CLITICS IN FOUR DIALECTS OF MODERN GREEK: A DYNAMIC ACCOUNT

Thesis Submitted in Partial Fulfillment
of the Requirements for the Degree of
Doctor of Philosophy

Stergios Chatzikyriakidis
September 2010

University of London
Department of Philosophy
King's College London

Abstract

This thesis is a study of the clitic systems of four dialects of Modern Greek: Standard Modern, Cypriot, Grecia Salentina and Pontic Greek. An account of the positioning system of each dialect is given, firstly from a synchronic perspective. Then, these synchronic accounts will be shown to be grounded in diachronic considerations. Specifically, the transition from Koine Greek to the medieval dialects of Greek and from these medieval dialects to the dialects under consideration is discussed and it is argued that the positioning system of each dialect is the result of the same processes driving syntactic change, namely routinization (Pickering & Garrod, 2004; Bouzouita, 2008a,b) and the emergence of generalized parsing triggers. These two phenomena will basically suffice to explain why the four dialects under consideration exhibit such diverse positioning systems. Furthermore, this thesis will deal with a number of other problematic clitic-related phenomena like Clitic Climbing, Doubling and the Person Case Constraint. Clitic Climbing will be shown to be straightforwardly captured assuming an analysis where restructuring verbs are treated as auxiliary-like verbs in the sense of Cann (forthcoming). Such an analysis will directly give us the Clitic Climbing facts. Doubling in the dialects under consideration will be also discussed, arguing that the availability of Clitic Doubling in a given language depends on whether clitics in this language have lost their bottom restriction or not. Lastly, this thesis provides crucial evidence supporting the account given by Cann & Kempson (2008) and Chatzikyriakidis & Kempson (2009) as regards the Person Case Constraint, according to which person case restrictions are the result of a very general treegrowth restriction, a hard wired tree-logic constraint that basically bans the projection of more than one node with the same underspecified address.

Acknowledgements

I would like to take this opportunity to convey my gratitude to a number of different people that have helped me in a variety of ways during the course of this PhD. First of all, I would like to thank my supervisor Ruth Kempson for her immense support on all levels. It is really impossible to accurately describe the amount of help she has given me on both an intellectual and a personal level. Furthermore, I would like to thank a number of people from the Philosophy Department at King's College, London for helping me develop my educational background. Wilfried Meyer-Viol and Shalom Lappin are the first people that come to my mind. The former, also being my co-supervisor, provided me with a solid background in formal semantics and has helped me with a number of formal issues throughout the course of this PhD. Shalom Lappin has also helped me in building a formal logic background and made me appreciate and understand the importance of interdisciplinary work. Furthermore, Eleni Gregoromichelaki is thanked for invaluable help throughout my years at King's College, especially my first year as an MSc student. She is also thanked for elucidating parts of the DS formalism in an overall simple and transparent way. Co-students Miriam Bouzouita and Andrew Gargett, both Drs now, are also greatly thanked. Firstly, for being friends. Furthermore, Miriam Bouzouita for numerous discussions on clitic related issues and for proofreading this thesis, and Andrew Gargett for numerous discussions on any kind of linguistic subject. Co-MSc student Aarne Talman is also thanked for being a friend and formal logic expert. People working outside King's College but within the DS project, notably Lutz Marten and Ronnie Cann are also thanked for providing me with intellectually stimulating discussions and valuable suggestions on various parts of this thesis. A number of other people, both academics and students, are thanked for providing me with help and advice during these years. Melita Stavrou, Janis Veloudis at Aristotle University,

Ioanna Sitaridou, Ian Roberts, Dimitris Michelioudakis and the Greek Dialectal Syntax reading group at Cambridge University, Stavroula Tsiplakou and Kleanthes Grohmann at the University of Cyprus, Panayiotis Pappas at Simon Fraser University, Anna Roussou at the University of Patras. The Arts and Humanities Research Council (AHRC) and the Leventis foundation are gratefully thanked for providing partial funding during the course of this PhD. Furthermore, I want to thank a number of people for providing me with data as well as help in order to find native speakers in both Grecia Salentina, Italy and Western and Central Macedonia, Greece. As regards Grecia Salentina Greek Paolo Dimitri, Luigi Tomassi, Salvatore Tomassi, Antonio Milano in Calimera, Medico de Pascalis and Giuseppe de Pascalis in Martano, Giorgio Filieri and his mother, Vincenzo Marti, Pantaleo Greco, Luigi Ingrosso, Giuseppe Pino, Gaetano Mastrolia, Cosimo Pantaleo Mastrolia in Sternatia. As regards the Pontic Greek data, my grandmother Efthimia Chatzikyriakidou in Palia Likogianni Imathias, Haris Stafilidis, Despoina Stafilidou, Anastasia Triantafillidou, Elli Toutountzidou, Eustathios Horiatzidis, Sofia Paraskeuaidou in Nea Nikomideia Imathias and Giorgos Davididis and Paulos Davididis in Mikrokleisoura Grevenon. Furthermore, Giorgos Stratis, Marilena Paraskeua, Theoni Neokleous, Melina Karaolia and Ivi Petrou for providing me with their intutions on Cypriot Greek.

On a personal level, a number of friends and relatives honored me with their love and friendship. Among them, my friends from Grevena: Jannis 'Bobby Marley' Panagakis (chaos and complexity) for the chaotic alcoholic nights, numerous discussions about everything that can or cannot be talked about. Tolis Moisiadis for getting in trouble (to gnosto mpleskimo) all the time, Vasilis Harizopoulos for providing me with tones of electronic books in the middle of my PhD (possibly I should not be thanking him about that but anyway), Georgia Tegou for being tsiormanos, George Mpakas and Vasilis Mitsiopoulos, Ilias Tsampardoukas for being a pigeon, the Almodobar team (Petsoulis, Zafeiropoulos, Tziros, Liakos, Bikas), Tsitso Veras, Mirto Potika, George Koutsotolis, Hristos Apostolidis, the 'ti taha ti' team (Chris Tolis, Papanik, Kostas Pirros, Fonias, Spyratos et alia), Tolikas for making the mistake of appointing me his best man, and many others which for space reasons I have to omit. Furthermore, my friends in Thessaloniki, studio co-owner Ilias Girbas, the people at the Vakhou building (Alexis Xanthopoulos, Thanasis Jinkovic, Stergios Koias), Spyros 'the elephant' Rassias, fellow students at Aristotle University Nikos

Tolias and Tasos Chatzikonstantinou (the latter is also greatly thanked for proofreading the biggest part of this thesis). My good friends, fellow musicians and flatmates in London Alex Kraniou and Vasilis Moschas. The IDIOTS (Alexandros Miaris, Dimitris Koufogiorgos, Alex Deligiannidis and me!), Foteini Androni aka “paoktsaki”, Sarah “0.1 times 20 equals 1” Anani. Musicians co-workers in London’s premier doghouse Arizona, Thodoros Paronis and Panos Markos, owner Omiros.

Lastly, a big thanks to my parents Konstantinos and Efthimia, my brother George, grandmothers Konstantinia and Efthimia, my aunts Maria and Carmen. This thesis is dedicated to them.

Στους γονείς μου Κωνσταντίνο και Ευθυμία και στη θεία μου Μαρία.

Contents

Abstract **2**

Acknowledgements **3**

1 Introduction **14**

1.1 Main Objectives and Structure of the Dissertation 14

1.2 Clitic Positioning in MG Dialects 17

2 Introducing the Dynamic Syntax Framework **21**

2.1 Basic Assumptions 21

2.1.1 Binary Trees and the Logic of Finite Trees 22

2.1.2 Treenode Decorations 28

2.1.2.1 The predicates *Fo*, *Ty*, *Tn* 29

2.1.2.2 Other Treenode Decorations 30

2.1.3 Requirements 31

2.1.4 The Pointer Function 32

2.1.5 CONTROL Features 33

2.1.6 Metavariables 35

2.1.7 Lexical Entries 36

2.1.8 Computational Rules 40

2.1.8.1 INTRODUCTION AND PREDICTION 41

2.1.8.2 COMPLETION 44

2.1.8.3 ANTICIPATION 45

2.1.8.4 THINNING 46

2.1.8.5 ELIMINATION 47
2.1.8.6 Structural Underspecification - The ADJUNCTION Rules 49
2.1.9 Complete Parses 57
2.2 More Advanced Concepts 66
2.2.1 Quantification 66
2.2.2 Parsing in Context - LINK Structures 73
2.2.2.1 Relative Clauses 73
2.2.2.2 Hanging Topic Left Dislocation 80
2.2.2.3 Coordinating Structures 82
2.2.2.4 Tense and Aspect 83
2.3 Parsing and the Availability of Different Parsing Strategies 85

3 Standard Modern and Grecia Salentina Greek 88
3.1 Introductory Remarks 88
3.1.1 Standard Modern Greek 88
3.1.2 Grecia Salentina Greek 89
3.2 The Data 89
3.3 Existing Approaches 94
3.3.1 GB/Minimalism 94
3.3.1.1 Philippaki 1977, 1987: Base Generated Clitics 95
3.3.1.2 Anagnostopoulou, 1999; Terzi, 1999: Clitics as Functional Heads 95
3.3.1.3 Terzi, 1999b: Functional Heads and Clitic Ordering . . . 99
3.3.1.4 Condoravdi & Kiparsky, 2002: A Unitary Account of Clitics in Greek? 102
3.3.1.5 Philippaki et al., 2002: Clitics as words revisited 104
3.3.1.6 Mavrogiorgos, 2006: Clitics within the Theory of Phases 105
3.3.2 General remarks on GB/Minimalist Approaches 107
3.3.3 HPSG Approaches 108
3.3.3.1 Alexopoulou, 1999 108
3.3.3.2 Alexopoulou and Kolliakou, 2002 109

3.3.3.3 Remarks on the HPSG analyses 110
3.4 A DS Analysis of Clitics in SMG and GSG 112
3.4.1 Single Clitic Constructions 112
3.4.2 Sequences of Clitics 131
3.4.3 Clitic Climbing 138
3.4.3.1 Approaching Clitic Climbing 141
3.4.3.2 Clitic Climbing in DS 143
3.5 Conclusions 161

4 Cypriot Greek **163**
4.1 Historical and Sociological Preliminaries 163
4.2 The Data 164
4.2.1 Enclitic and Proclitic Environments 164
4.2.2 Variation Environments 169
4.2.3 Unexpected Enclitic Environments 171
4.2.4 Clitic Ordering in Clitic Clusters 172
4.3 Existing Approaches 173
4.3.1 Terzi, 1999a 173
4.3.2 Agouraki, 2001 176
4.3.3 Revithiadou, 2006 178
4.4 A DS Analysis 180
4.5 The Enclitic Environments 180
4.6 The Proclitic Environments 181
4.6.1 Unfixed Nodes as Proclitic Triggers 181
4.6.2 Subordinating Conjunctions, Modality/Tense markers 184
4.6.3 The Complementizer *pu* 199
4.6.4 Negation 203
4.7 Variation environments 207
4.7.1 The Case of *oti* 207
4.7.2 The Case of epiδi 213
4.7.3 Complex Conjunctions/marker with *tze* 218

4.8 Gluing the Pieces Together 224
4.9 Clitic clustering and ordering in CG 228
4.10 Conclusions 231

5 Pontic Greek 233
5.1 Some Historical and Sociological Notes 233
5.2 The Data 234
5.3 Existing Approaches 239
5.3.1 Condoravdi and Kiparsky, 2002 239
5.4 A DS Analysis 241
5.4.1 Clitic Clusters 243
5.4.2 A Repair Strategy? 249
5.4.3 Are PG Clusters Problematic for Current Linguistic Theory? 250
5.4.4 Romeyka Pontic - Michelioudakis & Sitaridou, 2010 252
5.4.5 PG Clitics: Are These Really Clitics? 254
5.5 Conclusions 256

6 General Clitic Phenomena Found in all Dialects 258
6.1 The Person Case Constraint 258
6.1.1 Minimalist Accounts of the PCC 261
6.1.1.1 Nevins, 2007 266
6.1.1.2 Ormazabal & Romero, 2007 268
6.2 A DS analysis 271
6.2.1 The Martins' Observation and Latin Scrambling 271
6.2.2 The PCC as a Restriction on Underspecification 275
6.2.3 The PCC in the other Dialects 281
6.2.4 Ethical Datives 284
6.3 Doubling 293
6.3.1 Clitic Left Dislocation and Clitic Doubling in SMG 295
6.3.2 Some Thoughts on Doubling in Cypriot and Pontic Greek 315
6.3.2.1 Cypriot Greek 315
6.3.2.2 Pontic Greek 321

6.4 Conclusions 327

7 The Diachronic Development: A First Sketch 330

7.1 Introductory Remarks 330

7.2 The Koine Greek Clitic System: Competing Grammars? 331

7.2.1 Competing Grammars or a Pragmatically Governed System? 335

7.2.2 The Transition from Koine to Medieval Greek 344

7.2.3 From Medieval to Modern Greek 353

7.3 Conclusions 362

8 Conclusions 364

8.1 Conclusion Summary 364

8.1.1 General Notes as Regards Clitics 364

8.1.2 Clitic Positioning and Ordering 365

8.1.3 Clitic Climbing and Restructuring 368

8.1.4 The PCC and Person Restrictions in General 370

8.1.5 Clitic Doubling 371

8.2 Further Research 372

8.2.1 Scope over Coordination 372

8.2.2 Sequences of Two Dative Clitics 373

A Sources and Texts Consulted 394

B Additional SMG Doubling Examples 395

B.1 Indefinite Clitic Doubling in SMG 395

B.2 Bare Quantifier CD in SMG 396

B.3 Partitive CD in SMG 396

List of Abbreviations

ACC	Accusative
Appl	Applicative
ARG-STR	Argument Structure
CAG	Cappadocian Greek
CC	Clitic Climbing
CD	Clitic Doubling
CLLD	Clitic Left Dislocation
CG	Cypriot Greek
CL	Clitic
COMPS	Complements
CP	Complement Clause
CRD	Clitic Right Dislocation
DAT	Dative
DEP-STR	Dependency Structure
DO	Direct Object
DP	Determiner Phrase
DS	Dynamic Syntax
EPP	Extended Projection Principle
FP	Functional Phrase
GEN	Genitive
IO	Indirect Object
IP	Inflectional Phrase
GSG	Grecia Salentina Greek
GB	Government and Binding
HPSG	Head Driven Phrase Structure Grammar
HTLD	Hanging Topic Left Dislocation
LCA	Linear Correspondence Axiom
LFG	Lexical Functional Grammar
KG	Koiné
MP	Mood Phrase
MCG	Medieval Cypriot Greek

MMG	Medieval Mainland Greek
ModSp	Modern Spanish
MPG	Medieval Pontic Greek
NegP	Negation Phrase
PCC	Person Case Constraint
SUB	Subject
SUBJ	Subjunctive
PNP	Perfective non-past
PG	Pontic Greek
ReSp	Rennaisance Spanish
ΣP	Sigma Phrase
Spec	Specifier
TP	Tense Phrase

Chapter 1

Introduction

1.1 Main Objectives and Structure of the Dissertation

One of the primary objectives of this thesis is to provide an adequate synchronic, albeit one that is grounded in historical considerations, account of the distributional properties of clitics in four dialects of Modern Greek (MG), i.e Standard Modern Greek (SMG), Cypriot Greek (CG), Grecia Salentina Greek (GSG) and Pontic Greek (PG), using the framework of Dynamic Syntax. In that sense, even though the synchronic properties of the clitic systems under consideration costitute systems that can be potentially analyzed without reference to previous stages of the relevant dialects, these will be argued to be better understood and further vindicated once diachronic data are taken into consideration. The current state of affairs in these clitic systems will be argued to involve the same underlying mechanisms of change, namely routinization (in the sense of Pickering and Garrod, 1994; Bouzouita, 2008a,b) and the generalization of parsing strategies as lexical triggers.

Another objective of this thesis is to provide an account of a number of problematic clitic phenomena widespread across a number of languages (Person Case Constraint, Clitic Doubling) that will be relevant to both the dialects investigated and the rest of the languages or dialects exhibiting similar phenomena (e.g. Romance, Slavic or even Bantu languages).

The account will be formalized within the framework of Dynamic Syntax (DS, Kempson et al., 2001; Cann et al., 2005), a parsing oriented framework where syntax is taken

to be the result of growth of transparent semantic representations during a left to right, incremental parsing process. Given this use of a dynamic model of syntax, another objective will be to provide evidence for a dynamic, parsing-oriented framework for natural language syntax where parsing and processing considerations are seriously taken into account.

Lastly, this thesis aspires to contribute to the typological study of clitic systems of MG dialects by providing data with respect to the clitic systems of all the dialects under consideration previously unattended in the literature for MG dialects. These data will be based either on data from independent fieldwork research (GSG and PG) or are the result of consulting a number of native speakers of the dialect under consideration in each case (CG and SMG).

The dissertation is structured as follows: in chapter 2, the framework that is going to be used in this thesis, i.e. Dynamic Syntax, is introduced in a semi-formal but quite detailed way. Then, in chapter 3, the clitic systems of SMG and GSG are examined on a par. Positioning restrictions of the two clitic systems are defined as restrictions on the current parse state and as such two triggers are posited, one capturing proclisis with non-imperatives and one enclisis with imperatives verbs. At the end of the chapter Clitic Climbing (CC) in GSG, a phenomenon found in no other MG dialect, is discussed. Following Cann's (forthcoming) analysis of auxiliaries, restructuring verbs are assumed to be parsed as auxiliary-like verbs. No verbal type is projected by restructuring verbs and their semantics are captured inside the complex situation argument node in the same sense as auxiliaries. This treatment of restructuring verbs as auxiliary-like will immediately get the CC facts right, since CC will be assumed to derive in the same sense clitics in languages like SMG or Italian appear proclitic to the auxiliary rather than the verb in perfect tense constructions.

CG is then discussed in chapter 4. It is argued that its complex positioning system can be explained assuming three generalized parsing triggers. The lexical entry for clitics in CG will involve grouping of the elements triggering proclisis according to the strategies (or one of the strategies) used in parsing them. Novel variation data are discussed and are argued to derive from the fact that alternative parsing strategies are possible for the elements in question.

Chapter 5 deals with the clitic system of PG. The strictly enclitic nature of PG clitics is neatly captured by assuming a trigger for clitics in PG that is satisfied only if a functor type

value is present in the tree structure, in effect if any kind of verbal element has already been parsed. Novel data as regards clitic clustering in PG are presented revealing an idiosyncratic property of the PG clitic system, i.e. the fact that clusters of two 3rd person clitics are not allowed, something unique across MG dialects. However, given that PG clitics are case syncretic across the board and as such they are assumed to project locally unfixed nodes, clusters of two 3rd person clitics are predicted to be impossible by a general tree-logic constraint of the system which disallows more than one unfixed node of the same type to be present at the tree structure. The analysis presented for the unavailability of 3rd person clusters in PG will be the same as the one given for the PCC. Thus, the idiosyncratic ban on 3rd person clitics in PG receives an entirely general explanation.

In chapter 6, clitic phenomena found in all four dialects are examined, namely the PCC and Doubling constructions. Following Cann & Kempson (2008) and Chatzikyriakidis & Kempson (2009), the PCC is argued to be the reflex of the "no more than one unfixed node at a time constraint", already used to derive the PG facts in chapter 5. Such an account will unify the PCC with idiosyncratic restrictions like the unavailability of 3rd person clitic clusters in PG, since the explanation is based on a hard-wired tree-logic constraint and not feature driven. Then, Doubling is discussed and a modification of the analysis used for analyzing HTLD structures is proposed for CG CLLD. On the other hand, the fact that PG does not exhibit genuine CD but only CRD is easily explained assuming that PG clitics have retained a bottom restriction (something like a terminal node restriction) and as such cannot unify with an unfixed node where an NP has been parsed, since by definition the NP will involve more structure that will be precluded by any such unification.

In the last chapter, a first sketch of a diachronic account of the dialects under consideration is provided. The account starts by looking at the transition from Koine Greek (KG) to the respective medieval dialects. It is shown that such a transition can be understood as a process of routinization (Pickering & Garrod, 2004) where an earlier pragmatic governed system (the KG clitic system) ended up encoding these pragmatic preferences into pragmatic constraints (Bouzouita, 2008a,b). Then, the next step in the development of the modern dialects relies crucially on the form of their medieval counterparts. It will be shown that the system of SMG derives from a system (MMG) where proclisis was already generalized across a number of environments whereas the CG clitic system derives from

a system, that of MCG, with less proclitic environments than the medieval counterpart of SMG (MMG). Furthermore, an explanation for the strict enclitic nature of the PG clitic system is provided. The reason for the development of a strict enclitic system in the case of PG lies in the nature of the medieval Pontic (MPG) clitic system. MPG lexical entries involved the fewest proclitic environments but however no generalized proclitic trigger could emerge, since the environments inducing proclisis failed to generalize any parsing strategy and as such every proclitic environment had to be separately listed as a proclitic trigger in the lexical entry. Thus, even though MPG has the fewer proclitic environments, it however exhibits the most complex lexical entry due to the lack of generalized proclitic triggers. The fact that the entry for MPG clitics lists all the proclitic elements separately and thus could not develop any generalized proclitic trigger but instead involved one proclitic trigger for each element causing proclisis, led to the drop of all the proclitic triggers in PG and to the generalization of the most general trigger in the entry, i.e. the enclitic trigger.

1.2 Clitic Positioning in MG Dialects

MG dialects show a particular wealth of varying clitic systems, exhibiting a number of different phonological, morphological and syntactic properties. In this thesis, I will concentrate on the syntactic properies of the clitic systems of four dialects that are representative of this wealth. In particular, the thesis will, to a large extent, concentrate on clitic positioning, even though a number of other syntactic properties of clitics are going to be discussed in depth (PCC, Clitic Climbing, Doubling).

MG dialects can be classified into three major groups according to their positioning systems:

(i) Dialects exhibiting strict proclisis except for imperatives and gerunds where strict enclisis is found.

(ii) Dialects exhibiting a mixed positioning system where enclisis is the default unless a number of elements appear at the left periphery, in which case proclisis obtains.

(iii) Dialects exhibiting strict enclisis in all syntactic environments.[1]

[1] A 4th type of clitic positioning might be argued to occur in a number of Northern dialects of Thessaly or Macedonia in which clitics can appear inside the verbal form, since they do precede the plural imperative

Dialects falling under category (i) involve the dominant variety (SMG) and the Southern Italian dialects.[2] In these dialects clitics always appear before the verb in indicative environments (examples 1.1-1.4) but follow the verb in imperatival or gerundival ones (examples 1.5-1.10):

(1.1) *Ton* *aγapá*
him.CL-ACC love.3SG
'S/He/It loves him.' [SMG]

(1.2) **Aγápa* *ton*
love.3SG him.CL-ACC
'S/He/It loves him.' [SMG]

(1.3) *Ton* *gapa*
him.CL-ACC love.3SG
'S/He/It loves him.' [GSG]

(1.4) **Gapa* *ton*
love.3sg him.CL-ACC
'S/He/It loves him.' [GSG]

(1.5) Γ*rafe* *to*
write.IMP-2SG it.CL-ACC
'Write it!' [SMG]

(1.6) **To* *γrafe*
it.CL-ACC write.IMP-2SG
'Write it!' [SMG]

(1.7) *Grafe* *to*
write.IMP-2SG it.CL-ACC
'Write it!' [GSG]

ending -*ti*. These cases will not be dealt with here. The interested reader is directed to Thavoris (1977) and Joseph (1989) for the data plus discussion.

[2]The list is far bigger since most mainland Greek and Ionian islands varieties follow this pattern (Ralli, 2006).

(1.8) **To grafe*
it.CL-ACC write.IMP-2SG
'Write it!' [GSG]

(1.9) Γ*rafontas to*
write.GER it.CL-ACC
'By writing it...' [SMG]

(1.10) **To γrafontas*
it.CL-ACC write.GER
'By writing it...' [SMG]

Among dialects exhibiting the second pattern in clitic positioning we find dialects like CG (Terzi, 1999b; Agouraki, 2001; Revythiadou, 2006), Cappadocian Greek (CAG, Janse, 1994, 1998), Cretan Greek (Ralli, 2006). In these dialects clitics are in general postverbal but preverbal in case a number of elements including various functional words and fronted constituents appear at the left periphery:[3]

(1.11) *I Maria peθima (ton)*
the.NOM Mary.NOM miss.3SG him.CL-ACC
'Maria misses him.' [CG]

(1.12) Δ*ixni se to*
show.3SG you.CL it.CL
'He/She/It shows it to you.' [CAG - Janse, 1998]

(1.13) *Roto se iδes tone*
ask.1SG you.CL-ACC saw.2SG him
'I ask you. Did you see him?'[CRG - Ralli, 2006 apud Contossopoulos, 1994]

(1.14) *En ton iksero*
NEG him.CL-ACC know.1SG
'I do not know him.'[CG]

(1.15) *Ti to pikes?*
what him.CL did.2SG
'What have you done to him?' [CAG, Janse, 1998: 261]

[3]The detailed list of these elements for CG will be discussed later on in this thesis in the chapter discussing CG.

(1.16) *∆en* *exo* *ti* *na* *to* *kamo*
NEG have what SUBJ it.CL-ACC do.1SG
'I do not know what to do with it.' [CRG - Ralli, 2006 apud Contossopoulos, 1994]

In the last group of clitic systems we find PG.[4] In this dialect clitics are always postverbal no matter the form of the verb or the nature of any preceding element (see chapter 5 for more data):

(1.17) *Ekser* *aton*
know.1SG him.CL
'I know him.'

(1.18) *Píos* *entoken* *aton?*
who hit.3SG him.CL
'Who hit him?'

The dialects discussed in this thesis cover the whole range of positioning possibilities of MG and a proper account of these will also shed light to the clitic systems of other MG dialects with similar properties. Furthermore, such an account will also be extremely relevant to languages exhibiting similar clitic systems, notably Romance and/or Slavic languages. Without further ado, I will begin the exploration and analysis of the clitic systems of the dialects under consideration, starting with SMG. But before this is done, let us first introduce the framework that is going to be used in this thesis.

[4] Also Romeyka Pontic (RP), see chapter 5.

Chapter 2

Introducing the Dynamic Syntax Framework

2.1 Basic Assumptions

The Dynamic Syntax (DS) framework (Kempson et al. 2001; Cann et al. 2005) is a processing oriented framework. The basic assumption behind DS is that natural language syntax can be seen as the progressive accumulation of transparent semantic representations with the upper goal being the construction of a logical propositional formula (a formula of type *t*). Such a process is driven by means of monotonic tree growth, representing the attempt to model the way information is processed in a time-linear, incremental, word-to-word manner. However, tree structures in DS are considerably different from those found in derivational or declarative frameworks like minimalism or HPSG respectively, in that they are not inhabited by words as such, but rather by the representations of these words (Fodor, 1975). Furthermore, the tree structure corresponding to the representation of the end result of parsing a natural language string is a semantic representation assigned to this natural language string with respect to some context. This semantic representation does not correspond to word order but rather represents argument structure. However, the incremental left-to-right parsing via an array of successive, monotonically growing tree structures, handles word order through the mere definition of incremental parsing. The partial tree structures or the history of parsing stages are used to capture word order phenomena, since

parsing stages are totally dependent on the way words are ordered. The core feature of such a cumulative process builds on the idea that both hearers and speakers build these semantic representations in a left-to-right, incremental, monotonic, time-linear fashion. Under such a view, both parsing and production are assumed to involve the exact same mechanisms (see Kempson et al., 2001; Purver & Kempson, 2004; Cann et al., 2005; Purver et al., 2006 and Cann et al., 2007).[1] The process of building transparent semantic representations is done using labelled binary trees underpinned by tree-logic modalities. Treenodes are inhabited by labels expressing Type/Formula value information, tree-logic modalities and unfulfilled requirements of all the latter sort. The end result of every successful parse of a given natural language string involves a binary tree, in which all the nodes have type and formula values and no outstanding requirements exist. This tree end product is taken to be the result of multiple updates of partial trees. A number of mechanisms are responsible for updating or monotonically enriching these partial structures into more complete ones. In what follows, a semi-formal introduction to the basic components of DS is given.

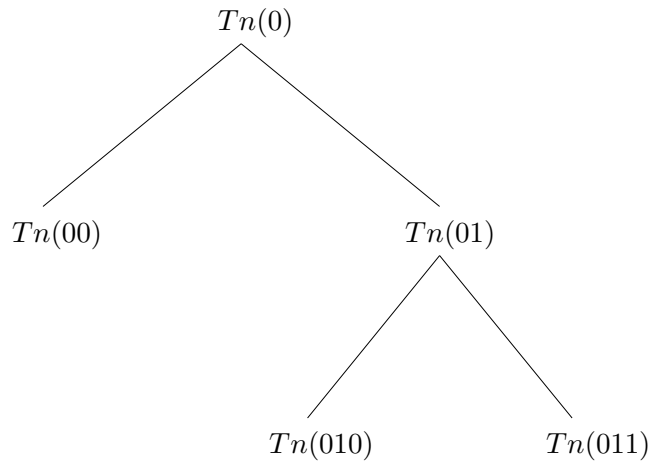
2.1.1 Binary Trees and the Logic of Finite Trees

In DS, the parsing process is represented as binary branching structure growth. Binary trees, as already mentioned, do not encode word order but rather argument structure. However, word order can be easily traced via the history of parsing stages, as we shall see later on. The whole tree system is underpinned by the Logic of Finite Trees (LOFT, Blackburn and Meyer-Viol, 1994), an expressive modal logic to talk about trees. LOFT is a modal language that allows one to describe or to identify one node in relation to any other node in the same tree. LOFT uses two basic modalities, $\downarrow$ and $\uparrow$, corresponding to the daughter and mother relation respectively. Left nodes are addressed as 0 nodes, whereas right nodes as 1 nodes. Conventionally, nodes on the left correspond to the argument nodes, i.e. the nodes in which the arguments will be represented, whereas the 1 nodes correspond to the functor nodes, i.e. the nodes in which all the various types of predicates will be represented. The rootnode, defined as the sole node that does not have a mother node, is given the treenode

[1] The only difference between the two is that in production there is also a goal tree which is used as a subsumption check. Production is rendered successful in case an identical tree to the goal tree is produced (see Purver & Otsuka, 2003; Purver & Kempson, 2004; Purver et al., 2006.)

address 0.[2] The treenode address of every treenode can be seen as the way one needs to traverse starting from the rootnode in order to get to the node in question. An easy algorithm for finding the treenode address of a given node is to look at the treenode address of its mother and then add 0 or 1 depending on whether the node is an argument or a functor node. A binary tree where all the nodes have treenode addresses is shown below:

(2.1)

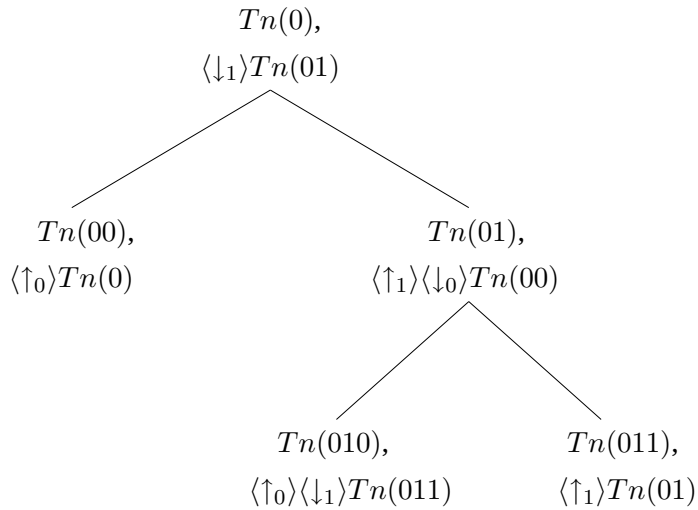


The two basic LOFT modalities ($\downarrow$ and $\uparrow$) are used in either their existential or universal version. The two different versions are distinguished by the use of angle ($\langle\rangle$) or square ($[]$) brackets, representing the existential and universal reading respectively. A tree modality inside angle brackets carries an existential statement. For example, the tree modality $\langle\downarrow\rangle$ will read as: there is a node if you go down the daughter relation, i.e. there is a daughter node. The same modality in square brackets ($[\downarrow]$) involves a universal statement, where no existential commitment is made. $[\downarrow]$ reads as: for all nodes found if you go down the daughter relation. The different properties of the two versions of brackets (the existential and the universal) will be exemplified in more detail as soon the *Fo* and *Ty* predicates are introduced. The two basic tree modalities can be further specified or can be more underspecified. For example, a given tree modality might further specify the daughter node it is referring to. In this sense, the daughter modality can be further represented as $\langle\downarrow_0\rangle$ or $\langle\downarrow_1\rangle$, depending on which of the two daughter nodes the tree modality is referring to. The

[2]Note that treenodes can also be given an arbitrary value, e.g. *Tn(a)*. When such a value is provided, this is unique in the sense that no other fixed node in the tree can carry the same value.

example below illustrates the flexibility of LOFT by showing a binary tree where different nodes are addressed from the perspective of other nodes using treenode modalities:

(2.2) The LOFT modalities in action



In the above tree, all nodes have a treenode address and a further statement identifying another node in the tree. For example, the statement $\langle\uparrow_0\rangle\langle\downarrow_1\rangle Tn(011)$ found in the 010 node reads as: you will find treenode 011 if you take a step across the 0 mother relation followed by a step across the 1 daughter relation. The LOFT system is a flexible way to refer to any node or any decoration in any node from the perspective of any node.

We have already seen that the two basic modalities can be further specified and refer to either the argument (node on the left) or the functor node (node on the right). Up to now the modalities we have used deal with immediate dominance relations. However, LOFT modalities can be recursively combined to give us non-immediate dominance relations. For example, the granddaughter relation can be easily defined by recursively using the mother relation, i.e. $\langle\downarrow\rangle\langle\downarrow\rangle$. The same holds for the grandmother relation. Using recursion in the LOFT system we can refer to any dominance relation. However, the power and flexibility of the LOFT system lies in the fact that the system also allows underspecified dominance relations where the exact dominance relation is not yet known to be expressed. LOFT makes use of the Kleene operators (star and plus, * and + respectively), in order to denote dominance relations that may be underspecified. The * operator expresses the general notion of dominance plus reflexiveness. For example, the modality $\langle\downarrow^*\rangle$ reads as

follows: there is a node if you go down zero or more steps across the daughter relation. This node can be any node below the current one or the current node itself since the number of steps can always be empty. On the other hand, the Kleene + operator is used to encode the pure notion of dominance. The Kleene + operator is equivalent to the Kleene * operator without the empty set.[3] In that sense, the modality $\langle\downarrow^{+}\rangle$ reads as follows: there is a node if you go down one or more steps across the daughter relation. The node in question can be any node below the current one, but not the current one.

Another very powerful characteristic of the LOFT system is its ability to refer not only to any node from the point of view of any given node but also to any decoration present on any node in the tree. For example, the statement $\langle\downarrow\rangle$X reads as: there is a daughter node in which X holds, in effect referring to a decoration holding at a daughter node. The up and down modalities ($\langle\uparrow\rangle$ and $\langle\downarrow\rangle$), in combination with their existential and universal interpretations plus the two Kleene operators (* and +) give us an extremely wide range of possibilities. The table given in (2.3) summarizes these possibilities:

[3]Formally $\langle\downarrow^{+}\rangle = \langle\downarrow\rangle\ \langle\downarrow^{*}\rangle$.

(2.3) LOFT combinations

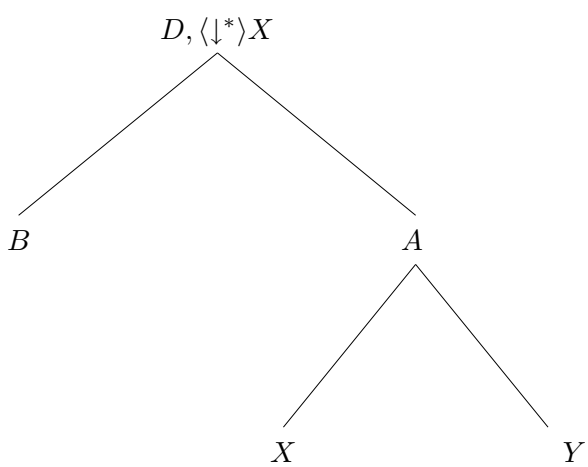
LOFT STATEMENT	MEANING
$\langle\downarrow\rangle X$	X holds at the daughter node
$\langle\downarrow_0\rangle X$	X holds at the 0 daughter
$\langle\downarrow_1\rangle X$	X holds at the 1 daughter
$\langle\downarrow^*\rangle X$	X holds at a node somewhere below or at the current node
$\langle\downarrow^+\rangle X$	X holds at a node somewhere below
$[\downarrow]X$	X holds at all daughter nodes
$[\downarrow_0]X$	X holds at all 0 daughter nodes
$[\downarrow_1]X$	X holds at all 1 daughter nodes
$[\downarrow^*]X$	X holds at all nodes below and the current one
$[\downarrow^+]X$	X holds at all nodes below
$\langle\uparrow\rangle X$	X holds at the mother node
$\langle\uparrow_0\rangle X$	X holds at the 0 mother node
$\langle\uparrow_1\rangle X$	X holds at the 1 mother node
$\langle\uparrow^*\rangle X$	X holds at a node above or at the current one
$\langle\uparrow^+\rangle X$	X holds at a node nodes above
$[\uparrow]X$	X holds at all mother nodes
$[\uparrow_0]X$	X holds at all 0 mother nodes
$[\uparrow_1]X$	X holds at all 1 mother nodes
$[\uparrow^*]X$	X holds at all nodes above and the current one
$[\uparrow^+]X$	X holds at all nodes above

It is quite interesting to note that, at first sight, a number of the above statements seem redundant, given standard properties of binary trees. For example the statements $\langle\uparrow_0\rangle X$ and $\langle\uparrow_1\rangle X$ state that X holds at the mother node if someone traverses the 0 or 1 relation. The 0 and 1 specifications might be claimed to be redundant, given that each node has only one mother node. However, such statements are not redundant at all since these identify the nature of the daughter node rather compared the mother node. In that respect, the relation $\langle\uparrow_1\rangle X$ can be better thought of as "mother of a functor node" relation, and as such carries

distinctive non-redundant information compared to e.g. the modality $\langle\uparrow_0\rangle X$ (mother of an argument daughter). In the same sense, the following three statements $[\uparrow]X$, $[\uparrow_0]X$ and $[\uparrow_1]X$ also seem redundant. Given that each node has only one mother node, it seems that these three statements can be reduced to their existential forms, i.e. $\langle\uparrow\rangle X$, $\langle\uparrow_0\rangle X$ and $\langle\uparrow_1\rangle X$. Again, this is not the case since the angle brackets carry an existential statement whereas square brackets do not. Thus, a statement of the form $[\uparrow_1]X$ will be true even in case there is no mother node (let us say we are at the top node), whereas $\langle\uparrow_1\rangle X$ will be true only if there is such a node where X holds. In the same vein, the statement $[\downarrow]X$ has a fixed number of nodes it refers to. Given that we are dealing with binary trees, this number is always 2. Again, such a statement is not redundant and not reducible to its existential counterpart, since in order to express its content, two statements involving angle brackets are needed, namely $\langle\downarrow_0\rangle X \wedge \langle\downarrow_1\rangle X$.[4]

The modalities presented so far allow one to refer to underspecified relations within partial trees. However, such structures are not underspecified descriptions of a specified situation. The statement $\langle\downarrow^*\rangle X$ says that there is an X in case one or more steps down the daughter relation are taken. Such a statement will not be true in case the number of steps is known but we just want to give an underspecified description of this fully specified situation. In the example below the statement $\langle\downarrow^*\rangle X$ will not be true:

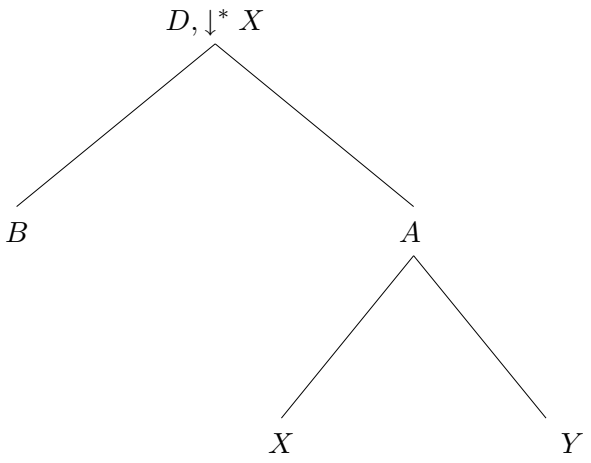
(2.4)



[4]Note that the statement $\langle\downarrow\rangle X$ cannot capture the meaning of the statement $[\downarrow]X$. The first statement refers to X holding at one or both of the daughter nodes while $[\downarrow]X$ is only compatible with a situation where X holds at both daughter nodes or no daughter nodes exist at all.

DS uses another kind of modality to express underspecified descriptions of fully specified situations. These different set of modalities are called external modalities, while the ones we have seen so far are called internal. The external existential modality is represented without the angle brackets, i.e. only with the arrow modality (e.g.$\downarrow^* X$). This new type of modality can be used in a situation like the one depicted in the previous tree, where one wants an underspecified description of a fully specified tree:

(2.5) Underspecification using external modalities



It is trivial to note that the external/internal modality distinction will collapse in the presence of a fully specified modality like e.g. $\langle\downarrow_0\rangle$. It is worth noting that the universal version of the external modality has not been defined in neither Kempson et al. (2001) or Cann et al. (2005). I will not define this modality here either, since no use of this kind of modality is going to be made in this thesis. With this last note, I end the discussion of the LOFT system redirecting the interested reader to Blackburn & Meyer-Viol (1994) and Kempson et al. (2001: chapter 9).

2.1.2 Treenode Decorations

Treenode decorations are pieces of information that describe a number of facts being true of a node. Treenode decorations along with binary trees and the LOFT language comprise the declarative structure of the DS model, the *DU* language. The *DU* language is the language

the DS formalism uses to describe trees and their decorations. A subset of this language has been already presented, i.e. the LOFT language. Besides the LOFT language, *DU* makes use of a number of other predicates in order to describe trees. The most important of which are presented below.

2.1.2.1 The predicates *Fo, Ty, Tn*

Nodes in DS are assumed to be inhabited by abstract representations rather than words themselves. In that sense, every node carries a Language of Thought (LOT) concept in the sense of Fodor (1975, 1981, 1983). This LOT concept is represented using the predicate *Fo* followed by the concept in question. For example the *Fo* value of a verb like *hit*, would be $Fo(\lambda x.\lambda y.hit'(x)(y))$. Note that the prime is used in order to denote that the concept and not the word itself is used. The general form of the monadic predicate is then *Fo(x')*, where the values of *x* are typed lambda terms. Every node is assumed to have a formula value. Formula values of argument nodes combine with those of functor nodes via means of functional application, as we will see shortly.

In addition to the *Fo* value, every node further carries a type (*Ty*) value. *Ty* refers to the semantic type of the word in question following the usual *e,t* notation widely used in formal semantics (see Dowty, 1981; Cann, 1994 and Heim & Kratzer, 1997 among others). *Ty* is also a monadic predicate, with the general form *Ty(x)*, where *x* ranges over a number of possible semantic types. However, the difference between semantic typing as used in DS and most of the formal semantic theories is that in DS the set of semantic types is a closed set. In that sense, no type-shifting or type recursion is available.[5] In addition, DS assumes a number of types not commonly found in the semantic literature. These always involve the type *cn* (standing for common noun) as one of their subparts.[6] For a list of the possible types in DS, consult Cann et al. (2005: 36). Semantic types of argument and functor nodes combine via modus ponens. For example combining an argument node of type *e* with a functor node of type $Ty(e \rightarrow t)$, derives a mother node of type *t*, following the rules of natural deduction:

[5]Even though type recursion would be needed assuming an optional argument analysis of adjuncts in the sense of Marten (2002).

[6]The meaning of this type will be discussed later on inthis chapter.

(2.6) $$\frac{e, e \rightarrow t}{t}$$

DS also makes use of a monadic predicate to represent treenode addresses. This monadic predicate has the form *Tn(x)*, with *Tn* standing for treenode and *x* ranging over possible LOFT treenode addresses. For example the initial node of a binary tree will have the value *Tn(0)*. In general, the *Tn* value of a given node is its LOFT treenode address. This treenode address is computed in the way sketched in our discussion of the LOFT language in section 2.1.1 or can have an arbitrary value, e.g *Tn(a)*. It is worth noting that the reason the monadic predicates *Fo*, *Ty* and *Tn* were dealt with together is that they share a very common property in the DS system. All three of them must obligatorily be present in any node. In that respect, every node in DS must have a proper *Fo*, *Ty* and *Tn* value. We will see the consequences of this assumption later on, when we will discuss underspecification, both structural and semantic.

2.1.2.2 Other Treenode Decorations

In addition to the three basic decorations *Fo*, *Ty*, *Tn* that are relevant for any node, DS makes use of a number of other decorations that can vary from language to language. For example in the analysis presented in this thesis as regards enclisis in SMG and GSG imperatives, the feature *Mood(x)* is going to be used. This feature can be also seen as a monadic predicate, where *x* ranges over a set of possible mood values. Another example of such decorations is the *[NEG+]* feature used both in Bouzouita (2008a,b) and in this thesis to encode the presence of negation. It is worth noting however that these additional features are used as diacritics in DS in case an analysis of the phenomena these labels are taken to describe is pending in the system. Nothing restricts the number of these labels in principle. In that sense, all different kinds of labels can appear in a given DS tree, depending on the language or the analysis itself. I will not get into a discussion of how many additional decorations besides the basic three monadic predicates *Fo*, *Ty*, *Tn* are needed. However, I should note that as a general strategy, DS strives to eliminate any diacritic-like decorations by replacing them with more general analyses deriving from the incrementality of tree-growth. This is the general stance I am going to take in this thesis, in effect trying to use additional features as sparingly as possible. Nonetheless, it should be kept in mind that there is no formal

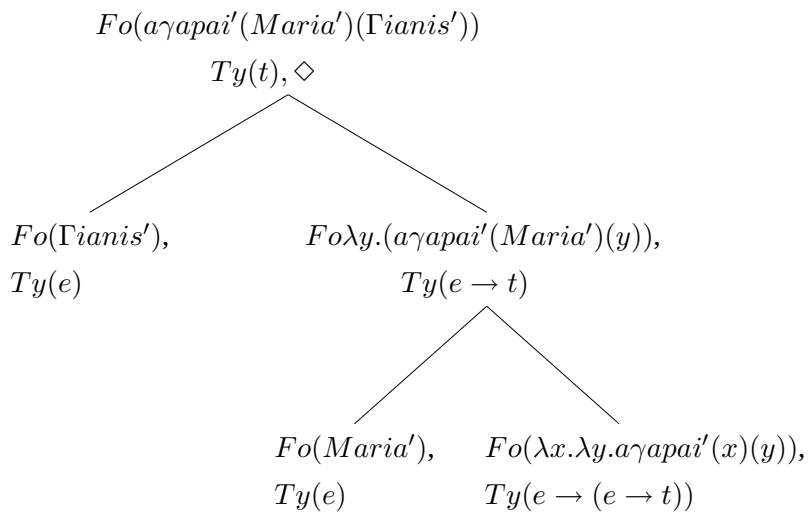
restriction in the system on how many of these decorations can be encoded in the system.

