detail in Tsakali & Anagnostopoulou's contribution in this volume. However, given the importance of this approach for the subject matter of this volume, we present the main points of this analysis also here, with a view to highlighting its elegance, as well as the questions that it leaves open and those that it gives rise to.

10. Sportiche labels the clitic projections "Voice Phrases", which in view of other connotations (concerning the structure of verbal projections – cf. Kratzer 1996) having nothing to do with the issues at hand, we have taken the liberty to replace with ClP (for Clitic Phrase) throughout.

- (8) *Clitic Criterion*  
At LF,  
a. A clitic must be in a spec-head relationship with a [+F] XP.  
b. A [+F] XP must be in a spec-head relationship with a clitic.
- (9) *Generalized Licensing Criterion*  
At LF,  
a. A [+F] head must be in a spec-head relationship with a [+F] XP.  
b. A [+F] XP must be in a spec-head relationship with a [+F] head.

Further, Sportiche sets the following clitic parameters:

- (10) *Clitic Parameters*  
a. Movement of XP\* to XP^ occurs overtly or covertly.  
b. Head (Cl) is overt or covert.  
c. XP\* is overt or covert.

The combination of these independent parameters captures the different construction types in (11):

- (11) a. Undoubled clitic constructions (as in French or Italian) arise when a covert XP\* moves overtly or covertly to XP^ with Cl overt.  
b. Clitic doubling constructions (as in the Balkan languages and Spanish) arise when an overt XP\* moves covertly with an overt Cl.  
c. Scrambling in Dutch/German arises when an overt XP\* moves overtly with a covert Cl.  
d. Clitic left dislocation (as in Italian etc., see Cinque 1977, 1983, 1990) arises when an overt XP\* moves overtly with an overt Cl to Spec of ClP and then beyond.

Turning to the content of the (F) feature that the clitic head licenses in the specifier of the phrase it heads, which in turn bears directly on the issue of the function of clitics, Sportiche claims that while direct object clitics license *specificity* in the specifier position of the phrase they head, indirect object doubling clitics function as pure (object) agreement markers.

In spite of its obvious appeal, Sportiche's influential analysis of clitic constructions raises two major issues. First, while the literature on clitic doubling in various languages converges on the view that direct object clitic doubling is much more restricted than indirect object clitic doubling (Suñer 1988; Sportiche 1996 for Spanish; Kallulli 1998, 2000 for Albanian; Anagnostopoulou 2003 for MG; Farkas 1978; Dobrovie-Sorin 1994 for Romanian), also under Sportiche's structurally uniform account, direct and indirect object clitic doubling remain essentially distinct phenomena triggered by distinct grammatical factors. Even abstracting away from the fact that concerning direct object clitic doubling other semantic properties in addition to (or instead of) specificity (e.g., *prominence*, *referential stability*, *presuppositionality*,

*familiarity, definiteness, topichood/givenness*) have been highlighted by different researchers as relevant across various languages (as we discuss in section 2.2), a fundamental question remains open: if the nature of agreement and specificity markers is indeed substantially different, why are clitics employed as means for fulfilling such different functions?

Secondly, if as Sportiche argues, clitic doubling of direct objects licenses specificity, then it should always be possible to clitic double direct objects instantiated by definite DPs under the standard assumption that definite DPs are presuppositional and therefore specific (Enç 1991; Diesing 1992).<sup>11</sup> That is, even if direct object clitic doubling is viewed as a sufficient but not necessary means of licensing specificity, Sportiche predicts that we should not be able to find cases in which clitic doubling of definite direct object DPs is barred. However, as shown in Kallulli (2000) for Albanian and Greek (but the pattern is more general, including Bulgarian and Romanian, as highlighted in the contributions in this volume), this prediction is not borne out (see also the discussion in section 2.2.3).<sup>12</sup>

Note also that, while many researchers have adopted Sportiche's analysis of clitic constructions, some of these very same researchers have argued for a right-dislocation analysis of what appears to be a clitic doubled DP in other contexts. For instance, while Anagnostopoulou (1994) and Alexiadou & Anagnostopoulou (2000) argue that clitic doubled definite DPs in MG constitute instances of genuine clitic doubling, in the same works they claim that indefinite DPs resumed by clitics are right-dislocated constituents.<sup>13</sup>

Furthermore, in addition to the facts that Romance and Balkan languages have at least partially in common, there also exists a variety of core data on clitic doubling that is only attested in Balkan languages, such as: asymmetries of the distribution of doubling clitics in restrictive relative clauses (Stavrou 1984, Alexiadou & Anagnostopoulou 2000 for MG; Dobrovie-Sorin 1990 for Romanian; Kallulli 2001, this volume for Albanian; Dimitrova-Vulchanova & Vulchanov, this volume, for Bulgarian); clitic doubling as a trigger of factivity in Albanian and MG (Kallulli 2006, this volume); putative clitic doubling within the DP (specifically, the co-occurrence of a possessive DP and a possessive clitic inside a nominal expression, cf. Avram & Coene and Giusti & Stavrou,

---

11. Note, however, that Enç (1991) acknowledges the existence of non-specific definites as in the example in (i) (attributed to D. Pesetsky):

(i) There is the following counterexample to Streck's theory.

12. Note also that the idea that clitic doubling induces specificity on its associate DP has been rejected in more recent work on this phenomenon in Spanish too (see e.g., Gutiérrez-Rexach 2000, 2001, 2002).

13. This claim has in turn been refuted in Kallulli (2000).

this volume), all of which cannot be readily accounted for under existing analyses, including Sportiche's.

## 2. The Balkan patterns and their theoretical significance

### 2.1 Genesis and the geographical distribution of clitic doubling patterns

In principle, clitic doubling in the Balkan languages is an innovative phenomenon that has arisen within the Balkan languages themselves, since nowhere is it to be found in their historically attested ancestors, at least not in its current shape. Even if the first manifestations of clitic doubling in the Balkan languages can be considered to stem from Vulgar Latin, and even if its (rare) appearance in New Testament Greek is due to previous contact of Greek with Vulgar Latin (Ilievski 1988 [1973] cited by Friedman, this volume), neither Vulgar Latin nor Ancient Greek disposed of a double series of third person pronouns (i.e., strong versus phonetically reduced clitic ones), with clitics similar in nature to the ones found in the daughter languages (see Salvi 2004, 182 ff. for a characterization of the Latin/Romance clitics). Clear-cut cases of (modern-type) clitic doubling are not attested in Old Church Slavonic either (but see Dimitrova-Vulchanova & Vulchanov, this volume), nor does clitic doubling occur in any of the non-Balkan Slavic languages. Since the development of clitic doubling in the Balkan languages is thus not genetically determined by an established historical source in any obvious way, it has in all likelihood spread from a well-defined center of innovation inside the region. Be that as it may, as recognized by Lopašov (1978), the resulting situation is far from homogeneous: clitic doubling seems subject to strict grammatical constraints in the West and the South of its expansion area, but gets increasingly conditioned by discourse-pragmatic factors towards the North and the East. This situation is represented through (12), with grammatically constrained clitic doubling on the left handside becoming freer and pragmatically significant as one proceeds to the right:

(12) Macedonian > Albanian > Romanian > Greek > Bulgarian

The representation in (12) conforms with the environments traditionally recognized to trigger clitic doubling in the specific languages, namely:

- for Macedonian, all definite direct objects and all indirect objects;
- for Albanian, all IOs, DOs instantiated by first and second person pronouns, and all non-focal/non-rhematic DO DPs;
- for Romanian, all full personal and definite pronouns, preverbal indirect objects and not [–specific] DPs, postverbal direct object DPs that are not [–specific] and are introduced by *pe*, and postverbal indirect object DPs which are not [–specific] and/or [–human] Goal;

- for Greek, no obligatory context, except with *olos* ‘all’;
- for Bulgarian, all objects that are interpreted as Experiencers and objects of *ima*, *njama* ‘there is (not)’.<sup>14</sup>

Lopašov (1978) considered the preverbal position of the object to be a trigger of clitic doubling, a situation that in current syntactic theory largely falls under the phenomenon of clitic left dislocation (see section 1.3). Secondly, Lopašov relates clitic doubling to the definiteness of the associate, acknowledging however that clitic doubling of a strong pronoun is more widespread than that of a non-pronominal DP. Finally, he particularly emphasizes the impossibility of doubling focalized objects (i.e., objects bearing logical accent in his terminology).

The contributions in this volume convey more detailed information with regard to the general picture in (12) and the language-specific peculiarities of clitic doubling across the languages under investigation, broadening the empirical domain by including varieties of Albanian, Serbo-Croatian, Bulgarian, Asia Minor Greek, Macedonian, Macedo-Romanian, Megleno-Romanian and Romani dialects. The vastly enlarged set of data presented and discussed in this volume confirms the view that the genesis of clitic doubling is to be sought in the region uniting Central Ghëg Albanian, Western Macedonian and Northern Macedo-Romanian (see in particular the contributions by Friedman and Tomić).