2.1.3 Requirements

The parsing process in DS is a goal-driven process, where tree-growth is incrementally unfolded via means of satisfying a number of conditions having the form of requirements. Requirements can be seen as goals that have not yet been achieved or more precisely as descriptions which do not hold at a given node but must do so at some point in order for the parsing process to be successful. Requirements have the general form *?Decoration*, where any kind of decoration is preceded by a question mark *(?Fo(x'), ?Tn(010)* etc.). The whole parsing process is based on requirements in DS. For example, the starting point for every parse called the *AXIOM* is a requirement to establish a propositional formula of type *t*. Requirements are used in a variety of ways and comprise the basic mechanism controlling well-formedness. A well-formed tree, representing the end result of successfully parsing a natural string must contain no outstanding requirements, i.e. all requirements must be satisfied in order for a given parse to be successful. In that sense, requirements can also be seen as a device explaining ungrammaticality.[7] The example below presents the result of successfully parsing the string *o Γianis aγapai ti Maria* 'John loves Mary' in SMG. Notice that no outstanding requirements exist and all nodes have type and formula values:

[7]Looking at tree-growth as the development of a number of parsing states using the $\leq$ relation, requirements must be present in both the minimal state ($T_$) and every non-maximal state ($T \notin PT_+$, where PT_+ is the set of maximal elements). Formally, $V_?(T_) \neq \emptyset$ and $\{ T \in PT \mid V_?(T) = \emptyset\} \subseteq PT_+$ (taken from Kempson et al., 2001: 270).

(2.7) A complete parse



Requirements are exceptional in DS, in that they are the only decorations that by definition must not persist. In that sense, requirements are exceptional in the DS system, in that they are non-monotonic unlike the rest of the decorations used (CONTROL features might have the same property, as we will discuss in a bit). Requirements persist until they are satisfied. Thus, a requirement holding at two states T and T' where T $\leq$ T', it must hold for all states in between T and T'.[8] However, by the mere definition of a requirement (as a goal to be achieved), the requirement will persist only until it is satisfied. Given that no outstanding requirements exist in a well-formed parse, no requirement must persist when the parsing process has ended in order to get a well-formed parse. For more information on the notion of requirements, the interested reader is referred to Kempson et al. (2001: 42-49; 270-273) and Cann et al. (2005: 36-38).

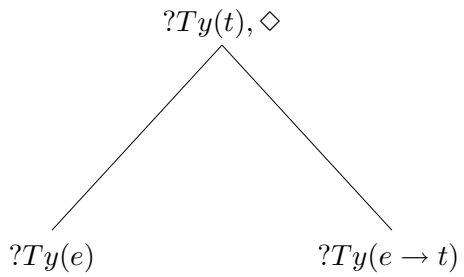
2.1.4 The Pointer Function

Given the dynamics of the system we have sketched so far where all information is growing incrementally via multiple updates of partial trees, what we need is a device that will enable us to track the exact node under development in the tree. This device is called the pointer

[8]Formally, if T $\leq$ T' and T' $\models ?\phi$, then $\forall$T"(T $\leq$ T" $\leq$ T' $\Rightarrow$ T" $\not\models \phi$) (Kempson et al., 2001: 270).

and is represented as $\diamond$. The basic function of the pointer is to track the current node under construction. In the example below, the pointer is at the initial type *t* requiring node:

(2.8) The pointer function



In the above tree the node under construction is the type *t* requiring node. This means that unless the pointer moves to another node, tree updating can be done only on the node bearing the pointer. Consequently, updating cannot be done for more than one node simultaneously, but has to proceed node by node. An indirect consequence is that the pointer can be also used as a device explaining ungrammaticality. For example, assuming a lexical entry for an NP[9] which presupposes the pointer to be on a node with a type *e* requirement, parsing cannot continue in case the pointer is at a node that does not have that specified requirement (more on this in a while). Obviously, what we also need is a way to move the pointer from node to node. Pointer movement is done either via lexical or computational actions. The exact way in which this is done will be explained when discussing lexical and computational actions later on in this chapter.

2.1.5 CONTROL Features

CONTROL features in DS are features disjoint from the set of decorations (as defined in Kempson et al., 2001). Their difference to other treenode decorations is that control features do not have to be persistent. In that way, they are similar to requirements. However, they are not identical, since requirements by definition must not persist while CONTROL features may or may not persist. CONTROL features can be seen as "conditions guarding actions" (Kempson et al., 2001: 289-290). An example of a CONTROL feature would be

[9]Note that the term NP will correspond to the DP in current Minimalist literature. I will use the term DP only when discussing Minimalist analyses that make use of that term. Otherwise the term NP will be used.

the use of a statement referring to the current state of affairs of a partial tree that if satisfied will trigger parsing of a given word. This CONTROL feature can function as the triggering point of that word's lexical entry. Note that such a statement will not be present in the tree itself but will only be part of the lexical entry's triggering point. To give an example, assume an imaginary language where NPs marked for case *x*, must always come first. In that sense, in case another NP has been parsed first, an NP with case *x* cannot be parsed. A straightforward way to encode this fact in DS is to assume that the lexical entry for *x* case marked NPs has a triggering point that requires that no NPs have already been parsed. For example, one possible encoding of this triggering point would be a conditional statement of the form shown below:[10]

(2.9)

IF	$?Ty(t), \langle\downarrow^*\rangle Ty(e)\bot$
THEN	ACTIONS
ELSE	abort

The above triggering point ($?Ty(t), \langle\downarrow^*\rangle Ty(e)\bot$) states that if it is not the case that somewhere below the node you are there is an NP, then proceed to the actions. Such a triggering point, encoded as a statement describing part of the partial tree at the time the *x* case marked NP comes into parse, is a CONTROL feature, since the statement $\langle\downarrow^*\rangle Ty(e)$ acts as a conditional action only and is not by any means represented in the tree structure itself. Such a feature cannot be a decoration on a tree, since the situation described is impossible to persist, given that any form of utterance will involve at least one type *e* argument (implicit or explicit). Note that triggering points in lexical entries are not obligatorily CONTROL features. The triggering point for a lexical entry, as we will shall se later on, might involve a statement that it is indeed part of the tree itself, e.g. a statement referring to a requirement or decoration being present at a given node. Such a statement is not a CONTROL feature but a regular decoration used as a triggering point. A number of CONTROL features will be used when clitics in the different dialects of MG are going to be examined. No explicit lexical entries will be given at the moment, since these have not yet been introduced. Once

[10]The above example is formulated, in anticipation of the discussion on lexical entries, using the standard DS schema for lexical entries. A detailed explanation of the basic IF THEN ELSE algorithmic schema used in lexical entries follows in section 2.1.7.

lexical actions are introduced, it will be fairly easy to understand how CONTROL features function within the system.

2.1.6 Metavariables

One very basic aspect of the DS formalism is the use of metavariables. These metavariables act as content placeholders and need to be substituted by a proper value of the predicate that the metavariable appears as an argument of at some point during the parsing process. Metavariables can appear in all the monadic predicates we have discussed and are represented by capitals, e.g. *Fo(U)*, *Ty(V)*. Metavariables comprise one of the basic mechanisms capturing underspecification in DS. A formula metavariable *Fo(U)* does not represent any logical formula but rather stands as a substitution site for potential formula values. Metavariables can carry restrictions narrowing down the domain of potential substitutents. For example, assuming that we want the formula value substituting a metavariable of type *e* to be specified as male, we encode a presupposition denoted as a subscript on the metavariable ($Fo(U_{male'})$). Such a move will ensure that only Formulas that satisfy the presupposition 'male' can be potential substituents for the metavariable. Given that metavariables are only content placeholders and not proper values, and given that every node in DS must come with at least an *Fo*, a *Ty* and a *Tn* value, we need to ensure that the metavariable will get updated at some point during the parsing process. We do that by positing a requirement for a proper value of the predicate that the metavariable is an argument of to be found on the same node the metavariable is posited. For example, if a formula value metavariable *Fo(U)* is present on a node, then a requirement for a proper formula value to be found on that same treenode must be present. This requirement is encoded in the form of an existential statement preceded by the question mark sign. For example this statement will have the form $?\exists\mathbf{x}.Fo(\mathbf{x})$ in case of a *Fo* metavariable, $?\exists\mathbf{x}.Ty(\mathbf{x})$ in case of a *Ty* metavariable and so on.[11] These statements read as: there is a requirement that a proper *Fo/Ty* value is found in the treenode. This requirement must be satisfied sometime until the end of the parsing process, otherwise the parse cannot be successfull. Non-substituted metavariables can thus lead to ungrammaticality in DS.

[11]The boldface variables are used to distinguish between variables that range over values of D_{Fo}, and regular variables, that can be proper values in treenode decorations (*Fo(x)*)

2.1.7 Lexical Entries

Treenode decorations, requirements and CONTROL features comprise the declarative part of the DS system. However, DS further employs a number of procedural mechanisms that project the features of the declarative system. The basic devices behind the DS procedural system, are lexical and transition/computational rules (henceforth computational rules). The first are language-specific rules projected from the words of a given language. These are the lexical entries each word is assumed to involve. On the other hand, computational rules are language-general rules and, as such, are assumed to hold for every language. Lexical actions in DS are encoded in a simple algorithmic IF THEN ELSE format. The IF part contains a statement(s) that should be true or false with respect to the partial tree at the time the word comes into parse. If this statement is satisfied, the actions encoded in the THEN part of the entry are projected. There is an elsewhere statement (ELSE) that might contain alternative information in case the first triggering point is not satisfied. The examples below present a sample entry for the proper noun *Bill* in English:

(2.10) Lexical entry for the proper noun *Bill* in English

IF	$?Ty(e)$
THEN	$\texttt{put}((Ty(e), Fo(Bill'), [\downarrow]\bot)$
ELSE	abort

The above example reads as follows: if you are at a node that has a type *e* requirement, then decorate this node with a type *e* value and a formula value representing the concept *'Bill'*. In any other case abort. In that sense, a proper noun like *Bill* in English will be able to get parsed as soon as a node has a requirement for a type *e*. This will allow a word like *Bill* to be parsed either as a subject or as an object in English. Languages with overt NP case marking are assumed to involve further restrictions, encoded as output filters, that will ensure that an NP marked for structural accusative for example will end up on the direct object node. Such an entry is shown below:

(2.11) Lexical entry for an NP marked for structural accusative

IF	$?Ty(e)$
THEN	$\texttt{put}((Ty(e), Fo(x'), ?\langle\uparrow_0\rangle Ty(e \rightarrow t), [\downarrow]\bot)$
ELSE	abort

In the above lexical entry, the accusative marked NP posits a requirement that the predicate node ($Ty(e \rightarrow t)$) is found in case a step across the zero daughter relation is taken from the current node, in effect identifying the NP's position with the direct object node. Such an encoding of case has the advantage of not having to posit case features. Case under such an account is seen as a restriction on the place in the overall configuration at which the term will end up.[12] We will see the relevance of such case filters when we will discuss the PCC in chapter 6. Besides projecting information on nodes already present in the partial tree, lexical entries can also construct nodes. For example, a lexical entry for a monotransitive verb in a typical pro-drop language like SMG will involve unfolding of the whole propositional template. The verb starting from the type *t* node will build the whole propositional template and further decorate the 011 node with a type and formula value, representing the adicity and the LOT concept of the verb in question. The pointer is then assumed to be returned to the type *t* requiring node. This will allow parsing of both VSO and VOS structures, since the pointer can move to both the subject and the object node by ANTICIPATION as we shall see shortly. The lexical entry for a monotransitive verb in SMG is shown below:

(2.12) Lexical entry for a monotransitive verb in SMG[13]

IF $?Ty(t)$

THEN $\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \texttt{put}(?Ty(e \rightarrow t);$
$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$
$\texttt{put}(Fo(\lambda x.\lambda y.verb'(x)(y)), Ty(e \rightarrow e(\rightarrow t)), [\downarrow]\bot);$
$\texttt{go}(\langle\uparrow_1\rangle); \texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(?Ty(e));$
$\texttt{go}(\langle\uparrow_0\rangle\langle\uparrow_1\rangle); \texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(Ty(e), Fo(U_x)), gofirst(?Ty(t))$

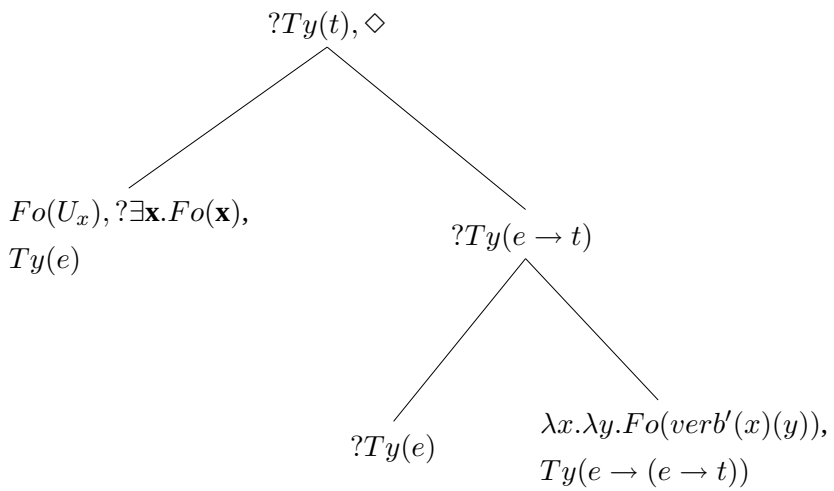
ELSE abort

The effect of parsing a monotransitive verb in SMG in tree notation is shown below:

[12] If one wants to be accurate, such an account of case functioning as an output filter will not disallow an NP from being parsed in other nodes that do not satisfy the case restriction. However, if the NP is parsed on these nodes the case filter requirement will never be satisfied and thus the parse will never be successful. Note that no account of case of determiners or adjectives is going to be given or discussed in this thesis.

[13] The subscript x on the U metavariable represents the restrictions that the subject metavariable will bear. I will not get into the details as regards these restrictions or subject agreement in general.

(2.13) Parsing a monotransitive verb in SMG



As mentioned, verbs in SMG are assumed to build the whole propositional structure and decorate the subject node with a type *e* value and a formula metavariable. This is done in order to capture the subject pro-drop properties of the language. Update of metavariables into proper Fo values can be done either by the context or by the natural language string itself. Update by the context is done using the rule of SUBSTITUTION, which substitutes an *Fo* metavariable by a proper *Fo* value from the context in case this *Fo* value does not appear anywhere in the local domain of the tree under construction.[14] On the other hand, update by the natural language string itself is done when the lexical entry of a word provides an *Fo* value for the *Fo* metavariable. Given these two possibilities of metavariable update, the subject can be overt or covert in SMG. Note that such an update from the context is not possible for the object node, since no metavariable and type value are projected in the object node but only a type requirement. In that sense, overt material is needed in order to fill the object slot, thus no object pro-drop being possible in SMG. There is a great deal of parametric variation with respect to the entries of lexical verbs in different languages in DS. For example, non pro-drop languages like English will project a requirement for a type *e* in both the subject node and object nodes (see Cann et al., 2005: 48), while full pro-drop languages like Japanese or presumably Latin (on the assumption that Latin is a full pro-drop language) will project a type value and a formula metavariable in both the subject and the object nodes (Kempson & Kurosawa, forthcoming: 7). Lexical actions are thus

[14]See Cann et al., 2004: chapter 2 for the exact definition of SUBSTITUTION.

the domain of parametric variation in DS. As such, syntactic differences across languages are assumed to derive from different lexical specifications, i.e. different lexicons. This fact will be of crucial relevance when we will discuss clitic positioning in a number of dialects of MG, since all variation in positioning will be derived via different lexical specifications by keeping the same set of computational rules.

The algorithmic format of lexical entries can be of arbitrary complexity. Embedded IF THEN schemas can appear within the initial IF THEN schema and so on. A lexical entry with the structure shown below is perfectly legitimate:

(2.14) An embedded IF THEN lexical entry format

```
IF          ...
THEN        IF      ...
            THEN    ...

                    ...
            ELSE    abort
ELSE  abort
```

A number of partial recursions on specific parts of the algorithm may be used. For example, an entry containing a disjunctive IF part will be denoted by separating the two disjunctive parts with the symbol |:

(2.15) An embedded IF THEN with a disjunctive IF part

```
IF          ...
THEN        IF      ...|
                    ...
            THEN    ...
            ELSE    abort
ELSE abort
```

Any addition to the algorithmic format just described will be introduced separately when necessary.

Finally, as already anticipated, triggering points in lexical entries can also involve CONTROL features. As we saw, CONTROL features are statements that guard the unfolding of actions in DS. CONTROL features do not need to persist, since they are used as a description of a given state of affairs at a given point during the parsing process. Being non-persistent, CONTROL features do not appear on the tree structure itself. An example of such a CONTROL feature is exemplified in the lexical entry shown below:

(2.16) Lexical entry with a CONTROL feature

IF	$?Ty(t), [\downarrow]\bot$
THEN	...
ELSE	abort

According to the statement in the IF part of the entry, the parse can continue to the actions only in case the statement $[\downarrow]\bot$ is true, i.e. only in case nothing holds below the type *t* requiring node. Such a statement can be true at some point(s) during the parsing process, but will definitely have to be false at some point(s) of the same process. Thus, such a statement will not persist. Therefore, it is crucial to distinguish between treenode decorations that appear on treenodes and should persist all the way through once they are introduced in the parsing process and CONTROL features that are not part of the treenode decorations themselves, but can however appear in the lexical entries as statements regulating the parsing process.

2.1.8 Computational Rules

Computational rules are general computational devices, comprising the basic tree construction mechanism. Formally, computational rules involve an input and an output description. The former designates where the pointer must be along with information about the node that the pointer is on as well as information about other nodes with respect to the pointer node, while the latter shows the transformation of the input in terms of requirements, adding nodes, pointer movement and so on. Computational rules are represented by a fraction-like schema, where the top part represents the input description, while the bottom part the output description:

(2.17) Computational rule schema

$$\frac{\text{Input Tree Description}}{\text{Output Tree Description}}$$

Computational rules are assumed to be a closed set of rules universally available to every language.[15] The most important of these are discussed below.

2.1.8.1 INTRODUCTION AND PREDICTION

I begin the discussion of computational rules with a rule that is going to be presented mainly for historical reasons, since the newest advances in the DS framework (Cann, forthcoming) abandon the use of this rule altogether and older accounts dealing with languages other than English either questioned the need for such rule (Chatzikyriakidis, 2006) or claimed that such a rule is not necessary for the languages examined (Bouzouita, 2008a for Medieval, Rennaisance and Modern Spanish).

Before I move to the actual INTRODUCTION and PREDICTION rule, it is essential to note at that point that the starting point of all parsing processes is a requirement to obtain a propositional formula of type *t*, i.e. a proposition. Given this, the parsing process begins with a tree containing only one node. The requirement to obtain a type *t* formula is encoded on that node. This one node tree description shown below is called the AXIOM:

(2.18) The AXIOM

$$?Ty(t), \Diamond$$

Returning to the rule of INTRODUCTION and PREDICTION, the first thing that needs to be said is that INTRODUCTION and PREDICTION is two rules rather than one. The first of the two rules, INTRODUCTION, splits a given goal expressed as a type requirement into two subgoals. Given that a node carries a type requirement, the rule of INTRODUCTION introduces two new subrequirements (type requirements) that will result in the initial requirement if modus ponens is applied between the two. The rule is shown below:

[15] Even though variations of rules, like for example the rule of LOCAL *ADJUNCTION might also be subject to parametric variation as we will see later on in this chapter.

(2.19) INTRODUCTION

$$\frac{\{...\{?Ty(Y)..., \Diamond\}\}}{\{...\{?Ty(Y), ?\langle\downarrow_0\rangle Ty(X), ?\langle\downarrow_0\rangle ?T(X \rightarrow Y)..., \Diamond\}\}}$$

Before getting into the actual content of the rule of INTRODUCTION shown above, it is important to explain the meaning of the ellipsis mark $(...)$. The ellipsis mark has two usages in the context of computational rules. If it is is found within a node carrying other kinds of information (e.g. $\{..., Fo(a), ...\}$, then the meaning of it is that other type of information might be present before and after the presented information. For example $\{..., Fo(a), ...\}$ reads as: there is a node that bears a $Fo(a)$ formula value and other kinds of decorations might be present before or after this formula value. However, if the ellipsis mark appears in a node containing only the ellipsis mark and furthermore precedes or follows another node $(\{...\{Fo(a)\}...\})$, then the meaning of it is that other nodes might precede or follow the node containing the $Fo(a)$ decoration. For example $\{..\{Fo(a)\}...\}$ states that other nodes might precede or follow the node that has the $Fo(a)$ decoration while on the other hand $\{\{Fo(a)\}\}$ means that no node precedes or follows the node that has the same decoration, i.e. there is only one node present in the tree structure.

Returning to the actual INTRODUCTION rule, we see that the rule adds two new requirements that must be found in the daughter nodes, but does not however build these nodes. The effect of the rule in tree notation is shown below:

(2.20) Applying INTRODUCTION[16]

$$?Ty(Y), ?\langle\downarrow_0\rangle Ty(X), ?\langle\downarrow_0\rangle Ty(X \rightarrow Y), \Diamond$$

The rule of PREDICTION builds the two daughter nodes, decorates them with the two subrequirements and leaves the pointer on the argument daughter node:

(2.21) PREDICTION

$$\frac{\{...\{...?Ty(Y), ?\langle\downarrow_0\rangle Ty(X), ?\langle\downarrow_0\rangle Ty(X \rightarrow Y)..., \Diamond\}\}}{\{...\{?Ty(Y), ?\langle\downarrow_0\rangle Ty(X), ?\langle\downarrow_0\rangle Ty(X \rightarrow Y)...\}\{?Ty(X), \Diamond\}\{?Ty(X \rightarrow Y)\}\}}$$

In tree notation:

[16]Notice that the structure below is also a partial binary tree even though there are no nodes. Imagine the structure below as the top node of a tree.

(2.22) Applying PREDICTION

$$\begin{array}{c} ?Ty(Y), ?\langle\downarrow_0\rangle Ty(X), ?\langle\downarrow_1\rangle Ty(X \rightarrow Y) \\ \diagup \qquad\qquad \diagdown \\ ?Ty(X), \diamond \qquad\qquad ?Ty(X \rightarrow Y) \end{array}$$

The rule as it is stated can apply at every node that has a type requirement. However, allowing the rule to apply for every node will cause a great deal of overgeneration, since by doing so one allows any type of node to be built from any type of node. In Cann et al. (2005: 43) the use of the rule was restricted to subject and predicate only.[17] This more specified version of INTRODUCTION and PREDICTION is shown below:

(2.23) INTRODUCTION - Subject and Predicate

$$\frac{\{...\{Tn(n), ?Ty(t)..., \diamond\}\}}{\{...\{Tn(n)?Ty(t), ?\langle\downarrow_0\rangle Ty(e), ?\langle\downarrow_0\rangle ?T(e \rightarrow t)..., \diamond\}\}}$$

(2.24) PREDICTION - Subject and Predicate

$$\frac{\{...\{...Tn(n), ?Ty(t), ?\langle\downarrow_0\rangle Ty(e), ?\langle\downarrow_1\rangle Ty(e \rightarrow t)..., \diamond\}\}}{\{...\{Tn(n), ?Ty(t), ?\langle\downarrow_0\rangle Ty(e), ?\langle\downarrow_1\rangle Ty(e \rightarrow t)...\}\{?Ty(e), \diamond\}\{?Ty(e \rightarrow t)\}\}}$$

The effect of the two rules in tree notation is shown below:

(2.25) Applying INTRODUCTION and PREDICTION

$$?Ty(t), ?\langle\downarrow_0\rangle Ty(e), ?\langle\downarrow_0\rangle Ty(e \rightarrow t) \quad \mapsto \quad \begin{array}{c} ?Ty(t), ?\langle\downarrow_0\rangle Ty(e), ?\langle\downarrow_1\rangle Ty(e \rightarrow t) \\ \diagup \qquad\qquad \diagdown \\ ?Ty(e), \diamond \qquad\qquad ?Ty(e \rightarrow t) \end{array}$$

Notice that under this new formulation, the rule of INTRODUCTION and PREDICTION can apply only in case no other nodes exist below the type *t* requiring node. This implies a strict SVO language where parsing of the subject is always prior to the verb or

[17] However, see McCormack (2008) for a more general use of INTRODUCTION and PREDICTION.

the object, specifically English.[18] This language-specific flavor of the rule eventually led to the elimination of the rule altogether (Cann, forthcoming), striving for a universal set of computational rules.

2.1.8.2 COMPLETION

The second computational rule we are going to look at is the rule of COMPLETION. The function of COMPLETION is to move the pointer ($\diamond$) from a daughter node that has a satisfied type requirement to its mother node. When pointer movement is done, the information that some requirement has been satisfied in one of the daughter nodes is encoded on the mother node. The rule is a pointer movement rule that facilitates the parsing process by moving the pointer to the mother node in case type satisfaction occurs in one of the daughter nodes. Formally the rule has the following form:

(2.26) COMPLETION

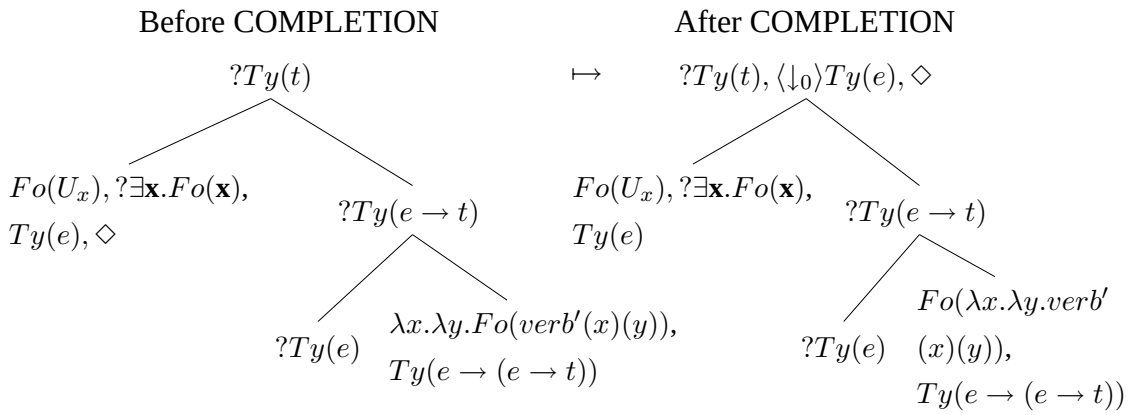
$$\frac{\{...\{Tn(n),...\},\{<\uparrow_i>,Tn(n),...,Ty(X)..,\diamond\}...\}}{\{...\{Tn(n),...,<\downarrow_i>Ty(X),...,\diamond\},\{<\uparrow_i>Tn(n),...,Ty(X),...\}...\}}$$

Where i $\in$(0,1,*)

For example, let us assume a situation where a monotransitive verb has been parsed in SMG and the pointer is at the subject node (00 node). At this point COMPLETION can apply, moving the pointer to the mother node, i.e. the type *t* requiring node, where it records the fact that a type value is present on its argument daughter node (00 node):

[18]Of course, OSV is also possible in English, notably in object contrastive focus constructions. In these cases parsing of the dislocated object is done via means of an unfixed node. See Cann et al. (2005: 59-65) and the discussion on unfixed nodes in this chapter.

(2.27) Applying COMPLETION



2.1.8.3 ANTICIPATION

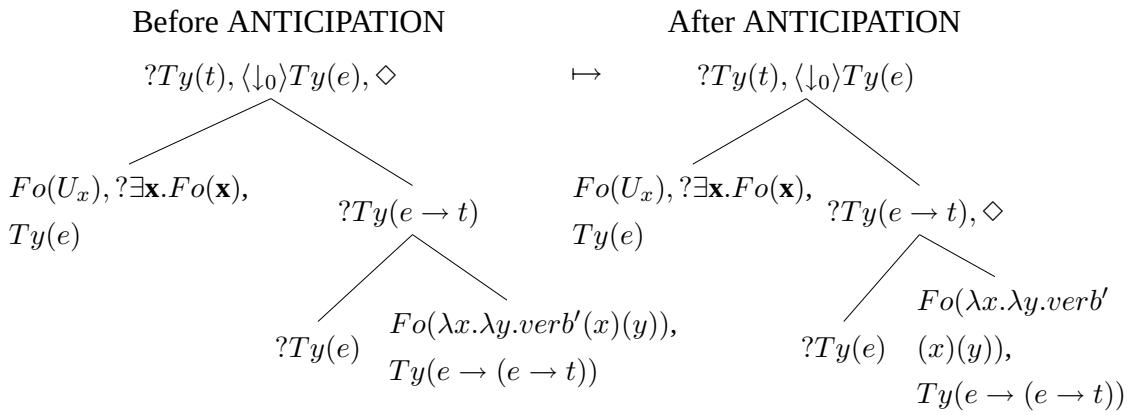
In our discussion so far, we have emphasized that DS is a goal driven syntactic framework. In that sense, the syntactic process is driven by requirements that need to be fulfilled. The rule of ANTICIPATION facilitates this goal driven process by moving the pointer from a mother to daughter node in case any unsatisfied requirement is present in that daughter node:

(2.28) ANTICIPATION

$$\frac{\{...\{Tn(n), ..., \diamond\}, \{\langle\uparrow\rangle Tn(n), ?X...\}...\}}{\{...\{Tn(n), ...\}, \{\langle\uparrow\rangle Tn(n), ?X..., \diamond\}...\}}$$

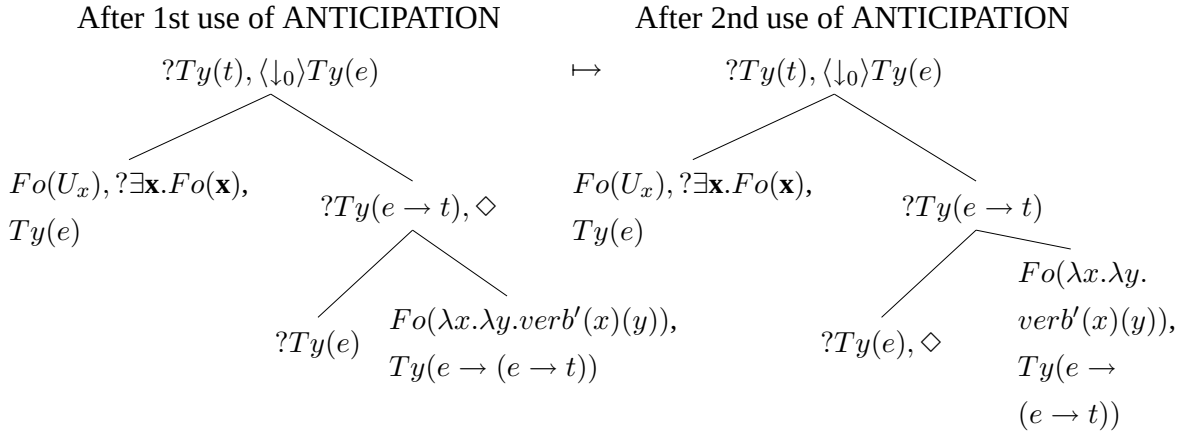
For example assuming the tree in (2.27) where COMPLETION has applied and the pointer is at the type *t* requiring node, ANTICIPATION can apply, moving the pointer to any of the two daughter nodes, since both of them carry unsatisfied requirements. In the example below, the pointer is moved to the functor node:

(2.29) Applying ANTICIPATION



Note that all the rules can apply in a recursive fashion. In that sense, it is possible to reach the lower type *e* requiring node in (2.27) by applying ANTICIPATION twice:

(2.30) Recursive ANTICIPATION



2.1.8.4 THINNING

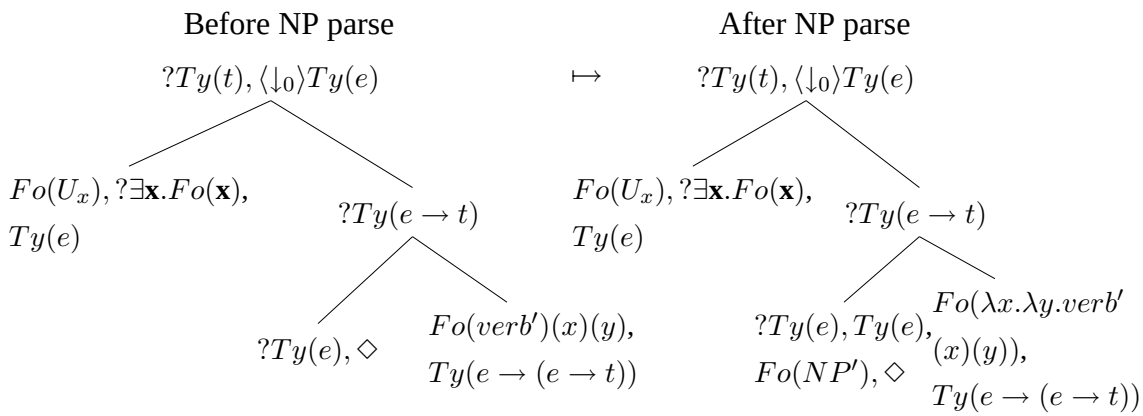
As we have already said, requirements by definition will not persist. Therefore, all requirements must be satisfied assuming the parse is successful. This means that at a certain point requirements must become statements. Thus, we need a rule that will eliminate requirements once these are satisfied. This is what the rule of THINNING does. Given a node carrying both a requirement for a value and the value itself, the rule of THINNING eliminates the requirement:

(2.31) THINNING

$$\frac{\{...\{...,X,...,?X,...,\Diamond\}...\}}{\{...\{...,X,...,\Diamond\}...\}}$$

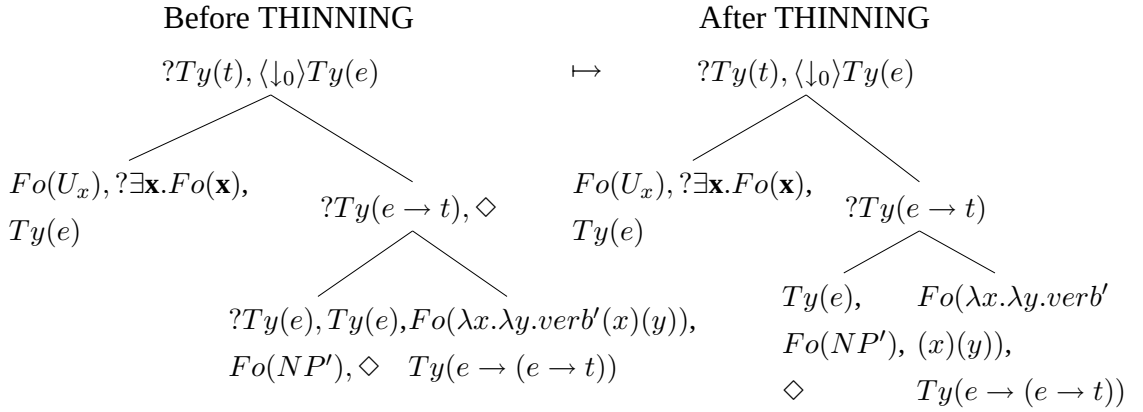
With the pointer left at the direct object in 2.30, an NP can be parsed, decorating the direct object node with a type and a formula value:

(2.32) Parsing the NP



At that point THINNING applies, eliminating the type *e* requirement in the presence of a proper type *e* value:

(2.33) THINNING



2.1.8.5 ELIMINATION

The rule of ELIMINATION operates when both daughter nodes have satisfied type and formula values. In case this is true, the rule deduces a new type by applying modus ponens

to the type values of the daughter nodes and performs functional application on the formula values of the same nodes. The rule includes a further condition that no outstanding requirements exist in any of the two daughter nodes. In case the latter is not true, the rule cannot apply:

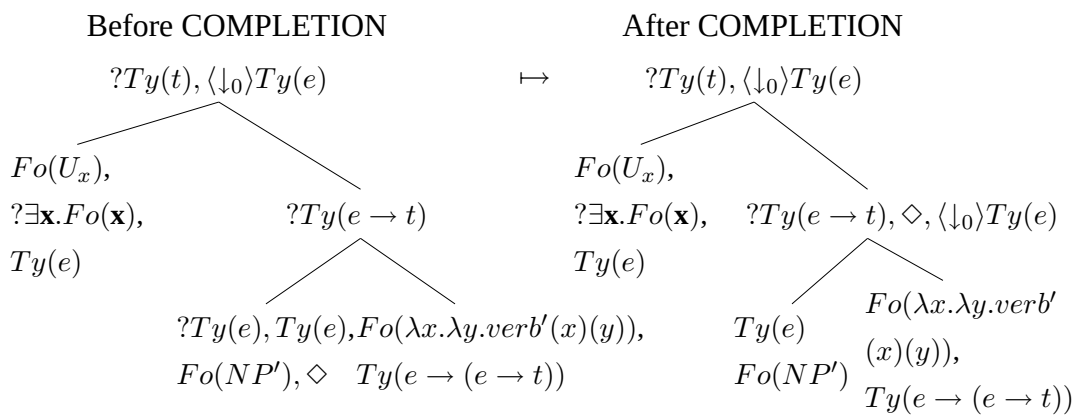
(2.34) ELIMINATION

$$\frac{\{...\{\langle\downarrow_0\rangle(Fo(a), Ty(X)), \langle\downarrow_1\rangle(Fo(b), Ty(X \rightarrow Y), ..., \diamond)\}...\}}{\{...\{Fo(b(a)), Ty(Y), \langle\downarrow_0\rangle(Fo(a), Ty(X)), \langle\downarrow_1\rangle(Fo(b), Ty(X \rightarrow Y), ..., \diamond)\}...\}}$$

Condition: $\langle\downarrow_i\rangle$ $?\phi$ does not hold and i $\in \{0, 1\}$

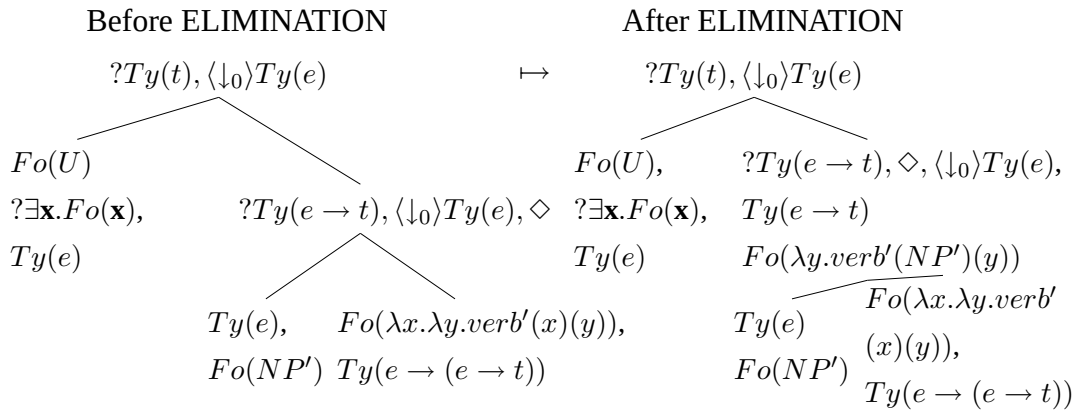
In (2.33), the pointer is at the direct object node. At that point, COMPLETION applies and moves the pointer to the mother node, since a type *e* value is present on the 010 node:

(2.35) COMPLETION



Elimination can now apply, in effect applying modus ponens and functional application for types and formula values respectively between the 010 and the 011 node:

(2.36) ELIMINATION



2.1.8.6 Structural Underspecification - The ADJUNCTION Rules

A central concept in DS is the assumption that natural languages are to a large extent underspecified regarding both content and structure. And while content underspecification has been largely employed within the formal semantics literature of the past 30 years, no attempts to move underspecification into the area of syntax have been made.[19] We have already seen how DS deals with semantic underspecification by employing metavariables as content placeholders. Structural underspecification on the other hand, is encoded in DS using a family of computational rules, the ADJUNCTION rules. The function of these rules is to introduce unfixed nodes, i.e. nodes that have not found their fixed position in the tree at the time of their introduction. The first of the ADJUNCTION rules we are going to see is the rule of *ADJUNCTION. According to this rule, an unfixed node marked with a type *e* requirement is projected from a type *t* requiring node:

(2.37) *ADJUNCTION

$$\frac{\{...\{Tn(n), ?Ty(t), \Diamond\}\}}{\{...\{Tn(n), ?Ty(t)\}, \{\langle\uparrow^*\rangle Tn(n), ?Ty(e), ?\exists\mathbf{x}.Tn(\mathbf{x}), \Diamond\}\}}$$

The effect of the above rule in tree notation is shown below:

[19] With the exception possibly being the notion of functional uncertainty, formalized within Lexical-Functional Grammar (Bresnan, 2001).

(2.38) The effect of the *ADJUNCTION rule

$$Tn(n), ?Ty(t)$$

$$\langle\uparrow^*\rangle Tn(n),$$
$$?Ty(e),$$
$$?\exists\mathbf{x}.Tn(\mathbf{x}), \Diamond$$

The rule leaves the pointer at the type *e* requiring node. The node is structurally underspecified since it does not carry a fixed treenode address. The only thing the node "knows" as regards its treenode position is that somewhere up above or at the current node, *Tn(n)* must be found.[20] The pointer is left at the type *e* requiring unfixed node. At that point, given that a lexical entry for an NP will have a type *e* requiring trigger, it can be parsed on that unfixed node. This will be on a node whose structural position in the tree is not yet known. The *ADJUNCTION rule works neatly for left dislocated structures like OV focus structures in SMG. In these cases, the preposed object is assumed to be parsed on an unfixed node, given that NPs in SMG will involve a type *e* requiring node as their trigger, which is indeed what I assume:

(2.39) Parsing *ton* Γ*iani* 'John' in *ton* Γ*iani xtipise* 'S/He/It hit John' on an unfixed node[21]

$$?Ty(t), Tn(n)$$

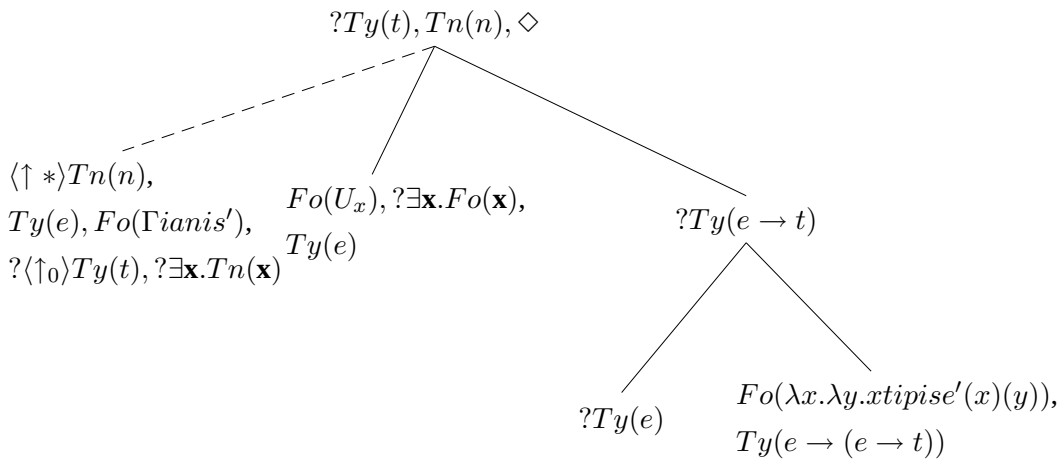
$$\langle\uparrow^*\rangle Tn(n),$$
$$Ty(e), Fo(\Gamma iani'),$$
$$?\exists\mathbf{x}.Tn(\mathbf{x}), \Diamond$$

The pointer moves up to the type *t* requiring node via COMPLETION and the verb can now be parsed. The structure we get after parsing the verb, given the entry for monotransitive verbs shown in (2.12), is the following:

[20]Notice that the reflexive satisfaction of the * in which the node unifies with its host trivially is not possible. *Tn(n)* is the treenode address of the type *t* requiring node. In that sense, the only way such a reflexive satisfaction will hold is in case the unfixed node unifies with the type t requiring node. However, such unification is impossible, given the incompatible specifications of the respective type requirements (*?e* and *?t*).

[21]I ignore determiners for the moment.

(2.40) Parsing *xtipise* ‘hit’ in *ton* Γ*iani xtipise* ‘S/He/It hit John’



In the above structure there is an unfixed node with a type and a formula value and an open slot decorated with a type *e* requirement (the 010 node). It is at that point that a process of unification between the unfixed node and the fixed object node (010) can take place using MERGE. MERGE is a computational rule which unifies two nodes just in case one of the two updates the treenode address of the other. The notion of update is defined by treenode address entailment. If a treenode address entails another treenode address, then the former can be seen as an update of the latter. In that sense an underspecified address like $\langle\downarrow^*\rangle$Tn(a) can be updated to a more specified address like $\langle\downarrow\rangle$Tn(a) but not vice versa. Furthermore, the two nodes must not bear any conflicting specifications. Formally, the rule of MERGE has the form below:

(2.41) MERGE

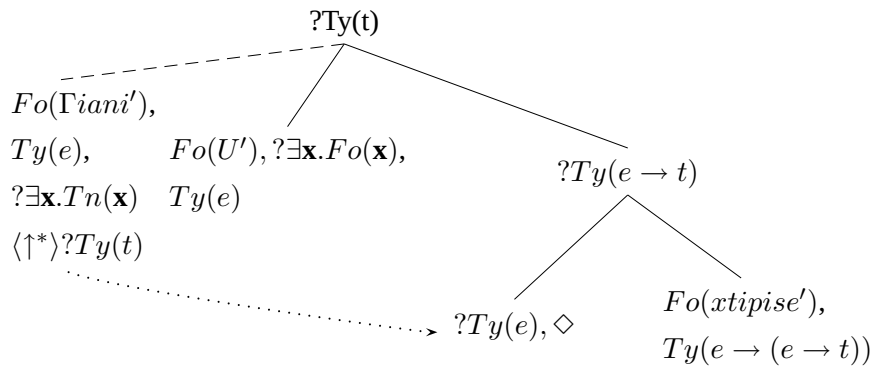
$$\frac{\{...\{..., DU, DU', ...\}...\}}{\{...\{..., DU \sqcup DU', ...\}...\}}$$

Where $\Diamond \in$ *DU’* and *DU* $\cup$ *DU’* is consistent

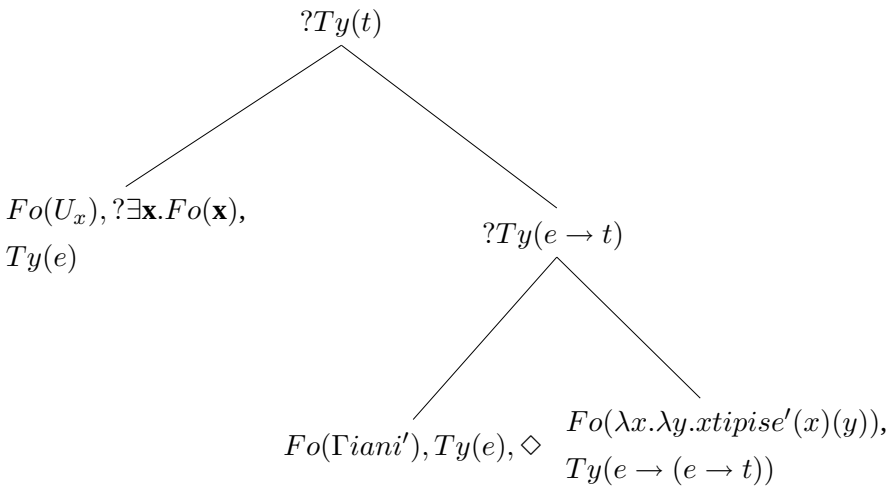
In example (2.40), the treenode address of the direct object node (010) can be a proper update of the underspecified address the unfixed node carries. Furthermore, the fact that the 010 node has a proper treenode address will eliminate the requirement of the unfixed node that a fixed treenode address should be found ($?\exists \mathbf{x} Tn(\mathbf{x})$). Note that no conflicting specifications exist, since the fixed node (DU’) has a type *e* requirement and the unfixed

node (*DU*) a type *e* value. Given that a type *e* value is an update of the type *e* requirement, no conflict arises. The trees below display the tree structure before and after MERGE has applied:[22]

(2.42) Before and after MERGE of the unfixed node



(2.43)



Notice that the rule of *ADJUNCTION will also work for cases where the left dislocated element is embedded in a complement clause. The modality associated with the *ADJUNCTION rule makes MERGE of the unfixed node with a type *e* node that is deeply nested possible. The only requirement is that this node must be found within the same tree structure.

[22]The steps of THINNING, eliminating the type *e* requirement as well as the requirement $?\exists\mathbf{x}Tn(\mathbf{x})$ are not shown.

We have seen that an unfixed node projected via the rule of *ADJUNCTION can be fixed at any level of embedding within the tree. Assuming that we want to localize the fixing site of the unfixed node, we cannot use that rule anymore, since the rule will overgenerate given that it can extend beyond the local domain. However, DS also makes use of a local variant of the *ADJUNCTION rule, called LOCAL *ADJUNCTION. This localized version restricts the potential fixing sites of the unfixed node within the local domain. In order to achieve this, a more specified underspecified modality is used:

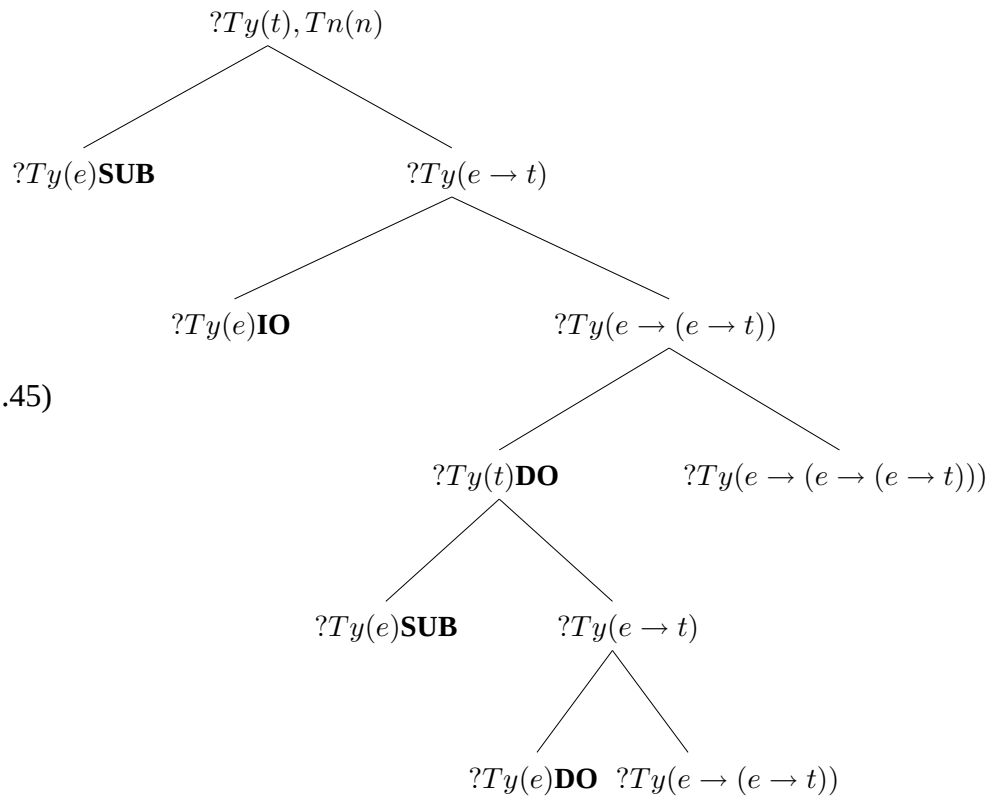
(2.44) LOCAL *ADJUNCTION

$$\frac{\{...\{Tn(a), ?Ty(t), \Diamond\}\}}{\{...\{Tn(a), ?Ty(t)\}, \{\langle\uparrow_0\rangle\langle\uparrow_1^*\rangle Tn(a), ?Ty(e), ?\exists\mathbf{x}.Tn(\mathbf{x})\Diamond\}\}}$$

The new modality $\langle\uparrow_0\rangle\langle\uparrow^*\rangle$ will ensure that the node gets fixed inside the local domain. MERGE of the NP into an embedded type *e* requiring node is impossible, since in that case more than one 0 steps across the mother relation will be needed. Let us explain. Take a look at the structure below. It depicts the skeletal structure of a ditransitive verb whose direct object is a verbal complement:[23]

[23]For illustration purposes, I have put the DO object lower than the IO. In general, the IO in DS is the lowest argument in the sense that it is the first argument combining with the verb. However, in this case the DO object being a verbal complement has a propositional structure on its own. It is much easier to read the tree in that format, rather than in a format where the DO is higher than the IO.

(2.45)



Assuming a locally unfixed node exists, MERGE of that node will only be possible in the 00, 010 and 0110 node, although in the latter case unsuccessfully given the conflicting requirements (type *e* and *t* respectively). The type *e* requiring nodes of the embedded domain cannot MERGE with the locally unfixed node, since more than one step across the 0 relation is to be taken in order to reach the top node (*Tn(n)*). In that sense, none of these nodes' treenode addresses are entailed by the modality $\langle\uparrow_0\rangle\langle\uparrow^*\rangle$. We will see later on how the rule of LOCAL *ADJUNCTION can straightforwardly capture the distributional properties of syncretized clitics in the dialects we are examining (see chapters 3, 4, 5 and 6). LOCAL *ADJUNCTION is an operation which reflects the tree relation when the structural representation of a clitic (in the dialects we are considering) is underspecified. However, this effect will not be created by use of the rule of LOCAL *ADJUNCTION itself. Rather, it will come directly from the lexical entries of the clitics in question. It is an open matter whether Greek in all its guises makes use of this rule as a general computational device. For this thesis, I will assume that the rule of LOCAL *ADJUNCTION is not operative in the dialects under examination as a general computational rule but only lexically encoded

in the entries of syncretized clitics.

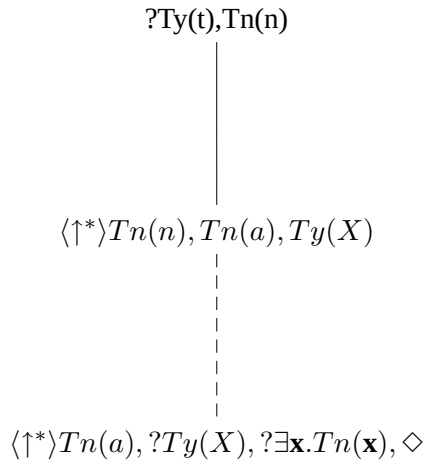
The last variant of the ADJUNCTION rules I am going to use in this thesis is the rule of LATE *ADJUNCTION. The rule of LATE *ADJUNCTION is different from the other two ADJUNCTION variants in that it projects an unfixed node from a type complete node to a node requiring the same type. The rule of LATE *ADJUNCTION directly specifies the structural context in which it is parsed, since application of the rule is only possible given that a node has a type value, i.e. in case an element has projected some information on that node. The rule is shown below:

(2.46) *LATE ADJUNCTION

$$\frac{\{...\{Tn(n), ?Ty(t), ..., \}\{\uparrow^* Tn(n), Tn(a), ..., Ty(X), \diamond\}, ...\}}{\{...\{Tn(n), ..., \}\{\langle\uparrow^*\rangle Tn(n), Tn(a), ..., Ty(X)\}, \{\langle\uparrow^*\rangle Tn(a), ?Ty(X), ?\exists\mathbf{x}.Tn(\mathbf{x}), \diamond\}, ...\}}$$

The effect of the rule in tree notation is shown below:

(2.47) Applying LATE *ADJUNCTION



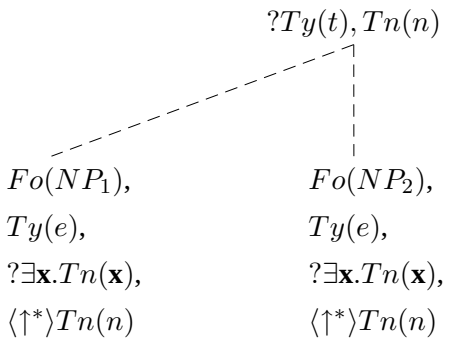
The above rule will prove to be crucial in constructions where late placement of an already introduced formula value will be needed, e.g. in VSO or CD constructions (see section 2.1.9 for parsing VSO sentences and chapter 6 for parsing CD constructions).[24]

Before I conclude the discussion of the basic computational rules, it is worth mentioning a tree growth constraint, which will be extremely useful in explaining various person case restrictions as we will see later on (PCC, unavailability of two 3rd person clitic clusters

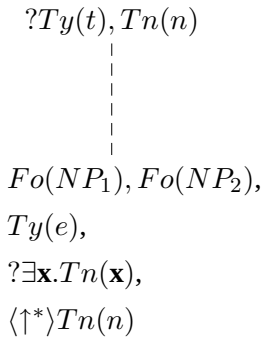
[24]See also Cann et al., 2005 for an account of extraposition using LATE *ADJUNCTION

in PG). We have seen that DS makes use of unfixed nodes, i.e. nodes that are structurally underspecified as regards their position in the tree. The tree-logic of the system imposes a hard-wired constraint on the number of unfixed nodes that can be used at the same time in a given tree structure. Specifically, the tree-logic of the system does not allow more than one unfixed node of the same type to be present in the same tree structure. The reason for this is that any two unfixed nodes of the same type will collapse into being the same one by means of treenode identity. In that sense, unfixed nodes proceed one at a time. However, note that different kinds of unfixed nodes are not subject to this constraint, given that the underspecified modality they carry is different. Therefore, a situation where both an unfixed and a locally unfixed node are present in a tree structure is licit. The first example shows the result of parsing two unfixed nodes of the same type. In the first stage, they are both introduced as separate nodes but they immediately collapse into being the same node:

(2.48) Parsing two NPs on unfixed nodes



(2.49) The two nodes collapse into one



Note that we ended up with one node that however carries two incompatible *Fo* decorations ($Fo(NP_1)$ and $Fo(NP_2)$). Given that every node has only one formula value, the

parse cannot be successful assuming the situation depicted in the second tree above.

The computational rules presented, comprise the basic set of rules DS uses. There are a number of other rules, notably LINK rules, that will be presented later on in this chapter. For the moment, I exemplify the use of all the rules introduced by going in detail through a number of complete parses.

2.1.9 Complete Parses

A complete parse of a SVO and a VSO sentence in SMG will be shown in this chapter. The sentences to be parsed are shown below:

(2.50)

O	*Γiorγos*	*xtipise*	*to*	*Γiani*
the.NOM	George.NOM	hit.3SG	the.ACC	John.ACC

(2.51)

Xtipise	*o*	*Γiorγos*	*to*	*Γiani*
hit.3SG	the.NOM	George.NOM	the.ACC	John.ACC

'George hit John'

In the first example, the subject is the first element that comes into parse. For the sake of the introduction and in order not to complicate things I will not deal with determiners at the moment (see the discussion on determiners in the section on quantification, this chapter: 2.2.1). For the moment I will assume that a NP in SMG involves a type *e* requiring trigger in order to get parsed. A nominative case-marked NP will further contribute a case-filter that will posit that the NP in question must be a subject. The lexical entry is shown below:

(2.52) Lexical entry for a nominative case-marked NP in SMG

IF	$?Ty(e)$
THEN	$\texttt{put}((Ty(e), Fo(NP'), ?\langle\uparrow_0\rangle Ty(t))$
ELSE	Abort

We start from the AXIOM, which is a requirement to obtain a propositional formula of type *t*:

(2.53) The AXIOM

$$?Ty(t), \Diamond$$

The subject comes into parse. However, its triggering point is not satisfied given that only a type *t* requirement exists whereas a type *e* requirement is needed. However, the rule of *ADJUNCTION can first apply, projecting an unfixed node bearing a type *e* requirement. Then, the NP can be parsed on that unfixed node:[25]

(2.54) *ADJUNCTION and parsing of the subject

$$Tn(n), ?Ty(t)$$

$$\langle\uparrow^*\rangle Tn(n),$$
$$?Ty(e), Ty(e), Fo(\Gamma ior \gamma os'),$$
$$?\langle\uparrow_0\rangle Ty(t), ?\exists \mathbf{x}.Tn(\mathbf{x}), \Diamond$$

Notice that the unfixed node carries both a requirement for a given type ($?Ty(e)$) as well as its satisfaction ($Ty(e)$). THINNING can thus apply eliminating the requirement. Furthermore, COMPLETION moves the pointer to the type *t* requiring node where the information that a given type has been satisfied in a daughter node is encoded:

(2.55) THINNING and COMPLETION

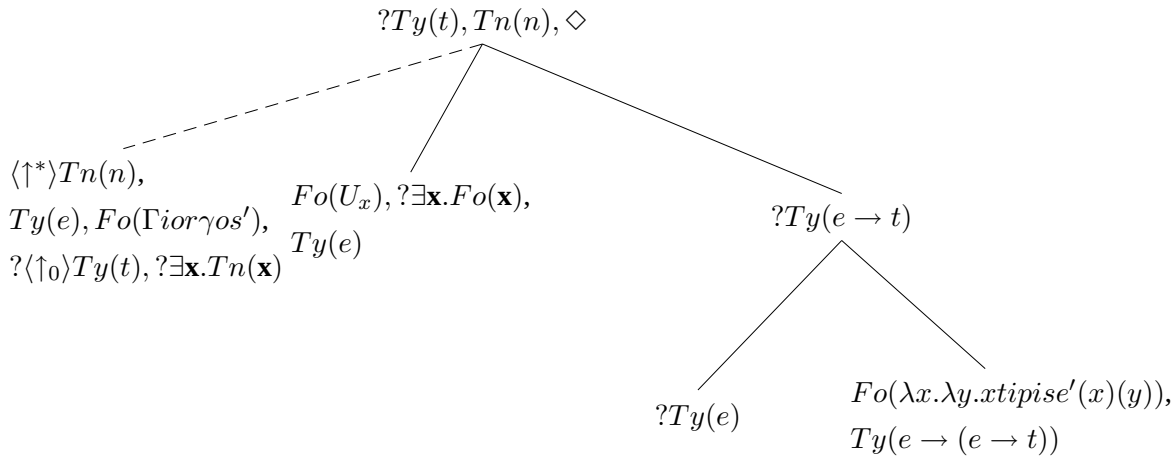
$$Tn(n), ?Ty(t), \langle\downarrow^*\rangle Ty(e), \Diamond$$

$$\langle\uparrow^*\rangle Tn(n),$$
$$Ty(e), Fo(\Gamma ior \gamma os'),$$
$$?\langle\uparrow_0\rangle Ty(t), ?\exists \mathbf{x}.Tn(\mathbf{x})$$

With the pointer at the type *t* requiring node the verb can be parsed:

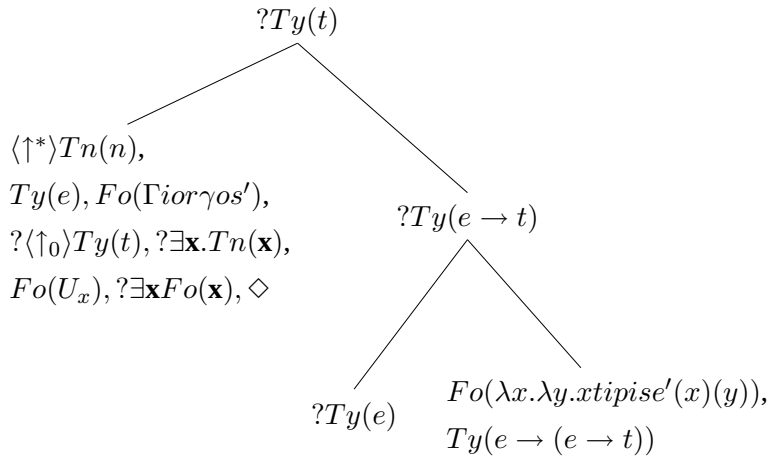
[25]Note that this is not the only option of parsing subjects. See 2.3 for an alternative way of how this can be done (involving a LINK structure).

(2.56) Parsing the verb



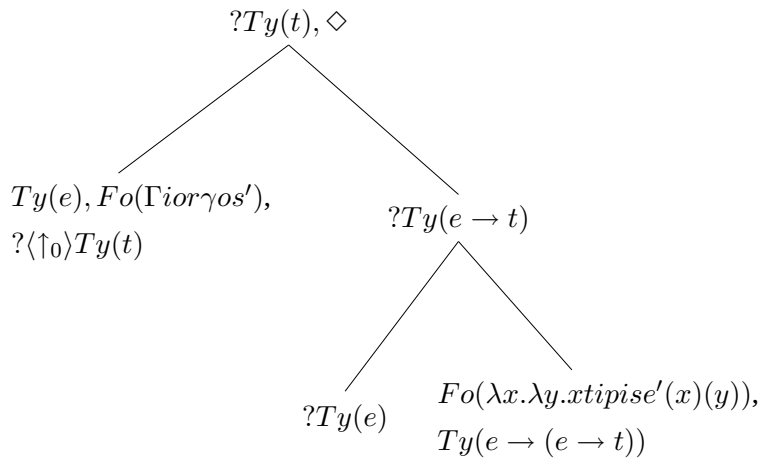
The verb projects the whole propositional template and decorates the object node with a type *e* requirement and the subject node with a type value and a formula metavariable. The pointer is at the type *t* requiring node and is moved to the subject node via ANTICIPATION since an unsatisfied requirement exists in that node ($?\exists\mathbf{x}.Fo(\mathbf{x})$). At that point MERGE takes place, unifying the unfixed node with the subject node. The result of MERGE is shown below:

(2.57) MERGE



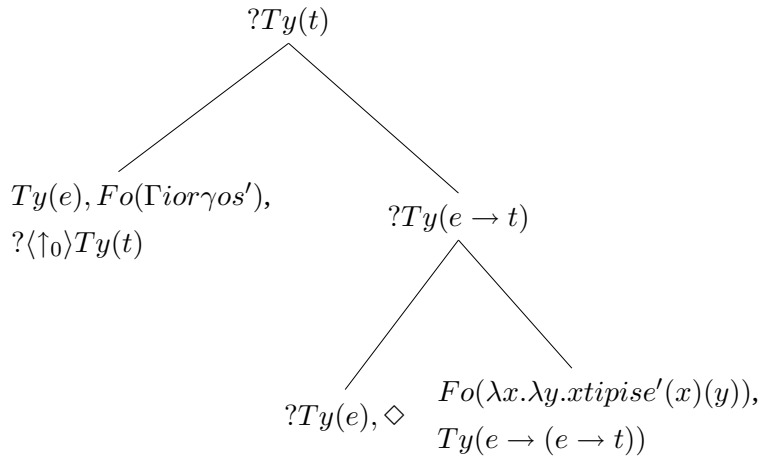
THINNING applies eliminating the requirements for a proper *Fo* value and a proper *Tn* address. The requirement that the mother node must bear a type *t* value is not yet satisfied though, since no such type exists yet in the mother node. The pointer moves up to the type *t* requiring node via COMPLETION.

(2.58) THINNING and COMPLETION



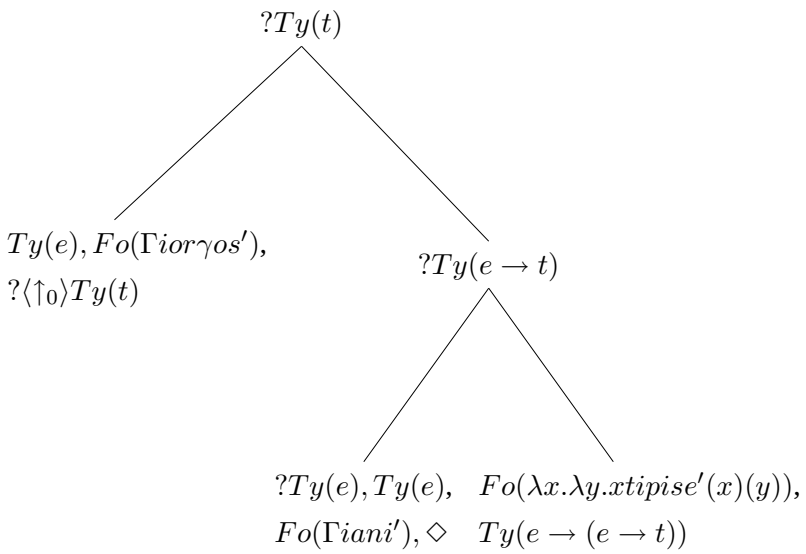
The next step is moving the pointer to the object node (010 node). We accomplish this by applying the rule of ANTICIPATION twice. The first application of the rule moves the pointer to the 01 node given that a requirement is present on that node, while the second application moves the pointer further down to the object node, again due to the presence of a requirement on that node:

(2.59) ANTICIPATION



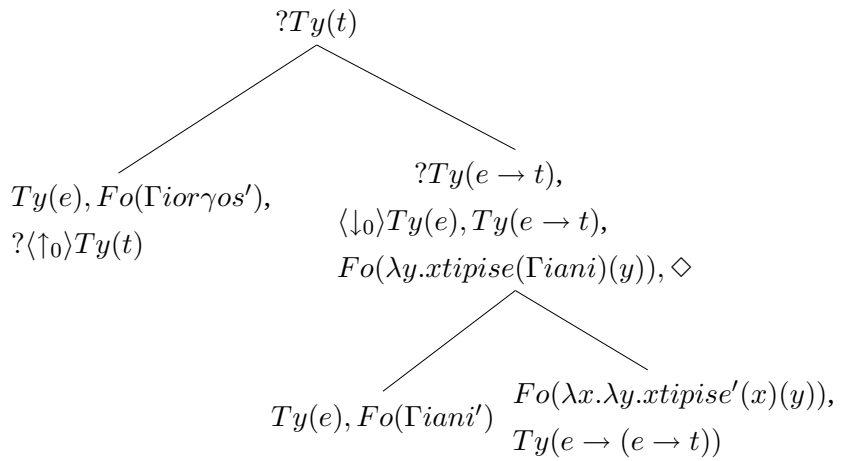
At that point, the object NP comes into parse. The result of parsing the NP is shown below:

(2.60) Parsing the object



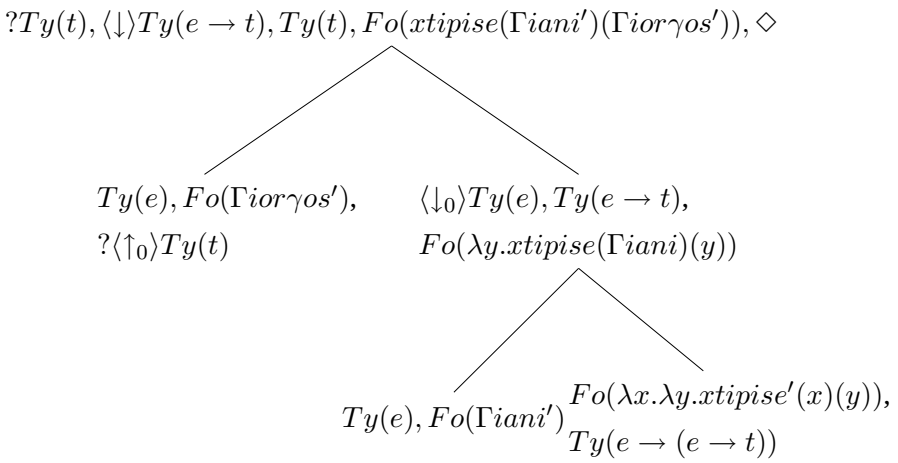
Now, THINNING applies in order to eliminate the type *e* requirement. Then COMPLETION moves the pointer up to the 01 node where ELIMINATION applies:

(2.61) THINNING, COMPLETION and ELIMINATION



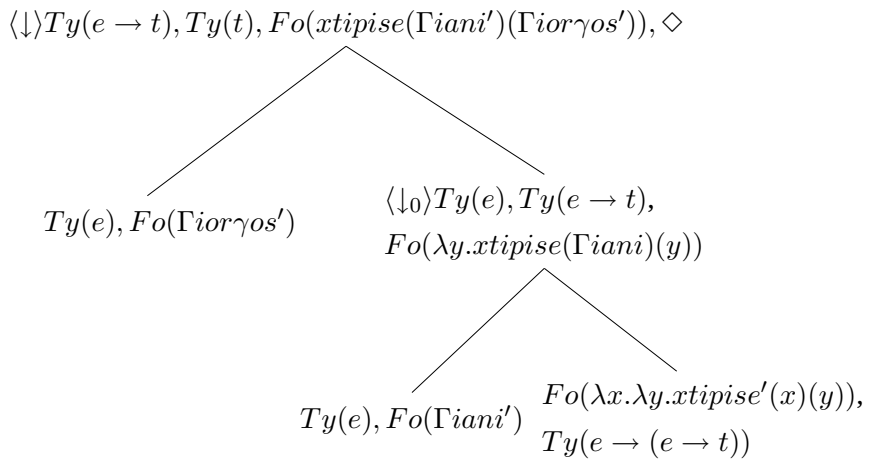
The same step is repeated, i.e. THINNING, COMPLETION and ELIMINATION:

(2.62) THINNING, COMPLETION and ELIMINATION



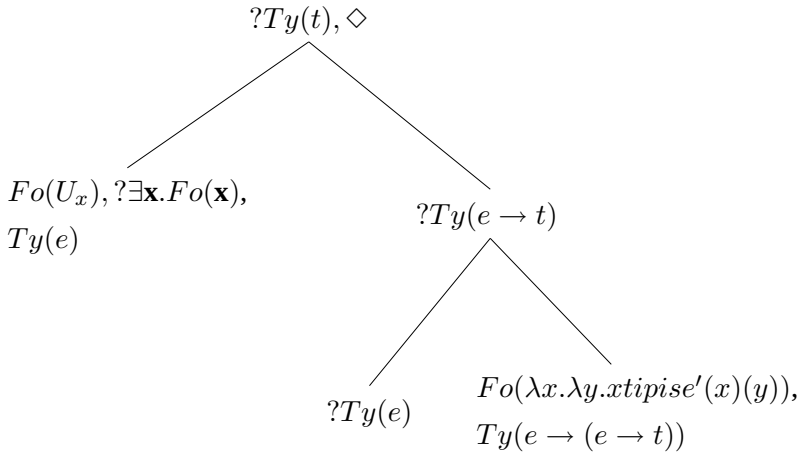
THINNING applies eliminating the type *t* requirement. The pointer goes down to the subject node via ANTICIPATION where THINNING applies eliminating the case filter requirement and comes up again via COMPLETION. The parse is now complete:

(2.63) Completing the parse



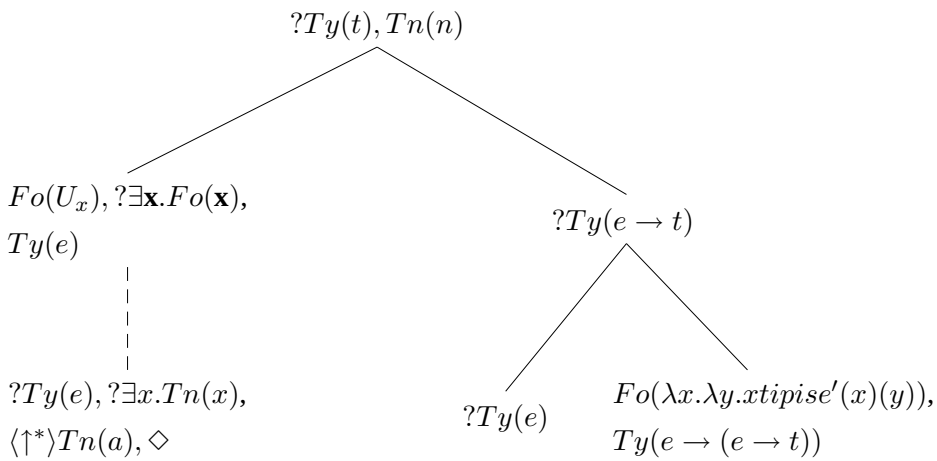
Now, in case we want to parse a VSO sentence, like e.g. the sentence in example (2.51), we begin by first parsing the verb starting from the AXIOM:

(2.64) Parsing the verb



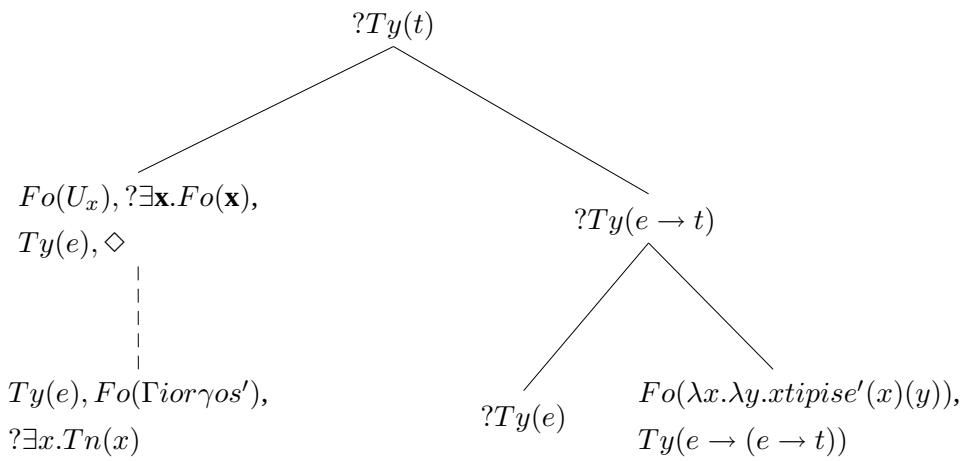
The pointer is moved to the subject node via ANTICIPATION, where LATE *ADJUNCTION is applied and projects an unfixed node with a requirement for a same type as the node where it begins from, i.e. a type *e* requirement:

(2.65) ANTICIPATION and LATE *ADJUNCTION



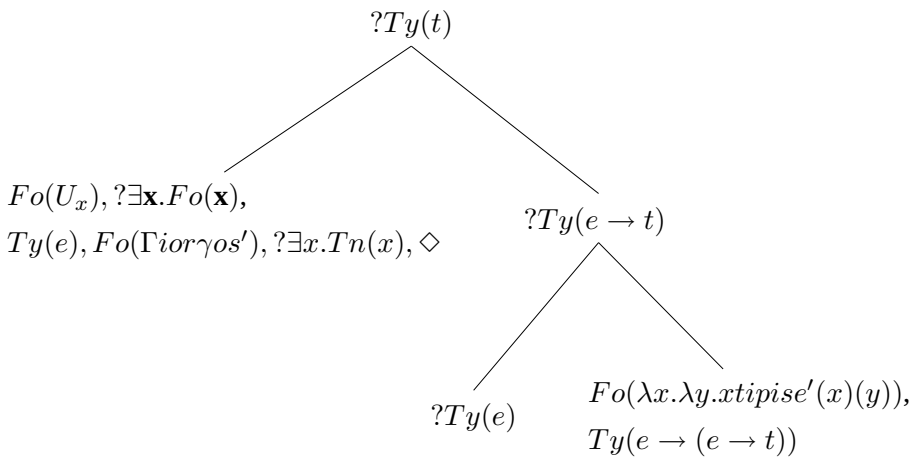
The subject is parsed on the unfixed node. Then, THINNING applies to eliminate the type *e* requirement of the unfixed node and the pointer is moved via COMPLETION to the 010 node:

(2.66) Parsing of the subject, THINNING and COMPLETION



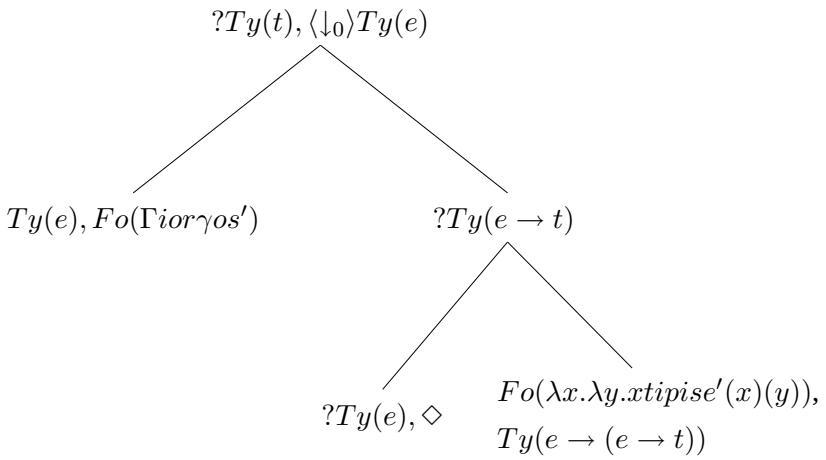
At that point the subject node (00) unifies with the unfixed node via MERGE:

(2.67) MERGE



At that point, THINNING applies eliminating the requirement for a proper treenode address to be found.[26] The pointer is then moved up via COMPLETION. Then, two steps of ANTICIPATION move the pointer to the object node:

(2.68) THINNING, COMPLETION and two steps of ANTICIPATION

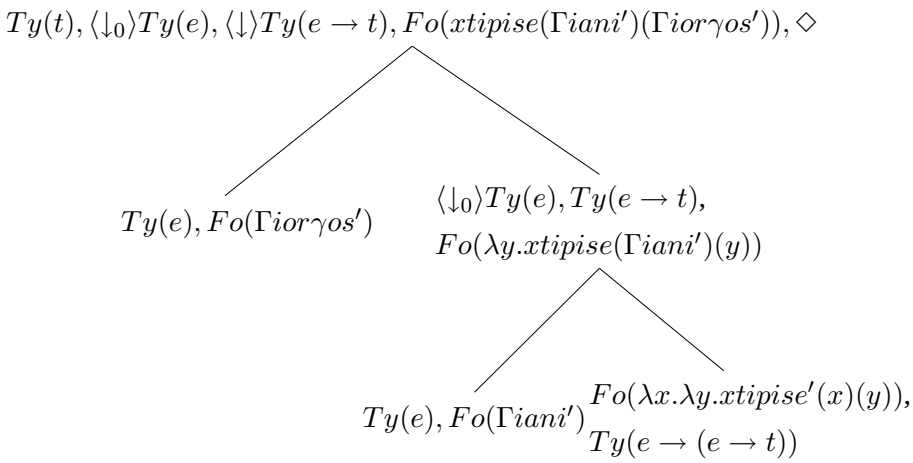


The object is parsed, and then THINNING gets rid of the type *e* requirement while COMPLETION moves the pointer to the 01 node where ELIMINATION applies. Applying

[26]If we wanted to be meticulous at this point, we would have to provide a *Tn* decoration for the fixed subject node, since the rule of THINNING applies in the presence of a decoration satisfying the requirement. However, the implicit assumption here is that every fixed node has a treenode address. Given this assumption, I do not encode the *Tn* decoration in the tree itself, except in cases where it is crucial for the application of a rule, e.g. the *ADJUNCTION rule. However, it should be kept in mind that all fixed nodes are implicitly assumed to bear a *Tn* value.

THINNING, COMPLETION and ELIMINATION again will give us a well-formed parse:

(2.69) Completing the parse



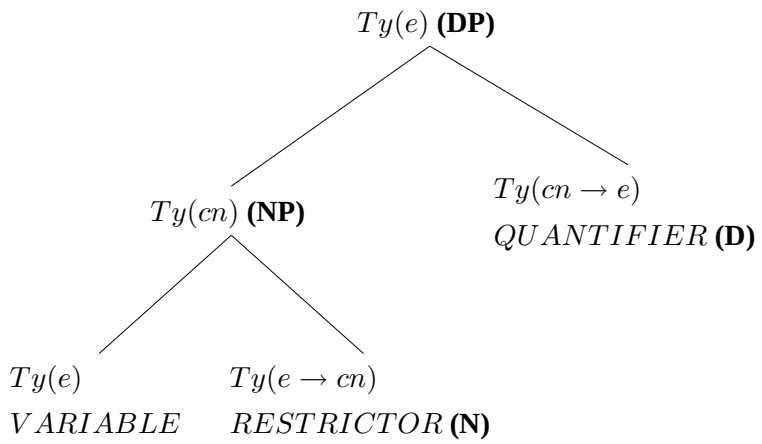
2.2 More Advanced Concepts

2.2.1 Quantification

We have already seen that the semantic type assumed for noun phrases in DS is a lower type *e*, contrary to what most of the formal semantic theories starting from Montague (1974) assume. In all these approaches (see Montague, 1974; Dowty, 1981; Morrill, 1994; Kratzer and Heim, 1997 among others), all noun phrases, both quantified and non-quantified, are given a higher semantic type ($((e \rightarrow t) \rightarrow t)$). The reason for this higher semantic type for NPs was that common nouns were assumed (and still are) to be of type $e \rightarrow t$. Given that an intransitive verb (or a transitive verb that has saturated one of its arguments) will be of the same type, the question that arises is what the type of the NP in a sentence like *A man cried* is. The solution proposed by Montague (1974) was to attribute a higher type to quantified determiners, while he assumed that the NP rather than the verb is the functor which takes the verb as its argument. The new higher type proposed for quantified determiners was $((e \rightarrow t) \rightarrow ((e \rightarrow t) \rightarrow t))$. This type combined with the common noun type $(e \rightarrow t)$ results in the semantic type $((e \rightarrow t) \rightarrow t))$, the generalized quantifier type. Then, the verb is assumed to be the argument and not the functor anymore, and combines with the quantified NP to return a t semantic type. Unlike all semantic analyses based

on generalized quantifiers, DS assumes NPs to be of the lowest semantic type possible, namely type *e*. Then, the complexity quantified NPs exhibit is captured assuming that NPs, even though being of type *e*, involve complex structure. The structure assumed for NPs in DS is the one depicted below. Note that this internal structure corresponds roughly to the mainstream view on the structure of the DP in GB/Minimalism (Abney, 1987). The equivalences to DP structure are shown in parentheses:[27]

(2.70) The structure of NPs in DS



As can be seen from the above tree structure, NPs in DS, even though assumed to be of the lowest type possible (*Ty(e)*), involve additional structure. Let us see what this additional structure stands for. The highest node is the node where the result of compiling all the other nodes via modus ponens and functional application is encoded, roughly the DP node in GB/Minimalism. The node indicated as "QUANTIFIER" is the node which will contribute the information on the form of quantification involved in each case. This node combined with the *Ty(cn)* (common noun) node will give us the higher *e* node. The RESTRICTOR node provides the binding domain of the variable introduced in the lower type *e* node (the variable node). NP content is expressed using the epsilon calculus. The epsilon calculus is a quantifier free predicate calculus equivalent logical system originally developed by Hilbert & Bernays (1939).[28] In a nutschell the epsilon calculus is based on

[27]DS, as already mentioned, does not represent word order in its semantic trees. Therefore, the structure assumed for NPs does not encode word order.

[28]The epsilon calculus can be seen as the formal study of the arbitrary names in natural deduction proof systems. See Lemmon (1965) for an introduction to logic using natural deduction.

a quantifier free predicate calculus language by adding two new operators ϵ and its dual τ (the epsilon and tau operator). Epsilon calculus formulas are then constructed in the following way: assume a predicate logic formula, like e.g. *Book'*. In order to turn this into an epsilon calculus formula we first add a variable, standing as the predicate's argument, i.e. *Book'x*. This is called the RESTRICTOR and its function is to indicate the domain quantified over. The next step is providing the operator relevant in each case along with the variable it binds. The new term constructed is of the following form: ϵ*,x,Book'x*. This formula denotes an entity satisfying the restrictor *Book'*. In case such an entity exists, this entity is chosen as the argument of the predicate, as a witness of *Book'*. I will not go into the formal details of how interpretation is assigned in epsilon calculus formulas. I just note that the ϵ and τ operators are equivalent to existential and universal quantifiers in predicate logic formulas respectively:[29]

(2.71) Epsilon calculus and predicate logic equivalences

$$\exists x F(x) \leftrightarrow F(\epsilon, x, Book'x)$$

$$\forall x F(x) \leftrightarrow F(\tau, x, Book'x)$$

Returning to the actual way NPs are represented in DS, let us explain how a quantified phrase, say *a man*, is parsed. Determiners in DS are assumed to contribute the form of quantification involved. In that sense, the lexical entry for a determiner, having a type *e* requiring node as its triggering point, will build the QUANTIFIER node and will further decorate it with a type and a formula value. The formula value will correspond to the form of quantification involved, while the type value will be of type $e \rightarrow cn$. Furthermore, the COMMON NOUN node will be also built and a requirement for a *cn* type will be projected in the same node, anticipating in a way the fact that a noun is going to follow. The lexical entry for indefinite *a* is shown below:

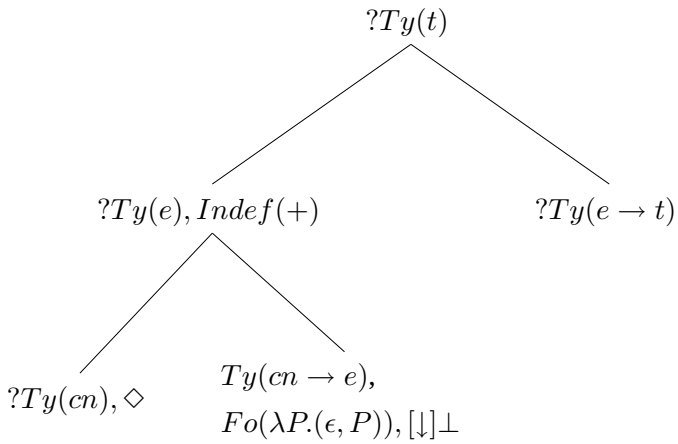
[29] For more information on the epsilon calculus see Hilbert & Bernays (1939), Meyer-Viol (1995).

(2.72) Lexical entry for indefinite *a*

IF	$?Ty(e)$
THEN	$\texttt{put}(Indef(+)); \texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$
	$\texttt{put}(Fo(\lambda P.(\epsilon, P)), Ty(cn \rightarrow e), [\downarrow]\bot);$
	$\texttt{go}(\langle\uparrow_1\rangle); \texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(?Ty(cn))$
ELSE	abort

In tree notation:

(2.73) Parsing the indefinite *a* [30]



The pointer is left at the type *cn* requiring node. At that point, the common noun *man* comes into parse. The lexical entry for a common noun like *man* builds the internal type *e* node and provides it with a variable. It further builds the RESTRICTOR node, where the formula value standing for the restrictor of the epsilon term is provided. The lexical entry for *man* is shown below:

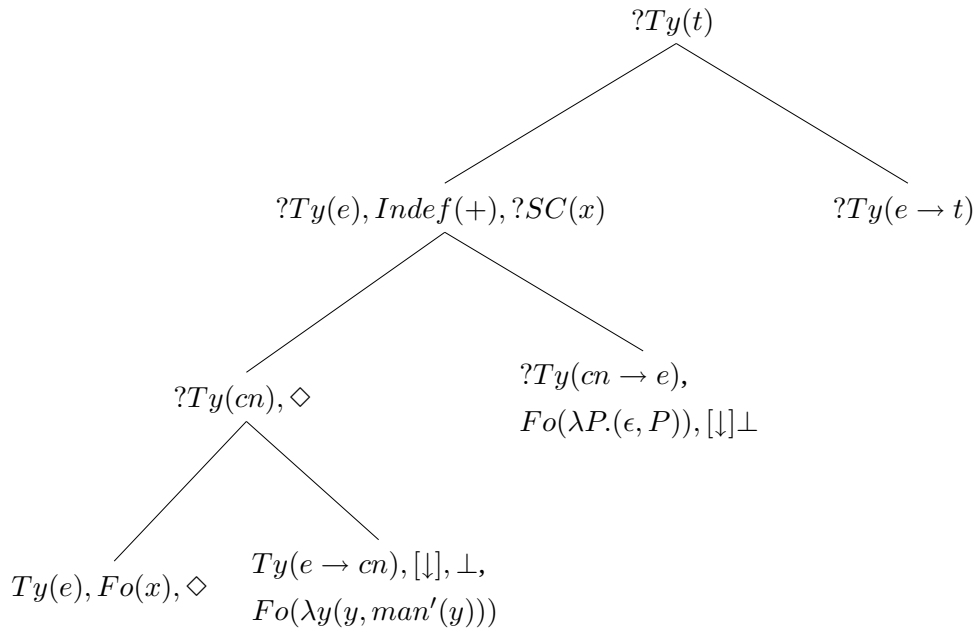
[30]The feature $Indef(+)$ is used in order to distinguish between indefinites and non-indefinites for the purpose of scopal relations. See Cann et al., 2005: chapter 3.