## 2.2 Semantic effects

One of the most perplexing aspects of clitic doubling is the fact that across languages, doubling clitics affect interpretation in ways subject to various idiosyncratic constraints that make it very hard, if not altogether impossible, to define their function in a unitary manner. To illustrate, early generative (and non-generative) studies described clitic doubling as sensitive to the feature *humanness* in Rumanian and *animacy* in Spanish (Jaeggli 1986, Borer 1984, Dobrovie-Sorin 1990), a view that was already untenable for particular varieties of Romance in the presence of examples like those in (5) above, and also for the languages of the Balkan Sprachbund at large. With the latter coming into the focus of research on the topic, other semantic properties such as *prominence*, *specificity*, *presuppositionality*, *familiarity*, *definiteness* and *topicality* have increasingly been scrutinized as to their relevance for the phenomenon of clitic doubling.

### 2.2.1 Clitic doubling, definiteness and specificity

Among the factors characterizing clitic doubling across languages has been the ban on doubling indefinite and/or non-specific DPs. However, the generalization that clitic

---

14. See the contributions by Guentchéva and Krapova & Cinque in this volume.

doubling is subject to the definiteness and/or specificity of the associate DP has proved to be in need of further qualification for various languages. For instance, as pointed out by Friedman (this volume, example (2) and (3)), ‘a sweater’ in a sentence such as ‘I am knitting you a sweater’ can be clitic doubled in Greek but not in Macedonian, where in general the definite (enclitic) article seems to be a precondition for direct object clitic doubling. However, as shown by Tomić (this volume), Macedonian appears to tolerate some exceptions to this generalization: the [-definite] direct object is not incompatible with clitic doubling provided that it refers to a specific member from a definite/given set (as in her (4) ‘she managed to her<sub>i</sub> marry one<sub>i</sub> of her daughters [e.g., Jana]’, or as in her (5a) ‘I saw him<sub>i</sub> a man<sub>i</sub> how he<sub>i</sub> removed the boulder’, where a [+human] DP complement of a perception verb combines with a CP). In addition, if an indefinite object DP occurring clause-initially is contrastive, and therefore D-linked, clitic doubling is obligatory (as in her (7c) ‘another village<sub>i</sub> they burnt it<sub>i</sub>’). Indirect objects on the other hand, whether or not accompanied by a definite article, are clitic doubled if they are specific. However, this does not entail that specific indirect objects must be doubled: clitic doubling can always fail, independently of topicalization, definiteness or specificity (see Tomić’s (18a), literally: ‘I am giving (them<sub>i</sub>) books to the children<sub>i</sub>’; (18b), literally: ‘TO THE CHILDREN<sub>i</sub> I am giving (them<sub>i</sub>) books’; and (18c), literally: ‘To the children<sub>i</sub> I am giving (them<sub>i</sub>) BOOKS’).

In sum, in Macedonian, definite and/or specific D-linked direct objects are clitic doubled, non-specific indirect objects are not clitic doubled, other situations tending to free variation.<sup>15</sup>

This state of affairs leads to the prediction that at least non-D-linked direct objects instantiated by bare indefinites must be incompatible with doubling clitics, a prediction that generally holds: *Riba ne jadam* (literally: Fish not I-eat, ‘I don’t eat fish’ – see note 4 in Tomić) is an example of a non-doubled bare (generic) topicalized direct object. This being the case, it is not obvious how to account for data where clitic doubling in similar contexts seems licit. Such are for instance Leafgren’s example from Bulgarian (cited under (16) in Friedman, this volume): *Banan<sub>i</sub> ne običam da go<sub>i</sub> jam* (literally: Banana not like-I to it eat, ‘I don’t like bananas’), and particularly the counterexample collected in a Macedonian dialect of Northern Greece by Topolinjska (note 9 in Tomić, this volume): *i ko si dojde vremeto, go<sub>i</sub> rodi dete<sub>i</sub>* (literally: and when to-REFL came time-the, him<sub>i</sub> bore-she child<sub>i</sub> ‘and when the time came, she gave birth to a child’).

These data challenge the view that the doubled indefinites here could count as specific and D-linked, if as is standardly assumed bare singulars are property-denoting

---

15. The facts are even less clear than portrayed here, since even ‘Jana gave him<sub>i</sub> the letter to [a] child<sub>i</sub>’ is acceptable (see e.g., (13) in Tomić (this volume) and Tomić’s comment on that).

expressions (and therefore non-specific) and non-D-linked (and if, by hypothesis, D-linking only applies to specific DPs). How then can they ever behave as specific expressions?

Though the definition of specificity is notoriously unspecific, for the sake of exposition let us assume that while definiteness entails uniqueness within a given domain, specificity boils down to pinpointing a single specimen from among many of the same kind (Enç 1991). As observed by Farkas & von Heusinger (2003), specificity as a semantic parameter playing a crucial role in clitic doubling is connected to the referential stability of the associate (which, in turn, brings specificity close to definiteness, even if the distinction specific/non-specific only applies to indefinites, as pointed out by Ioup (1977)): clitic doubling requires a sufficiently high degree of referential stability of its associate, and referential stability is distributed differently depending on the semantic type of the associate expression involved (see Hill & Tasmowski, this volume, for an application to Old and Modern Romanian). This kind of approach explains the preferential doubling of strong personal pronouns and of names of animate beings, and further, its contextual variability with respect to definite descriptions and its possibility with partitively interpreted indefinites.

A further way of guaranteeing 'sufficient' specificity consists in the use of (indefinite) expressions such as 'a certain' and its counterparts in other languages, which point to supposed means in the speaker's mind for distinguishing the associate from the other members of the set it belongs to, and possibly also to hints in the mind of the hearer, the two cognitive agents having distinct bases of knowledge and different states of knowledge concerning the intended referent (see Martin 2005 for elaboration along these lines).

But why should specificity as an epistemic notion be limited to ear-marked expressions at all? From a pragmatic point of view, specificity can be based on perception (as we conjecture is the case in Tomić's example (5a) given above),<sup>16</sup> or on private awareness or expectation (which is arguably the case with the Greek example 'I am knitting you a sweater' and the Macedonian 'she gave birth to a – i.e., 'the' or 'her' – child'), or implicit common ground (as in example (9) in Friedman, this volume: '... language is one of the chief elements that it<sub>i</sub> characterizes a community<sub>p</sub>, an ethnic community'), with no other marking than the doubling clitic itself.

Analogous difficulties may be expected with a characterization in terms of D-linking once D-linking is construed to cover not only previous mention in the discourse but also a broader context, since in this latter case D-linking can be extended

---

16. See also Guentchéva's (1994: 116) Bulgarian example *kučeto ja goni edna kotka* (literally: dog-the it<sub>i</sub> is chasing a cat<sub>p</sub>, 'the dog is chasing a cat') [description of what is seen on a picture], where perception is still necessarily indicated explicitly by the use of *edin* (fem. *edna*) 'a'.

to immediate common background knowledge, or on the contrary be restricted to private knowledge.<sup>17</sup>

As witnessed by the contributions in this volume, languages differ in their admittance of what counts as (sufficiently) specific/specified/specifiable for clitic doubling, to a point where the semantic/pragmatic constraints become so allusive that free variation, alternatively null variation and complete grammaticalization, obtains. This seems actually to be the situation concerning clitic doubling of indirect objects, at least in the junction region of Ohrid Macedonian, Macedo-Romanian, Megleno-Romanian and Standard Albanian, where clitic doubling is always possible, or obligatory. This, in turn, substantiates the different behaviour of direct and indirect objects with respect to clitic doubling already signalled in section 1.2: At least in Macedonian and Albanian, clitic doubling of indirect objects can be equated with object agreement (see Franco 2000 and section 2.4).

### 2.2.2 *Clitic doubling and topicality*

An alternative view is that the opposition between the Macedonian and the Bulgarian examples *Riba ne jadam* and *Banan<sub>p</sub> ne običam da go<sub>i</sub> jam* mentioned above is tied to the notion of *topic* rather than to, or in addition to, that of *specificity*. However, in view of the massive varieties of opinions regarding what topics are, the notion of topic, just like the notion of specificity, does not prove very safe to use either.

The relation established for Modern Greek between the given/new information continuum and clitic doubling as “the device that removes the object from the comment (new part) of the sentence and renders it part of the background (known) information” (quoted from Holton, Mackridge & Philippaki-Warbuton (1997) by Janse, this volume), boils down to a link between topicality and state of knowledge not alien to the discussion in section 2.2.1: various authors defend the view that a topic is always specific, as one cannot produce information about something that is entirely inaccessible (Cohen & Erteschik-Shir 2002; Endriss 2006). If so, either *riba* and *banan* cannot be topics and clitic doubling of *banan* should be excluded (contrary to fact), or *riba* and *banan* should be recognized as *bona fide* topics because they refer to a kind, an object of ⟨e⟩-type (i.e., individual) denotation, in which case *riba* should be clitic doubled (again contrary to fact). One way out of this conundrum would be by drawing finer-grained distinctions – that is, by distinguishing between several types of topics.

From a psychological point of view, different degrees of topicality can be distinguished (Chafe 1994), corresponding to the degree to which the referent of an expression is supposed to actually be activated in the hearer’s mind (that is, whether it is active, semi-active or inactive). This difference in degree of activation could then be

---

17. See also Rochemont’s (1986) notion of *c-construability*.

said to correlate to the different word orders as given in (13), where # marks an intonational break (see the contributions in this volume by de Boel, Tsakali & Anagnostopoulou, Guentchéva, Krapova & Cinque, and particularly Janse).<sup>18</sup>

- (13) a. {V, Clitic} Object  
 b. Object {V, Clitic}  
 c. {V, Clitic} # Object  
 d. Object # {V, Clitic}

In case-marking languages, (13d) is characteristically a *nominativus pendens*, in non-case marking languages it is what has come to be known as a *Hanging Topic* (i.e., an argument appearing outside the minimal propositional unit without case-marking signalling its purported function inside the sentence).<sup>19</sup> A hanging topic announces in a loose syntactic manner what the minimal propositional unit, grammatically complete on its own, is about.<sup>20</sup> One diagnostic of hanging topics is that they exhibit no connectedness effects (see especially the contributions in this volume of Krapova & Cinque, Giusti & Stavrou and Tsakali & Anagnostopoulou). But in non-case marking languages, there exist in fact few clues to distinguish (13b) from (13d). As both are able to fulfil the requirement of the Tobler-Mussafia (and the Wackernagel) ban on clitics in first position, the two constructions involve a constituent located at least inside the CP, and in the main clause, (13b) takes the appearance of a variant of (13d), the most noticeable difference between the two being reduced to the length of the following prosodic pause, a criterion notoriously difficult to handle.

Anagnostopoulou has identified a series of contexts (referenced in Tsakali & Anagnostopoulou, this volume) that legitimize the distinction of (13c) from (13a). Contrary to the structure in (13d), the object in (13c) (dubbed “afterthought” by Hyman

---

18. The superficial notation {V, Clitic} is to be interpreted contextually, either as Clitic+V or V+Clitic, abstracting away from a discussion about precise syntactic description.

19. We are using the label ‘minimal propositional unit’ here in a purely descriptive, non-theoretic sense, to denote the canonical string of constituents making up a proposition. Its precise syntactic constitution is of course dependent on the analysis that one assumes for the proposition (and consequently for hanging topic).

20. For spoken French, Claire Blanche-Benveniste (University of Aix-en-Provence) provided us with examples such as (i) with no clitic inside the clause at all (out of the corpus of spoken French gathered by the University of Aix (GARS)). The example in (i) clearly shows that the relation between a hanging topic, here ‘this hotel’, and the following proposition is not syntactic, but relies on lexical and/or world knowledge.

(i) *Cet hôtel, on ne peut pas dormir.*  
 ‘This hotel, one cannot sleep in.’

(1975) and “anti-topic” by Lambrecht (1981)) brings some qualification, often affective (see de Boel, this volume, and Guentchéva, this volume) about the already intended topic, and can thus never introduce a novel topic into the discourse (Lambrecht 1996: 203 ff.). The difference between *Riba ne jadam* and *Banan ne obiçam da go jam* then is due to the fact that only in the latter, the bare noun is a hanging topic, with the construction falling thus outside the realm of clitic doubling. (We resume the discussion of bare nouns and the (im)possibility of clitic doubling them in section 2.3.)

Historical material presented in this volume for Greek and Bulgarian (de Boel, Janse, Dimitrova-Vulchanova & Vulchanov) shows that clitic doubling first arose in contexts of type (13c) and (13d): for Old Bulgarian none of the examples of the Suprasliensis manuscript can be taken to instantiate the structure in (13a), whereas for Greek, the first attestations of (13a) do not seem to be anterior to the 14th century. But following our tentative definition in section 1, it is essentially (13a) that represents the structure of clitic doubling, with the doubling clitic providing a way of backgrounding its associate by lifting it out of the focus domain (the “new”/“non-given” information part of the sentence, cf. Kallulli 2000, this volume, as well as the discussion in section 2.2.4).

### 2.2.3 Further details: Core but lesser known data

As mentioned earlier, in some languages (e.g., Albanian and at least some varieties of Macedonian), datives are invariably doubled by clitics (see Kallulli 2000, this volume, Friedman, this volume, Tomić, this volume). Furthermore, in several languages (Albanian, Bulgarian, MG, Romanian), so-called “quirky subjects” are also obligatorily clitic doubled not only when they bear dative/genitive case, thus reminiscent of the Spanish and Romanian examples in (3), but also when they bear accusative case (and irrespective of whether they are definite or indefinite) as illustrated in (14a) and (14b) (cf. also Anagnostopoulou 1994, 2003; Krapova & Cinque, this volume).<sup>21</sup>

- (14) a. *Benin/ një djalë \*(e) shqetëson sjellja e tij.* (Albanian)  
 Ben.the<sub>ACC</sub> a boy CL,ACC,3S bothers behaviour<sub>NOM</sub> AGR his  
 ‘His (own) behaviour bothers Ben / a boy.’

21. The accusative DPs in (14) qualify as quirky subjects by virtue of the fact that they bind the pronouns inside the nominative themes. See also Bulgarian and Romanian (i) and (ii), where the subject ‘head’ is inalienable part of the object ‘the child’:

- (i) *Deteto<sub>i</sub> go<sub>i</sub> boli glavata<sub>r</sub>.* (Bulgarian)  
 child-the him hurts head-the  
 (ii) *Pe copil<sub>i</sub> îl<sub>i</sub> doare capul<sub>r</sub>.* (Romanian)  
 on child him hurts head-the  
 ‘The child has a headache.’

- b. *Ton Yánni \*(ton) ponái to kefáli tu.* (MG)  
 the Yánnis<sub>ACC</sub> him<sub>CL,ACC</sub> hurts the head<sub>NOM</sub> his  
 ‘Yannis has a headache.’

In fact, Krapova & Cinque (this volume) call such cases “true” instances of clitic doubling, free from any further grammatical or pragmatic requirement (but see section 2.2.4).

As to (non-quirky) non-pronominal direct object DPs, in some languages (e.g., MG) they *can* and in some others (e.g., Albanian) they *must* be clitic doubled if and only if they are outside the focus domain (irrespective of their [ $\pm$ definite] status), as illustrated by the Albanian data in (15) through (18); note here the complementarity of felicity conditions between the minimally different pairs with and without doubling clitics in the given contexts (examples adapted from Kallulli 2000).<sup>22</sup> That is, direct object doubling clitics trigger a topical (in the sense: given, or anti-focal) interpretation of their associate, a conclusion that is corroborated by the fact that in Albanian and MG existential constructions, the (non-topical) direct objects of ‘have’ cannot be doubled, as in (19a) for Albanian (from Kallulli, this volume) and (19b) for MG (adapted from Anagnostopoulou 1994: 19 (29a)).

- (15) a. What did Ana do?  
 b. *Ana \*(e) lexoi librin / një libër.* (Alb)  
 A.the<sub>NOM</sub> 3S,CL,ACC read book.the<sub>ACC</sub> / a book  
 ‘Ana read the book.’
- (16) a. What did Ana read?  
 b. *Ana \*(e) lexoi librin / një libër.*
- (17) a. Who read the/a book?  
 b. *Ana \*(e) lexoi librin / një libër.*
- (18) a. What did Ana do with the/a book?  
 b. *Ana \*(e) lexoi librin / një libër.*
- (19) a. *(\*I) kishte minj në gjithë apartamentin.* (Albanian)  
 CL,ACC,3PL had mice<sub>ACC</sub> in all apartment.the  
 ‘There were mice all over the apartment.’  
 b. *(\*Ta) ixe pontikia se olo to diamerisma.* (MG)  
 CL,ACC,3PL had mice<sub>ACC</sub> in all the apartment  
 ‘There were mice all over the apartment.’

22. Anagnostopoulou (1994), while acknowledging counterexamples, argues that direct object clitic doubling is contingent on definiteness. See Kallulli (2000, this volume) for discussion of this point.

However, even within one and the same language (here: Albanian) the broader picture is more complex than presented in (15) through (18), as revealed by the fact that direct objects instantiated by first and second person full pronouns are always clitic doubled (Kallulli 1995, this volume), as are direct objects instantiated by the quantifier ‘all’, a fact that also holds for (Argentinian) Spanish, Romanian and MG (see also Friedman, this volume), as is shown in (20).

- (20) a. *\*(I) pashë të gjithë.* (Albanian)  
 b. *\*(Los) vi todos.* (Argentinian Spanish)  
 c. *\*(Tus) idha olus.* (MG)  
 d. *\*(I-) am văzut pe toți.* (Romanian)  
 them<sub>ACC</sub> saw.1s all<sub>ACC</sub>  
 ‘I saw them all.’

Adding to this, clitic doubling of CP-complements in Albanian and MG triggers factivity, as in (21) and (22) (Kallulli 2006, this volume).