(2.74) Lexical entry for the common noun *man*

IF	$?Ty(cn)$
THEN	$\texttt{go}(\langle\uparrow_1\rangle); \texttt{put}(?Sc(x));$
	$\texttt{go}(\langle\downarrow_0\rangle); \texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$
	$\texttt{put}(Fo(\lambda y(y, man'(y))), Ty(e \rightarrow cn), [\downarrow]\bot);$
	$\texttt{go}(\langle\uparrow_1\rangle); \texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$
	$freshput(x, Fo(x)); \texttt{put}(Ty(e))$
ELSE	abort

The effect of the lexical entry in tree notation:

(2.75) Parsing *man* in *a man* shouted

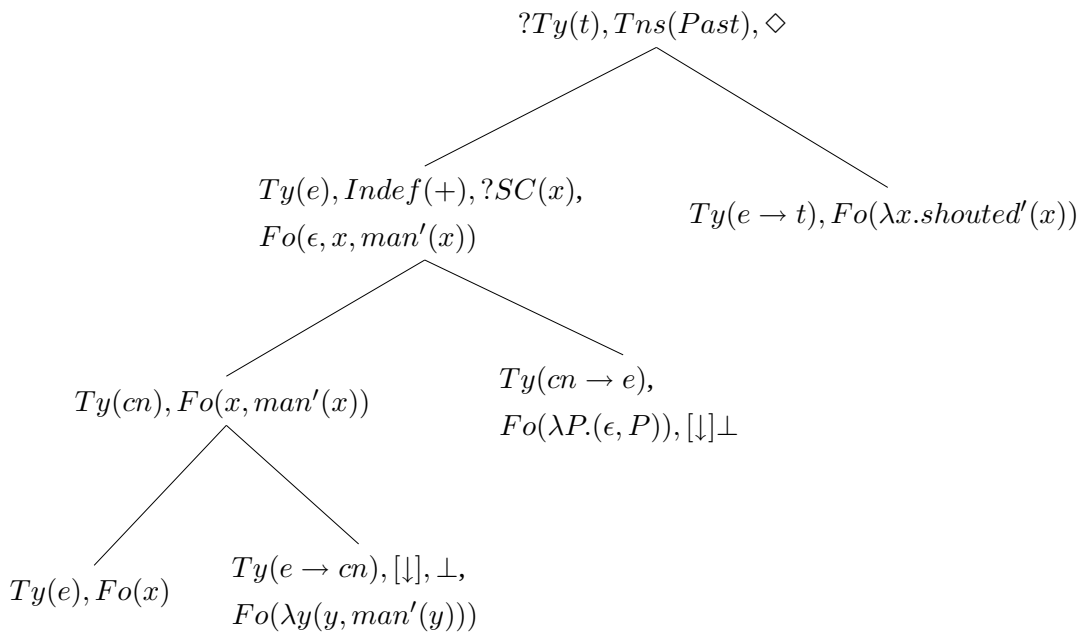


Note that the common noun projects a further requirement in the higher type *e* node, i.e. $?Sc(x)$. The requirement $?Sc(x)$ is a shorthand for a more complex modal requirement that ensures that the variable introduced in the lower type *e* node, participates in a scopal statement in the most local propositional node:

(2.76) $?SC(x) =_{def} ?\langle\uparrow_0\rangle\langle\uparrow_1^*\rangle\exists y.Scope(y < x) \vee Scope(x < y)$

With the pointer at the internal type node, COMPLETION can apply moving the pointer to the mother node. At this point ELIMINATION applies providing the 000 node with a type and a formula value. The same process is repeated and the higher type *e* requiring node also gets a formula and a type value. One step of COMPLETION and one step of ANTICIPATION moves the pointer to the $?Ty(e \rightarrow t)$ node where at that point a verb can come into parse:[31]

(2.77) Parsing *shouted* in *a man shouted*

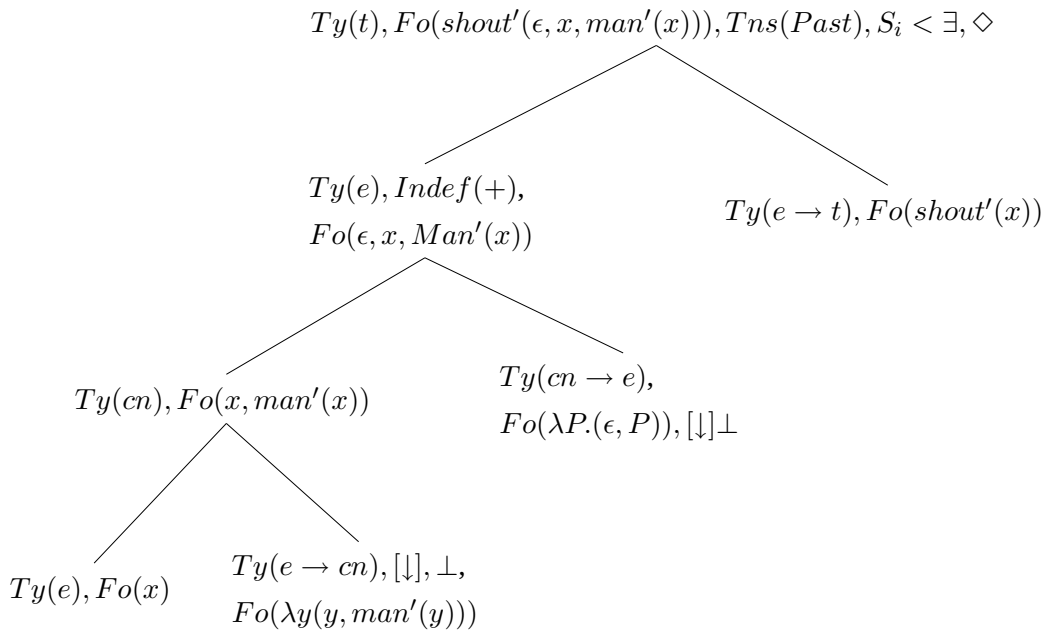


With the pointer at the top node, ELIMINATION applies and a well-formed parse obtains. It should be noted that a step of scope creation in the type *t* node needs to be taken. Since I will not deal with scope in this thesis, I will not get into detail here. It suffices to say that this scope statement collects all the variables appearing on NP nodes and orders them with respect to an index of evaluation S_i, which is the time at which the formula at the type *t* node is said to hold. The different scopal orderings (wide-narrow scope) follow from the form of quantification involved in each case. The interested reader is referred to Cann et al. (2005: Chapter 3) and Gregoromichelaki (2006, 2010) for more information

[31]The triggering point for verbs in English is assumed to be a type $e \rightarrow t$ requiring node in contrast to languages like SMG where the triggering point for verbs is assumed to be a type *t* requiring node.

on scope and scope calculation in DS. Returning to our actual example a scopal statement is created at the top node, with the index of evaluation taking wide scope over the indefinite. The creation of a scopal statement eliminates the requirement for a scopal statement to be present in some local type *t* node in the higher type *e* node and the parse is rendered successful:

(2.78) Completing the parse of *a man shouted*



The complex structure of NPs will be represented only when necessary in this thesis. In most cases, I will represent an NP having complex structure as involving just one type *e* node followed by the specification $[\downarrow]\top$ which reads as: there is structure below me. The reader should however bear in mind that all NPs (not pronominals) will involve complex structure. The complex structure assumed for NPs will prove to be the reason some languages allow CD constructions (SMG) and some others do not (PG). It will be shown in chapter 6 that availability of CD is due to whether clitics in a given language have a bottom restriction ($[\downarrow]\bot$) or not.

2.2.2 Parsing in Context - LINK Structures

Besides the tree structures in which each sentence involves a single tree (regardless of tree embedding), DS also makes use of pairs of trees which are linked to each other via a relation called LINK. LINK structures involve two separate tree structures linked by means of an arrow relation (LINK), that share in most of the cases a term. The node from which the LINK starts can be seen as setting the context in which the LINKed tree is going to be parsed. Examples of LINK relations include relative clauses, in which case the relative clause is parsed within the context of the head or HTLD constructions in which case the HTLD sentence is parsed within the context of having parsed the left-dislocated element first. LINK structures have a variety of uses in DS. Below I refer to the ones most relevant to the scope of this thesis.

2.2.2.1 Relative Clauses

LINK structures have been used by Kempson et al. (2001) and Cann et al. (2005) to analyze relative clauses. According to this analysis, the relative head is linked with the relative clause via a LINK relation. This LINK relation links a type *e* node (the relative head node) with a node that has a requirement for a type *t* (the LINKed tree in which the relative clause will be parsed). In defining LINK, two new modal relations must be introduced, $\langle L \rangle$ and $\langle L^{-1} \rangle$. The first of the two relations traverses a LINK relation from the node where the LINK starts to the LINKed node, while the second relation is its inverse, i.e. it traverses the LINK relation from the LINKed node to the node where the LINK starts. Returning to relative clauses, the rule that creates the LINK relation from the head noun node (type *e*) to a type *t* requiring node, called the LINK ADJUNCTION rule, is shown below:

(2.79) LINK ADJUNCTION

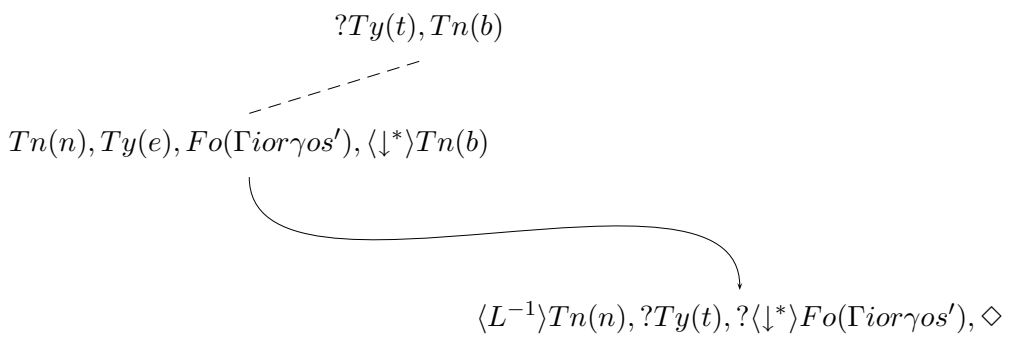
$$\frac{\{...\{Tn(a), Fo(\alpha), Ty(e), \Diamond\}...\}}{\{...\{Tn(a), Fo(\alpha), Ty(e), \Diamond\}...\}\{\langle L^{-1}\rangle Tn(a), ?Ty(t), ?\langle\downarrow^*\rangle Fo(\alpha), \Diamond\}}$$

Given this rule, let us see how relative clauses are parsed. I first begin with the so-called non-restrictive relatives which in languages like English or SMG have (or at least they are assumed to have) an intonational break between the head and the relative clause:

(2.80) *O* *Γiorγos,* *pu* *traguδai,* *irthe*
the.NOM George.NOM that sing.3SG came.3SG
'George who sings came.'

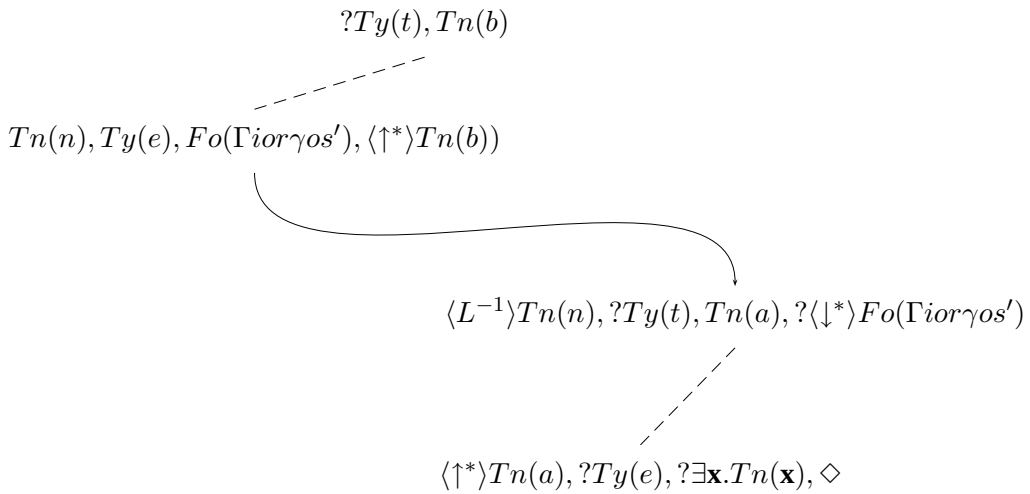
The first step involves parsing of the subject on an unfixed node. Then from that unfixed node LINK ADJUNCTION takes effect. The result is shown below:

(2.81) Parsing of the subject on an unfixed node and LINK ADJUNCTION



With the pointer at the type *t* requiring node of the LINKed tree, the rule of *ADJUNCTION takes effect.[32] The result is shown below:

(2.82) *ADJUNCTION



[32]Note that no problem is created given that another unfixed node is present, since the two unfixed nodes are in different trees.

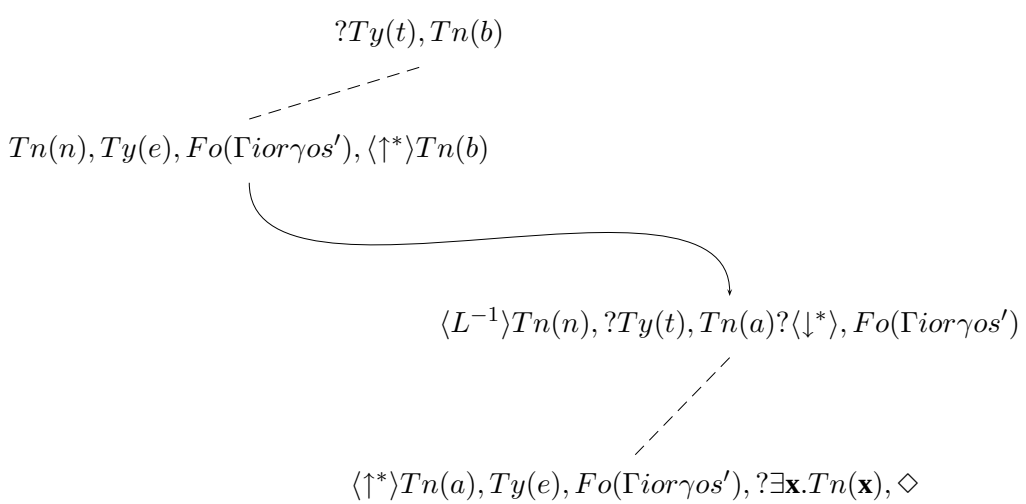
The next element that comes into parse is the relativizer *pu* 'who/which'. Following Cann et al. (2005), I assume that the relativizing element has the function of providing the value for the copy of the head noun. More precisely, its lexical entry specifies the configuration which is induced by the opening expression of a relative clause sequence. Specifically, the lexical entry for *pu* requires the presence of an unfixed node inside a LINKed tree in order to be parsed. If this condition is satisfied, then it provides the copy found in the node where the LINK starts. The lexical entry is shown below:

(2.83) Lexical entry for relativizing *pu*

IF	$?Ty(e), ?\exists\mathbf{x}.Tn(\mathbf{x}), \langle\uparrow^*\rangle\langle L^{-1}\rangle Fo(\mathbf{x})$
THEN	$\texttt{put}(Fo(\mathbf{x}), Ty(e), [\downarrow]\bot)$
ELSE	abort

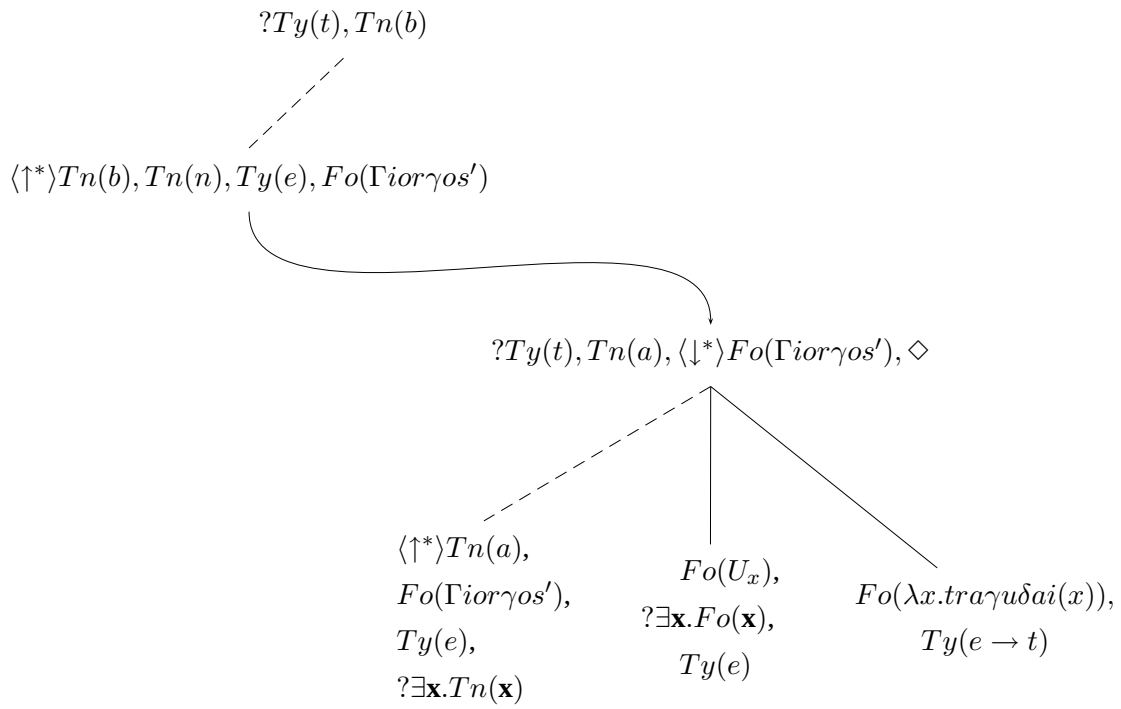
The above entry states that if the relativizer is on a type *e* requiring node that is unfixed ($?\exists\mathbf{x}.Tn(\mathbf{x})$) and furthermore there is a type value in the node where the LINK starts ($\langle\uparrow^*\rangle\langle L^{-1}\rangle Fo(\mathbf{x})$), then the relativizer proceeds and projects the same formula value along with a type *e* value on the unfixed node:

(2.84) Parsing of *pu*



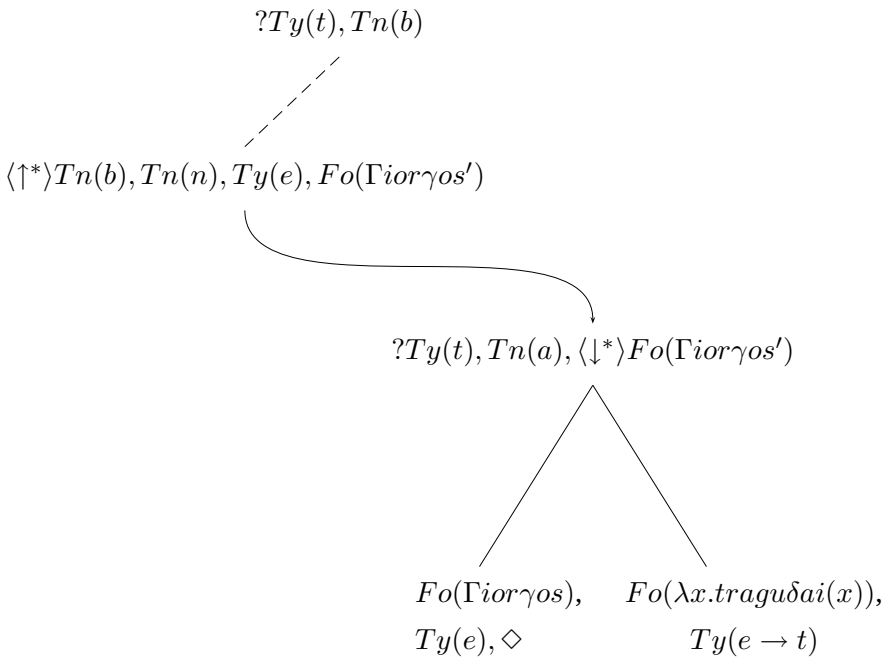
After COMPLETION, the verb comes into parse, projecting the following structure :

(2.85) Parsing the verb



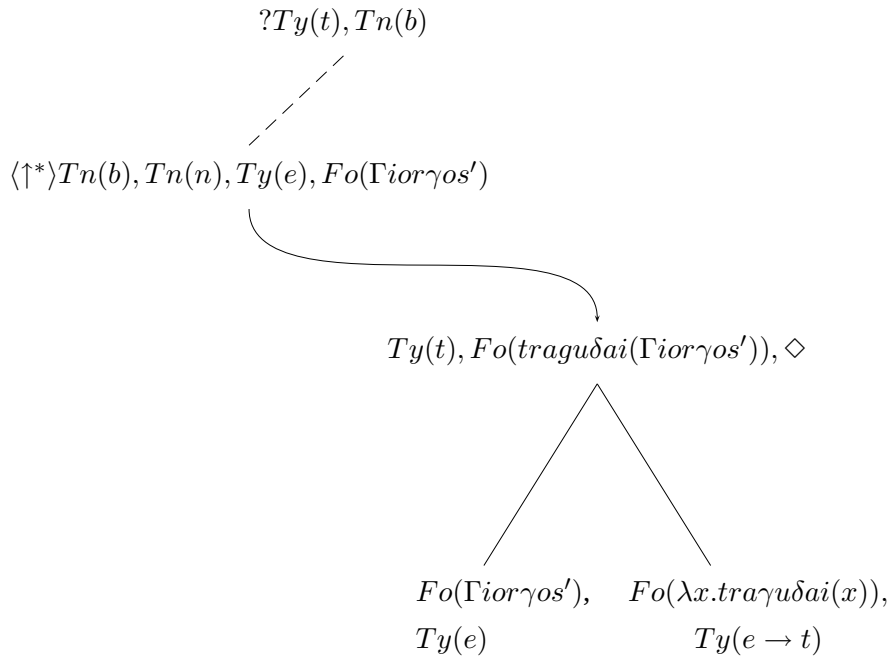
The unfixed node unifies in this case with the subject node via MERGE. Notice how this is another case where a metavariable is assigned a value as a side effect of a structure-building operation:

(2.86) After MERGE of the unfixed node



At this point the LINKed tree can be completed following the usual steps:

(2.87) Compiling the LINKed tree



Now, we want the pointer to return to the initial tree. However, COMPLETION cannot apply since the definition for COMPLETION does not include the inverse LINK relation in its set. A small modification of the rule will solve this problem:

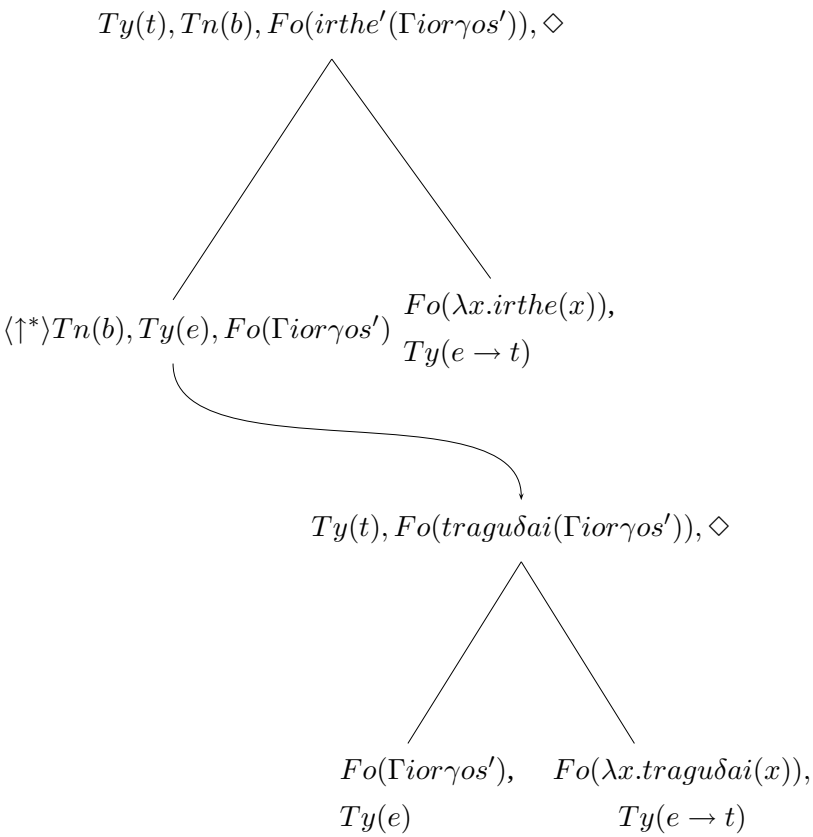
(2.88) COMPLETION revised

$$\frac{\{...\{Tn(n),...\},\{\langle\mu^{-1}\rangle Tn(n),...,Ty(X)..,\Diamond\}...\}}{\{...\{Tn(n),...,\langle\mu\rangle Ty(X),...,\Diamond\},\{\langle\mu^{-1}\rangle Tn(n),...,Ty(X),...\}...\}}$$

Where $\mu^{-1} \in \{\uparrow_0, \uparrow_1, \uparrow^*, L^{-1}\}, \mu \in \{\downarrow_0, \downarrow_1, \downarrow^*, L\}$.

With this rule COMPLETION can apply and return the pointer to the LINKed tree. Applying COMPLETION again will return the pointer to the type *t* requiring node of the main node, where the verb can now be parsed and again following the usual rules we compile the main tree as well:

(2.89) Compiling the main tree



The last step in completing the parse is a way to semantically interpret structures like the above. DS uses a number of rules that evaluate LINK trees according to each case. In our example, we need a rule that will interpret non-restrictive relative clauses. The LINK evaluation rule is shown below:

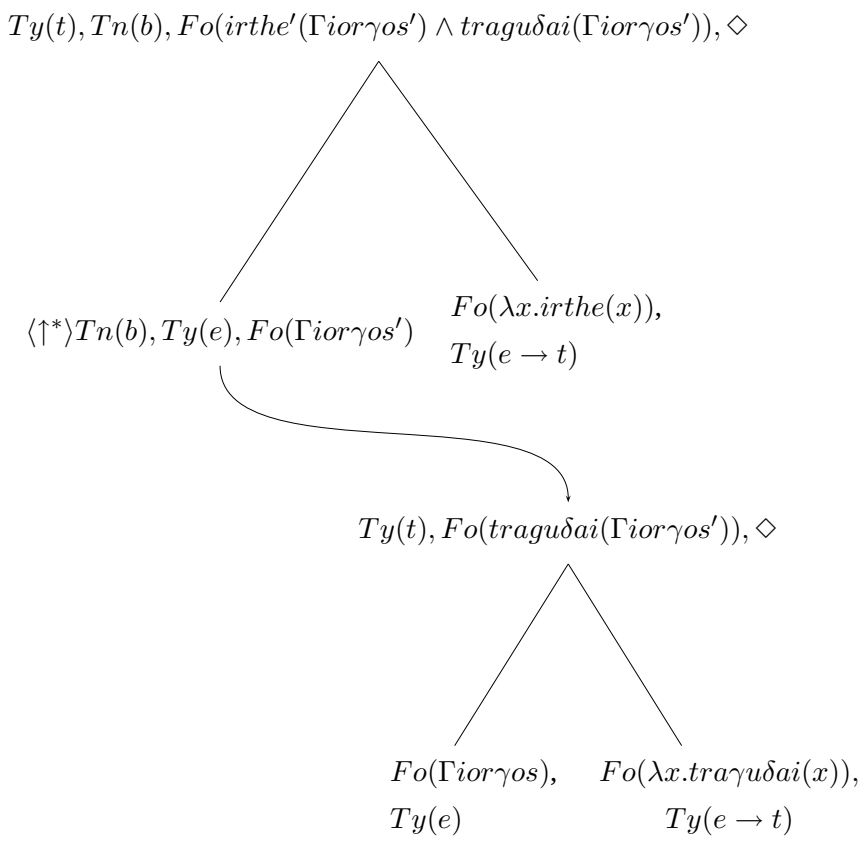
(2.90) LINK evaluation (non-restrictives)

$$\frac{\{...\{Tn(a), Fo(a), Ty(t), ..., \Diamond\}...\{\langle L^{-1}\rangle MOD(Tn(a)), Fo(b), Ty(t)\}...\}}{\{...\{Tn(a), Fo(a \wedge b), Ty(t), \Diamond\}...\{\langle L^{-1}\rangle MOD(Tn(a)), Fo(b), Ty(t)\}...\}}$$

Where MOD $\in \{\langle\uparrow_0\rangle, \langle\uparrow_1\rangle\}^*$

LINK evaluation conjoins the two type *t* formulas found in the main and LINKed tree respectively. Applying LINK evaluation in (2.91) results in the following:

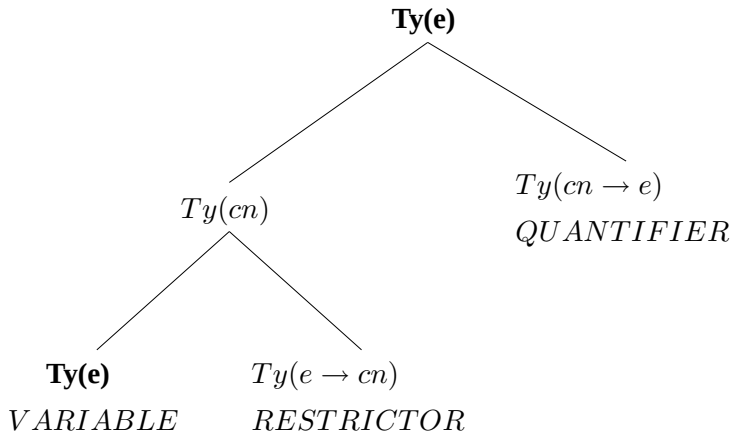
(2.91) After LINK evaluation



Turning to restrictive relative clauses, DS has a very natural mechanism to distinguish between non-restrictive and restrictive relatives. Remember that NPs are assumed to project

complex structure. This complex structure has two type *e* slots, the lower one where the variable that will combine with the restrictor functor node and the higher one representing the result of combining the common noun node with the binder node. The difference between non-restrictive and restrictive relatives will be accounted assuming that the LINK relation can in principle start from any of the two type nodes. This fact will give us the difference of restrictive and non-restrictive readings:

(2.92) The two possible e nodes that the LINK will start from (indicated in bold)



The exact details of the restrictive relatives analysis are not relevant here. The interested reader is directed to Cann et al. (2005) for the exact analysis of restrictive relatives as well as to Kula & Marten (forthcoming) for an analysis of restrictive and non-restrictive relatives in Bemba. What is relevant for us is that parsing of the relativizing element is done inside an unfixed node. This fact will become crucial when discussing proclisis in CG relative clauses.

2.2.2.2 Hanging Topic Left Dislocation

LINK structures are also used to model HTLD constructions. Cann et al. (2005) define two rules which link a type *e* node where the dislocated element is parsed to a type t requiring node, where the rest of the HTLD structure is parsed. The first rule is called TOPIC STRUCTURE INTRODUCTION and introduces a LINK transition from a type *e* requiring node to a type *t* requiring node, while leaving the pointer in the first of the two:

(2.93) TOPIC STRUCTURE INTRODUCTION

$$\frac{\{\{Tn(0), ?Ty(t), \diamond\}\}}{\{\{Tn(0), ?Ty(t)\}\}, \{\langle L\rangle Tn(0), ?Ty(e), \diamond\}}$$

Notice that the above rule does not mention anything about a shared term. That is because there is no shared term at the time the rule applies. The requirement for a shared term is introduced via the second rule, TOPIC STRUCTURE REQUIREMENT. This rule applies as soon as the dislocated argument is parsed. The rule is shown below:

(2.94) TOPIC STRUCTURE REQUIREMENT

$$\frac{\{\{Tn(0), ?Ty(t)\}, \{\langle L\rangle Tn(0), Fo(a), Ty(e), \diamond\}\}}{\{\{Tn(0), ?Ty(t), ?\langle D\rangle Fo(a), \diamond\}\}, \{\langle L\rangle Tn(0), Fo(a), Ty(e)\}\}}$$

Where D $\in \{\downarrow_0, \downarrow_1, \downarrow^*, L\}$

Let us say we want to parse the HTLD structure shown below:

(2.95) *O* *Γiorγos,* *ton* *γnorizo*
the.NOM George.NOM him.CL-ACC know.1SG
'I know George'

First, TOPIC STRUCTURE INTRODUCTION applies introducing the type *e* requiring node. The NP is parsed on that node and then TOPIC STRUCTURE REQUIREMENT takes effect moving the pointer to the type *t* requiring node and positing (on the same node) a requirement that a copy of the formula found in the node where the LINK begins must be found somewhere in the LINKed tree or to a tree LINked to the LINKed tree ($?\langle D\rangle Fo(\Gamma ior\gamma os')$):

(2.96)

$$\begin{array}{ll} \langle L\rangle Tn(0), & \langle L^{-1}\rangle Tn(n), ?Ty(t), ?\langle D\rangle Fo(\Gamma iorg\gamma os') \\ Fo(\Gamma iorg\gamma os'), [\downarrow]\top, & \\ Ty(e) & \end{array}$$

The crucial thing about this rule is that a formula value needs to be shared between the two trees. This is the reason that HTLD constructions in English require a resumptive element. SMG on the other, requires such an element only for the object cases. An HTLD construction where the left dislocated element is interpreted as the subject does not need such an element as witness the example below:

(2.97) *O* *Γiorγos,* *γnorizi* *ti* *Maria*
the.NOM George.NOM know.3SG the.ACC Mary.ACC
'George knows Mary'

However, such structures are not problematic at all for the account proposed, since in SMG verbs are assumed to provide a type *e* value and a formula metavariable in the subject node. In other words, the shared formula can be provided by the context in pro-drop languages like SMG, assuming that the metavariable gets updated to the *Fo* value that needs to be shared. In English there is no such option, since verbs in this language do not decorate the subject node with a type value and a formula metavariable, given the non pro-drop properties of the language.

2.2.2.3 Coordinating Structures

LINK structures are also used in order to account for coordination of any kind. The basic idea behind the DS analysis of coordination is that it involves a LINK transition from a type complete node to a node with a requirement for the same type. For example, the coordinating conjunction *and* has been given the following lexical entry in Cann et al. (2005):

(2.98) Lexical entry for coordinate conjunction *and*

IF	$Ty(X)$
THEN	$\texttt{make}(L); \texttt{go}(L); \texttt{put}(?Ty(X))$
ELSE	abort

The rule will apply in case a complete type is present. In that sense, the rule is general enough to account for NP coordination (*Ty(e)*), sentential coordination (*Ty(t)*) or even VP coordination ($Ty(e \rightarrow t)$).

There are a number of additional uses of LINK that are going to be used in this thesis but these will be introduced in the relevant chapters. The interested reader is referred to Cann et al. (2005) and Kula & Marten (forthcoming) for more information on relative clauses, and Cann et al. (2005) and chapter 6 of this thesis for more information on HTLD.

2.2.2.4 Tense and Aspect

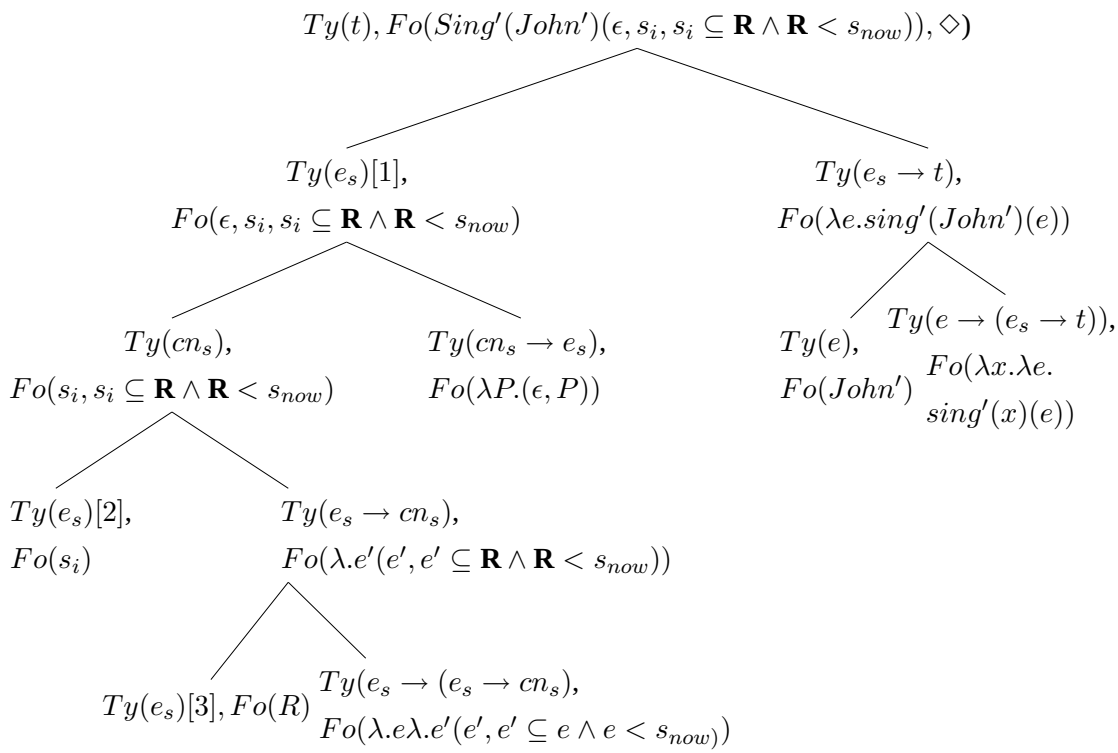
In both Kempson et al. (2001) and Cann et al. (2005), no attempt to address tense or aspect was made. Tense was encoded as a diacritic in Cann et al. (2005) (e.g. $Tns(Past)$), noting that a proper analysis of tense is pending in the framework. Recent advances within DS however assume a treatment of tense based on the introduction of an explicit situation argument, introducing a situation in which the formula value the proposition expresses will be true.[33] This situation argument node is then the locus of all tense and aspect information. In order for this situation argument node to be represented two additional nodes are added to the standard DS propositional spine. A node standing for the situation argument, which is assumed to be of type e_s (with *s* standing for situation),[34] and a functor node of type $e_s \rightarrow t$. The situation argument node in line with quantified NPs is assumed to involve complex structure. The example below shows the structure assumed in Cann (forthcoming) after the complete parse of *John sang*:[35]

[33]The situation argument node in DS was first introduced in Gregoromichelaki (2006). See Heim (1990), Chierchia (1995) and Gregoromichelaki (2006) among others for argumentation on the need to encode an explicit situation argument in representing propositions. Also see Tenny & Pustejovsky (2000) and references therein, for a history of event/situations in linguistic theory.

[34]In line with Gregoromichelaki (2006: 196) I assume that D_{Ty} involves *Ty(e)* as a general type with subtypes e_i for individuals, e_s for situations etc. Cann (forthcoming) uses the notation e_{sit} which is nothing more than a notational variant.

[35]Due to the number of situation e_s type nodes, I have provided a top-down numbering system from 1 to 3 for the three type e_s nodes of the complex situation argument node. From now on, the type e_s will be referred with their respective number to make communication with the reader easier.

(2.99) Parsing *John sang*



In the above structure, the intransitive verb *sing* is taken to be a two-place predicate, subcategorizing for both a subject and an event/situation argument. Let us see what the complex situation argument does. In the lowest e_s [3] node, the reference time metavariable **R** is introduced. This will combine with the semantic specifications given for the past tense in the lowest functor node ($Fo(\lambda.e\lambda.e'(e', e' \subseteq e \wedge e < s_{now}))$), to return a formula value in which the first lambda bound variable (*e*) is substituted by **R** ($Fo(\lambda.e'(e', e' \subseteq \mathbf{R} \wedge \mathbf{R} < s_{now}))$). This new formula states that the remaining lambda bound variable *e'* is contained within or holds at **R** ($e' \subseteq \mathbf{R}$) and that the reference time precedes the utterance time ($\mathbf{R} < s_{now}$)). In the intemediate e_s [2] node, a situation s_i is introduced. This situation will substitute the remaining lambda bound variable (*e'*) to return the formula value $Fo(s_i, s_i \subseteq \mathbf{R} \wedge \mathbf{R} < s_{now})$, i.e. provides a situation that will bear the given tense/aspect specifications. The penultimate step involves binding the last formula we have obtained by the ϵ operator. In the example above, the situation with the given specifications is bound by the ϵ operator and returns a formula value which roughly states that a past situation exists

($Fo(\epsilon, s_i, s_i \subseteq \mathbf{R} \wedge \mathbf{R} < s_{now})$). The very last step involves substituting this last formula for the lambda bound *e* found in the formula value for *sing* to get the well-formed type *t* formula $Ty(t), Fo(Sing'(John')(\epsilon, s_i, s_i \subseteq \mathbf{R} \wedge \mathbf{R} < s_{now}))$.

Tense/aspect information are assumed to be projected mainly from verbs, both auxiliary and content. However, a number of other elements can be taken to provide such information, like modality/tense particles/markers or even negation markers (see chapters 3 and 4 of this thesis).

2.3 Parsing and the Availability of Different Parsing Strategies

Until now we have discussed a number of different strategies that DS uses, such as unfixed nodes, LINK structures, parsing on a previously built fixed node. The availability of these different parsing strategies as general computational rules of the system allows these different strategies to be available for the one and the same element in some cases. For example, parsing of the subject in a typical SMG SVO sentence can be done in more than one way. The first way is the one used in 2.1.9 exemplifying the complete parses of the SVO sentence. Using this way, the rule of *ADJUNCTION applies first followed by parsing of the subject on that unfixed node:

(2.100) *ADJUNCTION and parsing of the subject ‘o Γiorγos’

$$Tn(n), ?Ty(t)$$

$$\langle\uparrow *\rangle Tn(n),$$
$$?Ty(e), Ty(e), Fo(\Gamma ior\gamma os'),$$
$$?\langle\uparrow_0\rangle Ty(t), ?\exists \mathbf{x}.Tn(\mathbf{x}), \diamond$$

The parse continues as shown in section 2.1.9 and a well-formed parse is obtained. However, remember that DS also makes use of LINK structures. In particular, the rules of

TOPIC STRUCTURE INTRODUCTION and TOPIC STRUCTURE REQUIREMENT are used in left dislocated topic constructions. First, TOPIC STRUCTURE INTRODUCTION applies and a LINK relation is induced from a type *e* requiring node to a type *t* requiring node. The subject can be parsed on that type *e* node. Then, TOPIC STRUCTURE REQUIREMENT posits that a requirement for the same formula value as the one found in the type *e* node where the LINK begins should be found in the LINKed type *t* requiring node:

(2.101) Parsing the subject *o* Γ*iorγos* 'the George.NOM' as a LINK structure

$$\langle L\rangle Tn(0),\ Fo(\Gamma ior\gamma os'),\ [\downarrow]\top,\ Ty(e) \qquad \langle L^{-1}\rangle Tn(n),\ ?Ty(t),\ ?\langle D\rangle Fo(\Gamma ior\gamma os'),\ \Diamond$$

The verb comes into parse and projects a formula metavariable and a type *e* value on the subject node and a type *e* requirement on the object node. The object can be parsed in the object node, and at that point assuming update of the formula metavariable of the subject with the formula value found in the node where the LINK begins, i.e. $Fo(\Gamma ior\gamma os')$, the parse is rendered successful. In that sense, parsing of the subject in a SVO construction in SMG is possible either via using an unfixed node or via using a LINK relation. This availability of different parsing strategies will be crucial when attempting a diachronic account of the clitic systems under consideration. It will be argued, following Bouzouita (2008a,b), that this availability is one of the key factors driving syntactic change, since the possibility of parsing one and the same element using different parsing strategies will be argued to cause speaker/hearer mismatches (see chapter 7). Such a mismatch obtains when the speaker's structure is misanalyzed by the hearer as involving a different parsing strategy. Notice that the hearer is also led to a well-formed result but the route to get to that well-formed result is different than the one used by the speaker. For example, such a situation will obtain when a speaker is uttering a typical SVO sentence and "uses" *ADJUNCTION for the subject, while the hearer parses this sentence "using" a LINK structure for the same element.

With this last note, I conclude the discussion on the DS framework. The interested reader is directed to Kempson et al. (2001) and Cann et al. (2005) for more details on the

DS framework.

Chapter 3

Standard Modern and Grecia Salentina Greek

3.1 Introductory Remarks

3.1.1 Standard Modern Greek

Standard Modern Greek (SMG) is the term used to refer to the official language of the Greek and Cypriot state, i.e. the form of Greek taught in schools in both Greece and Cyprus. SMG according to Mackridge (1985) has its basis in the Peloponnese Greek dialects. The reason for this can be traced back to the Greek war of Independence (1821-9) and the fact that one of the first areas to be liberated was the Peloponnese. Furthermore, the fact that the Peloponnese dialects (with the exception of Tsakonian) were the dialects closer to the written language (Mackridge, 1985: 4) also played a part in the use of Peloponesse Greek as the basis for SMG. This newly formed basis was further enriched by elements of the dialects of dominant Greek communities of the time, notably Greeks from Istanbul and the Ionian islands (Mackridge, 1985: 5). A number of other factors notably compulsory education and military service as well as the advent of radio and television "have made Greece (and especially Athens and Salonica) into a melting-pot in which speakers of various kinds of Greek have gradually sunk their linguistic differences" (Mackridge, 1985: 5). The form of Greek derived out of all these processes is what we today call SMG. The SMG data used

in this thesis will be based on various kinds of sources ranging from constructed examples to examples from the internet. The source of the data will be noted when necessary.

3.1.2 Grecia Salentina Greek

Grecia Salentina Greek (GSG) is the term that is going to be used in this thesis to refer to one of the varieties of the Southern Italian Greek dialect Grico (or Griko/Grekanico) that is spoken in the area of Grecia Salentina, Southern Italy.[1] Grico is comprised of GSG and the variety spoken in the area of Calabria. The reason GSG and not Grico in general is going to be analyzed is because the existing data on the clitic system of Calabrian Greek (Rohlfs, 1950, 1977; Katsoyannou, 1995) are not adequate enough to guarantee an analysis. GSG has been claimed to be spoken in 9 villages in the area of Grecia Salentina, Southeast of Lecce, suggesting a number of 20000 speakers (Horrocks, 1997). However, such a description does not seem to be accurate at present. At least in two of the nine villages (Melpignano and Soleto, possibly Zollino as well) GSG is not spoken anymore. Furthermore, in the rest of the villages, most (all?) native speakers of GSG are over 60 and comprise a very small unit of the population of the area. In view of this, the number 20000 seems exaggerated at least. The data used in this thesis will be mainly drawn from a fieldwork visit undertaken in July 2007 in 3 of the villages of Grecia Salentina, namely Calimera, Sternatia and Martano. Additional data from other sources will be cited when used.

3.2 The Data

Before we proceed with the data on the two dialects, it should be noted that the reason the two dialects are examined together is because their clitic positioning systems, as we shall see, are to a wide extent similar. In that respect, I will examine the two clitic systems on a par, providing a unitary analysis for both systems in case of identical data, while giving separate analyses otherwise. The first similarity of the two systems concerns the morphological forms of clitics in the two dialects. Surprisingly, these are the extremely

[1] See Manolessou (2005) for a discussion on the origins of the Grico dialect.

similar for both dialects. The only difference is that GSG has different forms for 1st/2nd plural clitics and the 3rd person accusative neuter clitic:[2]

(3.1) Clitic morphological forms for SMG

	$1st$	$2nd$	$3rd$
Sg accusative	me	se	ton/tin/to
Pl accusative	mas	sas	tus/tis - tes/ ta
Sg genitive	mu	su	tu/tis
Pl genitive	mas	sas	tus

(3.2) Clitic morphological forms for GSG

	$1st$	$2nd$	$3rd$
Sg accusative	me	se	ton/tin/(t)o
Pl accusative	ma(s)	sa(s)	tus/tis - tes/ ta
Sg genitive	mu	su	tu/tis
Pl genitive	ma(s)	sa(s)	tus

Note the difference in syncretism between 1st/2nd person clitics on the one hand and 3rd person clitics on the other. 1st/2nd person clitics exhibit non-syncretized singular forms with distinct morphological forms for genitive/accusative (me/se.ACC, mu/su.GEN), while exhibiting syncretized forms for 1st/2nd plural clitics (mas/sas.ACC, mas/sas.GEN). On the contrary, 3rd person clitics are non-syncretized across the board, i.e. both singular and plural forms exhibit distinct morphological forms for genitive/accusative. Dative case is morphologically realized as genitive, since dative morphology has dissapeared from both dialects.[3]

[2]The same system is reported for Calabrian Greek in Katsoyannou (1995) the only difference being that the doublet form *o* representing the 3rd person accusative neuter clitic is absent in Katsoyannou's description. The system presented in Karanastasis (1997) seems to follow the one presented here as well. However, in Karanastasis no distinctions between tonic and non-tonic pronouns is made. In that sense, it is very difficult to distinguish which ones are the clitic forms and which are not (see Karanastasis, 1997: 66-67). The same system is reported in Rohlfs (1977).

[3]It should be noted that the terms dative and genitive will be used interchangeably in this thesis to denote the dative function.