- (21) a. *Besova se Beni shkoi (por në fakt ai nuk shkoi).* (Albanian)  
 believed-I that Ben left (but in fact he not left)  
 ‘I believed that Ben left (but in fact he didn’t).’  
 b. *E besova se Beni shkoi (\*por në fakt ai nuk shkoi).*  
 CL,ACC,3S believed-I that Ben left (but in fact he not left)  
 ‘I believed the fact that Ben left (\*but actually he didn’t).’
- (22) a. *Pistepsa oti o Janis efije*  
 believed-I that the Janis left  
*(ala stin pragmatikotita den ejine kati tetio).* (MG)  
 (but in.the reality neg happened something such)
- b. *To pistepsa oti o Janis efije*  
 it<sub>CL,ACC</sub> believed-I that the Janis left  
*(\*ala stin pragmatikotita den ejine kati tetio).*  
 (but in.the reality not happened something such)

This picture gives rise to the following questions. First, if the nature of agreement and topic markers is indeed substantially different, as assumed at the end of section 2.2.1, why are doubling clitics employed as means for fulfilling such different functions? Second, why are datives and direct objects instantiated by first and second person full pronouns always clitic doubled? Third, where in this picture and how does the fact fit in that ‘all’-quantifier direct objects are obligatorily doubled? Fourth, what is the relation between presuppositionality and topichood/givenness?

In the next section, we try to unravel the interrelation between the various features (specificity, prominence, referential stability, presuppositionality, topichood, givenness) highlighted as relevant in the context of clitic doubling in the various contributions to this volume.

#### 2.2.4 *Systemizing the data: Relating specificity and topicality*

As was seen in the previous section, direct object doubling clitics in Albanian and Greek have a restricted distribution and (semantic) operator-like properties in that they mark their associates as ‘given’. We take the notion of topic to be identical to that of ‘given’ (in the sense of Schwarzschild 1999) – that is, the complement of focus. This definition is rather different from the one in Reinhart (1981), according to whom topics do not have to be familiar to the hearer (i.e., given) but simply constitute the entity that the proposition is about. However, in view of Krifka’s (2001) arguments on the existence of so-called “non-novel” indefinites, the notions of topichood and givenness can be conflated without worrying about indefinite topics.

So what are non-novel indefinites? Contrary to standard accounts of indefinites as expressions (NPs or DPs, depending on the theory) that introduce new discourse referents (Heim 1982), Krifka (2001) argues that non-novel indefinites are a special class of indefinites that pick up discourse referents that exist in the input context. Crucially, according to Krifka non-novel indefinites must be deaccented, an idea that is in tune with Schwarzschild’s (1999) contention that givenness is indicated by deaccentuation.

For a discourse referent to exist in the input context, it must either have been mentioned before in the immediate context,<sup>23</sup> or its existence must in some way be presupposed (e.g., through sensory salience, via world knowledge, or through accommodation).<sup>24</sup>

Evidence for non-novel indefinites stems for instance from adverbial quantification, as in the examples in (23), which illustrate the so-called “requantification problem” (Rooth 1985, 1995, von Stechow 1994). Specifically, the domain of quantification in (23) is given by the deaccented indefinite, which forces the assumption that deaccented indefinites may pick up existing referents and “requantify” over them.

- (23) a. *A freshman* usually wears a *BÁSEBALL* cap. (from Krifka 2001)  
           ‘Most freshmen wear a baseball cap.’
- b. *A FRÉSHMAN* usually wears a baseball cap.  
           ‘Most wearers of baseball caps are freshmen.’

Thus, it is suggested that deaccenting goes with specificity, at least on some takes of this phenomenon, namely those that equate specificity with presuppositionality or familiarity (Ioup 1977, Diesing 1992, Enç 1991). Non-novel (and therefore deaccented) indefinites help thus establish the link between givenness/topichood and specificity (and of course deaccentuation).

23. But note that the notion of previous mention in the discourse naturally does not establish a referent as non-novel when the latter is in the scope of negation.

24. Importantly, Krifka (2001) argues that this presupposition is typically accommodated.

Interesting in this context is also the fact that clitic doubling necessarily correlates with deaccentuation of a CP-associate, and that doubling and/or deaccentuation express the presupposed status of an embedded proposition (Kallulli 2006, this volume), which in turn is the basic ingredient of factivity (Kiparsky & Kiparsky 1970); see (21b) and (22b).

The insight that not all indefinites introduce new discourse referents is important in the context of clitic doubling, as it immediately accounts for counterexamples (e.g., the ‘sweater’ example discussed in section 2.2.1) to analyses such as those proposed in Anagnostopoulou (1994, 1999) and Anagnostopoulou & Giannakidou (1995) for MG, where clitic doubling is crucially dependent on definiteness, which in turn is related to the notion of familiarity (i.e., non-novelty) in these works. In sum, a desirable result of the perspective opened by analysing certain indefinite expressions as non-novel is the conflation of specificity and topichood/givenness (see also Cresti 1995 and Endriss 2006 on the interrelation between topicality and specificity).

Likewise, an obvious connection can be made between non-novelty and the notion of referential stability (Farkas & von Stechow 2003): referentially stable expressions are easily identifiable, and of course the less novel a discourse referent is, the more identifiable/accommodatable (and therefore the more referentially stable) it is. This conclusion is especially important in the context of the phenomenon of (dative) experiencer and/or quirky subject clitic doubling discussed in several of the contributions in this volume (see e.g., Hill & Tasmowski, Kallulli, and Krapova & Cinque). Dative, experiencer and/or quirky subject arguments are referentially stable expressions in that they often are human (or at least animate) and presuppositional in a sense that direct objects instantiated by full (3rd person) DPs do not need to be. To see this, consider the examples in (24):

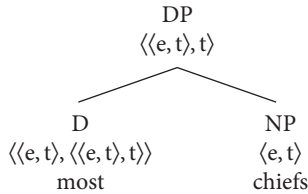
- (24) a. I bought my grand-daughter a Barbie.  
 b. I bought a Barbie for my grand-daughter.

Only (24b) but not (24a), which contains a (non-prepositional) dative object, is fine in a context in which the speaker has no grand-daughter. The moral to draw from this minimal pair is then that givenness does not necessarily entail previous mentioning of a referent in the discourse; it suffices that the existence of this referent can be inferred (for instance through accommodation).

Turning to the fact that direct object ‘all’-quantifiers in several languages are necessarily clitic doubled (as was shown in (20)), this can be derived by combining the view that its restrictor is always given (in the sense: non-novel) with the idea that doubling clitics trigger givenness and/or deaccentuation of their associates. We note, moreover, that in line with Matthewson’s (2001) analysis of quantification given in (25b), which differs from the traditional analysis of quantification (Barwise & Cooper 1981 and related literature) given in (25a) in that the generalized quantifier is formed

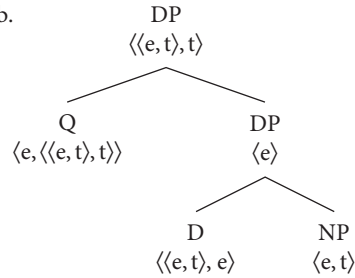
in two steps (crucially involving an intermediate DP-layer as the complement of the quantifier head), the restrictor of ‘all’ is always a DP, both when it is phonetically overt or null. As such, the clitic does not double the quantifier (i.e., the QP) but only its (non-novel/given/D-linked) DP-complement (which, as already mentioned, may be silent).

(25) a.



(Matthewson 2001: 146)

b.



(Matthewson 2001: 153)

In sum, it seems justified to state that the common denominator of the often vague notions specificity, topicality, prominence, presuppositionality is not ‘aboutness’ but ‘givenness’ or ‘d-linking’, which must be broadly understood as ‘existence in the input context’ (i.e., either previous mention in the immediate discourse, or presupposition through saliency or accommodation, which is in turn signalled by deaccentuation). This conflates the questions raised in section 2.2.1 and 2.2.2 above, which is precisely the path taken in the contributions in this volume.

### 2.3 Beyond information structure: The role of D in clitic doubling

As is well-known, with some exceptions as mentioned in section 2.2.1, across languages bare nouns of all sub-types (i.e., bare plurals, count bare singulars, mass nouns) cannot be clitic doubled (see *a.o.* Anagnostopoulou 1994, 1999, Casielles 1996, Kallulli 1998, 2000). This is a fact that, barring examples such as those quoted in 2.2.1, also holds when these bare nouns are construed as topical/given/D-linked. To see this, let us consider (26) from Farkas (1978).

- (26) a. *Popescu vrea pâine.*                      b. *Pâine are deja.*                      (Romanian)  
           ‘Popescu wants bread’                      bread has-he already

In the context in (26), the bare noun *pâine* ‘bread’ in the utterance of speaker B can only be interpreted as given. Yet, it cannot be clitic doubled in spite of its being left-dislocated, a fact that is important, since in Romanian also non-human objects that are therefore not preceded by the prepositional element *pe* (see Cornilescu & Dobrovie-Sorin, Hill & Tasmowski, this volume) can, or must, be clitic doubled in such a context. This and analogous examples can be replicated in Albanian, where

as mentioned earlier direct objects need to be clitic doubled in order to be marked as given. So then, why can't bare nouns be clitic doubled? Adopting proposals in Kallulli (1997, 1998, 2000) for count bare singulars, bare plurals and (bare) mass nouns, we suggest that bare nouns cannot be clitic doubled because they are not DPs with a morphologically null D, but NPs altogether lacking a D-layer. Clitics on the other hand are clearly D-heads (alternatively, they have a [+D] feature), which explains among other things their morphological identity or similarity (depending on the language) with determiners. That is, the impossibility of clitic doubling bare nouns be they singular, plural or mass ones, is due to a feature mismatch between the [+D] clitic head and the [-D] bare noun. As such, clitic doubling cannot serve as a means for expressing the anti-focal/topical/given status of a direct object bare noun.

This analysis also enables us to explain the possibility of clitic doubling *banan* in the Bulgarian example *Banan<sub>i</sub> ne običam da go<sub>i</sub> jam* and *dete* in the Macedonian example *i ko si dojde vremeto, go<sub>i</sub> rodi dete<sub>i</sub>* given in section 2.2.1. Specifically, on semantic grounds (Bulgarian and Macedonian seem to be able to drop the definite article in certain cases, such as the examples under consideration) we may assume that in these examples the seemingly bare nouns are not truly bare, in the sense that they must have a morphologically null D (recall our contention that the interpretation of the doubled expressions in both examples is definite, namely generic).<sup>25</sup>

#### 2.4 Interim summary: Clitic doubling between syntax, semantics and morphology

The clitic appearing in the clitic doubling construction by all means shares fundamental characteristics with the simple clitic: it appears on the same host, participates in the same clusters, integrates the phi-features of the associate or antecedent, and is refractory to lack of previous prominence. As to simple clitics, at least since the second half of the 20th century, some traditional grammarians in the field of Romance studies have claimed that they should be considered agreement markers of the verb with its direct or indirect object, similar in status to inflectional agreement of the verb with its subject, a view that has been integrated in the generative paradigm by the admittance of an AgrOP(rojection), parallel to AgrSP. Inflectional agreement with the subject is furthermore always compatible with the overt presence of an associate exhibiting the same appropriate phi-features.<sup>26</sup>

25. Note, however, that generic readings in most Balkan languages which are endowed with articles are impossible with bare indefinites (Kallulli 1997).

26. For a characterization and a universal typology of agreement markers, see Bossong (1998).

A novel argument concerning the status of clitics as agreement markers can be found in the contribution by Tsakali & Anagnostopoulou (this volume). Tsakali & Anagnostopoulou note that in a range of languages clitic doubling and participial agreement are in complementary distribution. In other words, a language has either participial agreement, or clitic doubling, but not both.

However, while clitic doubling constructions are by their very nature strongly reminiscent of object agreement constructions in languages like Basque, numerous studies have established that there are still essential differences between the two. As we have discussed above, these differences do not only bear on the typically affixal nature of the latter versus the typically non-affixal nature of the former in simple declarative clauses. One particular point that does not allow for a straightforward equation of clitic doubling and agreement is the fact mentioned earlier, namely that direct object clitic doubling is contextually constrained in ways that indirect object clitic doubling does not seem to be.

### 3. Further challenges, approaches and perspectives

#### 3.1 The interaction of clitic doubling with Differential Object Marking (DOM)

If doubling clitics mark their associates as topical, one does not expect clitic doubling to be compatible with environments where the associate of the doubling clitic is part of the focus domain. However, Romanian direct objects seem to contradict this prediction. For instance, in a context like the one in (27a) either of the sentences in (27b) are fine.

- (27) a. *Pe cine ai invitat?*      b. i. *L<sub>i</sub>-am invitat pe Radu<sub>i</sub>.*  
       ‘Who did you invite?’            ‘I invited Radu.’  
   ii. *Nu l<sub>i</sub>-am invitat decât pe Radu<sub>i</sub>.*  
   iii. *L<sub>i</sub>-am invitat numai pe Radu<sub>i</sub>.*  
   ‘I only invited Radu.’

In the sentences under (27b), *pe Radu* is not preceded by an intonational break and the direct object does not occupy a syntactic A'-position, as argued by Cornilescu & Dobrovie-Sorin (this volume).

Recall that in Romanian, clitic doubling is dependent on the presence of *pe* before a direct object DP. It is a DOM device similar to Spanish *a*, but with distinct effects (as discussed in Cornilescu & Dobrovie-Sorin, this volume) with respect to compatibility with clitic doubling. But if as suggested in section 2.2.4 prominence is to be accounted for in terms of non-novelty/givenness/D-linking, then the occurrence of the doubling clitics in the sentences in (27b) remains in need of explanation since their

associates here need not (and in fact cannot) be deaccentuated. Importantly however, the doubled expressions in the sentences in (27b) are contrastively focused. As such, they are necessarily D-linked in spite of the fact that they are not deaccentuated. And as is well-known at least since Rochemont (1986), contrastive focus is a special type of focus.<sup>27</sup> In this way, the view that clitic doubling is related to givenness can still be maintained. The conclusion then is that clitic doubling is not incompatible with identificational focus (Kiss 1998) in Romanian. In fact, it is feasible that the clitic in constructions such as those in (27b) doesn't double the *pe* expression, but rather a phonetically null DP (such as a bound variable *pro*), along the lines suggested in Kallulli (this volume) for clitic doubling of so-called D-linked *wh*-phrases in Albanian. This still does not account for (27bi) however. In (27bi), *pe Radu* is new information, but it is not contrastive. To maintain our general line of reasoning then, we are led to admit that clitic doubling is triggered by the pervasive presuppositionality tied to stability of reference (see Hill & Tasmowski, this volume).

### 3.2 Clitic doubling as a core syntax phenomenon

It is well-known that, across languages, so-called “given” information (as opposed to “new” information) systematically correlates with lack of phonetic prominence, or deaccentuation (Halliday 1967, Ladd 1996, Selkirk 1995, Schwarzschild 1999, among many others). Since, except for the discussed cases, clitic doubling systematically correlates with both deaccentuation and givenness of the doubled expressions (cf. Anagnostopoulou 1994, 1999 and especially Kallulli 2006, this volume), it could be hypothesised that doubling clitics function as a deaccenting device in that they affect the interpretation of their doubled associates in a way comparable with the deaccentuation of indefinites in English (cf. Krifka 2001). However, since PF and LF are two different components that do not communicate with each other (Chomsky 1995 et seq.), and since prosody is standardly assumed to be part of the PF component, the question arises how to model this correlation.

Inspired by Chomsky's (2004) “best case” scenario, according to which the components of the derivation of ⟨PHON, SEM⟩ proceed cyclically in parallel, the systematicity of the LF/PF correlation with respect to clitic doubling can be captured as in Kallulli (2006). One may think of prosodic information as encoded in the syntax, or indeed as part of the numeration itself, which is what syntax manipulates. Alternatively, the relevant syntactic feature (encoded in the form of a clitic pronoun) is interpretable at the LF interface (in terms of givenness), and also gets spelled out prosodically in a certain way at the PF interface, hence accounting for the correlation

---

27. Two hallmarks of contrastive focus are more pronounced accent and the possibility of using a cleft construction (instead).

of givenness and deaccentuation that was observed earlier. Crucially, however, under both views prosody can be treated as an abstract morpheme in the sense of Embick and Noyer (2007). As far as we can judge, both views may turn out to be empirically equivalent. Note also that either alternative can be made to comply with Chomsky's (2004: 107) architectural "best case":

[T]he best case is that there is a single cycle only [...] and the [...] components of the derivation of ⟨PHON, SEM⟩ proceed cyclically in parallel. L contains operations that transfer each unit to  $\Phi$  and to  $\Sigma$ . At the best case, these apply at the same stage of the cycle. Assume so.

In sum, clitic doubling is a narrow syntax operation, both under general lexicalist as well as non-lexicalist assumptions (see also Kallulli 2006, this volume). This, in turn, gives rise to non-trivial questions relating to the derivation-by-phase model (Chomsky 2001 et seq.), which future research will need to tackle.

#### 4. Conclusion

The Balkan clitic doubling patterns have raised important issues that have challenged even the most basic assumptions about the phenomenon of clitic doubling.<sup>28</sup> However, in spite of the substantial body of literature and important new empirical evidence, there exists to date no study that details the distinguishing peculiarities of this prevailing Balkan Sprachbund phenomenon across these languages, which would naturally lead to a better understanding of it. The present volume aims to fill this gap. It addresses specifically the question why clitic doubling appears at all, whether a uniform account of its various manifestations across (but also outside) the Balkan Sprachbund is feasible, the nature of micro-variation, and other issues in focus of present theoretical interest.

#### 5. The organization of the volume

The volume is organized according to the three major topics listed below, though naturally most of the contributions deal with more than one of the issues covered by these topics.

---

28. Important works in this context include Agouraki (1993), Anagnostopoulou (1994, 1999, 2003), Anagnostopoulou & Giannakidou (1995), Alexiadou & Anagnostopoulou (2000) for MG; Dobrovie-Sorin (1990, 1994), Gierling (1997), Farkas & von Heusinger (2003) for Romanian; Kallulli (1995, 1999a,b, 2000) for Albanian; Guentchéva (1994), Leafgren (1997), Rudin (1997), Dimitrova-Vulchanova & Hellan (1998), Arnaudova (2003) for Bulgarian, among many others.

- Distribution area of the phenomenon of clitic doubling, as well as its genesis and history within the Balkan continuum;
- Defining characteristics and/or distinctive factors of clitic doubling and related constructions (e.g., clitic left dislocation, clitic right dislocation, and hanging topic constructions) across the various languages, distinguishing between:
  - description of morpho-syntactic/semantic/pragmatic correlations; and
  - their formalization;
- Extension of the (initial) definition of the phenomenon.