As regards clitic positioning, SMG and GSG follow the pattern generally attested in mainland and Ionian Greek (Ralli, 2006). According to this pattern, clitics appear preverbally in all environments except with imperatives and gerunds:[4]

(3.3) *Ton* *aγapá*
him.CL-ACC love.3SG
'S/He/It loves him.' [SMG]

(3.4) **Aγápa* *ton*
love.3SG him.CL-ACC
'S/He/It loves him.' [SMG]

(3.5) *Ton* *gapa*
him.CL-ACC love.3SG
'S/He/It loves him.' [GSG]

(3.6) **Gapa* *ton*
love.3SG him.CL-ACC
'S/He/It loves him.' [GSG]

(3.7) Γ*rafe* *to*
write.IMP-2SG it.CL-ACC
'Write it!' [SMG]

(3.8) **To* *γrafe*
it.CL-ACC write.IMP-2SG
'Write it!' [SMG]

(3.9) *Grafe* *to*
write.IMP-2SG it.CL-ACC
'Write it!' [GSG]

(3.10) **To* *grafe*
it.CL-ACC write.IMP-2SG
'Write it!' [GSG]

[4] An analysis of gerunds is pending in the DS framework. For this reason clitic positioning with gerunds will not be dealt with in this thesis

(3.11) *Γrafontas* *to...*
write.GER it.CL-ACC
'By writing it...' [SMG]

(3.12) **To* *γrafontas* *...*
it.CL-ACC write.GER
'By writing it...' [SMG]

(3.13) *Pulonta* *o*
sell.GER it.CL-ACC
'By selling it...' [GSG]

(3.14) **(T)o* *pulonta*
it.CL-ACC sell.GER
'By selling it...' [GSG]

Note that SMG does not have infinitives. Subjunctive complementation is used instead in which case the clitic appears preverbally:[5]

(3.15) *Θelo* *na* *ton* *δo*
want.1SG SUBJ him.CL-ACC see.PNP-1SG
'I want to see him' [SMG]

(3.16) **Θelo* *na* *δo* *ton*
want.1SG SUBJ see.PNP-1SG him.CL-ACC
'I want to see him.' [SMG]

GSG makes restrictive use of infinitives with a very restricted number of verbs. However, clitics are incompatible with infinitives in GSG:[6]

(3.17) **Sotzo* *(to)* *vorasi* *(to)*
can.1SG it.CL-ACC buy.INF it.CL-ACC
'I can buy it.' [GSG]

[5]The Perfective Non-Past (PNP) form is a dependent form. It can only appear after the subjunctive marker *na*, *as* and the future particle *θ*a. For this reason the PNP has been often coined as a "subjunctive" (Philippaki & Veloudis, 1984; Giannakidou, 1995 among others).

[6]We will discuss the consequences of this fact when we examine Clitic Climbing in GSG later on in this chapter.

Subjunctive complementation as in SMG is the norm in GSG as well:[7]

(3.18) *Thelo* *no* *do*
want.1SG SUBJ-it.CL-ACC see.PNP-1SG
'I want to see it.' [GSG]

(3.19) **Thelo* *na* *do* *(t)o*
want.1SG SUBJ see.PNP-1SG it.CL-ACC
'I want to see it.' [GSG]

In a sequence of two clitics, ordering inside the clitic sequence is strictly DAT-ACC in non-imperative environments in both dialects:

(3.20) *Tu* *to* *eδosa*
him.CL-GEN it.CL-ACC gave.1SG
'I gave it to him.' [SMG]

(3.21) **To* *tu* *eδosa*
it.CL-ACC him.CL-GEN gave.1SG
'I gave it to him.' [SMG]

(3.22) *Tu* *to* *doka*
him.CL-GEN it.CL-ACC gave.1SG
'I gave it to him.'[GSG]

(3.23) **To* *tu* *doka*
it.CL-ACC him.CL-GEN gave.1SG
'I gave it to him.'[GSG]

In imperative environments both orderings are possible in SMG, whereas DAT-ACC ordering is again the only option for GSG:

(3.24) *Δos* *mu* *to*
give.IMP-2SG me.CL-GEN it.CL-ACC
'Give it to me.' [SMG]

[7]The form *no* is a cluster comprised of the subjunctive marker *na* and the clitic form *to* 'it'.

(3.25) *Δos* *to* *mu*
give.IMP-2SG it.CL-ACC me.CL-GEN
'Give it to me.' [SMG]

(3.26) *Do* *mu* *to*
give.IMP-2SG me.CL-GEN it.CL-ACC
'Give it to me.' [GSG]

(3.27) **Do* *to* *mu*
give.IMP-2SG it.CL-ACC me.CL-GEN
'Give it to me.' [GSG]

There are a number of other clitic related constructions in the two dialects, notably Clitic Climbing (CC) in GSG and Doubling constructions in both dialects. However, the data for these phenomena will be presented and discussed in the relevant sections (CC in this chapter: 3.5.3, Doubling in chapter 6).

3.3 Existing Approaches

3.3.1 GB/Minimalism

The accounts I am going to review with the exception of Condoravdi and Kiparksy (2002) are mainly accounts of the SMG clitic system. However, the similiarities between the two systems make these accounts relevant for GSG as well. The literature on SMG is quite vast (Sportiche, 1993; Anagnostopoulou, 1994, 1997, 1999, 2003, 2005; Philippaki 1977, 1987, 1994; Philippaki & Spyropoulos, 1999; Terzi, 1999; Papangeli, 2000; Androulakis, 2001; Philippaki, 2002; Mavrogiorgos, 2006 among others). All these accounts can be distinguished from each other depending on their stance on the following three issues : a) the affix vs word status of clitics b) the surface status of clitics, i.e. movement vs base-generation c) the argument vs adjunct status of clitics. All three issues are tightly connected to each other and the decision one makes on any of these, immediately influences the other two. For example, assuming that clitics are arguments of the verb implies a base-generation analysis where clitics are base generated in the object argument position. Within the same line of reasoning, deciding on the affix vs word status of clitics has further implications on

the nature of the account that is going to be pursued, since different rules of the system are going to be used in each case (movement of the clitic to the verb or vice versa). But let us see in more detail what these accounts actually propose.

3.3.1.1 Philippaki 1977, 1987: Base Generated Clitics

The earliest attempt to account for cliticization in SMG is Philippaki (1977). Philippaki (1977) following Kayne's (1975) classic analysis of French clitics assumes that clitics are base generated in the object argument position with subsequent syntactic movement to Infl/T. Syntactic movement is licit on the assumption that clitics are full words and not affixes. In Doubling constructions, Philippaki (1977, 1987) further assumes that the doubled NP (DP in contemporary terminology) is not an argument but rather a dislocated adjunct. A coindexation mechanism between the doubled NP/DP and the clitic is further postulated in order to explain their agreement in terms of phi-features and case. The account proposed by Philippaki has received a number of critiques especially with respect to the claim that doubled NPs/DPs constitute adjuncts rather than arguments of the clause. The following two accounts are both characteristic of such a critique.

3.3.1.2 Anagnostopoulou, 1999; Terzi, 1999: Clitics as Functional Heads

Anagnostopoulou (1999) disputes Philippaki's (1977, 1987) claim as regards the adjunct status of doubled DPs by providing a number of arguments that argue against such a view. Anagnostopoulou argues that, if Philippaki is right, then doubled DPs in these constructions are right-dislocated phrases. To begin with, Anagnostopoulou notes that in contrast to cases of clear right dislocations there is no necessary intonational break separating the doubled DP and the rest of the clause. Furthermore, doubled DPs can occur in positions where adjuncts do not seem to be tolerated, e.g. subjects of Small Clauses and ECM complements:

(3.28) *O* *Janis* *δen* *ti* *θeori* *ti* *Maria*
the.NOM John.NOM NEG her.CL-ACC consider.3SG the.ACC Mary.ACC
eksipni
intelligent.ACC
'John does not consider Mary to be clever.'

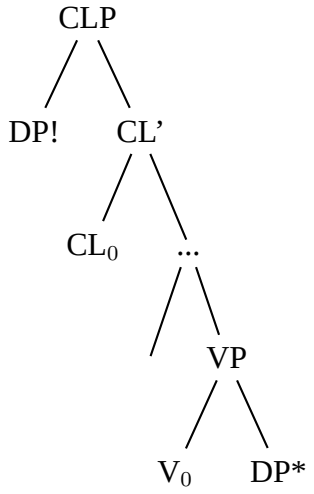
(3.29) *O* *Janis* *tin* *perimene* *ti* *Maria* *na*
the.NOM John.NOM her.CL-ACC expected.3SG the.ACC Mary.ACC SUBJ
paraponiete
complain.3SG
'John expected Mary to complain.'

Additionally, Anagnostopoulou (1999) claims that doubled DPs can appear before focused subjects, which are in general assumed to appear inside the VP:

(3.30) *Tin* *efaγe* *tin* *turta* *O* *JANIS*
her.CL-ACC ate.3SG the.ACC cake.ACC the.NOM John.NOM
'It was John that ate the cake'.

Anagnostopoulou, following Sportiche (1993), assumes that clitics are always base generated in pre-existing slots heading their own maximal projections, conventionally known as "clitic voices". Furthermore, clitics are not taken to be arguments of the verb. Contrary to Philippaki (1977,1987), the doubled DPs are considered to be the internal arguments. In the absence of a DP double, the internal argument is *pro*. The doubled DP moves to the SpecCL, establishing a Spec-Head relation between the double and the clitic, thus agreement between the two:

(3.31)



DP* is the base generated position of the DP while DP! is the position the DP moves to. Then, Sportiche (1993) identifies the following three clitic parameters:

(3.32)

a. Movement of DP* to DP! occurs overtly or covertly

b. Head (clitic) is overt or covert

c. XP* is overt or covert

The above parameters give rise to a number of different constructions. For example, assuming the Head is overt, DP* is covert and movement of the DP* to DP! is covert or overt, we get a situation where no doubling occurs while the clitic is phonologically present. Sportiche, argues that this option is actually taken up by languages like French. If the DP* is overt but its movement to DP! is covert, then a CD[8] construction in which the lower part of the chain created by covert movement of the DP* to DP! is pronounced is created. Specifically, in CD constructions the lower part of the chain created by covert movement of DP* to DP! is pronounced yielding the appropriate word order found in these constructions. Note that a direct consequence of the assumption that clitics do not constitute arguments of the verb is the postulation of a DP double in all cases even where no doubled XP exists. In non-doubling constructions, the DP* double is covert, moving to SpecCl where agreement is established via a Spec-Head relation.[9]

There are a number of things that are problematic under Anagnostopoulou's account. One of them has already been noted by Sportiche (1993: 32) and involves adjacency between the double and the clitic in constructions where an overt DP* moves overtly to DP!. If a Spec-Head relation is established between the double and the clitic, then, contrary to what we find in CLLD constructions, adjacency between the double and the clitic should be the case. The examples below from SMG show that adjacency is not required in CLLD:

(3.33) *Ton* *Jani* *ton* *ksero*
the John him.CL-ACC know
'I know John.'

[8]By the term clitic doubling, I refer to the Doubling construction whereby the clitic precedes the doubled DP. For a detailed discussion see chapter 6.

[9]In more detail, the clitic must establish a Spec-Head with an XP specified for [+F]. The property [F] is licensed in the XP by the clitic (Sportiche, 1993). These assumptions give rise to Sportiche's clitic criterion, according to which a clitic must be in a Spec-Head realtionship with an [+F]XP and vice versa.

(3.34) *Ton Jani δen ton ksero*
the John NEG him.CL-ACC know
'I do not know John.'

Ton Jani siγura δen ton ksero
the John surely NEG him.CL-ACC know
'Surely I do not know John.'

In order to side-step this problem Sportiche (1993) stipulates a filter similar to the doubly-filled COMP filter:

(3.35) *Doubly Filled Voice Filter*

*[$_{HP}$ XP[H...]]

Where H is a functional head licensing some property P and both XP and H overtly encode P

The relevant property is not specified (possibly case) but the above filter bans overt movement of an overt DP* to DP!. This is because assuming overt movement of the DP* to DP!, the filter proposed is violated since both the specifier XP (the moved DP*) and the H (the clitic) encode a common property P. Anagnostopoulou notes that such a filter does not really explain anything, noting that *"Obviously, a filter like that is stipulative; thus, the deep fact underlying the Clitic Doubling parameter remains unclear"* (Anagnostopoulou, 1999: 783).[10] However, no other account is proposed by her. Note that the account presented is based on the implicit assumption that clitics are bound affixes hosting functional projections. However, the arguments for the affixal status of clitics are not decisive and a number of good arguments to the contrary exist (Philippaki et al, 2004).

Another problematic part of Anagnostopoulou's analysis is the nature of the XP double (DP*) in case it is phonologically null (covert). Assuming the double is a *pro*, a number of difficulties arise, since this would immediately predict that object drop is in general possible in SMG contrary to fact. The same problem was noticed in Papangeli (2001),[11]

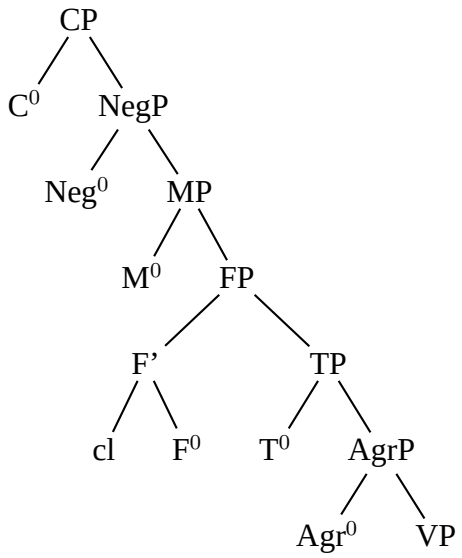
[10] See also Papangeli (2001) for a critique of the Doubly Filled Voice filter
[11] Even though formulated as a critique against Sportiche (1993).

which shows the problem in all its detail. On the other hand, if the covert double is not a *pro*, it is hard to see what else could this be.[12]

3.3.1.3 Terzi, 1999b: Functional Heads and Clitic Ordering

In her paper, Terzi (1999b) deals with free ordering of clitic clusters in SMG and CG. Terzi (1999b) following Sportiche (1993) assumes that clitics head their own functional projections (FPs) and are base generated in their surface position in indicatives. She argues that these so-called clitic heads are needed since multiple adjunction (in the case of a clitic cluster) to any verb related head would be problematic assuming Kayne's (1994) Linear Correspondence Axiom (LCA).[13] The following structure is assumed to be relevant for SMG in the presence of a single clitic:

(3.36)



The verb is assumed not to move past FP in finite constructions.[14] In non-finite constructions NegP is situated between CP and MP (Mood Phrase). This clausal structure then explains the ban on negative imperatives but the inclusion of negative gerundival constructions in SMG. Terzi further assumes that T^0 in imperatives and gerunds has defective

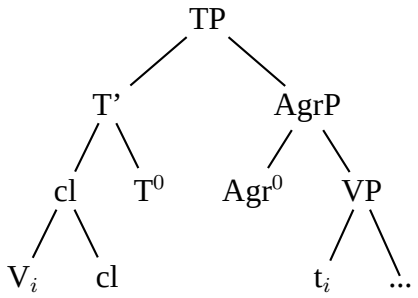
[12] Sportiche (1993: 25) himself argues against PRO on interpretational grounds.

[13] According to Kayne, multiple adjunction is banned since the adjuncted phrases symmetrically c-command each other, therefore they cannot be ordered for PF purposes. For more information see Kayne (1994: chapter 3).

[14] In particular, for SMG the verb is assumed to move as high as T^0 or Agr^0.

properties. The consequence of this is that the verb does not raise to adjoin to it overtly. In that respect, adjunction of the clitic to T^0 is licit according to the LCA (since no multiple adjunction configuration will be created) and thus no FP is needed to host the clitic in this case. The imperative verb moves to C^0, while gerunds to M^0 for checking reasons. It is at this point that a conflict with another established principle of the framework is created. Terzi is forced to assume that the imperative verb has to raise to C through movement to T^0 (after the clitic has adjoined to it) as a some kind of a 'by-product' of its movement to C^0 (or M^0 for gerunds), otherwise, i.e. if the verb directly moves to C_0, a Shortest Move Requirement (SMR) violation will incur.[15] Thus the verb (left-) incorporates into the clitic, hence the postverbal position of the clitic. What Terzi claims is that although multiple adjunctions are not possible before Spell-out they are licit at LF. In that respect she follows Chomsky (1995) rather than Kayne (1994), who assumes that the LCA applies to all levels. The structure after the verb has moved to adjoin to the clitic in T in imperatives is shown below:

(3.37)

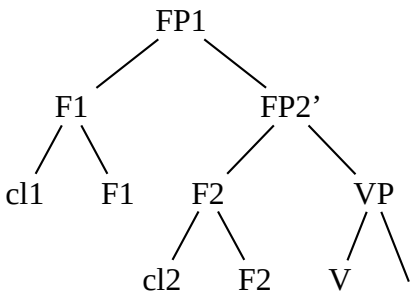


The verb plus clitic complex then moves to C in order for the verb to check its features. Recapitulating, the general idea in Terzi's argumentation is that in finite contexts the clitic cannot adjoin to T^0 due to the LCA. For this reason an independent abstract phrase is needed and this explains the preverbal word order. In non-finite contexts (imperatives/gerunds) because the verb does not need to raise to T^0 for feature-checking purposes before spell-out, there is no need for an independent projection to host the clitic which can be accommodated in T^0. Terzi then moves on and discusses constructions involving double object clitic clusters. Assuming the LCA, there are two possibilities for the clitics to appear

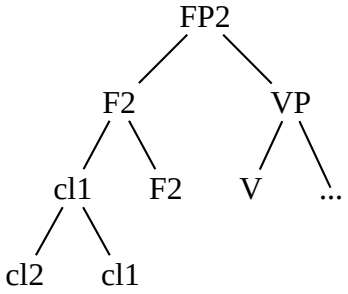
[15] SMR: A category moving to check feature(s) of a given type may not skip moving into an immediate relation with the closest c-commanding head which checks features of that type.

in the clausal structure. The first option involves the two clitics heading their own functional projection, i.e. each one attaching to a separate FP. The second option involves the two clitics adjoining to each other and then attaching as a cluster to an FP. The two relevant structures are shown below:

(3.38)



(3.39)



The relevant structure for SMG is argued to be the one in (3.38), where each clitic is accommodated in a separate functional projection. A further assumption made by Terzi (1999b) is that the dative clitic is located higher than the accusative one. What is rather idiosyncratic in this analysis is the fact that the two FPs do not have the same structural status. Whereas the two projections are independent from other verb-related elements when the verb is finite, in imperatives (or gerunds) the lower projection, FP2, is actually the TP so that, as explained earlier, the accusative clitic is supposed to adjoin to T^0. In case of ACC-DAT ordering in SMG imperatives the verb undergoes V-to-C movement. However, it cannot skip the FP2/TP due to the SMR. The verb checks its features in T^0 while left-incorporating into the accusative clitic, finally moving together to C, yielding the desired order. These assumptions are not adequate enough to explain the other order, i.e. DAT-ACC. For this case, Terzi employs the second possibility afforded by the LCA and argues that the clitics instantiate the second option where the dative clitic adjoins to the accusative one and

both move to C^0 as a complex (note that the relevant clitic projection, as in the previous case, is not FP but TP).

There are a number of issues raised regarding the account just sketched. Firstly, the stipulation of a new abstract 'last resort' type of functional projection FP, as in the case of Anagnostopoulou (1999), is rather ad hoc. There is no motivation (besides the LCA considerations) of why these functional projection should appear at all. Moreover, a major assumption on which this account builds is the defectiveness of case in imperatives. Such an assumption is however without justification and claims to the contrary exist (see Philippaki 1993, 1998; Philippaki et al., 2004). Furthermore, the nature of V-to-C movement is still quite unclear for the following reason: in the clausal structure Terzi (1999b) assumes for SMG (given in 3.36), a Mood Phrase (MP) is posited between Negation Phrase (NegP) and TP. This is the position where gerunds move according to Terzi. As its name suggests, this Mood Phrase should be some kind of locus of mood feature checking. If this is so, a plausible question to ask is why imperatives do not check their features in M but have to raise all the way up to C to do so. The answer is not given in Terzi (1999b). However, the reason behind such a move seems to be the avoidance of negated imperatives, since situating negation above M, movement of the imperative to M would predict negated imperatives to be licit contrary to fact. Thus, imperatival movement to C in the presence of M remains in my opinion a pure stipulation. What is furthermore left unexplained is the reason imperatives make use of both options for multiple clitic adjunction while in finite contexts only one option is possible. Lastly, similarly to the account proposed by Anagnostopoulou (1999), Terzi's analysis is also based on the implicit assumption that clitics are bound affixes rather than words, a claim that as already mentioned is not uncontroversial at all.

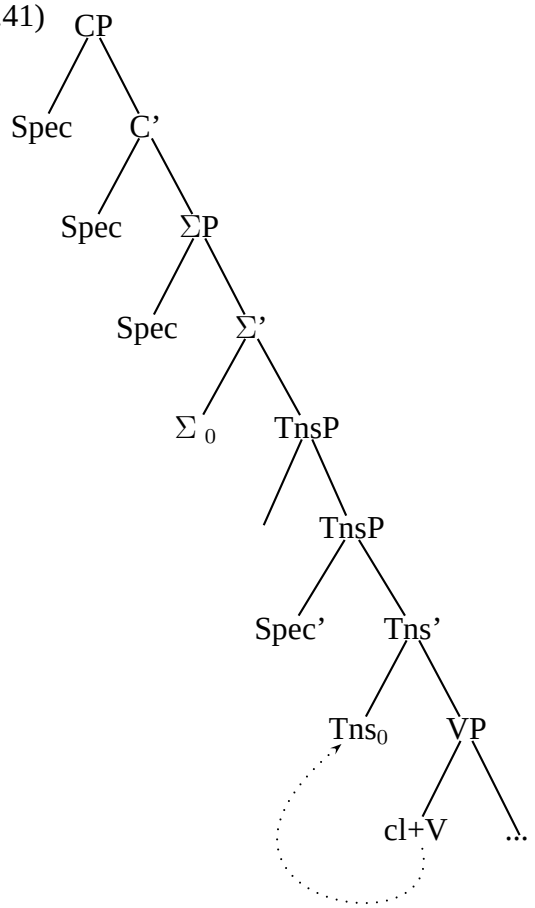
3.3.1.4 Condoravdi & Kiparsky, 2002: A Unitary Account of Clitics in Greek?

Condoravdi and Kiparky's paper is not a paper specific to SMG or GSG. It rather proposes an account which is argued to capture the whole range of clitic distribution found in MG dialects. Condoravdi and Kiparksy (2002) propose a tripartite classification of clitic systems for all MG dialects based on the status that clitics are assumed to have in each dialect:

(3.40) 1) Type A dialects: X^{max} clitics, syntactically adjoined to a maximal projection (Cappadocian Greek (CAG), CG).
2) Type B dialects: X_0 clitics, syntactically adjoined to a lexical head (PG).
3) Type C dialects: Lexical clitics, affixed words (SMG, GSG).

According to the above classification GSG and SMG clitics fall under category C. Clitics in that category, as argued by Condoravdi and Kiparsky (2002), are lexical affixes, word-to-word affixes to be more specific. The same phrase structure is assumed to underlie all MG dialects. Type C clitics, being word-to-word affixes, attach lexically to the left of a finite verb with subsequent head movement of the verb plus clitic cluster to Tns_0:

(3.41)



Condoravdi and Kiparksy (2002) adduce evidence that clitics are in fact agreement affixes in type C dialects by comparing them to subject agreement affixes. The latter are present only in finite verbal forms, a fact also true of clitics as Condoravdi and Kiparksy

(2002) argue. The exact reasoning is that clitics do not combine lexically with non-finite forms. However, this is not true since clitics are possible with non-finite verbal forms, namely imperatives and gerunds or only gerunds on the assumption that imperatives are actually finite. No matter the assumptions we make with respect to imperatives, such an analysis will give us the wrong results. Assuming that imperatives are non-finite verbal forms, we should explain why clitics are possible with these forms. Assuming they are finite verbal forms, the generalization that Condoravdi and Kiparsky (2002) propose, i.e clitics in C dialects attach lexically to the left of a finite verb, will not capture the imperative case. Regardless of what the analysis of clitics with non-finite forms is and regardless of what their stance on the imperative issue is, Condoravdi and Kiparsky will need two different assumptions to cover the whole range of phenomena. This will mean that clitics are affixes with finite verbal forms and something different with non-finite verbal forms, which seems rather unmotivated. Furthermore, there is no discussion about sequences of clitics and how more than one clitic lexically adjoin to the verb.

Summarizing, Condoravdi and Kiparksy (2002) propose an analysis that tries to capture the whole range of clitic phenomena found in the dialects of MG by classifying these into three major categories with respect to their linguistic status. However, their analysis does not go into the specifics of each dialect. As such, the peculiarities of each clitic system are not discussed. In our case, a number of issues regarding SMG and GSG clitics remain open.

3.3.1.5 Philippaki et al., 2002: Clitics as words revisited

Philippaki at al. (2002) attempt to reaffirm the Philippaki's (1977, 1987) analysis which, as we have seen, has been disputed by a number of researchers (Anagnostopoulou, 1999; Terzi, 1999b among others). Their analysis of clitics is based on the assumption that pronominal clitics in SMG have word-status. Specifically, they assume that clitics are phonologically deficient words. Following Kayne (1975) and Philippaki (1977, 1987), they claim that clitics provide the internal arguments of the verb, while the doubled DPs, when appearing, are functioning as adjuncts. In order to explain the differences between CD and CLLD constructions, they propose that DPs are adjoined to different phrase structure levels in these two constructions. Unlike left and right dislocations, which involve adjunction

at higher levels, CD and CLLD involve adjunction to vP and VP respectively. Following Chomsky (1995), according to whom clitics are both X_{max} and X_0, the clitic manages to move to I_0 in order to adjoin to the verb (or to the auxiliary *exo*, 'have', in perfect tense constructions) without violating HMC. The trace left after this movement operation makes postulation of a silent *pro* in this position unnecessary (unlike Sportiche, 1993 or Anagnostopoulou, 1999). However, the question that arises is how doubled DPs are assigned case when they are dislocated, given that case must have already been assigned to the clitic in argument position. To overcome this problem, Philippaki at al. (2002) assume that case of the doubled DP is assigned freely, following various proposals in the literature that argue for the free assignment of nominative case in SMG (Philippaki 1987; Tsimpli 1990; Alexiadou & Anagnostopoulou 1998). More specifically, they argue that if the assumption, according to which SMG can assign nominative case freely, is justified, then there is no reason not to assume that SMG can also freely assign accusative case. There is however, a far more general problem with respect to such an analysis. The account is entirely based on a pre-theoretical decision regarding the status of clitics as words. The word-like status of SMG clitics is by any means not fully justified and equally good arguments for the affixal status of SMG clitics also exist (notably Joseph, 1988, 2002 and Condoravdi & Kiparksy, 2002). In view of this, the whole spine of the account is based on a controversial assumption, a fact that considerably weakens its power.

3.3.1.6 Mavrogiorgos, 2006: Clitics within the Theory of Phases

In his paper, Mavrogiorgos (2006) argues that proclisis and enclisis are derived via exactly the same mechanisms. Clitics are assumed to be object agreement feature bundles (ϕPs in the sense of Cardinaletti and Starke, 1999). The ϕ-features present on the v* phase head probe the clitic, with susequent movement of the clitic to the v* head, motivated by an EPP feature in v*. Clitics merge with the relevant v* head projecting a specifier. In this Spec-Head relation, incorporation takes place. When movement of v* to T occurs, the clitic is moved along. Mavrogiorgos notes that finiteness is a crucial factor for the proclisis/enclisis alteration. It is argued that when person is not marked in T (argued to hold for imperatives and gerunds in SMG), enclisis takes place. The actual derivation of enclisis according to Mavrogiorgos is not a strictly syntactic phenomenon but rather a

PF/Morphological phenomenon. In that respect, he assumes that proclisis is always the basic syntactic derivation even in enclitic environments. Enclisis is then derived using the following mechanism: the clitic is first incorporated to v*. As a result of the latter, the verbal complex is assumed to be morphologically opaque, i.e. one element. When v* moves to T due to probing of T to v*, the verbal complex moves as a unit to T, in which stage the clitic is still proclitic to the verb. At that point, Mavrogiorgos argues that the T-v*-V complex is sent to PF where the v* along with the clitic is re-ordered in T's PF-slot, eventually giving rise to enclisis.

Even though the analysis presented in Mavrogiorgos (2006) tries to account for both proclisis and enclisis in a uniform way, a number of stipulations have to be posited in order for such uniformity to be reached. First of all, the assumption that SMG imperatives are non-finite, and as such non-person, which constitutes a major backbone in the analysis just sketched is not at all uncontroversial. There are a number of arguments in the literature that claim imperatives are in fact finite (Philippaki, 1993, 1998). However, even if one accepts that imperatives are non-finite, a major problem remains. What is rather unmotivated in such an account is why a re-ordering process takes place at PF. It is not clear why such a process should not be available with finite verb constructions. Furthermore, the claim that finiteness does not play a role in clitic systems like CG is not entirely correct (Mavrogiorgos, 2006: 51). The core purpose of this claim is to argue that enclisis in SMG is actually regulated by finiteness. Since both indicatives and imperatives pattern alike in CG, i.e. they both giving rise to enclisis, finiteness is argued not to be relevant for CG. However, the behaviour of CG imperatives and indicatives with respect to clitic positioning is not the same. Imperative verbs are strictly associated with enclisis, whereas indicatives can give rise to proclisis as long as a number of elements appear in the left periphery (see next chapter for a detailed list of these elements). For example, proclisis is impossible with imperatives even in the presence of a focused constituent in the left periphery (a classic case where proclisis would be obtained in a finite environment in CG):[16]

(3.42) *TORA* *(*mu)* *pe* *(mu)*
now me.CL-GEN say.IMP mu.CL-GEN
'Tell me now (and not any other time).'

[16]Capitals in the examples indicate contrastive focus intonation.

(3.43) *TUTO* *TO* *VIVLIO* *(*mu)* *δose* *(mu)*
this the.ACC book.BOOK me.CL-GEN give.IMP me.CL-GEN
'It is this book, I want you to give me.'

Another issue that is not mentioned in Mavrogiorgos' proposal is how the EPP in v* and T* are satisfied. Since enclisis involves the same derivation as proclisis, the clitic is probed by v* due to an EPP feature. The clitic is also assumed to satisfy the EPP feature of T in imperatives. The exact way in which this is achieved is something that is not discussed.

3.3.2 General remarks on GB/Minimalist Approaches

What seems striking regarding all the accounts presented is the need to pre-theoretically decide on the status of clitics as either words or affixes. There is a considerable ongoing debate on the issue (Joseph, 1988, 2002; Philippaki & Spyropoulos, 1999; Papangeli, 2000; Androulakis, 2001; Condoravdi & Kiparsky, 2001, Philippaki et al, 2002 among others), the arguments of which are summarized in Chatzikyriakidis (2006). The pre-theoretical decision one makes with respect to the status of clitics in both SMG and GSG will have immediate consequences on the type of analysis that is going to be pursued. For example, assuming that clitics have a word status will give rise to a movement approach where the clitic is base generated in argument position and subsequently moves to its surface position via movement. The nature of the account, i.e. a base-generated or a movement approach, will again dictate the type of approach with respect to the argument vs adjunct status of clitics. A movement approach generating the clitic in the argument position will immediately predict argument status for the clitic and adjunct status for the DP double, while a base generation approach will predict clitics to be adjuncts and the DP double an argument. Looking carefully at all these theoretical problems, one can notice that these do not constitute real problems but rather framework internal problems. The affix-word debate is important to decide on only in case the framework one makes use of makes a number of different assumptions and different rules available in each case. Indeed, this is the case for GB/Minimalism, thus the two lines of approaching the phenomenon, namely the movement vs base generation analyses. The adjunct vs argument debate is again a framework

dependent conception based on assumptions similar to the early GB θ-criterion principle.[17] However, working in a framework where no such assumptions are made, no such decision needs to be taken. Indeed this is the stance I am going to take, arguing that moving into a dynamic framework where parsing and incrementality are serioulsy taken into consideration, such pre-theoretical decisions can effectively be avoided. We will see how DS enables us to side-step the affix-word debate providing an analysis that is neutral to any stance on the debate, while on the other hand does not have to make an argument-adjunct distinction in doubling constructions. But before we move on and propose our DS analysis, let us first look at what researchers working within an alternative framework, that of HPSG, have assumed for SMG clitics.

3.3.3 HPSG Approaches

3.3.3.1 Alexopoulou, 1999

Alexopoulou (1999) attempts an analysis of SMG clitics within the HPSG framework. As in earlier HPSG approaches (Monachesi 1993, 1998a,b,1999 and Miller & Sag 1997, for Italian and French respectively), Alexopoulou claims clitics to be affixes rather than full words (see Alexopoulou, 1999 for the actual arguments). According to such an approach the verb plus the clitic will involve one lexical entry, since the two are assumed to form a single constituent. This constituent is of type *cl-wd* (clitic word) and is subject to a number of restrictions imposed by the general lexical entry given for *cl-wd* in Miller & Sag (1997):

(3.44) *Cl-wd* lexical entry (adapted from Miller & Sag, 1997)

$$\begin{bmatrix} \text{MORPH} \begin{bmatrix} \text{FORM} & \text{FPRAF}\ (\boxed{1}) \end{bmatrix} \\ \text{SYNSEM} \begin{bmatrix} \text{LOC—CAT} \begin{bmatrix} \text{VAL} \begin{bmatrix} \text{SUBJ} & \boxed{2} \\ \text{COMPS} & \boxed{3}\ \text{list(non-aff)} \end{bmatrix} \\ \text{ARG-STR} \begin{bmatrix} \boxed{2} \oplus \boxed{3} \bigcirc\ \text{nelist(aff)} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

In the lexical entry for *cl-wd* above, the ARG-STR list also contains elements which are not syntactically realized (aff). In that respect the ARG-STR list is the same as the

[17] To be more specific, based on the assumption that a given θ-role cannot be assigned more than once.

SUBJ and COMPS list combined, plus the arguments having the form of affixes if any.[18] The lexical entry for a verb plus clitic cluster is based on the restriction posited for the type *cl-wd.* As such, the clitic is not realized in the COMPS list but only on the ARG-STR list, on the assumption that it is an affix. The example shown below illustrates the lexical entry for the verb plus clitic complex *tin xtipise* 'her.CL-ACC hit.3SG':

(3.45) Lexical entry for *tin xtipise* 'her.CL-ACC hit.3SG'

$$\begin{bmatrix} \text{MORPH} & \begin{bmatrix} \text{FORM} & \boxed{3}\,\text{Tin-htipise} \\ \text{IFORM} & \text{xtipise} \end{bmatrix} \\ \text{SYNSEM} & \begin{bmatrix} \text{LOC—CAT} & \begin{bmatrix} \text{VAL} & \begin{bmatrix} \text{COMPS} & \langle \boxed{1} \rangle \end{bmatrix} \\ \text{ARG-STR} & \begin{bmatrix} \langle \boxed{1} \rangle \text{NP.NOM.3SG} \\ \langle \boxed{2} \rangle \text{NP.AFF.ACC.3SG} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

Moving on to clitic doubling, Alexopoulou firstly assumes that the doubled DP is an adjunct. A feature MOD (modifier) is used in order for adjuncts to choose their head (in doubling cases the head is the verbal cluster). Furthermore, the doubled DP is not encoded in the ARG-STR list, being an adjunct, but is however present in the COMPS list. There is a further feature, called DEP-STR (dependency structure), which encodes all the linguistic elements that are dependent on the word in discussion. The feature DEP-STR is the union of the features COMPS and ARG-STR. CLLD on the other hand is treated as argument extraction, with the adjunct being realized as a gap in the verb's lexical entry (see Alexopoulou, 1999: 176-177 for the exact formulation).

3.3.3.2 Alexopoulou and Kolliakou, 2002

Alexopoulou and Kolliakou (2002) discuss both topicalization and CLLD in Greek under an HPSG perspective. Their account of clitics departs from the one found in Alexopoulou

[18]The ◯ stands for the shuffle operation introduced by Reape (1994). The formal definition goes as follows (taken from Miller and Sag, 1995: footnote 20): Given a list A of length m and a list B of length n (disjoint from A) then A O B designates the family of lists of length m+n such that (1) the members of A ◯ B are the set union of the members of A and the members of B and (2) if X precedes Y in A or B, then X precedes Y in A ◯ B.

(1999), in that clitics in CLLD are not defined as gaps anymore. Instead, a new non-local featrure CL is used, subject to the same constraints governing all the other nonlocal features as regards feature propagation. The lexical entry for the non-canonical affix feature (see Alexopoulou and Kolliakou, 2002: 232 for the type inheritance tree) has a CL feature, which is a set that can be optionally but uniquely identified with the affix's HEAD value. This HEAD is taken to be of type noun, and thus specified for case and AGR (agreement). The lexical entry is shown below:

(3.46)
$$\begin{bmatrix} \textit{affix} & \\ \text{CAT—HEAD} & \boxed{1}\begin{bmatrix} \textit{noun} & \\ \text{CASE} & \text{c} \\ \text{AGR} & \begin{bmatrix} \text{PERSON} & \text{p} \\ \text{NUMBER} & \text{n} \\ \text{GENDER} & \text{g} \end{bmatrix} \\ \text{CLITIC} & (\boxed{1}) \end{bmatrix} \end{bmatrix}$$

Propagation of the CL value or the SLASH value in cliticless dislocated NP constructions is ensured via Ginzburg and Sag's (2000) Non-LOCAL Amalagmation Constraint (Alexopoulou & Kolliakou, 2002: 234) that defines the non-local features of a head as the concatenation of the non-local features of its daughters. Furthermore, two further constraints are posited with respect to SLASH and CL features. These involve the propagation of features from head daughter to mother. These two rules effectively override the Generalized Head Feature Principle (GHFP), according to which the SYNSEM of the head daughter node are by default the same as the SYNSEM of its mother (see Alexopoulou and Kolliakou, 2002: 236 for the actual rules).

3.3.3.3 Remarks on the HPSG analyses

The first thing one notices when looking at the way HPSG deals with syntactic phenomena is the highly lexicalized nature of the framework itself. Such a lexicalized nature is not a problem in itself. However, the way such a lexicalized framework is used to account

for certain phenomena by simply overloading the lexicon with pieces of information is in my view problematic. Frameworks like HPSG are extremely powerful (the mathematical formalism underlying HPSG is Turing-Machine equivalent in terms of formal power, see Carpenter, 1991) and the question that arises when faced with analyses that tend to overload the lexicon with arguably redundant information is whether such approaches show the ability of a given framework, or of a given account formalized in this framework, to provide an adequate account of the facts or rather the ability of the system to account for practically everything (since the expressiveness of HPSG can account for practically everything.). The next question that comes to mind within the context of having posed the previous question, is whether the accounts we have just seen are examples of lexicon overload. We have seen that clitics are analyzed as affixes and, as such, are part of the lexical entry of the entire verbal complex. This then means that someone will need a separate lexical entry for the verb and a separate lexical entry for the verbal cluster (verb plus clitic). Additionally, different lexical entries encoding all the different clitic forms that combine with the verb in question will be further needed (verb plus 1st person singular clitic, verb plus 2nd person plural clitic and so on.). Then, one also needs different lexical entries for the various clitic combinations of clitics with the verb plus a way to exclude impossible clitic combinations (e.g. PCC constructions, see chapter 6 in this thesis). Such an approach can hardly be called economical. Furthermore, the accounts just presented are not devoid of further stipulations. For example, it is rather unclear why clitics in Alexopoulou and Koliakou's (2002) account should be included in the ARG-STR but not in the COMPS list of the verb and vice versa for the doubling element in doubling constructions. What is puzzling in such an approach, is why an NP that has been attributed adjunct status appears in the COMPS list of the verb, i.e. the list encoding the complements of the verb. There are a number of other things that can be said regarding the technicalities of the HPSG analyses I have presented, but I will not discuss them here, since it is not my intention to further elaborate on these analyses or to decide whether minimalism or HPSG fairs better with respect to the clitic challenge. In that respect, I will stop the discussion here and move on to propose a DS analysis of clitics in SMG and GSG.

3.4 A DS Analysis of Clitics in SMG and GSG

3.4.1 Single Clitic Constructions

Being a lexicalized framework, DS assumes parametric variation in natural languages to reside in the lexicon. Therefore, one could plausibly point out that DS, like any other framework, is vulnerable to the affix-word debate, since a decision with respect to the status of elements like clitics will have further consequences on whether such elements will have their own lexical entry or rather be part of the lexical entry of a larger unit (the verbal complex in this case). However, this assumption is only superficially correct since in DS every element, no matter its traditional characterization (as a word or affix), can have its own lexical entry. The only requirement is that this element provides distinct procedural information on how the parse should/must or should not/must not proceed. This might sound controversial but becomes fairly straightforward assuming that for DS lexical entries are just pieces of information on how the parsing process proceeds. In that respect, every element providing such information can have its own lexical entry in DS. A classic illustration of that is the analysis of Japanese case suffixes put forth by Cann et al. (2005). In this analysis, Japanese case suffixes involve a separate lexical entry and are not part of the entry for the NP preceding these suffixes. Their role is to fix the locally unfixed node the preceding NP is parsed on, in order for multiple scrambling to be allowed (see Cann et al., 2005: 236-240 for details).[19] Given this view on lexical entries, the affix-word debate is largely irrelevant to DS, since any DS account can be compatible with any pre-theoretical decision on the affix-word debate. It is rather uncontroversial under this view that clitics will provide distinct procedural infromation with respect to the parsing process, thus involving their own lexical entries. This is the stance I am going to take in this thesis, following earlier approaches within the same framework notably Bouzouita (2002, 2008a,b), Cann et al. (2005), Chatzikyriakidis (2006, 2009a,b, forthcoming), Cann & Kempson (2008), Chatzikyriakidis & Kempson (2009), Gregoromichelaki (2010) among others.

I will begin our discussion on clitic positioning by examining the case of 3rd person

[19]This is due to the "no more than one unfixed node at a time constraint" discussed in chapter 1. See also chapter 6, this thesis.

accusative clitics in the two dialects. The first question that needs to be addressed before a lexical entry is given in DS, is what the triggering point of that entry will be, i.e. how the IF part of the algorithm specifying the conditions that need to hold on the current parse state in order for the word to be parsed is going to look like. NPs and strong pronominals in DS are treated as involving a type *e* trigger in their lexical entries (see Kempson er al. (2001); Cann et al. (2005); Bouzouita (2008a,b) and Chatzikyriakidis (2006, 2009a) among others). This allows NPs and strong pronominals to be parsed in the relevant *e* slots provided in the argument spine projected by the verb, or in a type *e* requiring unfixed node.[20] However, positing such a trigger for clitics will vastly overgenerate, since assuming the clitic is parsed on an unfixed node, fixing of the clitic outside the local domain will always be possible, contrary to the clitic facts in general (clitics must be interpreted locally). Such a key fact as regards the behaviour of clitics will then point to the direction of positing the initial type *t* requiring node as the triggering point for clitics. Indeed this is the stance taken by Cann et al. (2005), Chatzikyriakidis (2006, 2009a,b), Chatzikyriakidis & Kempson (2009) and Gregoromichelaki (2010) in giving lexical entries for SMG clitics. Following all these approaches, it will be assumed that the initial trigger for clitics in both GSG and SMG is the initial type *t* node. However, this restriction on its own is not enough to account for the distributional facts associated with SMG and GSG. In order to do that, we first need a way to ensure proclisis with non-imperatives verbs. There is a rather straightforward way to achieve this in DS. This involves an additional requirement in the entry for the clitic encoded as an embedded IF trigger that encodes that at the current parse state, no functor type has a type value (encoded as for all functor nodes, these have a type value), i.e. no verb of any type has already been parsed. The entry up to that point is shown below:

[20]The level of parametric variation with respect to NP positioning lies in the lexicon of a given language, constructive case and the availability of the *LOCAL ADJUNCTION rule being key factors with respect to word order in each language.

(3.47) Trigger ensuring proclisis

IF	$?Ty(t)$	
THEN	IF	$[\downarrow_1^+]?Ty(x)$
	THEN	...
	ELSE	abort
ELSE	abort	

Given the above entry, the clitic can be parsed only in case the pointer is at a type *t* requiring node and all functor nodes below this type *t* requiring node bear type requirements. Such a situation will be true only if a verb has not been parsed yet, since parsing of a verb will involve the projection of a type value in one of the functor nodes (which of the functor nodes will depend on the verb's adicity). The next thing we need to posit is the actual actions induced by the clitic, i.e. the THEN part of the entry. There are two types of analyses in the DS literature as regards the actions the lexical entries of clitics project. The first type of analysis assumes that the clitic builds a locally unfixed node and projects its information on that unfixed node while the second assumes that the clitic builds fixed structure and projects its information on a fixed node. The former type of analysis would predict clitics to be able to get fixed in more than one structural position in the local domain, while the latter would predict that clitics are associated with a fixed argument slot. The crucial question is which analysis is best suited for 3rd person accusative clitics in SMG and GSG. For GSG, things seem to be straightforward, since no double accusative verbs seem to exist and thus the assumption that 3rd person accusative clitics are always associated with the direct object can easily be sustained.[21] However, in SMG things are not that straightforward, since a limited set of verbs, the so-called double accusative verbs, subcategorize for two objects marked for accusative. The set includes verbs like *δiδasko, maθeno, kernao* 'teach, teach, buy':

(3.48) *O* *Γianis* *δiδaski* *aglika* *ti* *Maria*
the.NOM John.NOM teaches English.ACC the.ACC Mary.ACC
'John teaches Maria English'.

[21] Katsoyannou (1995) mentions that in the other main Grico variety, Calabrian Greek, double accusative constuctions are restricted to just one verb, the verb *matenno*, 'learn', and even in that case they are extremely rare (only one instance is found in her corpus).

(3.49) *O* *Γianis* *maθeni* *aglika* *ti* *Maria*
the.NOM John.NOM teaches Engish.ACC the.ACC Mary.ACC
'John teaches Maria English'.

(3.50) *O* *pateras* *kerase* *ta* *peδia* *paγoto*
the.NOM father.NOM bought the.ACC kids.ACC ice-cream.ACC
'The father bought ice-cream for the kids.'

In these constructions, the indirect object can also be a genitive or a PP construction as in regular ditransitives in SMG:

(3.51) *O* *Γianis* *δiδaski* *aglika* *sti* *Maria*
the.NOM John.NOM teaches Engish.ACC to-the.ACC Mary.ACC
'John teaches Maria English'.

(3.52) *O* *Γianis* *δiδaski* *aglika* *tis* *Marias*
the.NOM John.NOM teaches Engish.ACC the.GEN Mary.GEN
'John teaches Maria English'.

However, when substituting the full NPs with clitics a surprising restriction makes its appearance. It is impossible to substitute both NPs with two accusative clitics. In other words, a combination of two accusative clitics is illicit:

(3.53) **O* *pateras* *mu* *ton* *to* *δiδakse*
the.NOM father.NOM my him.ACC it.ACC taught.3SG
'My father taught it to him.'

(3.54) **O* *pateras* *mu* *me* *to* *δiδakse*
the.NOM father.NOM my me.ACC it.ACC taught.3SG
'My father taught it to me.'

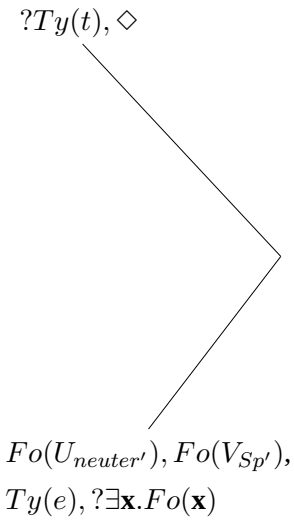
The only way to save the grammaticality of the above structures is to substitute the indirect accusative clitic with a genitive one:

(3.55) *O* *pateras* *mu* *tu* *to* *δiδakse*
the.NOM father.NOM my him.GEN it.ACC taught.3SG
'My father taught it to him.'

(3.56) *O* *pateras* *mu* *mu* *to* *δiδakse*
the.NOM father.NOM my me.GEN it.ACC taught.3SG
'My father taught it to me.'

The above facts, especially the exclusion of two accusative clitics, seem to imply that even in SMG accusative clitics receive a fixed position in the tree. Such an approach will immediately predict (3.53) and (3.54) to be ungrammatical, since parsing of the two clitics will result in a structure where both clitics occupy the same position. The parse can never be completed since the two clitics will involve incompatible formula metavariable presuppositions, making an update into a proper formula value satisfying the presuppositions of both clitics impossible:[22]

(3.57) Parsing *me to* on the direct object node



Double accusative verbs in SMG exhibit some further properties that distinguish them from regular ditransitives of the "give" class. As noted by Anagnostopoulou (2001), double accusative verbs can passivize indirect objects while failing to passivize direct objects

[22]The exact details of the presuppositions the different clitics will project will not be fleshed out here. It is however quite straightforward to understand that these preusuppositions will act as restrictors on the potential formula value substituends. Therefore, we need to ensure that a 3rd person clitic can never be updated by a formula standing for any of the discourse participants and vice versa. The presuppositions used in this thesis are only provisional and should not be taken as a serious attempt to deal with metavariable updating. As already said, I will encode the presusppositions associated with 3rd person clitics with a variable (e.g. x) whereas 1st/2nd clitics as Sp' and Hr' respectively. In case more precise information is needed as regards presuppositions further specifications might be provided for 3rd person clitics, e.g. *male*, *neuter* and so on.

contrary to what holds for regular ditransitives. The only way the direct object of a double accusative verb can be passivized is by marking the indirect object with genitive or by turning the NP into a PP headed by the preposition *se*, 'to', i.e. using one of the two strategies used in regular ditransitive constructions:

(3.58) *O* *Γiorγos* *δiδaxtike* *tin* *θeoria* *tis*
the.NOM George.NOM was.taught the.ACC theory.ACC the.GEN
sxetikotitas *apo* *ti* *Maria*
relativity.GEN from the.ACC Mary.ACC
'George was taught the theory of relativity from Mary.'

(3.59) **I* *θeoria* *tis* *sxetikotitas* *δiδaxtike* *ton*
the.NOM theory.NOM the.GEN relativity.GEN was.taught the.ACC
Γiorγo *apo* *ti* *Maria*
George.ACC from the.ACC Mary.ACC
'The theory of relativity was taught to George from Mary.'

(3.60) *I* *θeoria* *tis* *sxetikotitas* *δiδaxtike* *ston*
the.NOM theory.NOM the.GEN relativity.GEN was.taught to-the.ACC
Γiorγo *apo* *ti* *Maria*
George.ACC from the.ACC Mary.ACC
'The theory of relativity was taught to George from Mary.'

If only direct objects in ditransitive constructions can passivize in SMG, then the above data would be puzzling, since in the above constructions only the recipient (traditionally the indirect object) can be passivized when a double accusative construction is used (cf. 3.58 and 3.59). The theme can be passivized only in case a regular ditransitive strategy is used (i.e. a PP or a genitive, see 3.60). Given this data, and following Kordoni (2004) I argue that in double accusative constructions the recipient rather than the theme is the direct object.[23] Such an assumption is reinforced by the facts in (3.59) and (3.60), where the passivization of the theme is not possible when the recipient is marked for accusative while possible when the recipient is headed by the preposition *se* 'to' or marked for genitive case, i.e. when it conforms to the construction found in regular ditransitive clauses. Following the

[23]See Kordoni (2004) for more arguments. It is also worth noting that such an assumption is already implicitly made in Tzartzanos (1940).

reasoning just sketched, accusative clitics are always interpreted as direct objects, even in double accusative constructions. Assuming that the recipient is the direct object in double accusative constructions, the clitic always receives a fixed interpretation. Indeed, this is what is borne out from the facts:

(3.61) *Ton* *Γiorγo* *ton* *δiδaksa* *tin* *agliki*
the.ACC George.ACC him.CL-ACC taught.1SG the.ACC English.ACC
'I taught English to George.'

(3.62) **Tin* *agliki* *tin* *δiδaksa* *ton* *Γiorγo*
the.ACC English.ACC her.CL-ACC taught.1SG the.ACC George.ACC
'I taught English to George.'

Example (3.62) becomes grammatical (even though semantically infelicitous) if we assume the English language to be the recipient, e.g. in a possible world where languages are animate beings and human beings can be the subject of teaching:

(3.63) *Tin* *agliki* *tin* *δiδaksa* *ton* *Γiorγo*
the.ACC English.ACC her.CL-ACC taught.1SG the.ACC George.ACC
'I taught George to English.'

Thus, it seems that an analysis of SMG accusative clitics as receiving a fixed direct object interpretation seems plausible under the data presented. In that respect, the lexical entry I am going to propose for 3rd person accusative clitics in GSG and SMG will assume that the clitic builds the direct object node and decorates it with a type and a formula metavariable:

(3.64) Lexical entry for the 3rd person accusative clitic *to* 'it' in both GSG and SMG

IF $?Ty(t)$
THEN IF $[\downarrow_1^+]?Ty(x)$
THEN $\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle)$
$\texttt{put}(Ty(e), Fo(U_x)$
$?\exists \mathbf{x}.Fo(\mathbf{x})); gofirst(?Ty(t))$
ELSE abort
ELSE abort

The initial IF part combined with the embedded IF part, require a current parse state, in which the pointer is at a type *t* requiring node and below this type t requiring node, all functor nodes bear type requirements ($[\downarrow_1^+]?Ty(x)$). Such a triggering point will not only predict that a verb has not been parsed when the clitic comes into parse, but will further allow the clitic to be parsed after preverbal objects/subjects or modality/tense markers according to fact:

(3.65) *(*Ton)* *Θa* *ton* *skotoso* *(*ton)*
him.CL-ACC FUT him.CL-ACC kill.1SG him.CL-ACC
'I will kill him' [SMG]

(3.66) *Θelo* *(*ton)* *na* *ton* *δo* *(*ton)*
want.1SG him.CL-ACC SUBJ him.CL-ACC see.PNP-1SG him.CL-ACC
''I want to see him [SMG]

(3.67) *Telo* *(*ton)* *na* *(t)on* *do* *(*ton)*
want.1SG him.CL-ACC SUBJ him.CL-ACC see.PNP-1SG him.CL-ACC
''I want to see him [GSG]

As we see in the above examples, clitics are predicted to follow the subjunctive particle *na* or the future/modality marker *θa* . This is captured in the following manner: parsing of the two elements no matter the style of the analysis one follows, i.e. either a feature analysis where *na* and *θa* are just projected features on the initial type *t* node, or a situation node analysis in which both markers provide information in the situation node (a 0 node), this will still satisfy the restrictions posed by the clitic's triggering point. Assuming an analysis where these elements are just projected features on the type *t* requiring node, all functor nodes will bear type requirements (actually there will be no functor nodes so the universal statement will be trivially true), while assuming an analysis where *na* and *θa* provide tense and aspect information in the situation argument node (as I will propose in chapter 4), then again no functor nodes will be built by any of these two elements and thus again no functor nodes with type values will exist. Furthermore, cases where the clitic comes into parse after a number of preverbal NPs have already been parsed are also predicted to be possible, since the parse of any preverbal NP will always involve the introduction of an unfixed node or a LINK structure. The unfixed node case is trivially captured, since in case a preverbal

NP is parsed on an unfixed node no functor nodes will exist in the tree. On the other hand, assuming an NP has been parsed as a LINK (let us say in an HTLD construction), then a separate tree structure will be involved and thus again the statement $[\downarrow_1^+]?Ty(x)$ will be true for the LINKed tree, i.e. the main tree:

(3.68) After parsing a preverbal object

$$?Ty(t), Tn(a)$$

$$\langle\uparrow *\rangle Tn(a)),$$
$$Ty(e), Fo(\Gamma ior\gamma o'),$$
$$?\exists \mathbf{x}.Tn(\mathbf{x}), \Diamond$$

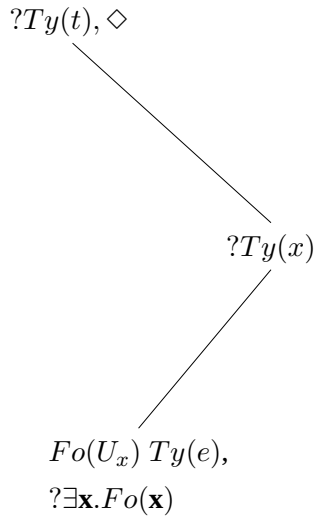
(3.69) After parsing a topic subject

$$\langle L\rangle Tn(0), \quad \langle L^{-1}\rangle Tn(n), ?Ty(t), ?\langle D\rangle Fo(\Gamma ior\gamma os'), \Diamond$$
$$Fo(\Gamma ior\gamma os'),$$
$$Ty(e)$$

In (3.68) the triggering point is satisfied and the clitic builds the 01 and 010 node and decorates the latter with a type value and a formula metavariable. This formula metavariable must be substituted later on via the context or via the natural language string itself. Notice that no bottom restrition is posited on this node. This will predict doubling to be possible in SMG and GSG, since MERGE of a full NP, which will have more structure than just the type *e* node (as sketched in the section 1), will be possible given that no bottom restriction will preclude it. However, such a restriction will work neatly for strong pronouns predicting that doubling with strong pronouns is not an option. The effect of parsing the clitic is shown schematically below:[24]

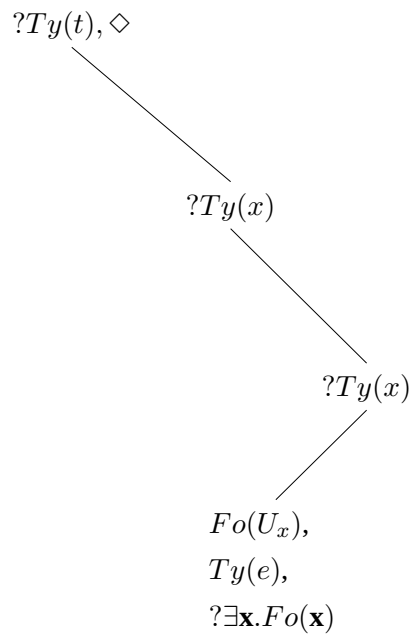
[24]Note that in the tree in (3.70) there is a $?Ty(x)$ statement in the 01 node that does not come from the entry for clitics. I assume that once a node is built, a type requirement is automatically put in that node. This **x** requirement might then later on be updated into a more specific type value requirement (e.g. *?Ty(e)*). This builds on the assumption that every node must end up with a type value by the end of the parsing process.

(3.70) After parsing the clitic



Before I proceed, it is worth noting that in the above structure the higher situation nodes are omitted for ease of exposition, since these are irrelevant for the kind of analysis pursued here. However, given that the higher situation argument nodes are always present given the newest developments in DS (Cann, forthcoming), one should have in mind that the direct object position will be lower than shown in case the situation nodes are present. This further means that the lexical entry for 3rd person accusative clitics must involve an additional application of the *make* and *go* function that will effectively build an additional 1 node. The lexical entry and the structure projected by the clitic given the presence of the higher situation nodes is shown below:

(3.71) The effect of parsing a 3rd person accusative clitic with the higher situation nodes present



(3.72) Lexical entry for the 3rd person accusative clitics in both GSG and SMG in the presence of the higher situation nodes

IF	$?Ty(t)$	
THEN	IF	$[\downarrow_1^+]?Ty(x)$
	THEN	$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$
		$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle)$
		$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle)$
		$\texttt{put}(Ty(e), Fo(U_x)$
		$?\exists \mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$
	ELSE	abort
ELSE	abort	

The situation nodes will be omitted for ease of exposition unless they constitute a vital part of the analysis. The lexical entries for fixed position clitics from now on will be formulated with one application of the functions *make* and *go* in bold parentheses, to indicate optionality of application depending on whether one assumes the higher situation nodes to be present or not.

What is now left is accounting for enclisis in imperatives. There is no proper analysis of imperatives in DS and I am not going to pursue any serious attempt to give an analysis of imperatives in general or imperatives in SMG and GSG specifically. For the moment, I will assume following Chatzikyriakidis (2006, 2009a,b), Bouzouita (2003, 2008a,b) and Chatzikyriakidis & Kempson (2009) that imperatives, besides the unfolding of propositional structure and projection of type and formula values, also project an imperatival feature *Mood(Imp)* in the initial type *t* requiring node that identifies them as imperatives. In that sense, accounting for enclisis in imperatives requires a second disjunctive trigger (encoded using the symbol | standing for inclusive disjunction) that will allow parsing to proceed just in case the current parse state has the pointer at a type *t* requiring node and an imperatival feature exists in that node.[25] The modified entry capturing enclisis with imperatives is shown below:

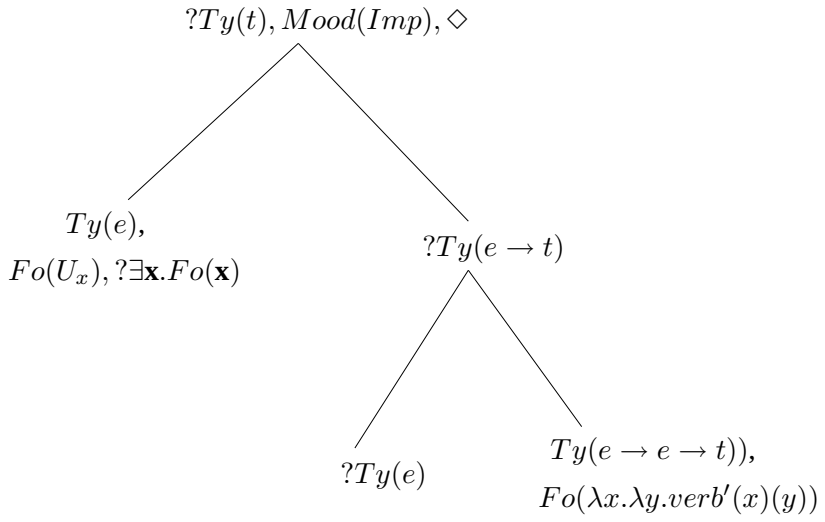
(3.73) Lexical entry for 3rd person accusative clitics in both SMG and GSG (imperatives included)

IF	$?Ty(t)$	
THEN	IF	$[\downarrow_1^+]?Ty(x)\vert$
		$Mood(Imp)$
	THEN	**(**$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$**)**
		$\texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle)$
		$\texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle)$
		$\texttt{put}(Ty(e), Fo(U_x), ?\exists\mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$
	ELSE	abort
ELSE	abort	

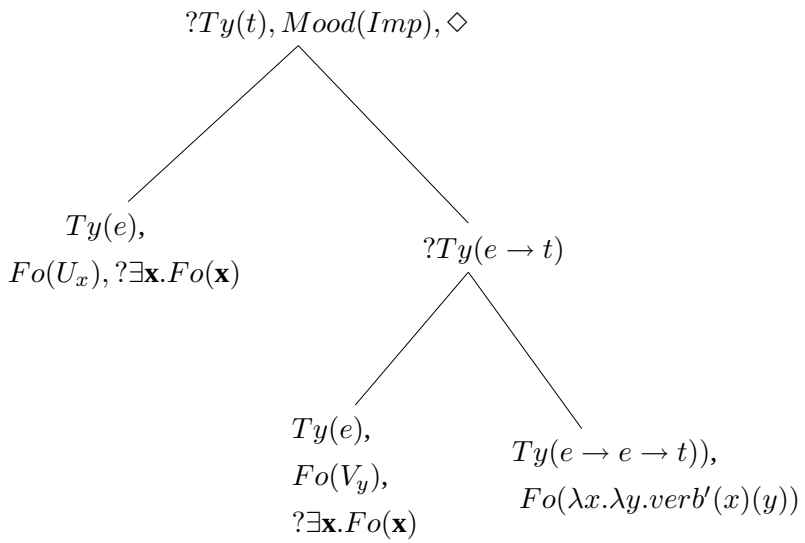
Assuming that an imperative verb has been parsed, a verbal type will exist in the 011 node. In that sense, the triggering point ($[\downarrow_1^+]?Ty(x)$) will not be satisfied. The algorithm then proceeds and checks the second trigger. An imperatival feature will be present in the type *t* requiring node, the trigger is satisfied and thus the algorithm proceeds to the THEN part:

[25]This imperatival feature *Mood(Imp)* should be taken as a shorthand for a proper analysis of imperatives that, as already mentioned, is pending in the system.

(3.74) Parsing a transitive imperative verb



(3.75) Parsing the clitic after the transitive imperative verb



The lexical entry given correctly captures the proclisis-enclisis alternation found in the two dialects. What is not blocked is a situation in which an imperative verb comes into parse after a clitic has already been parsed first. The account proposed captures enclisis with imperatives but however cannot block proclisis in the same environment. This is because assuming imperative verbs involve the same trigering restriction as indicatives (namely the initial type *t* requiring node), nothing will stop an imperative verb to get parsed

after a clitic has already done so. The clitic will leave the pointer at a type *t* requiring node, in which case nothing can stop an imperative verb from being parsed.

The solution we are going to propose in line with Chatzikyriakidis (2009a) is that imperative verbs have an additional restriction in their lexical entries that aborts in case any fixed nodes are present in the tree structure. In order to see the motivation behind such a triggering point, it is important to look briefly, before discussing it in more detail in chapter 7, at the development of the SMG (and arguably assume a similar development for GSG) clitic system from Medieval Mainland Greek (MMG).[26] MMG exhibits a clitic system where clitics in general follow the verb except when a number of elements appear at the left periphery, in which case proclisis instead of enclisis obtains. The elements comprising this list include functional elements like the subjunctive marker *na*, negation markers *δe(n)* and *mi(n)*, subordinating conjunctions, and a number of fronted elements at the left periphery (for a detailed list of triggers see Pappas 2004 and this thesis chapter 7). Imperatives in MMG, followed the same pattern, i.e. enclisis in neutral environments and proclisis in the presence of a proclitic trigger. The examples below exemplify the fact that proclisis was possible with imperatives in MMG:

(3.76) *Ala* *me* *ipe*
other.ACC me.CL-ACC say.IMP
'Tell me something else.' [*(Pappas 2004: 95)*]

(3.77) *Aγia* *tin* *ipe*
holy.ACC her.CL-ACC say.IMP
'Call her holy.' [*(Pappas 2004: 70)*]

The fronted constituent is the reason for proclisis in the above examples. On the contrary, enclisis with imperatives is categorical in both SMG and GSG. What I am going to propose is that the emergence of categorical restrictions in both SMG and GSG as regards the proclisis-enclisis alteration can be seen as the result of a routinization process (in the sense of Pickering and Garrod 2004) with the pragmatics atrophying over time within that process (Bouzouita 2008a,b). This routinization process involves the encoding of discource strategies into the area of syntax. Loose discourse strategies representing speakers'

[26]This term corresponds to the system described in Pappas (2004).

discourse preferences get stricter and are eventually encoded in the syntax through the lexicon.[27] For example looking at clitic placement in Koiné Greek (KG), Pappas (2006) notes, disagreeing with both Janse (1993) and Taylor (2002), that clitic positioning with a number of elements that are categorical (or near categorical) in MMG is variant between proclisis and enclisis. The following tables, adapted from Pappas (2006: 322) illustrate the latter fact:

(3.78) Pappas'(2006) table (slightly modified) for clitic positioning in the *Oxyrynchus Papyri* (vols. 1-56 - KG)

Environment	Preverbal	Postverbal
Clause initial	4	231
Infinitival complement	12	0
Adverbs	34	14
NP-object	21	15
NP-Subject	10	13
PP	13	37
Complementizers	17	37
Wh-expressions	18	16

[27] Discourse preferences or discourse considerations refer to a general effort minimizing process by which pronouns according to Cann & Kempson (2008), being anaphoric elements appear as early in the clause as possible. For example, for Latin weak pronouns, Bouzouita (2008a,b) argues that weak pronouns tend to appear after an emergent propositional domain has been signalled. See chapter 7, section 7.2.1 for a detailed discussion of such a proposal.

(3.79) Pappas (2004) table for MMG

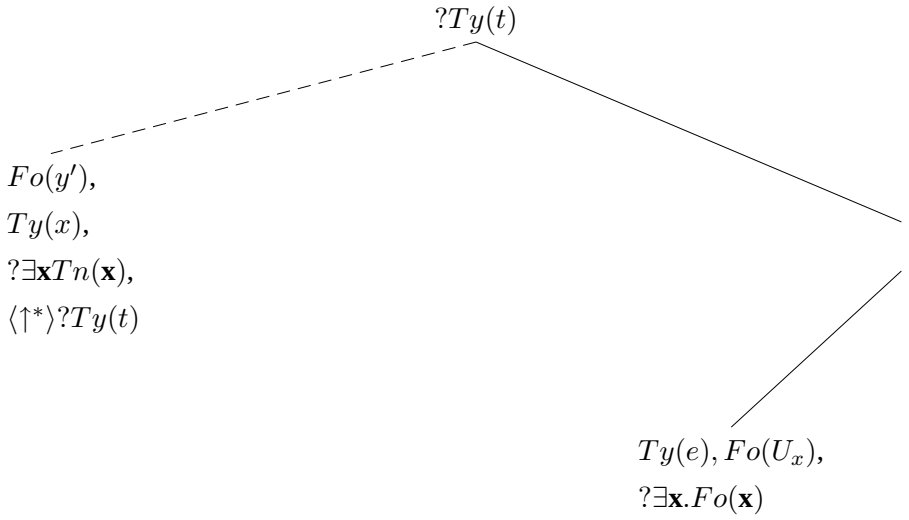
Environment	Preverbal	Postverbal
Clause initial	59	719
Coordinating conjunction	58	681
oti	8	30
δioti	4	0
Reduplicated object	39	79
Wh-elements	439	5
Negation	431	3
na,ina,as	1525	4
ean,an,pos	324	2
Object, PP, non-temporal adverb	898	90
Subject	334	130
Temporal expression	86	63
Imperatives	25	305

Comparing the KG with the MMG table, one notices a radical shift towards proclisis that tends to be or is categorical in a number of environments in MMG. For example, in the KG table, Wh-expressions seem to be balanced between proclisis and enclisis whereas in MMG proclisis is almost categorical in the same environment (proclisis: 439 tokens, enclisis: 5 tokens[28]). Thus, it seems that KG clitic positioning was variant in many environments, and proclitic environments were being triggered by discourse considerations and not by syntactic constraints. Once these discourse considerations became general calcified strategies and lost their pragmatic basis, speakers re-analyzed these discourse strategies as being syntactic. The end result is the MMG clitic system, where proclitic environments are encoded in the syntax. The next step towards the systems of SMG and GSG involves generalization of proclisis in all cases (see chapter 7 for the exact details). However, imperatives did not follow this direction. The reason for this is that generalization of proclisis was impossible to be achieved in imperatives, since most of the elements triggering proclisis in MMG are incompatible with imperatives. The only proclitic triggers

[28]The numbers are the numbers for Pappas' interrogative and relative pronoun category combined.

compatible with imperatives were fronted elements at the left periphery (objects/subjects, adverbs/temporal expressions). None of the function words, Wh-words or subordinate conjunctions were possible with imperative verbs. In that sense, linguistic input seems not to have been enough for such a generalization to occur. Thus, when the proclitic triggers collapsed and proclisis was generalized with finite verbs, imperatives followed the opposite route generalizing enclisis instead of proclisis. The collapse of proclitic triggers into a single trigger dictating proclisis with finite verbs, made proclisis with fronted constituents in imperatives impossible, since the trigger capturing proclisis with fronted constituents was not there anymore. In order for this exclusion of proclisis to occur, imperative verbs developed a trigger that would exclude proclitic environments. As we have already seen, fronted consituents were the only proclitic triggers possible for imperatives. In case a fronted constituent attracts the clitic to proclitic position, one ends up with the following partial tree after parsing both the fronted constituent and the clitic:[29]

(3.80) Parsing a clitic after a focused constituent

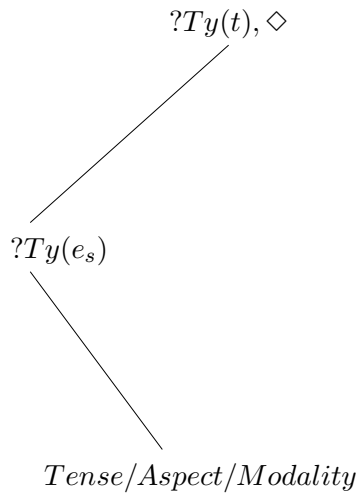


The assumption, as already said, is that at some point generalization of enclisis took place in imperatives, and such environments became exceptional enough to have the status of being avoided. In that respect, I assume that imperative verbs developed a restriction

[29]The type value of the fronted constituent is marked as *x*, since other elements with different semantic typing like adverbs or PPs can also be fronted. The exact range of the values x can take will not pursued here.

that blocks parsing of the imperative verbs in case a clitic has already been parsed. The triggering restriction proposed aborts in case any fixed node exists in the tree. In the above tree, the clitic projects two fixed nodes, so under the new triggering restriction, imperatives will not be possible in the above situation. The intuition behind this restriction is that imperatives must be the first element to project fixed structure in the tree in both SMG and GSG. This intuition will further produce a number of welcomed results with respect to structures that are incompatible with imperatives. First of all, imperatives will be predicted to be impossible after the subjunctive marker *na* and the future marker θa. Assuming an analysis of tense/mood markers where these build the situation argument node and provide tense and aspect information within the situation argument node complex (see next chapter for discussion and formalization), imperatives will not be able to get parsed after any of these markers has already been parsed first, since at least one fixed node will exist in the tree (the situation argument node):

(3.81) After parsing parsing a tense/modality marker (excluding formal detail)



The same holds for auxiliary verbs. Parsing an auxiliary will involve the projection of tense and aspect information in the complex situation argument node (see 3.5.3 for a detailed discussion of auxiliary verbs), and as such imperative verbs cannot be parsed after an auxiliary verb. The entry below exhibits the triggering restrictions of imperative verbs (the actions part is not formulated). It requires all nodes below the type t requiring one

to carry a requirement for a fixed node to be found, i.e. all nodes must be unfixed when imperatives come into parse:

(3.82) Lexical entry for imperatives

IF	$?Ty(t)$	
THEN	IF	$[\downarrow^{+}]?\exists\mathbf{x}.Tn(\mathbf{x})$
	THEN	...
	ELSE	abort
ELSE	abort	

Additional justification for such a trigger might come from the fact that negated imperatives are not possible in SMG or GSG. Negated subjunctives are used instead:

(3.83) **Mi* *δos* *mu* *to*
NEG give.SG-IMP me.CL-GEN it.CL-ACC

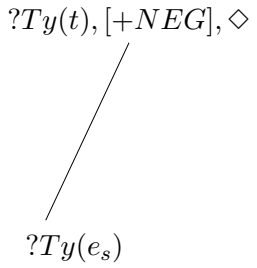
Mi *mu* *to* *δosis*
NEG me.CL-GEN it.CL-ACC give.2SG-PNP
'Do not give it to me.' [SMG]

(3.84) **Mi* *do* *mu* *to*
NEG give.SG-IMP me.CL-GEN it.CL-ACC

Mi *mu* *to* *doki*
NEG me.CL-GEN it.CL-ACC give.2SG-PNP
'Do not give it to me.' [GSG]

Assuming an analysis where negation builds the situation argument node and decorates it with a type e_s requirement, such an account account correctly predicts that negated imperatives should not be possible. This is because parsing of negation will have built a fixed node and in that sense the trigger we posited for imperatives will not be satisfied in case negation has already been parsed:

(3.85) After parsing *en*



Given the suggested entry for imperatives, examples like (3.14) repeated below will be predicted to be ungrammatical:

(3.86) **To grafe*
it.CL-ACC write.IMP-2SG
'Write it!' [GSG]

Thus, the account given correctly predicts the distribution of one clitic constructions. The next step is to give an analysis of clitic sequences in SMG and GSG.

3.4.2 Sequences of Clitics

One of the crucial differences between the clitic systems of SMG and GSG is the ordering inside clitic clusters in imperative environments. While SMG displays ordering flexibility in those environments allowing both DAT-ACC and ACC-DAT ordering, GSG follows the pattern found in all other MG dialects and allows only DAT-ACC ordering.[30] The relevant data are repeated below:

(3.87) *Δos mu to*
give.IMP me.CL-GEN it.CL-ACC
'Give it to me.' [SMG]

(3.88) *Δos to mu*
give.IMP it.CL-ACC me.CL-GEN
'Give it to me.' [SMG]

[30]There is no MG dialect, at least to my knowledge, that exhibits ordering flexibility in imperative environments besides SMG. It seems in that respect that free ordering in imperative environments is not a general characteristic of MG but an idiosyncrasy of SMG.

(3.89) *Do* *mu* *to*
give.IMP me.CL-GEN it.CL-ACC
'Give it to me.' [GSG]

(3.90) **Do* *to* *mu*
give.IMP it.CL-ACC me.CL-GEN
'Give it to me.' [GSG]

The first step in giving an analysis of clitic sequences is to see how genitive clitics can be approached in DS. Unlike accusative clitics, which were shown to receive a fixed interpretation in the tree structure, genitives are structurally underspecified, since besides their basic argumental function, i.e. indirect objects, these can further function as direct objects with a number of verbs like *milo, tilefono* 'talk', 'call':[31]

(3.91) *Tu* *milo*
him.CL-GEN talk
'I talk to him.' [GSG, SMG]

(3.92) *Tu* *tilefonisa*
him.CL-GEN called
'I telephoned him.' [SMG]

One might argue that genitive arguments in the above structures are not direct objects, as a number of diagnostics for distinguishing direct objects in Greek do not hold for these verbs. The basic diagnostic in distinguishing the direct object in Greek is passivization. Passivization is possible with direct objects but impossible with indirect objects. The verbs in question cannot passivize, alluding to the fact that they might not be direct objects:

(3.93) *O* *Γiorγos* *milise* *tu* *Γiani*
the.NOM George.NOM talked.3SG the.GEN John.GEN
'George talked to John'

(3.94) **O* *Γianis* *milithike* *apo* *to* *Γiorγo*
the.NOM John.NOM was-talked by the.GEN George.GEN

[31]There is of course the possessive use of genitive in both dialects, which is the function par excellence of genitives in Greek dialects. This aspect of genitives is not going to be dealt with in this thesis.

Note that the impossibility of giving an English translation for (3.94) points to the fact that passivization of the verb *talk to* in English, which allows indirect object passivization is also not possible. In that sense, the argument of *talk to* cannot be an indirect object either, at least in English and given that the passivization diagnostic is valid. The discussion is vast and it is not my intention to decide whether genitive single arguments are direct, indirect objects or something completely different. In DS terms however, constructing a dative clitic as the sole internal argument of a monotransitive verb will mean that dative clitics can occupy the structural position that in general direct objects occupy. But even if we assume that the sole arguments of these verbs are indirect objects and as such occupy the indirect object position (something highly implausible since there are no data pointing to that direction), there are constructions in which the genitive clitic acts as a possessive, benefactive/malefative or as an ethical dative that will not be captured under a fixed node analysis of genitives:[32]

(3.95) *Mu* *to* *arostisan* *to* *peδi*
me.CL-GEN it.CL-ACC made.sick the.ACC child.ACC
a. 'They made the child sick (and I'm worried/I'm upset about it).'
b. 'They made my child sick'.

(3.96) *Pai* *ce* *mu* *to* *δinei* *sto* *giorgo* *anti* *na* *to*
goes and me.-GEN it.ACC gives to-the George instead SUBJ it.ACC
δosei *se* *mena*
give to me
'He gives it to George instead of giving it to me. (and I'm angry with that)'.

Given that the PCC holds only between arguments, the fact that the PCC is active in all the above constructions, indicates that possessive, benefactive/malefactive or ethical dative clitics are interpreted as arguments (albeit optional) in SMG (see chapter 6 for an extensive discussion of the interaction between ethical datives and the PCC):

(3.97) **Mu* *se* *arostisan*
me.CL-GEN you.CL-ACC made-sick
'They made you sick (and I'm worried/I'm upset about it).'

[32]More on ethical datives in chapter 6.

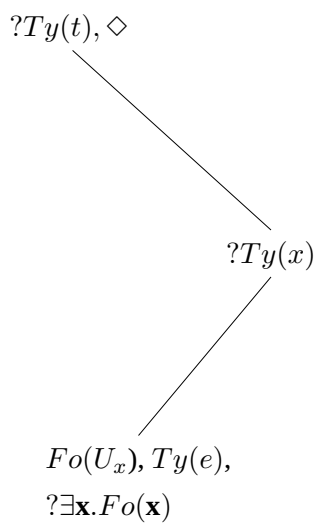
(3.98) **Pai ke mu se δinei sto giorgo anti na se*
goes and me.-GEN you.ACC gives to-the George instead SUBJ you.ACC
δosei se mena
give to me
'He gives you to George instead of giving you to me (and I'm angry with that).'

Thus, a fixed node analysis of genitive clitics will not work, since fixing the genitive clitic in a specific position will predict that genitive clitics must be always interpreted as indirect objects, contrary to fact. In that respect, it seems plausible to use underspecification for genitive clitics. However, we will have to be careful in choosing the kind of underspecification involved in order to avoid overgeneration. First of all, clitics are always interpreted in their local domain.[33] This means that encoding the rule of *ADJUNCTION in the entry for any kind of clitic will not work, since such a move will predict that clitics can be interpreted anywhere in the tree structure, thus violating locality. The perfect candidate that will avoid such locality violations but will still encode underspecification is the rule of LOCAL *ADJUNCTION which will posit that the clitic is underspecified locally, i.e. within the local domain. However, there is a further restriction associated with clitics in SMG and GSG that will not be captured even with the encoding of LOCAL *ADJUNCTION. This restriction states that clitics in MG in general can never be interpreted as subjects. In that respect, the rule of LOCAL *ADJUNCTION will still overgenerate, since it will predict that genitive clitics will be possible in the subject node contrary to fact. It is rather straightforward to treat this overgeneration. A simple modification in the modality will treat this overgeneration. What we have to do is change the Kleene star operator to the Kleene plus operator. The new modality will then become $\downarrow_1^+\downarrow_0$ instead of $\downarrow_1^*\downarrow_0$ ($\downarrow_1\downarrow_1^+\downarrow_0$ in case the situation nodes are present). This minor modification will ensure that the argument node where the clitic will eventually be fixed can be found if at least one step down the functor spine is taken. In case the clitic is fixed in the subject position, i.e. the 00 node, no such step must be taken, thus the incompatibiity of the proposed modality with the subject position. The next thing that needs to be taken care of is strict DAT-ACC ordering in indicative

[33]There is the phenomenon of Clitic Climbing that seems to violate this restriction. However, most of the analyses of CC assume a monoclausal rather than a biclausal structure to be involved in CC (Rizzi 1982; Cardinaletti and Shlonsky 2004; Cinque 2006; Chatzikyriakidis, 2010 among others). See next section for a discussion plus a DS analysis of CC.

environments in both SMG and GSG. Assuming a similar restriction as posited for imperative verbs, i.e. all nodes below must be unfixed, we get the right results since in case a 3rd person accusative clitic is parsed, two fixed nodes will exist, thus excluding ACC-DAT:

(3.99) After parsing the 3rd person accusative clitic



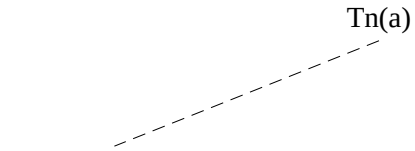
The presence of the 01 node will block parsing of the genitive clitic. In that way, ACC-DAT order is predicted to be impossible. With all these assumptions in place, the proposed lexical entry for genitive clitics is the one shown below:

(3.100) Lexical entry for genitive clitics in SMG and GSG (imperatives excluded)

IF $?Ty(t), Tn(a)$
THEN IF $[\downarrow^+]?\exists \mathbf{x}.Tn(\mathbf{x})$
THEN **(** $\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$ **)**
$\texttt{make}(\langle\downarrow_1^+\rangle); \texttt{go}(\langle\downarrow_1^+\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a));$
$\texttt{put}(Ty(e), Fo(U), ?\exists \mathbf{x}.Fo(\mathbf{x}));$
$\texttt{put}(?\exists \mathbf{x}.Tn(\mathbf{x}), gofirst(?Ty(t)))$
ELSE abort
ELSE abort

DAT-ACC ordering is on the contrary possible. Assuming that a genitive clitic has been parsed what we obtain is the following:

(3.101) After parsing a genitive clitic



$\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a),$
$?Ty(e), ?\exists\mathbf{x}.Tn(\mathbf{x}), \Diamond$

In the above partial tree, all functor nodes are fixed (basically there are no functor nodes, so the universal statement is trivially satisfied), thus the 3rd person accusative clitic can be parsed after a genitive clitic has already done so.

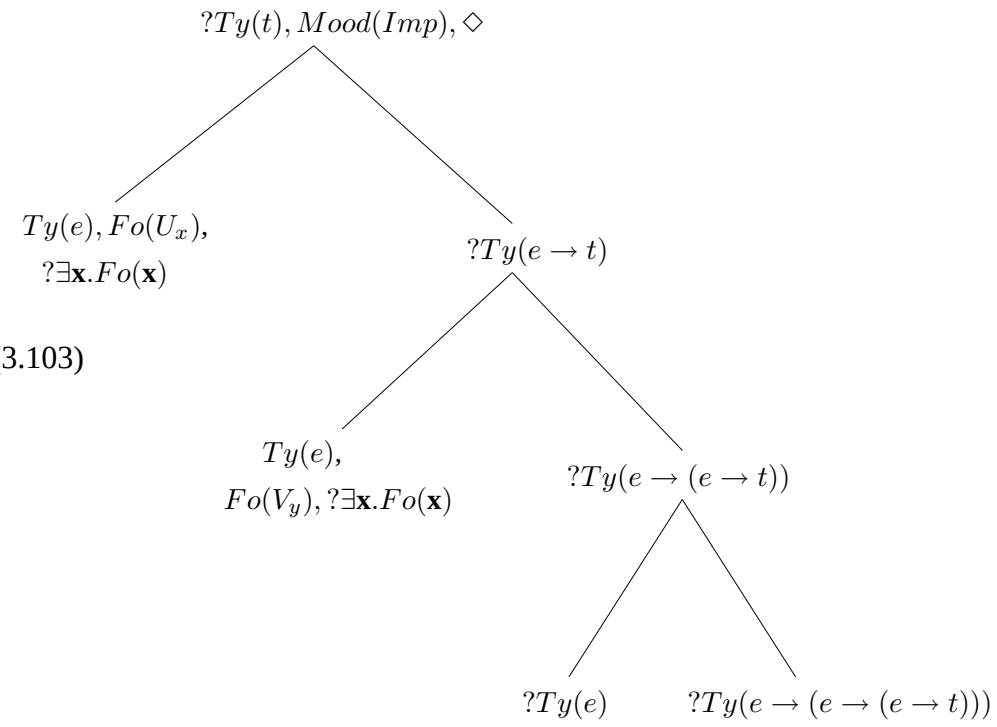
The next step is to look at ordering of sequences of clitics in imperatives. Remember that this is the point where the two dialects differ. On the one hand, SMG allows free ordering of the clitic sequence while GSG exhibits only the order DAT-ACC. For SMG, what we need is just a second disjunctive trigger encoding the presence of an imperatival feature as the condition for parsing the clitic. This condition will allow both orderings to be captured:

(3.102) Lexical entry for genitive clitics in SMG only (imperatives included)

IF $?Ty(t), Tn(a)$
THEN IF $[\downarrow^+]?\exists\mathbf{x}.Tn(\mathbf{x})|$
$Mood(Imp)$
THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1^+\rangle); \texttt{go}(\langle\downarrow_1^+\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a));$
$\texttt{put}(Ty(e), Fo(U), ?\exists\mathbf{x}.Fo(\mathbf{x});$
$?\exists\mathbf{x}.Tn(\mathbf{x})); gofirst(?Ty(t))$
ELSE abort
ELSE abort

Let us see how ACC-DAT order is captured in imperative environments. The accusative clitic is parsed after the imperative verb and fixes its position at the direct object node:

(3.103)



The genitive clitic is successfully parsed, since the $Mood(Imp)$ condition is satisfied. The same reasoning applies to the reverse DAT-ACC ordering.

Obviously, such an analysis will not work for GSG. For GSG, we need a way to exclude ACC-DAT ordering. In order to do that, we posit the restriction that the genitive clitic in imperatives should be the first fixed object. This restriction is a more specified version of the "first fixed node in the tree" restriction posited for imperatives and genitive clitic proclitic ordering. Such a restriction will abort in case the accusative clitic will be parsed first. However, assuming that the verb plus clitics are parsed as one chunk, constructions where an object has been parsed as an unfixed node and an imperative verb with a genitive clitic follow will be predicted to be possible, since the unfixed node will be able to fix its position in the tree only after the whole chunk has been parsed, i.e. after parsing the clitic.[34]

[34]The reason being that given parsing of the verb plus clitics as a chunk, application of general computational rules (in this case MERGE) will not be possible until lexical scanning of both the verb and the clitics has been done first. This assumption will prove to be crucial in order to capture the PCC with imperatives. See chapter 6 for a similar discussion.

In that respect, in case the genitive clitic comes into parse after an object in an unfixed node is already present, it will still be able to get parsed. The entry for GSG is shown below:

(3.104) Lexical entry for genitive clitics in GSG only (imperatives included)

IF $?Ty(t), Tn(a)$

THEN IF $[\downarrow^+]?\exists\mathbf{x}.Tn(\mathbf{x})|$
$Mood(Imp), [\downarrow_1^+][\downarrow_0]?Ty(x)$

THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1^+\rangle); \texttt{go}(\langle\downarrow_1^+\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a));$
$\texttt{put}(Ty(e), Fo(U_x), ?\exists\mathbf{x}.Fo(\mathbf{x});$
$?\exists\mathbf{x}.Tn(\mathbf{x})); gofirst(?Ty(t))$

ELSE Abort

Assuming that an accusative clitic has been parsed first, one of the object nodes will have a type value as in (3.103). This however will preclude a genitive clitic to be parsed since the restriction $[\downarrow_1^+][\downarrow_0]?Ty(x)$ will not be satisfied.[35]

3.4.3 Clitic Climbing

An idiosyncatic property of the GSG clitic system is the existence of obligatory CC with verbs other than auxiliaries. Specifically, GSG allows clitic climbing with two verbs of the restructuring class,[36] the verbs *sotzo* and *spitseo*, 'can' and 'finish' respectively:[37]

[35]Notice that ordering phenomena are dealt with inside the lexicon, i.e. by triggering restrictions in the lexical entries and not via general computational rules. This is a conscious decision, since I do not believe that some sort of universal clitic ordering exists. However, we would expect a number of other languages with similar clitic phenomena to be dealt with the same analysis. The interested reader is however directed to Manzini and Savoia (2004) for an interesting discussion plus references on the clitic ordering issue. Also see Miller & Sag (1997) and Grimshaw (2001) for a discussion against universal clitic ordering.

[36]Restructuring verbs include modal, motion and aspectual verbs.

[37]Notice that the infinitive *torisi* 'see' is introduced with an element, marked as a complementizer, i.e. *tse*. Unfortunately, I have not found any other verbs that are introduced with this complementizer. It is thus very difficult for me to propose an analysis of what *tse* actually is. However, the fact that it is marked as a complementizer should not be taken as an indication of a biclausal structure. Notice that similar elements in Italian *a,ad,di*, also induce climbing when they are used with apsectual or motion verbs (e.g. *andare, finire,*

(3.105)

a. *To* *sotzume* *avorasi.*
it.CL-ACC can buy.INF

b. **Sotzume* *to* *avorasi.*
can it.CL-ACC buy.INF

c. **Sotzume* *avorasi* *to.*
can buy.INF it.CL-ACC
'We can buy it.'

(3.106)

a. *To* *spitseo* *tse* *torisi* *avri*
it.CL-ACC finish.1SG COMP see.INF tomorrow

b. **Spitseo* *tse* *to* *torisi* *avri*
finish.1SG COMP it.CL-ACC see.INF tomorrow

c. **Spitseo* *tse* *torisi* *to* *avri*
finish.1SG COMP see.INF it.CL-ACC tomorrow
'I will finish seeing it tomorrow.'

It is crucial to note that these two verbs are the only ones of the restructuring class that still subcategorize for an infinitive. Such a fact is essential in understanding the unavailability of climbing with the rest of the verbs of the same class, since all the other restructuring verbs use the subjnctive marker *na* as a complementation strategy:

(3.107)

a. *Telume* *no(na-to)* *avorasume.*
want SUBJ-it.CL-ACC buy.1PL

cominciare). In that sense, one might say that *tse* should be treated accordingly. However, the lack of data do not allow any judgment at the moment. In the analysis proposed, *tse* will not be assumed to be associated with a biclausal domain. It is however certain that more data about *tse* should be gathered in order to have a clearer picture of its exact properties. Lastly, it should be also kept in mind that the term complementizer does not carry any framework dependent assumptions, like for example the presence of a full CP. Such notions are irrelevant to the account to be proposed or DS in general.

b. **To telume na avorasume.*
it.CL-ACCwant SUBJ buy.1PL
'We want to buy it.'

(3.108)

a. *Ancigneo na to toro*
begin.1SG SUBJ it.CL-ACC see.INF

b. **To Ancigneo na toro*
it.CL-ACC begin.1SG SUBJ see.INF
'I begin to see it.'

It seems that the unavailability of the infinitive strategy in the rest of the verbs is at least one of the reasons that climbing is only available for the two verbs of the class that subcategorize for an infinitive. Assuming monoclausality is what really lies behind CC constructions, the subjunctive strategy is incompatible with monoclausal interpretations since disjoint reference is always possible given that the verb following the subjunctive marker is finite. What other semantic or historical reasons may be behind this rather idiosyncratic property of GSG remains unresolved. It does not seem however that something semantic is really at play here. There is no generalization across classes and verbs similar in meaning (compare *spitseo* and *ancigneo* for example) behave differently with respect to climbing. Surprisingly, Romanian which like GSG uses the subjunctive strategy but retains the infinitive with a limited number of verbs, does not allow climbing except with the verb *a putea* 'can' (see Monachesi, 1998a):

(3.109)

a. *O pot vedea*
her.CL-ACC can see.INF

b. **Pot vedea o*
can see.INF her.CL-ACC
'I can see her.'

In what follows, I will try to provide an account of obligatoriness of CC in GSG. But before I do that, it will be good to take a brief look at some of the most prominent accounts as regards CC in general.