Victor Friedman's lead-in article presents a broad descriptive perspective on clitic doubling, pulling together a wide range of facts from an unusually comprehensive list of Balkan languages and their dialects. A central point of the paper is that the geographical distribution of clitic doubling and its degree of integration into the various Balkan linguistic systems arguably reflects different diachronic stages as well as different synchronic systems. As the author puts it "[o]f particular importance is the fact that the phenomenon shows varying degrees of encoding (as pragmatic or grammatical devices) on the basis of areal rather than genealogical relations". Friedman argues that clitic doubling is more highly grammaticalized in the West Balkans than in the East, and that the variations shown by Balkan Slavic, Balkan Romance, and Albanian point to the areality of this feature.

Olga Mišeska Tomić investigates clitic doubling in the Balkan Slavic languages, whose dialects form a dialectal continuum also with respect to clitic doubling. She argues that clitic doubling in these languages shows gradual variation along a vertical north-south axis and a horizontal east-west axis. On the north-south axis, there is variation with respect to the categories that can be clitic doubled. On the east-west axis, languages/dialects vary with respect to the conditions on clitic doubling, with almost total dependence on discourse factors in the easternmost dialects in the area and relatively complete grammaticalization in the westernmost ones. In the majority of the Macedonian dialects, discourse factors do not play any role and all definite direct objects and all specific indirect objects are clitic doubled. In Western Macedonia, the two vertical north-south and the horizontal east-west axes (along which clitic doubling variation in Balkan Slavic moves) intersect, so that in the Western Macedonian dialects, as well as in Standard Macedonian, clitic doubling is obligatory with all definite direct and all specific indirect objects. In the case of indirect objects, the specificity effects disappear. The paper also describes clitic doubling in Aromanian and Megleno-Romanian, two Romance Balkan languages in close contact with the Western Macedonian dialects.

Gunnar de Boel's contribution focuses on the emergence of clitic doubling in Ancient Greek and its evolution into Medieval Greek. According to de Boel, the construction was felt to be vulgar, as the oldest isolated examples can be found in the (private) papyri of the Hellenistic age, but they hardly appear in literary texts.

Clitic doubling thus existed in its modern form in the oldest texts that use the demotic language (from the twelfth century onwards).

The article by Mila Dimitrova-Vulchanova and Valentin Vulchanov addresses the origin of the phenomenon of clitic doubling in the Balkan linguistic continuum based on evidence from the earliest Old Bulgarian manuscripts, thus contributing rare diachronic data to the problem of clitic doubling. The authors argue that clitic doubling in Old Bulgarian as in Modern Bulgarian are exclusively sanctioned by a specific information structure value that they define as [+contrastive topic] and which they claim is a sufficient trigger for the surfacing of doubling clitics.

In their contribution, Virginia Hill and Liliane Tasmowski offer a semantic-pragmatic approach to Romanian clitic doubling, paying special attention to three peculiarities of this construction in this language, namely (i) the obligatory preposition *pe* preceding the direct object DP; (ii) the interaction between this type of DOM and the doubling clitic; and (iii) the nature of the asymmetry between direct and indirect object clitic doubling. Specifically, the authors argue that indirect objects meet the pragmatic prominence conditions for clitic doubling due to their morpho-semantics, whereas direct objects need categorial marking (i.e., *pe*) to qualify for this process. The analysis is corroborated by historical material from Old Romanian.

Mark Janse surveys changes in the positioning and discourse function of clitics throughout the history of Greek, focusing on the set of clitic doubling constructions and on the analysis of relevant data from Asia Minor Greek in particular, thus providing a range of novel data and showing that in many ways these varieties continue a situation obtaining in Medieval Greek.

Zlatka Guentchéva's article on clitic doubling in Modern Bulgarian focuses on the distinction between clitic doubling and 'left-dislocated' constructions. On this basis, it then examines the interaction of clitic doubling with constituent order and information structure. The author underscores the role of clitic doubling as a topic marker, and investigates the effects of the construction in the structuring of texts.

Dalina Kallulli presents a unified account of direct and indirect object clitic doubling in Albanian, which rests on the claim that clitic doubling is always a spell-out of agreement with a topical argument, which in turn entails referential stability. The paper also extends the analysis in Kallulli (2000) to account for certain problematic aspects (such as clitic doubling asymmetries in relative clauses and *wh*-constructions) not covered in previous work. In addition, this paper presents an investigation of the semantic effects of clitic doubling of clausal arguments (i.e., CPs), which also corroborates the author's basic claim. This is the only paper of the volume that also brings prosodic evidence to bear on the morpho-syntax and semantics of doubling clitics.

Iliyana Krapova and Guglielmo Cinque's article focuses on what the authors claim to instantiate "clitic doubling proper" in Bulgarian, which in their perspective is restricted to clauses with psychological and physical perception predicates. Several properties

of this pattern are identified, among which obligatory doubling of quantifiers, *wh*-phrases and focus phrases. Other patterns of clitic doubling as they appear throughout the Balkan languages (including Bulgarian dialects) are also identified.

The article by Alexandra Cornilescu and Carmen Dobrovie-Sorin examines a contrast between Spanish and Romanian, namely that only in the former but not in the latter clitic doubling of direct object DPs is blocked by contrastive focus and quantificational features. The paper offers a formal analysis of this contrast that relies on the following theoretical ingredients: (i) Head-Movement is Head-to-Head Merge; (ii) clitic placement is a Spec-Head agreement configuration with a null pronoun *pro* sitting in the Spec of (the complex head containing) the clitic; (iii) clitic doubling results from an interarboreal operation that merges a complex head  $Cl+Vv+T(ense)$  with the VP containing the clitic doubled DP; (iv) the contrasts between Romanian and (River Plate) Spanish is due to the fact that in Spanish, Spec of CP is distinct from Spec of  $Cl+Vv+T$ , whereas in Romanian, Comp is part of the complex functional head clustering around T, and correlatively, Spec of C is not distinct from, but rather a slot inside the Spec of the complex head  $Comp+Cl+Vv+T$ .

The contribution by Vina Tsakali and Elena Anagnostopoulou focuses on the parameters that regulate the cross-linguistic distribution of clitic doubling. The central claim of this paper is that the presence of participial agreement determines the availability of clitic doubling: participle agreement excludes clitic doubling, and vice versa (language internally as well as construction-specifically). More specifically, the authors argue that the presence of participle agreement in clitic-languages induces split-checking of phi-features, which forces the associate of the clitic to be phonetically null (*pro*). When no split-checking ensues, a language may optionally be a clitic-doubling language. This paper also brings acquisition data to bear on the phenomenon and its analysis by the authors.

Finally, two contributions in this volume, one by Larisa Avram and Martine Coene and another by Giuliana Giusti and Melita Stavrou, deal with the phenomenon of clitic doubling within the DP.

Avram and Coene question the view according to which Romanian dative/genitive possessive clitics can be placed both DP-internally and DP-externally, and argue instead that the clitics in the two constructions are only superficially identical. The clitic within the DP is a possessive clitic, valued genitive, which does not move out of the DP and cannot be doubled. The one outside the DP, at the left periphery of the clause, is an indirect object base-generated inside the vP, and valued dative. Its possessive interpretation is context-dependent, being semantically (or pragmatically) determined. The difference with respect to the availability of possessive clitic raising and doubling is accounted for within a derivation-by-phase framework (Chomsky 2001). The authors' basic proposal is that DP-internal clitics are "frozen" within the DP-phase and can therefore not move to the left periphery of the clause. Both the impossibility

of their moving out of the DP to the clausal domain and the ambiguity of the sentences containing clausal dative clitics are accounted for in terms of the Attract Closest condition redefined in terms of phases.

Giusti and Stavrou aim to explain two awkward differences between Bulgarian and Greek doubling clitics within the DP: (i) only in Bulgarian but not in Greek a Topicalized Possessor is resumed (or doubled) by the corresponding dative (or genitive, depending on the language) clitic; (ii) only in Greek but not in Bulgarian a strong pronoun Possessor is resumed by the corresponding genitive/dative clitic. The authors propose that only Bulgarian but not Greek has a topic position available within its (extended) DP-structure, and what seems to be clitic doubling inside the DP is in fact an instance of clitic left dislocation. As to the doubling of strong pronoun Possessors in Greek, the authors contend that it is due to a deficiency of the strong pronoun and that it should be altogether distinguished from clitic doubling. The authors conclude that neither Bulgarian nor Greek have clitic doubling inside the DP.

## References

- Adger, D. 2003. *Core Syntax*. Oxford: OUP.
- Agouraki, Y. 1993. Spec-Head Licensing: The Scope of the Theory. Ph.D. dissertation, University College London.
- Alexiadou, A. & Anagnostopoulou, E. 2000. Asymmetries in the distribution of clitics: The case of Greek restrictive relatives. In Beukema, F. & den Dikken, M. (Eds), 47–70.
- Anagnostopoulou, E. 1994. Clitic Dependencies in Modern Greek. Ph.D. dissertation, University of Salzburg.
- Anagnostopoulou, E. 1999. Conditions on clitic doubling in Greek. In van Riemsdijk, H. (Ed.), 761–798.
- Anagnostopoulou, E. 2003. *The Syntax of Ditransitives: Evidence from Clitics*. Berlin: Mouton de Gruyter.
- Anagnostopoulou, E. & Giannakidou, A. 1995. Clitics and prominence, or why specificity is not enough. *Papers from the Chicago Linguistics Society* 31: 1–14. University of Chicago.
- Arnauodova, O. 2003. Focus and Bulgarian Clause Structure. Ph.D. dissertation, University of Ottawa.
- Barwise, J. & Cooper, R. 1981. Generalized quantifiers and natural language. *Linguistics and Philosophy* 4: 159–219.
- Beukema, F. & den Dikken, M. (Eds). 1999. *Clitic Phenomena in European Languages*. Amsterdam: John Benjamins.
- Borer, H. 1981. Parametric Variation in Clitic Constructions. Ph.D. dissertation, MIT.
- Borer, H. 1984. *Parametric Syntax: Case Studies in Semitic and Romance Languages*. Dordrecht: Foris.
- Bosson, G. 1998. Vers une typologie des indices actanciels. Les clitics romans dans une perspective comparative. In Ramat, P. & Roma, E. (Eds), *Atti del 30 congresso internazionale della Società Linguistica Italiana*. 9–43. Roma: Bulzoni.

- Caragiu Marioțeanu, M. 1975. *Compendiu de dialectologie română*. București: Ed. Științifică și Enciclopedică.
- Casielles, E. 1996. On the misbehavior of bare nouns in Spanish. In Parodi, C. et al. (Eds), *Aspects of Romance Linguistics. Selected Papers from the 24th Linguistic Symposium on Romance Languages*. 135–148. Washington DC: Georgetown University Press.
- Cecchetto, C. 2000. Doubling structures and reconstruction. *Probus* 12: 1–34.
- Chafe, W. L. 1994. *Discourse, Consciousness and Time*. Chicago IL: Chicago University Press.
- Chomsky, N. 1981. *Lectures on Government and Binding*. Dordrecht: Kluwer.
- Chomsky, N. 1995. *The Minimalist Program*. Cambridge MA: The MIT Press.
- Chomsky, N. 2001. Derivation by phase. In Kenstowicz, M. (Ed.), *Ken Hale: A Life in Language*. 1–52. Cambridge MA: The MIT Press.
- Chomsky, N. 2004. Beyond explanatory adequacy. In Belletti, A. (Ed.), *Structures and Beyond*. 104–131. Oxford: OUP.
- Cohen, A. & Erteschik-Shir, N. 2002. Topic, focus, and the interpretation of bare plurals. *Natural Language Semantics* 10: 125–165.
- Cresti, D. 1995. Indefinite Topics. Ph.D. dissertation, MIT.
- Diesing, M. 1992. *Indefinites*. Cambridge MA: The MIT Press.
- Dimitrova-Vulchanova, M. & Hellan, L. 1998. Introduction. In Dimitrova-Vulchanova, M. & Hellan, L. (Eds), *Topics in South Slavic Syntax and Semantics ix–xxvii*. Amsterdam: John Benjamins.
- Dobrovie-Sorin, C. 1990. Clitic doubling, wh-movement and quantification in Romanian. *Linguistic Inquiry* 21: 351–397.
- Dobrovie-Sorin, C. 1994. *The Syntax of Romanian*. Berlin: Mouton de Gruyter.
- Embick, D. & Noyer, R. 2007. Distributed morphology and the syntax/morphology interface. In Ramchand, G. & Reiss, C. (Eds), *The Oxford Handbook of Linguistic Interfaces*. 289–324. Oxford: OUP.
- Enç, M. 1991. The semantics of specificity. *Linguistic Inquiry* 22: 1–25.
- Endriss, C. 2006. Quantificational Topics. Ph.D. dissertation, University of Potsdam.
- Farkas, D. 1978. Direct and indirect object reduplication in Romanian. In Farkas, D., Jacobsen, W. & Todrys, K. (Eds), *Papers from the Fourteenth Regional Meeting of the Chicago Linguistic Society*. 88–97.
- Farkas, D. & von Stechow, K. 2003. Stability of reference and object marking in Romanian. Handout at the ESSLI Workshop on Direct Reference and Specificity. Vienna, August.
- Franco, J. 1993. On Object Agreement in Spanish. Ph.D. dissertation, University of Southern California.
- Franco, J. 2000. Agreement as a continuum. The case of Spanish pronominal clitics. In Beukema, F. & den Dikken, M. (Eds), 147–189.
- Gierling, D. 1997. Clitic doubling, specificity and focus in Romanian. In Black, J. & Motapanyane, V. (Eds), *Clitics, Pronouns and Movement*. 63–85. Amsterdam: John Benjamins.
- Guentchéva, Z. 1994. *Thématisation de l'objet en bulgare*. Frankfurt: Peter Lang.
- Gutiérrez-Rexach, J. 2000. The formal semantics of clitic doubling. *Journal of Semantics* 16: 315–380.
- Gutiérrez-Rexach, J. 2001. Interface conditions and the semantics of argument clitics. In Gutiérrez-Rexach, J. & Silva-Villar, L. (Eds), *Current Issues in Spanish Syntax and Semantics*. 107–142. Berlin: Mouton de Gruyter.
- Gutiérrez-Rexach, J. 2002. Constraint interaction at the semantics/pragmatics interface. In Jaszczolt, K. & Turner, K. (Eds), *Meanings in Contrast. The Cambridge Papers 335–354*. Amsterdam: John Benjamins.

- Halliday, M. 1967. *Intonation and Grammar in British English*. Berlin: Mouton de Gruyter.
- Heim, I. 1982. The Semantics of Definite and Indefinite Noun Phrases. Ph.D. dissertation, University of Massachusetts, Amherst.
- Hyman, L. 1975. On the change from SOV to SVO: Evidence from Niger-Congo. In Li, C.N. (Ed.), *Word Order and Word Order Change* 113–47. Austin TX: University of Texas Press.
- Ioup, G. 1977. Specificity and the interpretation of quantifiers. *Linguistics and Philosophy* 1: 233–245.
- Jaeggli, O. 1982. *Topics in Romance Syntax*. Dordrecht: Foris.
- Jaeggli, O. 1986. Three issues in the theory of clitics: Case, doubled NPs, and extraction. In Borer, H. (Ed.), *Syntax and Semantics 19: The Syntax of Pronominal Clitics* 15–42. New York NY: Academic Press.
- Kallulli, D. 1995. *Clitics in Albanian*. University of Trondheim Working Papers in Linguistics 24.
- Kallulli, D. 1997. NP predicates. In Blight, R. & Moosally, M. (Eds), *Proceedings of the 1997 Texas Linguistics Society conference "The Syntax and Semantics of Predication"*. Texas Linguistic Forum 38: 173–188.
- Kallulli, D. 1998. The common basis of DO clitic doubling and scrambling. In Cambier-Langeveld, T., Lipták, A. & Redford, M. (Eds), *Proceedings of ConSole VI* 103–121. Leiden: SOLE.
- Kallulli, D. 1999a. The argument-predicate distinction and the non-optionality of DO clitic doubling and scrambling. In Shahin, K.N., Blake, S.J. & Kim, E. (Eds), *Proceedings of the 17th West Coast Conference on Formal Linguistics (WCCFL)* 347–361. Stanford CA: CSLI.
- Kallulli, D. 1999b. The Comparative Syntax of Albanian. On the Contribution of Syntactic Types to Propositional Interpretation. Ph.D. dissertation, University of Durham.
- Kallulli, D. 2000. Direct object clitic doubling in Albanian and Greek. In Beukema, F. & den Dikken, M. (Eds), 209–248.
- Kallulli, Dalina. 2001. Restrictive relative clauses revisited. In Hirotani, M., Coetzee, A., Hall, N. & Kim, J. (Eds), *Proceedings of NELS 30*, 353–362. Amherst MA: GLSA.
- Kallulli, D. 2006. Triggering factivity: Prosodic evidence for syntactic structure. In Baumer, D., Montero, D. & Scanlon, M. (Eds), *Proceedings of the 25th West Coast Conference on Formal Linguistics (WCCFL)* 211–219. [Cascadilla Proceedings Project]. Somerville MA: Cascadilla.
- Kayne, R. 1975. *French Syntax*. Cambridge MA: The MIT Press.
- Kiparsky, P. & Kiparsky, C. 1970. Fact. In Bierwisch, M. & Heidolph, K. (Eds), *Progress in Linguistics*. 345–369. The Hague: Mouton.
- Kiss, K.É. 1998. Identificational focus versus information focus. *Language* 74: 245–268.
- Kratzer, A. 1996. Severing the external argument from its verb. In Rooryck, J. & Zaring, L. (Eds), *Phrase Structure and the Lexicon*. 109–137. Dordrecht: Foris.
- Krifka, M. 2001. Non-novel indefinites in adverbial quantification. In Condoravdi, C. & Renardel de Lavalette, G. (Eds), *Logical Perspectives on Language and Information* 1–40. Stanford CA: CSLI.
- Ladd, R. 1996. *Intonational Phonology*. Cambridge: CUP.
- Lambrecht, K. 1981. *Topic, Antitopic and Verb Agreement in Non-standard French*. Amsterdam: John Benjamins.
- Lambrecht, K. 1996. *Information Structure and Sentence Form*. Cambridge: Cambridge University Press.