3.4.3.1 Approaching Clitic Climbing

The first thing one must take into consideration in giving an analysis of CC is what is the problem with such constructions, i.e. what makes them problematic to linguistic theory. The problem can be simply stated as a locality violation, with clitics being attached to a verbal host other than the one they constitute arguments of. A number of different accounts have been proposed over the years in different grammatical frameworks. Earlier approaches within the GB/Minimalist tradition assume that CC is a case of restructuring. A restructuring rule is used, transforming a biclausal structure into a monoclausal one. Such an approach can be found for example in Rizzi (1982), where a restructuring rule is posited to account for CC (see also Manzini (1983) for a similar treatment):

(3.110) Rizzi's restructuring rule - My formalization

V (P) V.INF → V.CMLX

The above rule transforms a biclausal structure consisting of a verb and an infinitive into a monoclausal structure where the two verbs are assumed to form a verbal complex (V.CMLX). However, the fact that in some CC languages a number of adverbs can intervene between the verb and the infinitive caused serious problems to such accounts. On the other hand, Kayne (1989) working within the barriers framework claimed that CC is the result of clitic movement out of the infinitival VP. The exact reasoning is then that in non-CC languages the VP consitutes a barrier for movement and thus movement is debarred, while in CC languages the IP L-marks the VP and the latter is not considered a blocking category anymore (in the sense of Chomsky, 1986).[38]

Recent approaches within the minimalist program propose that CC occurs when one of the verbs appears not as a fully fledged verb heading its own VP, but rather as an instantiation of an FP within the richly articulated FP structure of the clause proposed by Cinque (1999). In this sense, optional climbing is caused when one verb can be inserted either as a functional or a lexical verb:

[38]This crucially requires that IP in these cases is a lexical category since only lexical categories can L-mark according to Chomsky (1986). IP is not considered a lexical category in Chomsky (1986).

(3.111) Functional and lexical instantiation of a verb

a) [CP...[FP...[FP V_{restr}[FP...[VP V]]]]] Climbing case

b) [CP...[FP...[FP[VP V_{restr} [CP...[FP...[FP[VP V]]]]]]] Non-climbing case

Cinque (2001) and Cardinaletti & Shlonsky (2004) propose that a) is involved in CC constructions and b) in non-CC constructions. Cinque (2006) on the other hand argues that a) is involved in both cases.

Within HPSG, CC has been considered to be an argument sharing phenomenon (Miller & Sag 1997; Monachesi, 1993, 1998a,b, 1999 among others). All the HPSG analyses concur on the latter claim . The assumption is that the climbing inducing verb subcategorizes for an infinitive plus its arguments:

(3.112)
$$\begin{bmatrix} \text{HEAD V} \\ \text{VCLASS modal} \vee \text{aspectual} \vee \text{motion} \\ SUBJ\langle NP \rangle \\ \text{COMPS L} \oplus \quad \langle \text{V} \begin{bmatrix} \text{CLTS}\{\} \\ \langle \text{NP} \rangle \\ \text{COMPS L} \end{bmatrix} \end{bmatrix}$$

Argument sharing explains why the clitic can climb in CC constructions but does not however have anything to say with respect to restructuring effects found in CC environments like for example unavailability of infinitival negation when CC has taken place (examples below from Italian):

(3.113) **Lo vuole non vedere*
it.CL-ACC want NEG buy.INF
'I want to not see it.'

(3.114) *Vuole non vederlo*
want NEG buy.INF it.CL-ACC
'I want to not see it.'

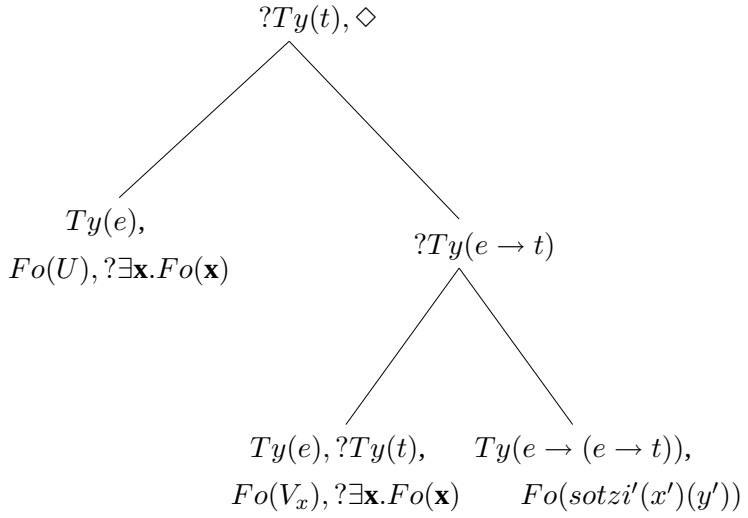
Furthermore, it is not clear what subcategorization for an infinitive plus its arguments means and furthermore why non-restructuring verbs are not able to do this kind of subcategorization.

In what follows I will present a DS analysis of CC, arguing that the functional-non-functional distinction assumed in the recent minimalist literature can receive better interpretation and formalization once a shift towards a parsing oriented framework has been done.

3.4.3.2 Clitic Climbing in DS

Given our account of clitics and assuming a biclausal structure is involved in CC constructions, the problem posed by CC in GSG, is that the clitic ends up projecting a type *e* value on the direct object node projected by the modal that also bears a type *t* requirement which represents the infinitival clause's type. The partial tree after the modal *sotzo*, 'can', in *to sotzo vorasi*, 'I can buy it', is parsed is shown below:

(3.115) Parsing *sotzi* in *to sotzi vorasi*



Notice that both a type *e* value and type *t* requirement exist on the direct object node. Such a partial tree cannot lead to a well-formed parse. The reason is simple: satisfying the type *t* requirement will lead to a situation where the node carries two distinct, incompatible type values, which is obviously not allowed by the system. Leaving the requirement unsatisfied will not do any better, since the outstanding requirement will remain in the tree. Since

a complete tree representing a successful parse, as already mentioned in the introduction to the framework, must not have any outstanding requirements, the parse will never be completed in case the type *t* requirement does not get satisfied. This is the situation we obtain if we assume that climbing inducing verbs are parsed like regular verbal complement verbs, i.e. when a biclausal structure is assumed to be involved in CC constructions. However, CC has been argued convincingly to involve a monoclausal rather than a biclausal structure (see Cinque 2006 for an extensive discussion). The next question that comes to mind is how DS can capture this fact, or more practically how this can be formalized in DS. The account I will propose assumes that climbing inducing verbs behave like auxiliary verbs in that they do not project a verbal type value, but rather project their semantics inside a complex situation argument node in the sense of Cann (forthcoming). According to Cann (forthcoming) English auxiliaries are taken to project a metavariable in the predicate node along with the relevant information as regards tense and aspect in the complex situation argument node. The same assumptions can be used to analyze auxiliaries *exo/ixa* 'have/had' in SMG.[39] In SMG these two auxiliaries are used to form the present and past perfect respectively. In the presence of a clitic in perfect constructions, climbing of the clitic before the auxiliary is obligatory:

(3.116)

a. *To* *exo* *δesi*
it.CL-ACC have.1SG tied.PST-PARTCPL
'I have it tied.' [SMG]

b. **Exo* *to* *δesi*
have.1SG it.CL-ACC tied.PST-PARTCPL
'I have it tied.' [SMG]

In line with Cann's (forthcoming) analysis for auxiliaries in English I assume that in SMG auxiliary *exo*, projects all the relevant semantic information in the complex situation argument node and thus no verbal type value needs to be projected from the auxiliary.

[39]There is a similar construction used in GSG involving the auxiliaries just mentioned and a past participle agreeing for gender case and number with the direct object of the verb. Such constructions, which are also possible in SMG, are not presented since, as shown in Iatridou (1991) for the SMG identical constructions, these do consitute cases of secondary predication rather than perfect tense structures.

However, unlike English which projects a formula metavariable (*Fo(U)*) and a type value containing a metavariable ($Ty(U \rightarrow (e_s \rightarrow t))$) in the 011 node, SMG auxiliaries do not project the 011 node at all (more on this in a bit). They do however project the subject node along with a formula metavariable and a type value. The reason for the latter move is that subject agreement in SMG is encoded in the auxiliary. Thus, all the relevant information that make pro-drop possible in SMG should be encoded in there. Auxiliaries eventually leave the pointer at the situation functor node as in the English case. The lexical entry plus the result of parsing *exo*, 'have', is shown below:

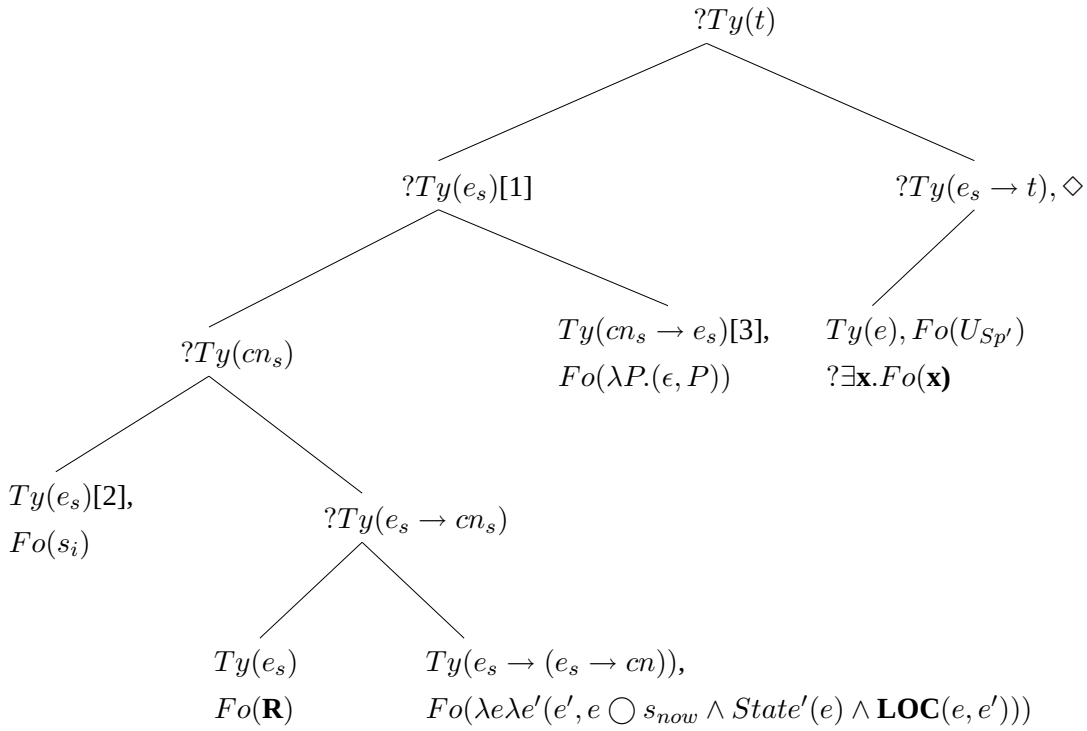
(3.117) Lexical entry for *exo* 'have.1SG' in SMG

IF $?Ty(t)$

THEN $\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(?Ty(e_s));$
$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \texttt{put}(Ty(cn_s \rightarrow e_s), Fo(\lambda P.(\epsilon, P)));$
$\texttt{go}(\langle\uparrow_1\rangle); \texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(?Ty(cn_s));$
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(Ty(e_s), freshput(s));$
$\texttt{go}(\langle\uparrow_0\rangle); \texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \texttt{put}(?Ty(e_s \rightarrow cn_s));$
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(Ty(e_s, Fo(\mathbf{R}));$
$\texttt{go}(\langle\uparrow_0\rangle); \texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \texttt{put}(Ty(e_s \rightarrow (e_s \rightarrow cn_s)));$
$\texttt{put}(Fo(\lambda e \lambda e'(e', e \bigcirc s_{now} \wedge State'(e) \wedge \mathbf{LOC}(e, e'))));$
$\texttt{go}(\langle\uparrow_1\rangle\langle\uparrow_1\rangle\langle\uparrow_0\rangle\langle\uparrow_0\rangle); \texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(Ty(e), Fo(U_{Speaker'}), ?\exists.\mathbf{x}.Fo(\mathbf{x}));$
$\texttt{go}(\langle\uparrow_0\rangle); \texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \texttt{put}(?Ty(e_s \rightarrow t))$

ELSE abort

(3.118) The effect of parsing *exo*



The auxiliary introduces both the reference time **R** and a fresh situation s_i in the type e_s [3] node. Furthermore, it introduces the tense/aspect specifications for the present perfect ($Fo(\lambda e \lambda e'(e', e \bigcirc s_{now} \wedge State'(e) \wedge \mathbf{LOC}(e, e')))$), where $\bigcirc$ stands for the overlap relation while **LOC** expresses an underspecified relation between the event and the reference points that enables the various perfect readings to be generated.[40] Additionally, the auxiliary projects the formula $Fo(\lambda P.(\epsilon, P))$ which will bind the situation that will emerge after combining the intemediate e_s [2] node with the $?Ty(e_s \rightarrow cn_s)$ node, given that the latter gets a type and a formula value. The rest of the structure projected is rather straightforward. Assuming this entry for auxiliaries and the entries we have given for clitics in SMG, climbing is predicted to be the only option with auxiliaries. Let us see why. Assuming a clitic has been parsed first, the pointer is left at the type *t* requiring node. This node is then the trigger for parsing the auxiliary (see lexical entry above) and the rest follows naturally. On the other hand, assuming that the auxiliary has been parsed first the pointer

[40] See Cann (forthcoming) for details and motivation of the **LOC** feature.

is left at the situation functor node. In case a clitic comes into parse, the parse will abort, since the initial triggering point of clitics, i.e. a type *t* requiring node, will not be satisfied given that the pointer will be at the situation functor node ($?Ty(e_s \rightarrow t)$). It should be noted that the pointer cannot move up via COMPLETION, since no type or formula will be satisfied in the situation functor node in parsing an auxiliary. Considering this, the only option for the clitic is to appear proclitic to the auxiliary according to fact. Furthermore, as already mentioned, the auxiliary *exo* 'have' in SMG, unlike its English counterpart (Cann forthcoming), does not project a formula value and a type value containing a metavariable on the 011 node. This is because if we assume a type value to be projected by the auxiliary in the 011 node, then VP ellipsis will be predicted to be possible with auxiliaries in SMG contrary to fact. Given the formula metavariable posited in the 011 node, the possibility of substituting this formula metavariable with a value provided by the context can give us a well formed sentence. This works nicely for English but will fallaciously predict VP ellipsis with auxiliaries to be possible in SMG. The examples below show the relevant facts for English and SMG:

(3.119) *Have you hit John? Yes, I have*

(3.120) *Exis xtipisi ton Jani? *Ne, exo*
have hit the$_{acc}$ John$_{acc}$ yes have.1SG
'Have you hit John? Yes, I have.' [SMG]

So far, so good. The question is how is such an account relevant to restructuring verb climbing? The answer is that an analogous account can straightforwardly be put forth for restructuring verbs by simply making the following two assumptions: a) climbing inducing verbs do not project any verbal type value and b) the semantics of restructuring verbs are projected in the complex situation argument node similarly to auxiliaries. Let us illustrate this claim using the GSG verb *sotzo* 'can'. Following Cann (forthcoming) I take modals to behave like auxiliaries in that they project their semantics in the situation argument node rather than projecting a verbal type. However, remember that modals are content verbs and more than tense and aspect information will be needed to capture their semantics. Fortunately, there is a way in which this can be done. Remember that aspect and tense

information are introduced in the $Ty(e_s \rightarrow (e_s \rightarrow cn_s))$ node of the complex situation argument node and percolate up to the $Ty(e_s \rightarrow cn_s)$ node, where they combine with the formula $Fo(s_i)$ provided by the e_s[2] node. Now, assuming that this situation ($Fo(s_i)$) can also be evaluated with respect to possible worlds we immediately get a solution to our problem. The only thing that we will have to further assume is that a 'world' parameter is also projected as part of a complex formula value involving a situation and a world parameter, both independently bound by the right operator of the epsilon calculus in each case. The assumption I am going to make is that modal verbs project such a complex formula value involving a situation and a world parameter on the type e_s[2] node. Then, the next step is the use of possible world semantics. For example, the lexical entry for *must* can be seen as specifying that the proposition expressed by the infinitive plus its arguments is true in all contextualy given possible worlds accessible from the default world.[41] The same can be argued to be the case for the ability modal *sotzo* 'can', the difference being that the domain of quantification in this case is ability worlds, a subset of the set of possible worlds rather than just possible worlds. *Sotzo* under such an approach, will project a complex *Fo* value in the e_s[3] node, encoding both a situation and a world parameter ($Fo(s_i, w_i)$). This world parameter must be a member of the set of ability contexts which in turn are a subset of the set of contextually accessible worlds.[42] The next thing we need to take care of is the form of quantification quantifying over these possible worlds. Since we are dealing with the set of all ability contexts, what we need is universal quantification. In that respect, we posit a tau term, instead of an epsilon term to capture the universal quantification properties of the world parameter projected by *sotzo*. Furthermore, tense/aspect specifications are also going to be included. Under the account just sketched, the only difference between modals and auxiliaries is that the former introduce a complex *Fo* value including both a situation and a world parameter in the e_s[2] node, whereas the latter projects only a situation parameter as

[41] A fact already suggested to me by Ronnie Cann (p.c).

[42] This is what Kratzer (1981, 1991) calls the modal base. However, as it is shown by Kratzer (1991) the modal base in itself is not enough to capture the semantics of modality (at least deontic modality) right and for that reason she furthermore used a partial ordering of the modal base *W* with respect to *w* depending on the closeness of each world in *W* to the default world *w*. I will not get into these details in this thesis. The reader should however have in mind, that such additions do not affect the account given and can be fairly easily formalized in DS. Fom more information on modal bases and the ordering source see Kratzer (1981, 1991). See also Zvolensky for a discussion on Kratzer (1981, 1991) and alternatives.

part of the *Fo* value. One further difference between the two is the node where the pointer is assumed to be left. We have seen that the pointer is left at the $e_s \rightarrow t$ node after an auxiliary is parsed. However, leaving the pointer at the same node in the case of restructuring verbs will predict that infinitives have two distinct parsing triggers, a type *t* and a type $e_s \rightarrow t$ requiring trigger. The type *t* requiring trigger is independently needed for constructions where the infinitive functions as the complement of a regular complement taking verb. In that case, and assuming that the complement taking verb will decorate the direct object node with a type *t* requirement and will leave the pointer there with no other nodes existing below that node, the trigger for the infitive must be the type *t* node. In order to avoid redundancy, I posit that the trigger for infinitives is a type *t* requiring node in all cases. [43] Lastly, I further assume that restructuring verbs in GSG further project the predicate node, in contrast to SMG auxiliaries and similarly to English auxiliaries and modals. The reason for this is that VP ellipsis is possible with restructuring verbs in GSG:

(3.121) *To* *sotzi* *vorasi?* *Ne,* *sotzi*
it.CL-ACC can.3SG buy.INF yes can.3SG
'Can he buy it? Yes, he can.' [GSG]

Thus, I further assume that the predicate node (011) is also projected via the entry of *sotzo* 'can.1SG'. This predicate node bears a type value and a formula metavariable. Formula metavariable substitution from context explains the grammaticality of structures like the above. Putting all these assumptions together one gets the lexical entry for *sotzo* 'can.1SG' shown below:

[43] A further welcomed result is that such treatment can further enable us to distinguish between infinitives and past participles without actually referring to any other of their properties. For example, under such an analysis, an infinitive will always be impossible after an auxiliary has been parsed first, since the pointer in that case will be at the situation functor node.

(3.122) Lexical entry for *sotzo*.1SG, 'can' in GSG

IF $?Ty(t)$

THEN $\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(?Ty(e_s));$

$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \texttt{put}(Ty(cn \rightarrow e_s), Fo(\lambda P \lambda R(\epsilon, P, \tau, R)));$

$\texttt{go}(\langle\uparrow_1\rangle); \texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(?Ty(cn_s));$

$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(Ty(e_s), freshput(w_i, s_i)); \texttt{go}(\langle\uparrow_0\rangle);$

$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \texttt{put}(?Ty(e_s \rightarrow cn_s));$

$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(Ty(e_s), Fo(R));$

$\texttt{go}(\langle\uparrow_0\rangle;)\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \texttt{put}(Ty(e_s \rightarrow (e_s \rightarrow cn_s)), Fo(\lambda.e(\lambda.e', \phi)$

$(e', e' \subseteq e \wedge e = snow) \wedge \phi \in W_{ab} \in W));$

$\texttt{go}(\langle\uparrow_1\rangle\langle\uparrow_1\rangle\langle\uparrow_0\rangle\langle\uparrow_0\rangle); \texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \texttt{put}(?Ty(e_s \rightarrow t));$

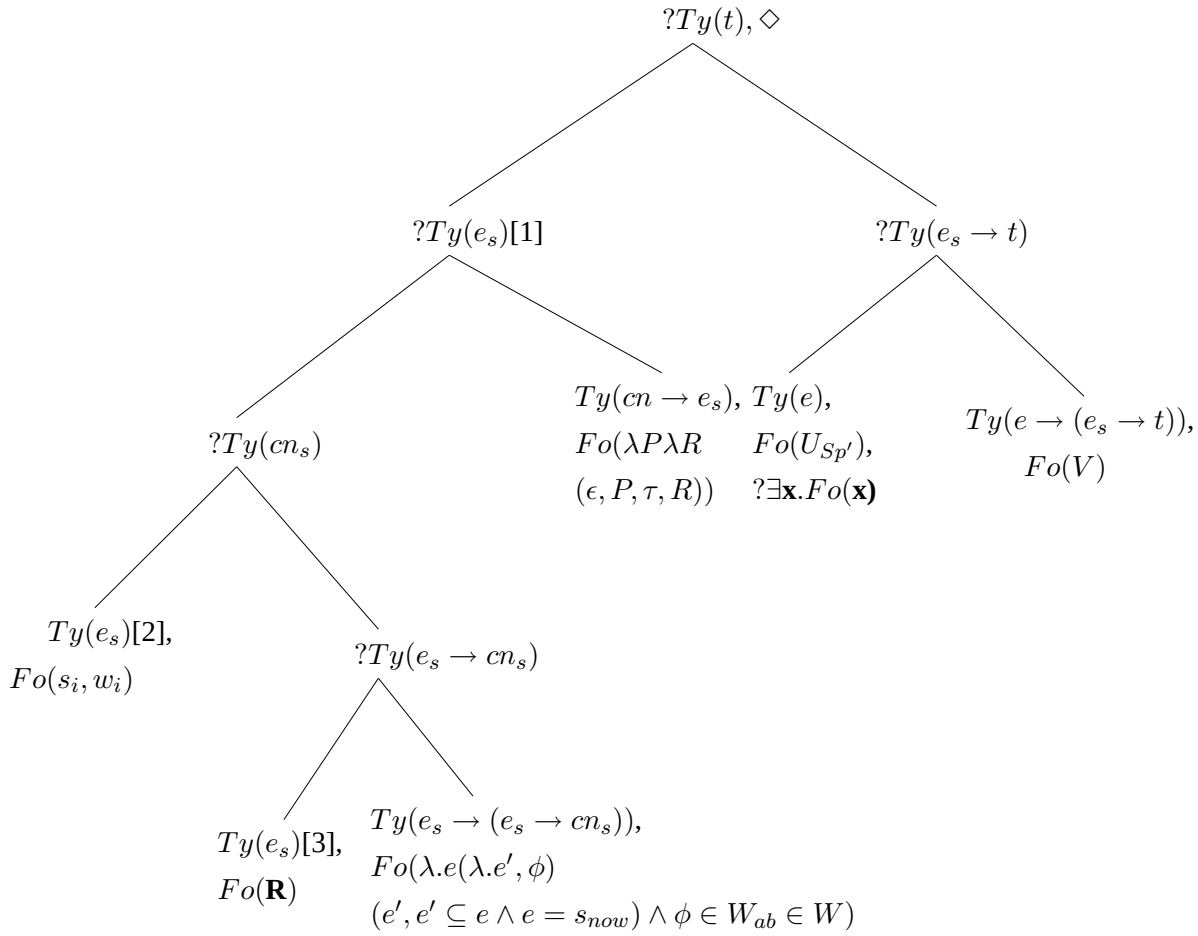
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(Ty(e), Fo(U_{Speaker'}), ?\exists.\mathbf{x}.Fo(\mathbf{x}));$

$\texttt{go}(\langle\uparrow_0\rangle)\texttt{make}(\langle\downarrow_1\rangle)\texttt{go}(\langle\downarrow_1\rangle); \texttt{put}(Ty(e_S \rightarrow (e_s \rightarrow e_s)), Fo(V)), gofirst(?Ty(t))$

ELSE abort

The result of parsing *sotzo* is shown below:

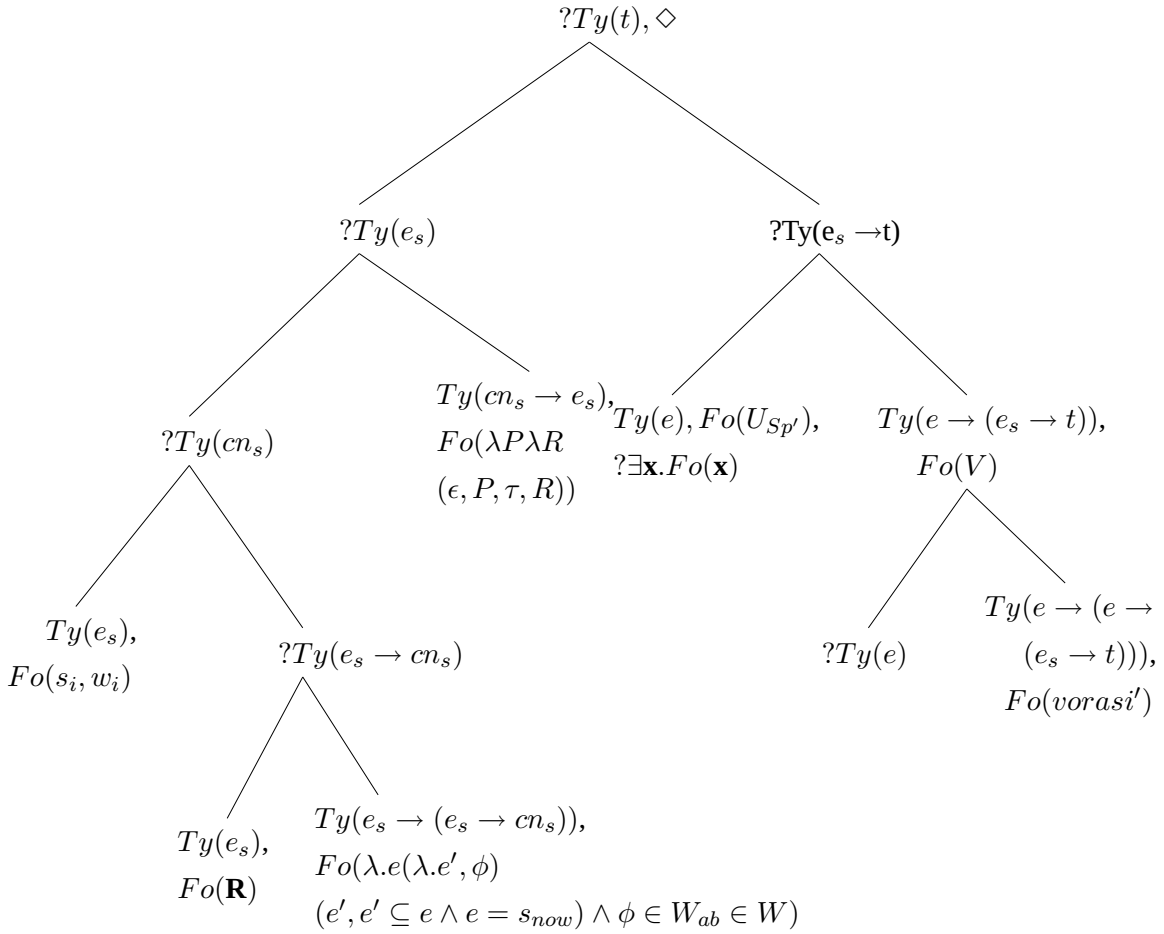
(3.123) Parsing *sotzo*



The intermediate Ty(e_s) [2] node has a complex formula value introducing both a situation and a world parameter ($Fo(s_i, w_i)$). The lowest functor node on the other hand ($Ty(e_s \rightarrow (e_s \rightarrow cn_s)$) contains three lambda bound variables ($Fo(\lambda.e(\lambda.e', \phi)(e', e' \subseteq e \wedge e = snow) \wedge \phi \in W_{ab} \in W)$). The first variable (*e*) is to be substituted by the reference time metavariable **R**. Then, the other two (*e'* and ϕ) stand for the two variables that will be substituted by the two parameters of the complex situation argument, s_i and w_i. In that sense, *e'* is taken to hold at a time *e* (**R** after substitution), where time *e* is the same as the utterance time s_{now}, while ϕ is taken to belong to the set of ability contexts, which in turn are a subset of the set of contextually accessible worlds *W* ($\phi \in W_{ab} \in W$). Then, at the $Ty(cn_s \rightarrow e_s)$ node the form of quantification for each of the arguments is introduced. The

situation parameter is associated with existential (ϵ) while the world parameter with universal (τ) quantification ($Fo(\lambda P \lambda R.(\epsilon, P, \tau, R))$). The subject node is further projected (010 node) and a type value and a formula metavariable are posited in the same node. Lastly, the predicate node is also built and decorated with a type value and a formula metavariable. The pointer is left at the type *t* requiring node. Given the structure projected by *sotzo* in (3.123), no clitic can be parsed after *sotzo* has already done so. The trigger for accusative clitics ($[\downarrow_1^+]?Ty(x)$) will not be satisfied because a functor node will bear a type value (the predicate node), while the trigger for genitive clitics ($[\downarrow^+]?\exists \mathbf{x}.Tn(\mathbf{x})$) will not be satisfied given that a number of fixed nodes will exist after parsing *sotzo*. Parsing of the clitic after the infinitive, is not possible for the same reasons. The infinitive projects the verbal type and formula value as well as the direct object node and returns the pointer to the type *t* requiring node (full lambda formulas in the functor nodes are omitted for reasons of space):

(3.124) Parsing *vorasi* ‘to buy’ in **sotzo vorasi to* ‘I can buy it’

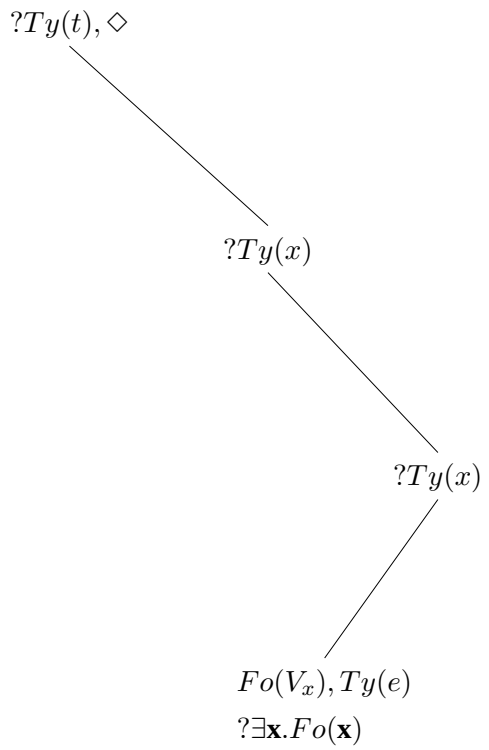


It is at that point that the clitic comes into parse. The lexical entry for 3rd person accusative clitics specifies that no functor node with a type value must exist in order for the parsing process to proceed ($[\downarrow_1^+]?Ty(x)$). However, this is not true in this case, since two functor nodes with type values exist (the predicate and the transitive predicate node, 011 and 0111 nodes respectively). On the other hand, genitive clitics cannot be parsed also given that a number of fixed nodes exist and therefore their triggering restriction is also not satisfied ($[\downarrow^+]?\exists \mathbf{x}.Tn(\mathbf{x})$). Thus, CC is the only option in the presence of *sotzo* in GSG. In a CC case like (3.125), the clitic is parsed first, building and decorating the direct object node with a type value and a formula metavariable:

(3.125) *To* *sotzo* *vorasi*
it.CL-ACC can.1SG buy.INF

'I can buy it.'

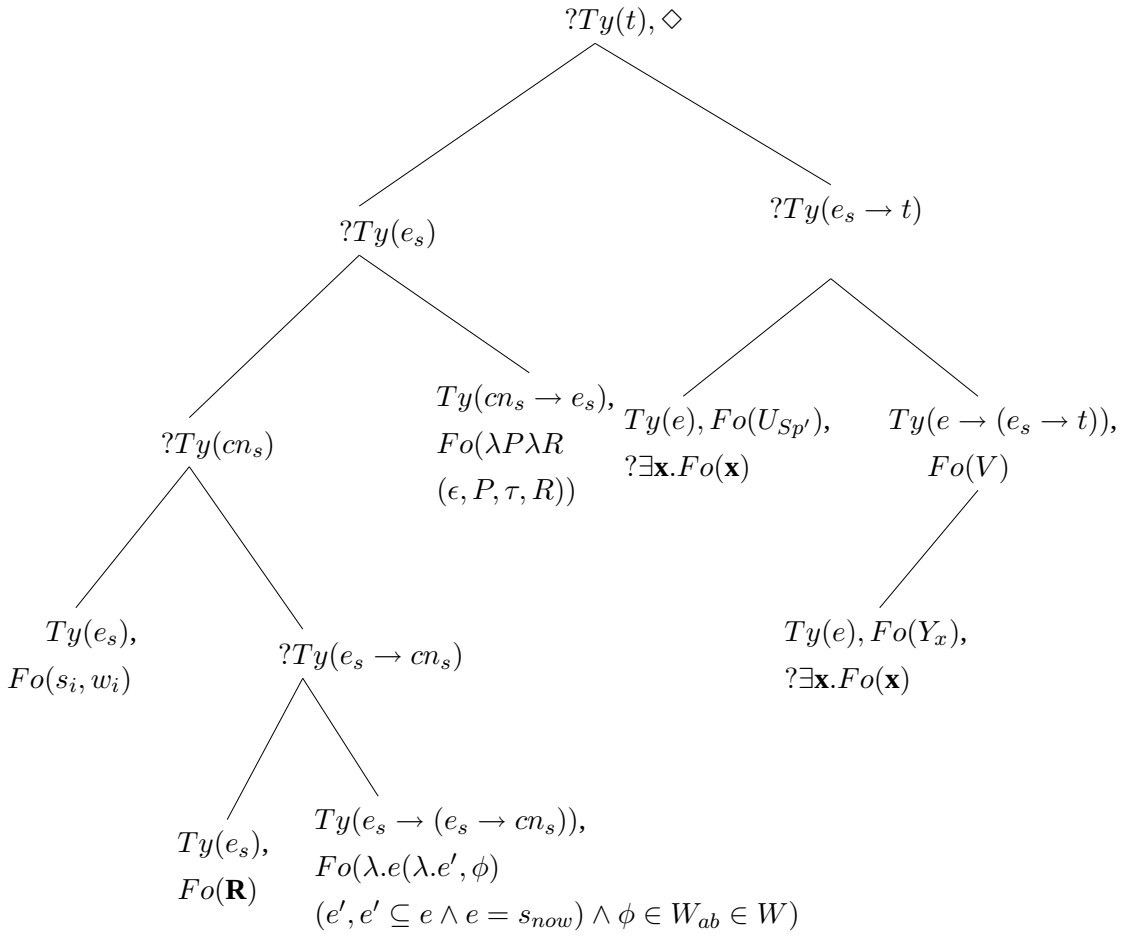
(3.126) Parsing *to* 'it' in *to sotzo vorasi* 'I can buy it.'



With the pointer being at the type *t* requiring node, *sotzo* comes into parse:[44]

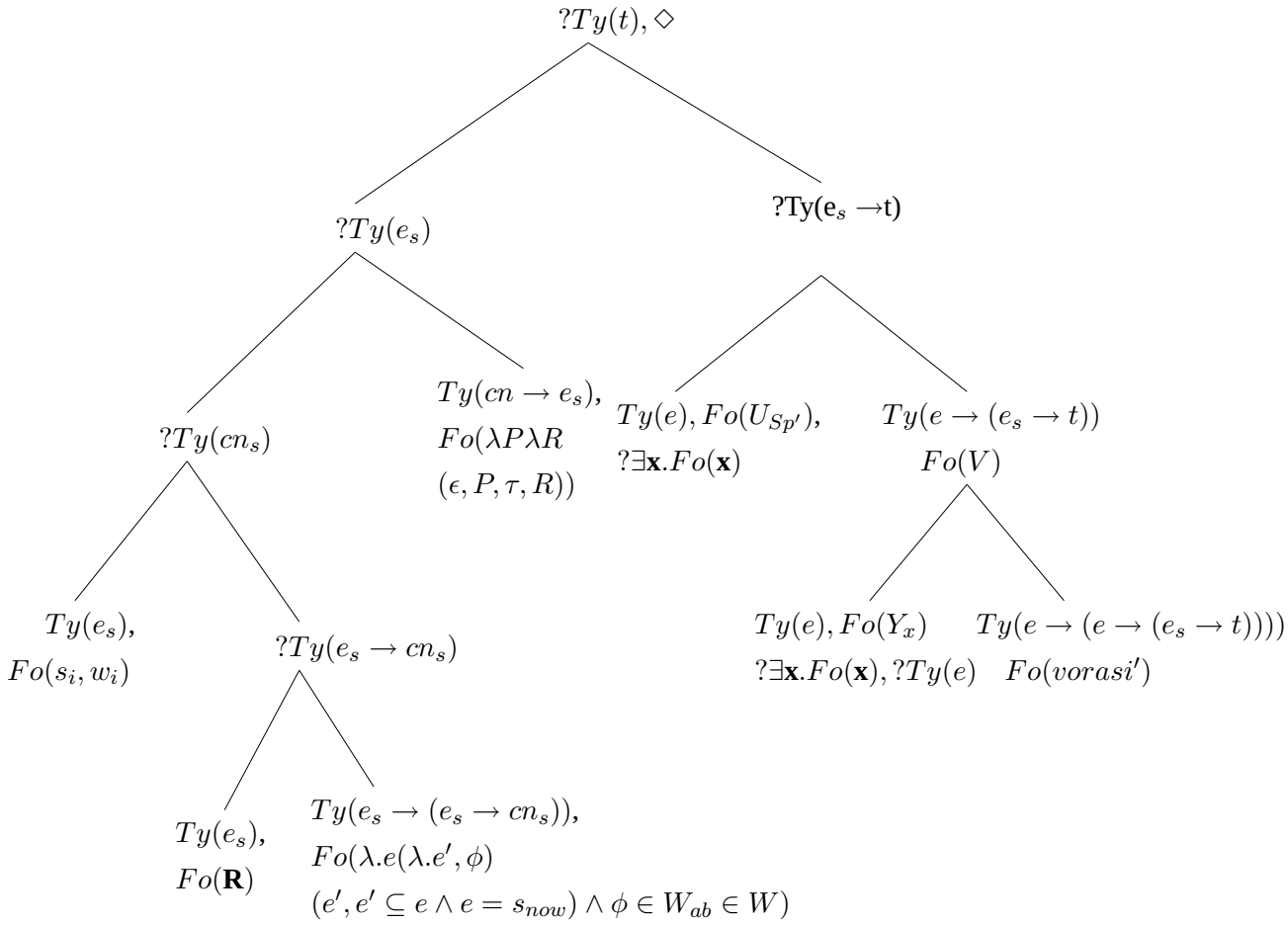
[44]Lambdas in the functor nodes are omitted in this tree and the following ones, for reasons of space.

(3.127) Parsing *sotzo* ‘can’ in *to sotzo vorasi* ‘I can buy it’



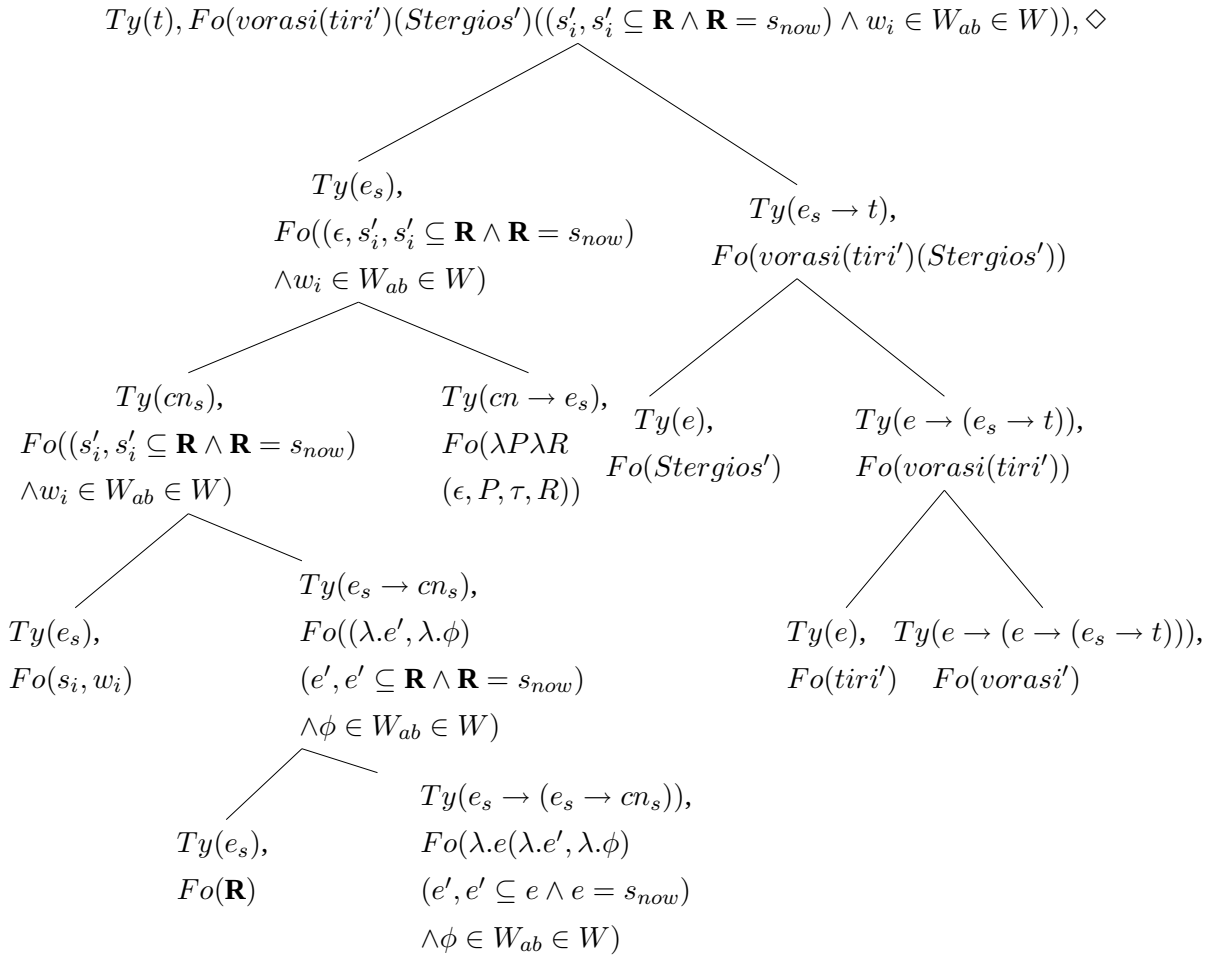
The pointer is again at the type *t* requiring node. The infinitive comes into parse, projecting a verbal type plus a formula value in the 0111 node. It further builds the direct object node and decorates it with type *e* requirement (this requirement will vanish via THINNING since a type *e* value projected from the clitic will already be there):

(3.128) Parsing *vorasi* ‘can’ in *to sotzo vorasi* ‘I can buy it’



At that point the usual rules of ELIMINATION, THINNING, COMPLETION and metavariable SUBSTITUTION will apply when neccessary to give us the well-formed output shown below (with *tiri* ‘cheese’ substituting the object metavariable and *Stergios* the subject one):

(3.129) Completing the parse



The difference between optional climbing languages, on the one hand, and GSG on the other, is that in the latter infinitives are incompatible with clitics whereas in optional climbing languages infinitives can host clitics. In that sense, the lexical entries of restructuring verbs in optional CC languages will have to involve a trigger that will capture enclitic positioning with infinitives as well (see Chatzikyriakidis, 2009c, forthcoming). One of the consequences of the above account (and any monoclausal account in general) is that restructuring verbs cannot be control verbs anymore. This is because no subject is controlled but rather both the verb and the infinitive share the same subject. This has already been noted in the literature by Cinque (2006: 21) who argues that even apparent control cases like *want*, inherit their subject from the embedded lexical verb and thus are not control

cases. In our case, no such inheritance is necessary, since both the restructuring verb and the infinitive share the same subject, but the intuition that restructuring does not involve control is common to both accounts.

As already mentioned, the two climbing inducing verbs in GSG can form multiple climbing constructions, in which the clitic climbs across two verbs:

(3.130) *(T)o* *sotzo* *spiccetsi* *tse* *di* *avri*
it.CL-ACC can finish.INF COMP see.INF tomorrow
'I can finish seeing it tomorrow.'

(3.131) **Sotzo* *spiccetsi* *tse* *to* *di* *avri*
can spiccetsi COMP it.CL-ACC see.INF tomorrow
'I can finish seeing it tomorrow.'

(3.132) **Sotzo* *spiccetsi* *tse* *di* *to* *avri*
can spiccetsi COMP see.INF it.CL-ACC tomorrow
'I can finish seeing it tomorrow.'

Climbing in the intermediate position is also banned in GSG, in contrast to languages like Italian where intermediate climbing appears to be permitted:[45]

(3.133) **Sotzo* *to* *spiccetsi* *tse* *di* *avri*
can it.CL-ACC finish.INF COMP see.INF tomorrow
'I can finish seeing it tomorrow.'