- Leafgren, J.R. 1997. Bulgarian clitic doubling: overt topicality. *Journal of Slavic Linguistics* 5: 117–144.
- Lopašov, J. 1978. *Mestoimennyye povtory dopolnenija v balkanskih jazykah*. Leningrad: Nauka.
- Martin, F. 2005. Les indéfinis spécifiques ‘un certain N’ et ‘un N bien précis’. In Vogeleer, S. & Tasmowski, L. (Eds), *Déterminants indéfinis en français*. *Travaux de linguistique* 50: 151–168.
- Matthewson, L. 2001. Quantification and the nature of crosslinguistic variation. *Natural Language Semantics* 9: 145–189.
- Miklosich, F. 1862. Die slavischen Elemente im Rumänischen. *Denkschriften der Kaiserlichen Akademie der Wissenschaften, Philosophisch-historische Klasse* 12: 1–70.
- Reinhart, T. 1981. Definite NP anaphora and c-command domains. *Linguistic Inquiry* 12(4): 605–635.
- Reinheimer, S. & Tasmowski, L. 2005. *Pratique des langues romanes*. II. *Les pronoms personnels*. Paris: l’Harmattan.
- Rivas, A. 1977. A Theory of Clitics. Ph.D. dissertation, MIT.
- Rizzi, L. 1991. Residual verb second and the wh-criterion. *Technical Reports in Formal and Computational Linguistics* 2. Geneva: University of Geneva.
- Rochemont, M. 1986. *Focus in Generative Grammar*. Amsterdam: John Benjamins.
- Rooth, M. 1985. Association with Focus. Ph.D. dissertation, University of Massachusetts at Amherst.
- Rooth, M. 1995. Indefinites, adverbs of quantification and focus semantics. In Carlson, G.N. & Pelletier, F.J. (Eds), *The Generic Book*. 265–291. Chicago IL: Chicago University Press.
- Rudin, C. 1997. AgrO and Bulgarian pronominal clitics. In Lindseth, M. & Franks, S. (Eds), *Formal Approaches to Slavic Linguistics* 5: 224–52. Ann Arbor MI: Michigan Slavic Publications.
- Salvi, G. 2004. *La formazione della struttura di frase romanza*. Tübingen: Niemeyer.
- Schwarzschild, R. 1999. GIVENness, AVOIDF and other constraints on the placement of accent. *Natural Language Semantics* 7: 141–177.
- Selkirk, E. 1995. Sentence prosody: Intonation, stress, and phrasing. In Goldsmith, J.A. (Ed.), *The Handbook of Phonological Theory*. 550–69. Oxford: Blackwell.
- Sportiche, D. 1992. Clitic constructions. MS, UCLA.
- Sportiche, D. 1996. Clitic constructions. In Rooryck, J. & Zaring, L. (Eds), *Phrase Structure and the Lexicon*. 213–276. Dordrecht: Kluwer.
- Stavrou, M. 1984. Ἡ κλιτική αντωνυμία στις περιοριστικές αναφορικές προτάσεις με εξάρτηση άμεσου αντικείμενου που εισαγονται με το “που”. Μελετες για την Ελληνική Γλωσσα: 121–136.
- Steriade, D. 1980. Clitic doubling in the Romanian wh-constructions and the analysis of topicalization. *Papers of the Chicago Linguistic Society*. 282–297.
- Strozer, J. 1976. Clitics in Spanish. Ph.D. dissertation, UCLA.
- Suñer, M. 1988. The role of agreement in clitic-doubled constructions. *Natural Language and Linguistic Theory* 6: 391–434.
- Tomić O. Mišeska. (Ed.) 2004. *Balkan Syntax and Semantics*. Amsterdam: John Benjamins.
- Tomić, O. Mišeska. 2006. *Balkan Sprachbund Morpho-Syntactic Features*. Dordrecht: Springer.
- Torrego, E. 1995. On the nature of clitic doubling. In Campos, H. & Kempchinsky, P. (Eds), *Evolution and Revolution in Linguistic Theory*. 399–418. Washington DC: Georgetown University Press.

Torrego, E. 1999. *The Dependencies of Objects*. Cambridge MA: The MIT Press.

Uriagereka, J. 1995. Aspects of the syntax of clitic placement in Western Romance. *Linguistic Inquiry* 26: 79–123.

van Riemsdijk, H. (Ed.). 1999. *Clitics in the Languages of Europe*. Berlin: Mouton de Gruyter.

von Stechow, K.U. 1994. Restrictions on Quantifier Domains. Ph.D. dissertation, University of Massachusetts.

Zwicky, A. 1977. *On Clitics*. Bloomington IN: Indiana University Linguistics Club.

PART I

## **Clitic doubling within the Balkan continuum**

Rise and spread



# Balkan object reduplication in areal and dialectological perspective\*

Victor A. Friedman  
University of Chicago

When examined in its Balkan context, object reduplication tells us a number of things about language contact phenomena in general and Balkan contact phenomena in particular. It provides a striking illustration of the way a pragmatic phenomenon becomes syntactic and finally grammaticalized almost to the point of morphologization. The areal distribution of reduplication and its degree of integration into the various Balkan linguistic systems arguably reflects different diachronic stages as well as different synchronic systems. Of particular importance is the fact that the phenomenon shows varying degrees of encoding (as pragmatic or grammatical devices) on the basis of areal rather than genealogical relations. At the same time, this geographic distribution is additional evidence for the manner in which grammatical change can be triggered by pragmatic devices: object reduplication is more highly grammaticalized in the west Balkans than in the east, and the variations shown by Balkan Slavic, Balkan Romance, and Albanian point to the areality of this feature. These data also demonstrate the importance of taking dialectology and less commonly examined Balkan languages (Aromanian, Romani) into account when discussing degrees of Balkanization.

## 1. Introduction

### 1.1 Reduplication as a term and as a phenomenon

In Balkan linguistic studies, *object reduplication* refers to the occurrence of a word or word-like unit (clitic, short or weak form pronoun) that has exactly the same grammatical role in a clause as another word, i.e., a clitic and long form of the same oblique

---

\*Much of the research for this article as well as an earlier draft were completed while I was a visiting fellow at the Research Center for Linguistic Typology at LaTrobe University, whose support I gratefully acknowledge here.

personal pronoun or a clitic agreeing in gender, number, and case with a substantival direct or indirect object.<sup>1</sup> The following example from Macedonian is typical:

- (1) *mu go davam moliv-ot na momče-to.* (Macedonian)  
 him.DAT it.ACC I.give pencil-DEF to boy-DEF  
 'I give the pencil to the boy.'

Object reduplication is characteristic of all the Indo-European Balkan languages, albeit its degree of grammaticalization – and thus its function – differs among these languages. This paper will give an overview of basic and more recent research on Balkan object reduplication, bringing together examples from a variety of sources including both published and unpublished dialect descriptions that have not heretofore been utilized in the study of object reduplication and the author's own field work (although I have utilized published examples in most instances, the Albanian, Aromanian, Macedonian, Megleno-Romanian and Romani data and the generalizations drawn therefrom were all checked with and/or collected from native speakers by me in Albania and Macedonia). In this survey of object reduplication in the Balkan languages I shall attempt to give a more nuanced account of its Balkan nature than has heretofore been available. Of particular importance in this regard will be the fact that reduplication phenomena show varying degrees of encoding (as pragmatic or grammatical devices) on the basis of areal rather than genealogical relation.<sup>2</sup> At the same time, this geographic distribution is additional evidence for the manner in which grammatical change can be triggered by pragmatic devices (cf. Mithun 1992; Friedman 1994; Fielder 1999; Wertheim forthcoming): object reduplication is more highly grammaticalized in the west Balkans than in the east, and the variation shown by Balkan Slavic, Balkan Romance, and Albanian points to the areality of this feature and southwestern Macedonia as the core zone. These data also demonstrate the importance of taking dialectology and less commonly examined Balkan languages (especially Aromanian, Megleno-Romanian, and Romani) into account when discussing degrees of Balkanization. By *Balkanization* I mean the development and grammaticalization of *Balkanisms* in the sense first used by Seliščev (1925). Regardless of whether the mechanism is understood to be what I will call here *intrusion* (the borrowing, interference, or imperfect learning and subsequent spread of a feature found in a contact language) or *feature selection* (the extension or increased development of

1. Subject reduplication, albeit a related type of phenomenon, will not be considered here. Complement clauses functioning as direct objects, however, will be discussed when relevant.

2. The term *genealogical* has come to be preferred to *genetic* among some linguists to refer to the relationship between languages that trace their primary origins to a common speech community. Although Sandfeld (1930) already demonstrated that Balkan reduplication differs from that found in the Romance languages of western Europe, nonetheless, the specifically Balkan nature of Balkan object reduplication remains to be investigated on its own.