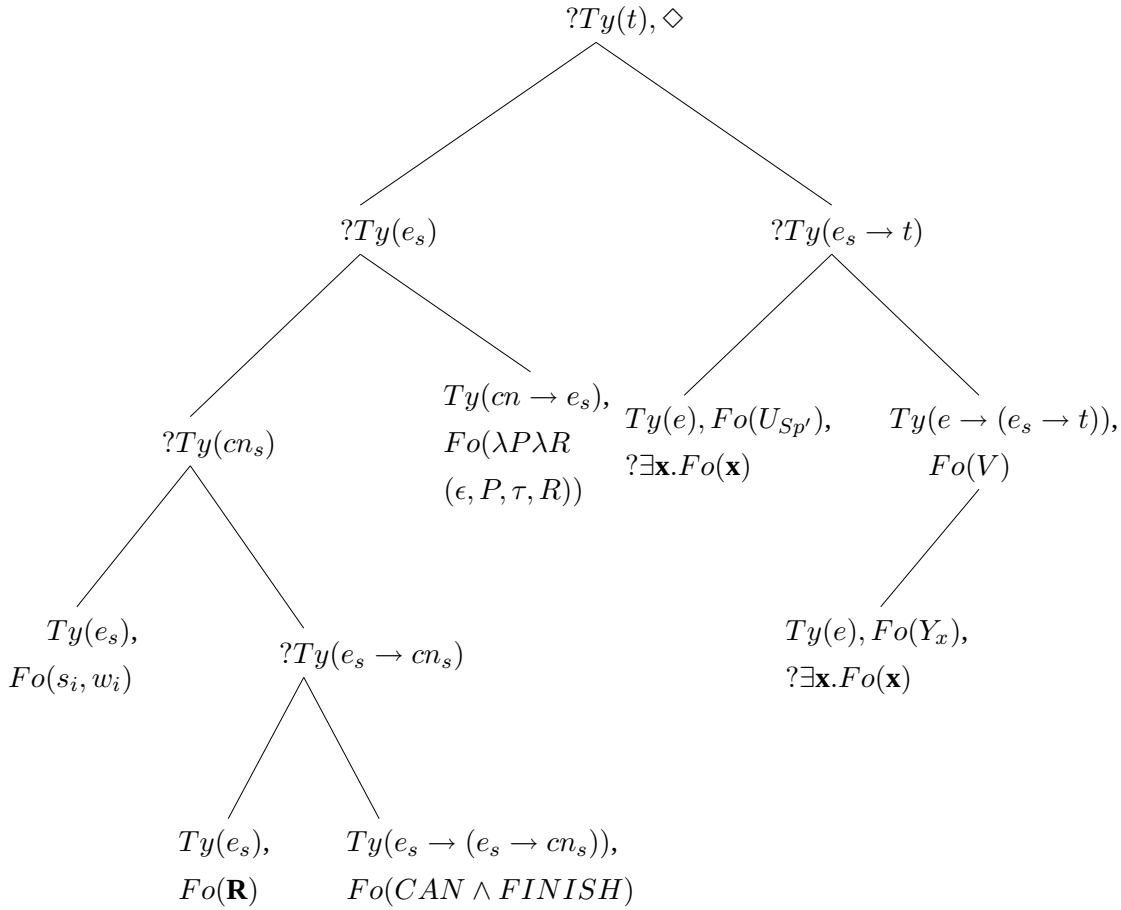
(3.134) **Sotzo* *spiccetsi* *to* *tse* *di* *avri*
can finish.INF it.CL-ACC COMP see.INF tomorrow
'I can finish seeing it tomorrow.'

The account sketched so far correctly predicts the above facts. Let us see why. The restructuring infinitive is assumed to be parsed in the same sense as *sotzo*, i.e. as encoding its semantics in the complex situation argument node. The difference between finite and infinitive restructuring verbs lies in the fact that infinitives in contrast to finite verbal forms will not introduce a freshput situation or world parameter in the type e_s[2] node but will

[45] A number of restrictions apply in case of multiple climbing in Italian. See Cardinaletti & Shlonsky (2004) for an extensive discussion on the issue.

rather depend on the situation or world already projected by the finite verb. Tense, aspect or world specifications will be projected on the relevant nodes. If these nodes already contain such specifications, a combination of the two specifications is done using a form of generalized conjunction in the sense of Cann (forthcoming) that combines two formulae of the same type. The example below illustrates the result of parsing *sotzo spiccetsi* 'can finish':[46]

(3.135) Parsing *sotzo spitsetsi*

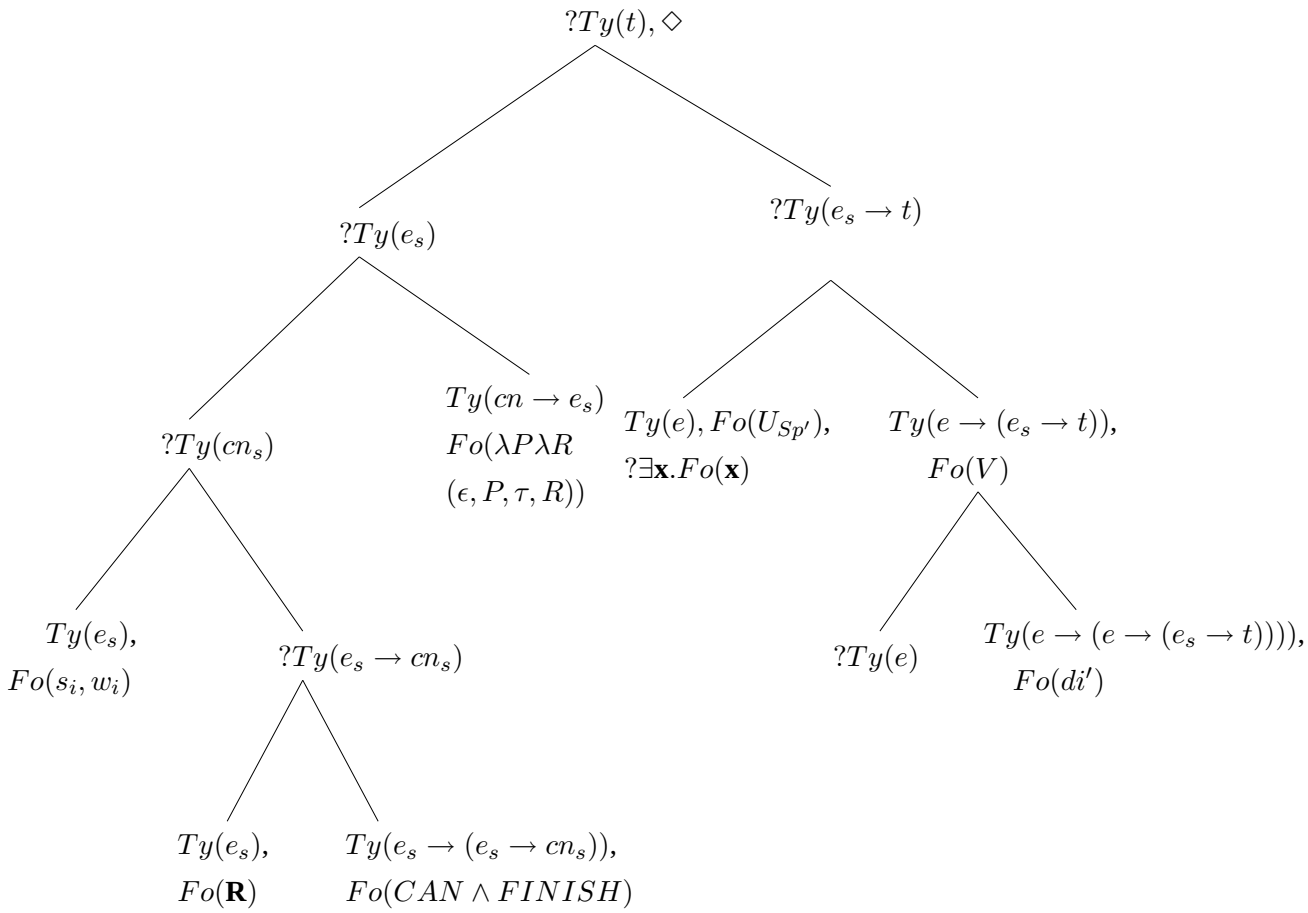


Example (3.131) is predicted to be ungrammatical, since no clitic trigger is again satisfied (a predicate type and a number of fixed nodes exist). The triggering restriction of the

[46]Note that the exact formal specifications of what both *sotzo* and *spitsetsi* contribute to the 0001 node are not shown. This is because the exact semantics have not been fully worked out yet in this case. The statement $CAN \wedge FINISH$ is only a diacritic for what the exact semantics projected by *sotzo* and *spitsetsi* will be.

clitic will not be satisfied and thus the parse will abort. In (3.132) the clitic will come into parse after the lexical infinitive will project its structure, i.e. the rest of the propositional template plus a verbal type in the 0111 node:

(3.136) After parsing *di* 'see.INF' in **sotzo spiccetsi tse di to avri*



The clitic cannot be parsed since in the above partial tree a functor type is present ($Ty(e \rightarrow (e \rightarrow (e_s \rightarrow t)))$), thus the triggering restrictions we have given for clitics do not get satisfied. The trigger for accusative clitics ($[\downarrow_1^+]?Ty(x)$) will fail due to the presence of a verbal type, while the trigger for genitives ($[\downarrow^+]?\exists \mathbf{x}.Tn(\mathbf{x})$) due to the presence of fixed nodes. Similarly in (3.133) and (3.134), parsing of a clitic is not possible, since again both a functor type and a number of fixed nodes exist. The only option in that respect under our account is multiple CC as indicated by the data.

The account proposed correctly predicts the climbing facts in GSG. Furthermore, CC with auxiliary verbs in SMG is also accounted for correctly by using the same reasoning. There are a number of other welcoming results this account has to offer for a number of other phenomena found in CC languages, notably unavailability of negating infinitives in CC contexts,[47] impossibility of using the same adverb twice when CC has occured (Cinque, 2006) and auxiliary switch in Italian. However, these phenomena will not be discussed here for reasons of relevance. The reader interested in the way DS can account for these is directed to Chatzikyriakidis (2009c, forthcoming).

3.5 Conclusions

In this section, an account of the distributional properties of SMG and GSG was provided. The proclisis/enclisis alternation in non-imperative and imperative environments was effectively captured assuming two parsing triggers in the entries for clitics in the two dialects, one capturing proclisis and the other enclisis. Accusative clitics were argued to be specified as regards their structural position, and the evidence from double accusative constructions were shown incapable of disproving such a claim. On the contrary, they do support it. On the other hand, genitive clitics are assumed to be structurally underspecified and as such they were argued to project locally unfixed nodes. Strict DAT-ACC ordering was captured assuming that genitive clitics involve a triggering restriction which disallows them from being parsed in case some fixed structure has already been established. Given that 3rd person accusative clitics, being structurally specified, will induce such fixed structure, parsing of a genitive clitic after an accusative one will be blocked. On the other hand, a 3rd person accusative clitic will be able to be parsed after a genitive clitic since no functor nodes with type values will exist after the genitive clitic is parsed. In imperative environments, the "first fixed node" restriction is modified to "first fixed object" restriction in GSG. Genitive clitics in GSG carry this restriction and as such they cannot be parsed in case a fixed object already exists. No such restriction is needed for SMG, given the free ordering exhibited in imperative environments.

[47] Such a fact is precluded independently for GSG, since infinitives cannot be negated in general in GSG or Calabrian Greek (Katsoyannou, 1995: 451).

CC in GSG was discussed and was shown that following an auxiliary-like analysis for restructuring verbs, in which these do not project any verbal type but rather project their semantics inside a complex situation argument node in the sense of Cann (forthcoming), CC is directly predicted in a rather straightforward and intuitive way, since restructuring verb climbing is analyzed in the same sense as auxiliary climbing in languages like SMG. Thus, such an approach is a dynamic and straightforward formalization of the idea that restructuring verbs are not full verbs but rather functional ones, argued but not formalized by Cinque (2001, 2006) and Cardinaletti & Shlonsky (2004).

Chapter 4

Cypriot Greek

4.1 Historical and Sociological Preliminaries

Cypriot Greek is the Greek dialect of the Greek Cypriots.[1] The first evidence of Cypriot Greek comes from a legal text dated in the 14th century (The Assizes, Horrocks 1997).[2] Despite Cypriot Greek being the native tongue of possibly most (if not all) Cypriot Greeks in both Cyprus and diaspora, the official language of the Cypriot Greek state is SMG (Newton, 1972; Tsiplakou 2009, among others), and as such modern day CG stands in a diglossic relationship to SMG (Tsiplakou, 2009, in press). Tsiplakou (2009) showed that this diglossic relationship in modern day Cyprus is between SMG and an emergent pancypriot Koiné. At the same time, all regional varieties contributing to the formation of this new form of Cypriot, i.e. the Koiné, are subject to levelling (Tsiplakou, 2009). The data used in this thesis are from native Cypriot Greek speakers between 20 to 35 from different areas of Cyprus. I do not know at the present whether differences in clitic positioning are found among individual regional varieties or whether the data gathered present a clitic positioning system which is influenced by SMG. Unfortunately, there is no way at present to test this. I believe that the data gathered, besides presenting some novel aspects on clitic positioning, are a fairly accurate description of the CG positioning system, at least in the way

[1] Although some older Turkish Cypriots are also speakers of the dialect.

[2] Note that the presence of Greek in Cyprus dates back to the 11-10th century BC (Arcadocypriot). However, Cypriot Greek is not an evolution of the Arcadocypriot dialect of Ancient Greek but rather evolved from Koiné Greek like the rest of the MG dialects, the only exception being perhaps Tsakonian (Horrocks 1997).

it is spoken today by young CG speakers. It is however wise to keep in mind that this thesis will not try to address regional differences with respect to clitic positioning (if any). Therefore, it should not be taken for granted that the clitic positioning system described in this thesis is uniform across all speakers of CG (which might be true but at the present it is just impossible to tell).

4.2 The Data

The CG clitic morphological forms are exactly the ones we find in SMG. The table below illustrates these forms:

(4.1) Clitic morphological forms for CG

	$1st$	$2nd$	$3rd$
Sg accusative	me	se	ton/tin/to
Pl accusative	mas	sas	tus/tis - tes/ ta
Sg genitive	mu	su	tu/tis
Pl genitive	mas	sas	tus

As regards clitic positioning, CG, unlike SMG and GSG, is not defined solely with respect to the verb but other linguistic elements other than verbs can affect clitic positioning. CG exhibits three different positioning environments: a) enclitic b) proclitic and c) variation environments. In what follows, these are presented.

4.2.1 Enclitic and Proclitic Environments

Clitics in CG are in general enclitic in both imperative and non-imperative contexts

(4.2)

a. *Poli* *anthropi* *kamnun* *to* *sosta*
many people do.3PL it.CL-ACC right

b. **Poli* *anthropi* *to* *kamnun* *sosta*
many people it.CL-ACC do.3PL right
'Many people people do it right.'

(4.3)

a. *Kamne* *to*
do.IMP it.CL-ACC

b. **To* *kamne*
it.CL-ACC do.IMP
'Do it.'

However, clitics appear preverbally in case a number of linguistic elements appear in the left periphery of the clause. These elements include function words, fronted constituents and subordinating conjunctions. In particular, the following elements trigger proclisis in CG:

1) Wh-elements/Modality-Tense markers:

Wh-elements in CG are associated with proclisis. The same holds for subjunctive markers *na/as*, negation particles *en/mi(n)* and the future particle *enna*:

(4.4) *Pios* *ton* *iδe* *(*ton)?*
who.NOM him.CL-ACC saw.3SG him.CL-ACC
'Who saw him? '

(4.5) *En* *ton* *iksero* *(*ton).*
NEG him.CL-ACC know.1SG him.CL-ACC
'I do not know him.'

(4.6) *Min* *ton* *δis* *(*ton)*
NEG him.CL-ACC see.2SG.PNP him.CL-ACC
'Do not see him.'

(4.7) *Θelo* *na* *ton* *δo* *(*ton).*
want.1SG SUBJ him.CL-ACC see.1SG him.CL-ACC
'I want to see him.'

(4.8) *Enna* *ton* *δo* *apopse* *(*ton).*
FUT him.CL-ACC see.1SG tonight him.CL-ACC
'I'm going to see him tonight.'

2) Conditional/Temporal conjunctions:

Subordinating conjunctions that introduce conditional/temporal clauses are associated with proclisis:

(4.9) *Molis* *ton* *iδa* *(*ton),...*
as-soon-as him.CL-ACC saw.1SG him.CL-ACC
'As soon as I saw him,...'

(4.10) *An* *ton* *δite* *(*ton),...*
if him.CL-ACC saw.2PL him.CL-ACC
'If you see him...'

(4.11) *Pu* *ton* *iδa* *(*ton),* *milisa* *tu.*
when him.CL-ACC saw.1SG him.CL-ACC spoke.1SG him.CL-GEN
'When I saw him, I spoke to him.'

(4.12) *Aman* *ton* *ipopsiazete* *(*ton),* *γorazi* *tis*
if/when him.CL-ACC is-suspicious him.CL-ACC buy.3SG her.CL-GEN
lulufkja/fjora
flowers
'If/When she is suspicious of him, he buys her flowers.'

3) Focused deictic objects, non-coreferential with the clitic:[3]

(4.13) *TUTO* *tu* *edokes* *(*tu)* *oi* *to* *alo*
this him.CL-GEN gave.1SG him.CL-GEN not the other
'You gave him this one, not the other one.'

Notice that when the fronted object is coreferential with the clitic, proclisis is no longer possible:

(4.14) *TUTO* *(*ton)* *ksero* *to*
this it.CL-ACC know.1SG it.CL-ACC
'I know this.'

[3]CG does not have OV focus structures but uses clefts to express these types of expressions (Tsiplakou et al., 2006). Deictic NPs are however permitted in OV focus structures.

The above structure can be either a CLLD or an HTLD construction, that is construed as a bona fide part of the construction or, in some sense, partially detached from it. In the latter case, an intonational break is needed. In both cases, enclisis obtains. Note that in an HTLD construction, case connectivity is no longer necessary:

(4.15) *O* *Γianis,* *(*ton)* *iδa* *ton* *extes*
the.NOM John.NOM him.CL-ACC saw.1SG him.CL-ACC yesterday
'I saw John yesterday.'

Examples like the above exhibiting a case mismatch are unambiguously HTLD cases. Now, if we want to be sure that CLLD is associated with enclisis, we should find an environment where only CLLD is possible. Such an environment obtains if we embed the clause exhibiting CLLD or HTLD inside a verbal complement. In this case, CLLD but not HTLD are possible:

(4.16) *Ipe* *oti* *ton* *Γiani* *ksero* *ton*
said.3SG that the.ACC John.ACC know.1SG him.CL-ACC
'S/He/It said that I know John.'

(4.17) **Ipe* *oti* *o* *Γianis,* *iδa* *ton* *extes*
said.3SG that the.NOM John.NOM saw.1SG him.CL-ACC yesterday
'S/He/It said that I saw John yesterday.'

We can thus be sure that the structure in (4.16) is a CLLD case. Proclisis renders the sentence ungrammatical in the same example:

(4.18) **Ipe* *oti* *ton* *Jani* *ton* *ksero*
said.3SG that the.ACC John.ACC him.CL-ACC know.1SG
'S/He/It said that I know John.'

CLLD structures are incompatible with a focused reading in general (see Kalulli 2000 among others). In that respect, focus that seems to be the reason for proclisis in examples like (4.13) is not compatible with CLLD. The unavailability of focus in CLLD structures might then be a potential reason for enclisis in CLLD structures.
4) Focused subjects and adverbs:

(4.19) *I* *MARIA* *ton* *ikseri* *oi* *i* *Ioana*
the.NOM Mary.NOM him.CL-ACC know.3SG not the.NOM Ioana.NOM
'Mary knows him, not Ioanna.'

(4.20) *XTES* *ton* *iδa*
yesterday him.CL-ACC saw.1SG
'It was yesterday that I saw him.'

(4.21) *EKI* *to* *θelo*
there it.CL-ACC want.1SG
'I want it there.'

Note that in the case of a CLLD construction which further contains a focused subject, the clitic appears preverbally:

(4.22) *Ton* *Jani* *I* *MARIA* *ton* *ikseri*
the.ACC John.ACC the.NOM Mary.NOM him.CL-ACC know.3SG
'It is Mary who knows John.'

Embedding the above sentence inside a complement, we still get the same results. Since CLLD is associated with enclisis, the reason for proclisis in this case is the focused subject:

(4.23) *Ipe* *mu* *oti* *ton* *Jani* *I* *MARIA*
said.3SG me.CL-GEN that the.ACC John.ACC the.NOM Mary.NOM
ton *ikseri*
him.CL-ACC know.3SG
'S/He/It said that it is Mary who knows John.'

5) The factive complementizer *pu*:

Complements of factive and semi-factive verbs introduced via the factive complementizer *pu* trigger proclisis:

(4.24) *Lipame* *pu* *ton* *iδes*
be-sorry.1SG COMP him.CL-ACC saw.2SG
'I'm sorry that you saw him.'

(4.25) *Θimume* *pu* *ton* *iδes*
remember.1SG COMP him.CL-ACC saw.2SG
'I remember that you saw him.'

Terzi (1999a: footnote 19) notes that some speakers she consulted accepted enclisis to be also possible with *pu*. In my corpus, one of the 5 speakers consulted also accepted enclisis. However, he marked enclisis being the less preferred option. We will see later on that an analysis that assumes variation in clitic positioning with the complementizer *pu* is possible given the account that is going to be proposed for variant positioning with the complementizer *oti*.

4.2.2 Variation Environments

An interesting fact regarding the distributional properties of the CG clitic positioning sytem that has in general gone unnoticed in the literature (with the exception of Revithiadou (2006) but only for *oti* clauses), is the existence of a number of variation environments, i.e. environments that allow either proclisis or enclisis. These include complements of non-factive verbs introduced with the non-factive complementizer *oti* and subordinates of cause introduced via *epiδi*, *γiati* 'because'. The standard assumption in the literature (Terzi, 1999a; Agouraki, 2001) before Revithiadou (2006) was that no variation environments existed. However, Revithiadou (2006) pointed out that the non-factive complementizer *oti* is associated with both proclisis and enclisis in CG:

(4.26)

a. *Ipen* *oti* *eθkiavasen* *to*
said.3SG COMP read.3SG it.CL-ACC

b. *Ipen* *oti* *to* *eθkiavasen*
said.3SG COMP it.CL-ACC read.3SG
'S/He/It said that s/he read it.'

My data confirm Revithiadou's claims as regards clitic positioning with *oti*. All speakers of CG consulted indicated both proclisis and enclisis to be possible with *oti*. Some of them further indicated enclisis to be prevalent in terms of preference.

Subordinates of cause introduced with *epiδi* , *γiati* also give rise to variation in clitic positioning, something which has not been mentioned in the literature. All speakers consulted

verified the previous claim:[4]

(4.27)

a.	*Epiδi*	*enevriases*	*me,*	*en*	*kamno*	*tipota*
	because	made-angry.2SG	me.CL-ACC	NEG	do.1SG	nothing
b.	*Epiδi*	*me*	*enevriases,*	*en*	*kamno*	*tipota*
	because	me.CL-ACC	made-angry.2SG	NEG	do.1SG	nothing

'I'm not doing anything because you made me angry.'

(4.28)

a.	*Γiati*	*kamni*	*me*	*tze*	*nioθo*	*orea*
	because	makes.3SG	me.CL-ACC	and	feel.1SG	nice
b.	*Γiati*	*me*	*kamni*	*tze*	*nioθo*	*orea*
	because	me.CL-ACC	makes.3SG	and	feel.1SG	nice

'Because s/he/it makes me feel nice.'

Lastly, the temporal complementizer *eno* 'while' is compatible with both proclisis and enclisis. However, this positioning variation depends on the semantics of the sentence it introduces. In its strict temporal use, the complementizer *eno* is associated with proclisis only:

(4.29)

a.	*Eno*	*to*	*eθkiavazen,*	*iδe*	*ton*	*Γiorko*
	while	it.CL-ACC	read.3SG	saw.3SG	the.ACC	George.ACC
b.	**Eno*	*eθkiavazen*	*to,*	*iδe*	*ton*	*Γiorko*
	while	read.3SG	it.CL-ACC	saw.3SG	the.ACC	George.ACC

'While s/he was reading, s/he saw George.'

Strangely enough, in contrastive *while* clauses introduced with *eno* the only possibility in the presence of a clitic is enclisis:

[4]Speaker judgments were elicited in one of the following ways: 1) by translating the equivalent SMG examples in CG or 2) by positing grammaticality judgments as regards a number of CG sentences involving subordinates of cause with varying clitic positioning.

(4.30)

a. *O* Γ*iorkos* *extipise* *ton* *eno* *o* Γ*ianis*
the.NOM George.NOM hit.3SG him.CL-ACC while the.NOM John.NOM
eksitimasen *ton*
sweared-at.3SG him.CL-ACC

b. **O* Γ*iorkos* *extipise* *ton* *eno* *o* Γ*ianis*
the.NOM George.NOM hit.3SG him.CL-ACC while the.NOM John.NOM
ton *eksitimasen*
him.CL-ACC sweared-at.3SG
'George hit him while John sweared at him.'

No other variation environments were found. The next issue that is worth looking at are environments that unexpectedly fail to trigger proclisis in CG.

4.2.3 Unexpected Enclitic Environments

A number of environments unexpectedly fail to trigger proclisis in CG. All these involve a marker-conjunction or a conjunction-conjunction complex, the second part of the complex always being *tze* 'and'.[5] Such behaviour is quite surprising given that the first element of the complex in each case triggers proclisis while furthermore the interpretation of the complex would also lead someone to expect proclisis. This is because all these complexes are either complexes introducing subordinate clauses (concessive, dubitative) or the emphatic negation complex *en tze*. However, in all these cases enclisis rather than proclisis obtains as witness the examples below:

(4.31)

a. *En* *tze* *iδa* *ton*
NEG and saw.1SG him.CL-ACC

b. **En* *tze* *ton* *iδa*
NEG and him.CL-ACC saw.1SG
'I did not see him.'

[5]This fact has already been noted in the literature on CG clitics (Agouraki 2001).

a. *An tze iδa ton...*
if and saw.1SG him.CL-ACC...

b. **An tze ton iδa*
if and him.CL-ACC saw.1SG
'Even though I saw him...'

a. *Foume mipos tze δo ton*
be-fraid.1SG DUB and see.1SG him.CL-ACC

b. **Foume mipos tze ton δo*
be-fraid.1SG DUB and him.CL-ACC see.1SG
'I'm afraid that I might see him.'

This is pretty much the positioning facts as regards positioning. Now I will examine the ordering facts inside clitic clusters in CG.

4.2.4 Clitic Ordering in Clitic Clusters

CG resembles SMG and GSG in that DAT-ACC order is the only option in non-imperative environments:

(4.32)

a. *Tu to ipa*
him.CL-GEN it.CL-ACC said.1SG

b. **To tu ipa*
it.CL-ACC him.CL-GEN said.1SG
'I said it to him.'

As regards imperatives, the standard assumption in the literature (Terzi, 1999b) is that CG, similarly to SMG, freely allows both orderings inside the cluster. The data below are from Terzi (1999b):

(4.33)

a. *Θkiavase* *mu* *to*
read.IMP me.CL-GEN it.CL-ACC

b. *Θkiavase* *to* *mu*
read.IMP it.CL-ACC me.CL-GEN
'Read it to me!'

My data do not concur with her claim however. All the speakers consulted indicated that only DAT-ACC is possible in both finite and imperative contexts, judging sentences like (b) above as severely ungrammatical. The same fact is reported in Paraskeua (2007). What is the reason behind this difference in grammaticality judgments remains unknown to me. The judgments I got were extremely uniform since all speakers reported severe ungrammaticality as regards imperative constructions with ACC-DAT clitic clusters. What I am going to argue for in this thesis is that clitic ordering is strict in all environments in CG. A direct implication of such a claim, given that GSG, as mentioned earlier, and PG, as we will see later on, exhibit strict ordering in both non-imperative and imperative environments is that free ordering of clitics in imperative environments is not a generally attested phenomenon in MG dialects, but rather specific to SMG.

With this last remark, I conclude the data section of CG, noting that a number of complementary data will be introduced as the CG analysis proceeds or as the analysis regarding phenomena common to all dialects under investigation (PCC, CLLD and CD) is presented.

4.3 Existing Approaches

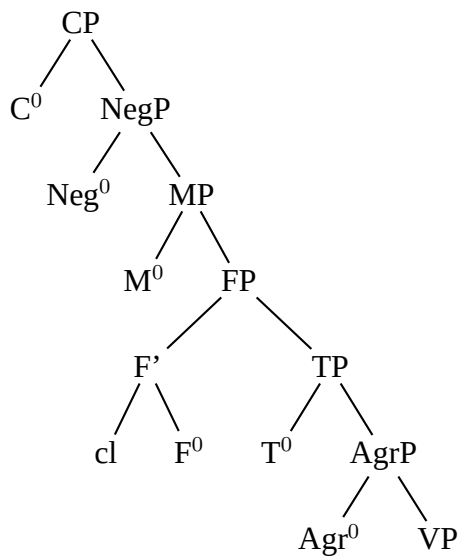
There are a couple of analyses regarding clitic positioning in CG, notably Agouraki (1997, 2001), Terzi (1999a, 1999b), Condoravdi & Kiparsky (2001) and Revithiadou (2006). In what follows, I briefly review these approaches.

4.3.1 Terzi, 1999a

Terzi (1999a) tries to give an account of the distributional properties of the CG clitic system by, at the same time, retaining the same clausal architecture for both SMG and CG. The

clausal structure assumed in Terzi (1999a) is effectively the same assumed for both SMG and CG in Terzi (1999b), and has been already presented in reviewing the literature on SMG and GSG (see 3.3.1.3):

(4.34)



Following Anagnostopoulou (1994), Terzi considers CG clitics to be X_0 elements adjoining to a featureless Functional head F_0 that takes TP as its complement. V-cl ordering in finite contexts is attributed to movement of the verb to M_0, while the same ordering in imperatives is attributed to movement of the verb to C. The exact trigger of this movement however remains rather unclear in this analysis. Even Terzi herself notes that it is difficult to see any morphological differences between the CG and the SMG verb, such that movement would be triggered for the former but not for the latter. She then argues that the difference between CG and SMG must be attributed to the level of representation that the clitics are formally licensed in the two varieties. Within this line of reasoning, it is claimed that CG clitics need to be licensed inside the domain of a functional projection with operator like properties before Spell-out. In the presence of such a functional head, i.e NegP, MP, FocP, the clitics surface preverbally since the functional head licenses the clitic inside its domain. However, in the absence of such a head, the verb has to move to M to get licensed.

There are a number of things that are problematic with respect to Terzi's (1999a) analysis. Firstly, as mentioned, is the nature of the trigger that causes movement of the verb to M_0 in CG but not in SMG. Secondly, the notion of a syntactic licenser with operator like

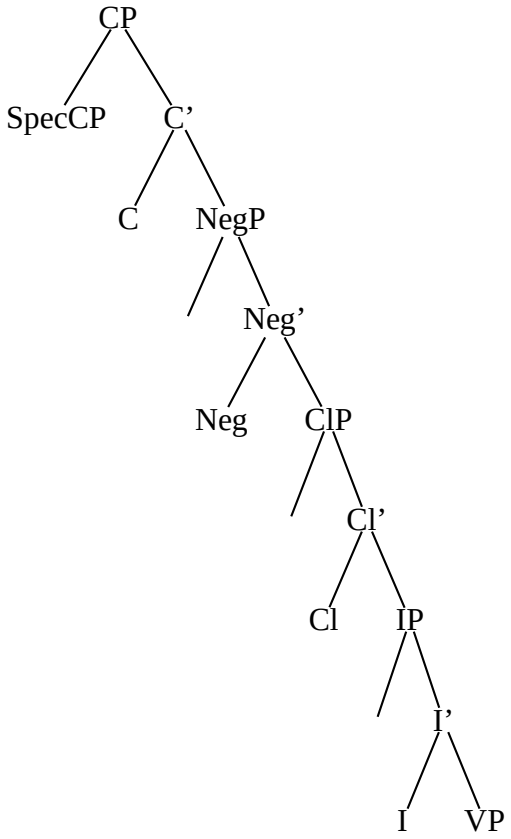
properties is also unclear. What constitutes an operator-like licenser is something that is not fully justified. For example, Terzi claims that the difference in positioning between the factive complementizer *pu* and non-factive *oti* with respect to clitic positioning is due to the operator-like properties of the former but not the latter. Leaving aside the issue of why *pu* is assumed to have operator-like properties while *oti* does not (see Roussou, 2000 for a discussion), the new data presented in Revithiadou (2006) and also in this thesis (see the data section, 4.2), which show variation rather than strict proclisis in environments involving the non-factive complementizer *oti*, pose a serious problem to Terzi's analysis, since under these new data Terzi would have to assume that *oti* exhibits operator-like properties in proclitic but not in enclitic constructions. Another major discrepancy with respect to Terzi's (1999a) analysis involves the motivation behind V-C movement in imperative constructions. According to Terzi (1999a), CG indicative verbs also move to M_0, which is situated between NegP and TP. The question pending is why imperatives do not move to M_0 as well (since this functional head, as its name suggests, licenses mood features) but rather move to C_0 instead. It seems worth emphasizing that the same argument was raised by Roussou (2000) in discussing the same analysis. The answer seems to be that by adopting a unified V-M movement approach for all verbs in CG and thus by further abandoning the assumption that imperatives move to C, then assuming in line with Terzi (1999a,b) that NegP is situated higher than MP, negated imperatives are expected to be licit contrary to fact.

Concluding, it must be said that Terzi (1999b) is not going to be reviewed here for two independent reasons. Firstly, because the account she gives as regards clitic clustering in CG is the same as the one assumed for the equivalent phenomena in SMG, an account already presented in the chapter discussing SMG and GSG. Furthermore, this account presents an analysis of clitic sequences in CG on the assumption that free ordering is possible in CG. Since, as already discussed, this claim does not seem to hold for CG (strict DAT-ACC is the case for imperatives as well), Terzi's analysis needs to be modified to account for the new data. However, the interested reader is directed to chapter 3 of this thesis for more information and critique regarding Terzi (1999b).

4.3.2 Agouraki, 2001

Departing from analyses like Terzi (1999a,b) which assume that enclisis with imperatives and indicatives involve two different landing sites for the verb, i.e. C_0 and M_0 respectively, Agouraki (2001) puts forth an account where all CG verbs move to C_0. This movement is attributed to a filled C_0 requirement that CG has. No MP is present in Agouraki's analysis. The clausal structure proposed is shown below:

(4.35) Structure proposed for CG



SpecCP hosts wh-phrases, topics and foci. In wh-questions or syntactic foci structures, Agouraki assumes the C_0 position to be filled by a null [WH] and [F] complementizer respectively. This stipulation is essential for the analysis to work, since in the presence of any of these complementizers V-C movement will be blocked. In case of blocking, the clitic stays in situ and thus proclisis obtains. A verb feature is further assumed to be present in C_0 in yes-no questions and emphatic sentences. This feature is specified

[+Interrogative]/[+Emphatic]. In contrast to null complementizers [Wh, F], this does not block V-C movement and the verb checks the relevant feature in the C head. With respect to negation markers, Agouraki proposes that these are either situated in a lower NegP (as shown in the tree) and move to C_0 or are base generated in C_0. There are a number of inconsistencies regarding Agouraki's analysis. Firstly, the stipulation of a number of null complementizers in C_0 does not seem to have any empirical grounding. But even if it were the case, it is not clear why wh-phrases and syntactic foci have their null complementizer counterparts while topics, which are situated in SpecCP as well, do not.

Furthermore, the analysis proposed for structures involving negation markers is also not devoid of problems. Assuming that negation is base generated in C_0, we will have to assume a lower NegP as well in order to capture cases where a subordinating conjunction is already filling the C position or where a wh-phrase or a syntactic focus phrase is present in SpecCP (in which case the C position is filled by covert complementizers [WH] and [F]). However, such a move is totally stipulatory and unmotivated, since no evidence arguing for two different types of NegP exists.[6] Assuming that NegP is not base-generated at C_0, but moves from a lower NegP to C_0 in complementizer-less clauses we side-step this stipulation. However, such a move is not unproblematic either. The reason is that we need a trigger for movement of negation to C_0. Agouraki argues that movement to C_0 is triggered by the presence of the morphological feature [Declarative]/[Interrogative]/[Emphasis]. Movement to C_0 must then check one of these features. However, the problem that arises is which of these features does negation check, since none of them is associated with negation. In that respect, movement of negation to C_0 is not motivated. What is more, as was the case for Terzi (1999a) as well, no account of variation environments is given. Therefore, the account does not seem adequate to capture the whole range of phenomena associated with CG clitic placement.

[6]Note that this is totally different to a situation where one of the two forms (the two forms being *en* and *min*, *en* used for indicative environments and *mi(n)* for subjunctive and imperatival ones) of negation will stay in situ while the other one will move. According to Agouraki, the same negation particle, i.e. *en*, must stay in situ in some cases (i.e. in the presence of a subordinating conjunction), while it must move to C in the absence of such a complementizer or a WH/Focus feature in C.

4.3.3 Revithiadou, 2006

Departing from 'all syntax' accounts of CG clitics, Revithiadou (2006) claims that clitic positioning is determined not by syntax but rather by phonology. According to this account, the syntax provides a number of well-formed structures to the phonological component, where, via constraint ranking, only the optimal ones survive. More specifically, differences in clitic positioning in different dialects arise from different constraint rankings. For CG, the ranking for prosodic structure proposed by Revithiadou is as follows:

(4.36) Ranking of prosodic structure in CG

FAITH(acc), EXH, WCON(L), NONREC >> PCON,WCON(R)

FAITH(acc) is a constraint which posits that inherent accent of the input must be preserved in the output (Revithiadou 2006: 87, citing Revithiadou 1999). EXH and NONREC are prosodic domination constraints. The former specifies that no C^i immediately dominates C^j for j<i-1 while the latter that no C^i dominates C^j for j=1. The rest of the constraints are alignment constraints (see Revithiadou, 2006: 85 for the formal details). Revithiadou follows Boskovic's (1995, 2001) analysis for Bulgarian and Serbo-Croatian second position clitics and assumes that in CG, like in Bulgarian and Serbo-croatian, clitics are generated in more than one position in the syntactic component. The copy to be pronounced is then decided by the prosodic ranking of each language. Revithiadou argues that constructions exhibiting variation are a major problem for all syntax accounts. However, there is no treatment of these constructions in her analysis. The only thing mentioned is that in *oti* constructions, the complementizer has two phrasing options. It can be either phrased with the verb of the main clause or alternatively with the verb of the embedded clause. The two different phrasing possibilities cause the two alternative patterns of clitic positioning. In case *oti* is phrased along with the main verb, then enclisis obtains. However, if *oti* is phrased along with the embedded verb then proclisis obtains. Revithiadou does not show how these variation data will be captured under her constraint ranking. For example, it must be shown that if both proclisis and enclisis are possible, then assuming that the syntax provides phonology with both copies of clitics, the two alternative patterns of clisis (proclisis, enclisis) must be shown to score the same number of violations. This is however not shown

in Revithiadou. But even if we assume that both enclisis and proclisis patterns indeed score the same number of violations and thus both patterns are equally well-formed, there are a number of facts that would immediately falsify such an account. The first counter-evidence for such an account is constructions that are the same as the *oti* constructions, the only difference being a different complementizer, e.g. the conditional conjunction *ean*. In these constructions variation is not possible, the only possible construction being proclisis:

(4.37)

a. **Ipen* *ean* *eθkiavasen* *to*
said.3SG COMP read.3SG him.CL-ACC

b. *Ipen* *ean* *to* *eθkiavasen*
said.3SG COMP him.CL-ACC read.3SG
'S/He said/asked if s/he read it.'

It is not clear how the above will be handled assuming a phrasing account is used to explain variation with *oti*. Even worse, there are constructions in which *oti* has only one phrasing option, i.e. being phrased with the embedded verb, and still both proclisis and enclisis are allowed. Such a situation results when the *oti* clause is left-dislocated, i.e. when complementizer *oti* appears as the first element of the clause:

(4.38)

a. *Oti* *ton* *ikseri,* *xtes* *mu* *to* *ipe*
COMP him.CL-ACC know.3SG yesterday me it said

b. *Oti* *kseri* *ton,* *xtes* *mu* *to* *ipe*
COMP know.3SG him.CL-ACC yesterday me it said
'Yesterday s/he told me that s/he knows him.'

It is not clear how cases like the above will be treated under an account which attributes variation to the different phrasing possibilities of *oti*. In conclusion, Revithiadou's account might provide a way to account for the proclisis/enclisis alteration in CG via means of constraint ranking but is inadequate when dealing with variation environments.

4.4 A DS Analysis

4.5 The Enclitic Environments

I will begin the account of CG clitics with enclitic positioning. Remember that enclisis is the default in CG, unless a number of elements in the left periphery of the clause "attract" the clitic to preverbal position. In contrast to SMG and GSG clitics, the trigger of which specifies the ABSENCE of any verbal type as a prerequisite of parsing the clitic, the trigger of CG clitics will have to specify the PRESENCE of such a verbal type instead. Such a triggering point capturing enclisis in CG can be formulated as follows:

(4.39) Trigger ensuring enclisis

IF	$?Ty(t)$	
THEN	IF	$\langle\downarrow_1^+\rangle Ty(x)$
	THEN	...
	ELSE	abort
ELSE	abort	

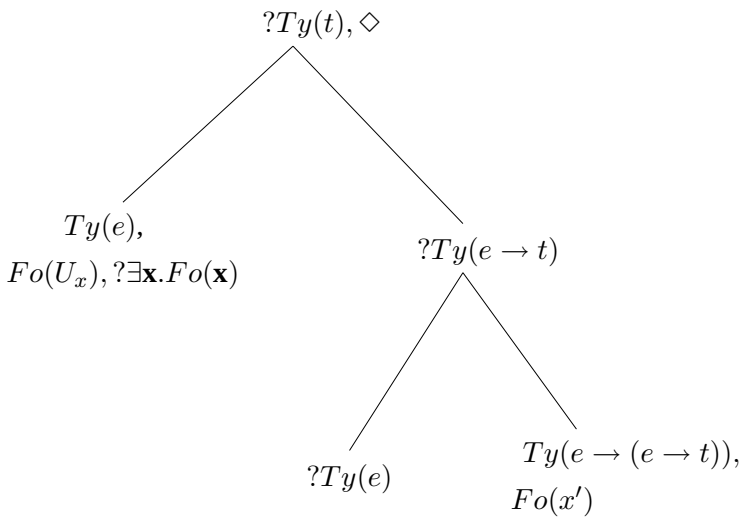
The above will correctly predict sentences like the ones shown below to be grammatical:

(4.40) *Poli* *anthropi* *kamnun* *to* *sosta*
many people do.3PL it.CL-ACC right
'Many people do it right.'

(4.41) *Kamne* *to*
do.IMP it.CL-ACC
'Do it.'

Assuming that a verb has been parsed, a verbal type will be projected in a functor node. Which functor node that will be depends on the verb's adicity. For example parsing of a transitive verb will result in the projection of a verbal type in the 011 node:

(4.42) Parsing a transitive verb



At this point the clitic comes into parse. Both its triggers are satisifed, since the pointer is at the type *t* requiring node and a type value exists in one of the functor nodes, and as such the parsing process can continue. The same mechanism is at play in parsing imperative verbs. Imperative verbs will also project a type value in a functor node. In that sense, enclitic positioning with imperatives is predicted in exactly the same way as with indicatives.

The entry proposed will suffice to capture enclisis in the absence of elements triggering proclisis at the left periphery.

4.6 The Proclitic Environments

The set of proclitic triggers found in CG is anything but homogeneous, as it consists of a number of differently functioning elements: modality/tense markers, wh-elements, focussed constituents and subordinating conjunctions among others. However, a number of generalizations can be captured by looking into the nature of these triggers more carefully.

4.6.1 Unfixed Nodes as Proclitic Triggers

In the introduction to the DS framework and in the discussion on SMG and GSG, extensive reference to the use of unfixed nodes was made. It was mentioned for example, that

a number of variants of the ADJUNCTION rule exist, while furthermore arguing for parametric variation with respect to how many of these ADJUNCTION rules each language has. It seems that the rule of *ADJUNCTION as a parsing strategy can serve as a basis for a generalized proclitic trigger. Bouzouita (2002, 2008a,b) in discussing clitic positioning in MedSp, argued that one of proclitic triggers in MedSp is the existence of an unfixed node. This was based on standard assumptions in DS, according to which Wh-elements and fronted constituents are or can be parsed on an unfixed node (Kempson et al., 2001; Cann et al., 2005). This assumption extends straightforwardly to CG, since the similarity of clitic positioning in the two systems is striking. In that sense, the first generalization that can be drawn is that the existence of an unfixed node is a trigger for proclisis. Note that this proclitic trigger does not refer to the actual content of the elements triggering proclisis but rather to the parsing strategy used (unfixed node). But let us see in detail the basis of this intuition. Let us start with Wh-elements:

(4.43) *Pios* *ton* *ikseri* *ton* *Γiorko?*
who.NOM him.CL-ACC knows.3SG the.ACC George.ACC
'Who knows George?'

Following Kempson et al. (2001), I take wh-elements to be projected as decorating a an unfixed node, marking the initial node with a Q feature, its role being the identification of the wh-element as an interrogative wh-element. The effect is that the unfixed node is decorated with a type value and a specialized formula metavariable WH.[7] The entry below is the lexical entry for *pios* 'who.NOM' in CG:

[7]WH is a specialized metavariable in the sense that update of the metavariable WH to a proper formula value is not obligatory in order to get a well-formed parse. Given that, WH is a distinguished metavariable, since all the other metavariables require update by a proper formula value at some point during the parsing process (expressed via the requirement $?\exists\mathbf{x}.Fo(\mathbf{x})$).

(4.44) Lexical entry for pios in CG

IF	$?Ty(e)$	
THEN	IF	$\langle\uparrow_*\rangle ?Ty(t), ?\exists\mathbf{x}.Tn(\mathbf{x})$
	THEN	go($\langle\uparrow_*\rangle$); put($Cat(Q)$);
		go($\langle\downarrow_*\rangle$); put(**WH**, $Ty(e)$,
		$?\langle\uparrow_0\rangle Ty(t), [\downarrow]\bot$)
	ELSE	abort
ELSE	abort	

Notice the existence of the case filter $?\langle\uparrow_0\rangle Ty(t)$, identifying the unfixed node as immediately dominated by a type *t* node, i.e. as being the subject node.

In the same vein, focus constructions are also modelled as decorating unfixed nodes.[8] A focused NP at the left periphery is parsed on an unfixed node, finding its position in the tree later on in the parse. The same reasoning is used for focused adverbs or any other focused element. Within this line of reasoning, we can formulate a trigger which will allow the clitic to be parsed in case an unfixed node is present in the tree structure. The formulation proposed is shown below:

(4.45) Unfixed node as a proclitic trigger

IF	$?Ty(t)$	
THEN	IF	$\langle\downarrow_*\rangle Ty(x), ?\exists\mathbf{x}.Tn(\mathbf{x})$
	THEN	...
	ELSE	abort
ELSE	abort	

Such a trigger will ensure that the clitic will appear before the verb in case an unfixed node is present in the tree structure. Note that the specification is not restricted to unfixed

[8]The intuition that wh and focus constructions share a number of properties is not something new in the literature. For example in the GB/Minimalist tradition it has been extensively argued that Wh movement and FOCUS movement target the same position in various languages (see Kenesei, 1993 for Hungarian; Rizzi, 1997 for Italian and Ouhalla, 1997 for Standard Arabic among others). Note however that DS does not use a specific focus feature or a specific focus device to account for focus effects. See Kempson et al. (2004) and Kempson et al. (forthcoming) for an argument that topic and focus derive directly from, actually they are an aspect of, the dynamics of utterance interpretation.

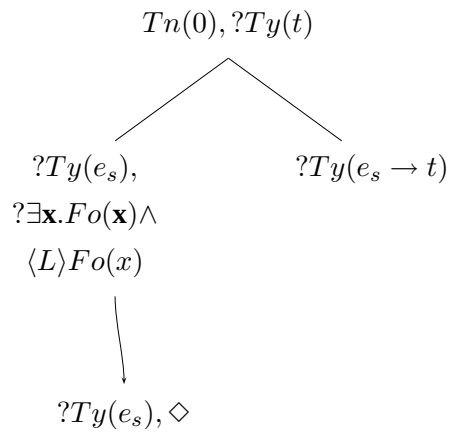
nodes that are type specific. Actually, the specification proposed (following the formulation found in Bouzouita (2008a)) does not refer to the type value of the unfixed node or to its formula value. Such a specification will allow us to account for cases where a strong pronoun is parsed on an unfixed node, in which case a metavariable instead of a proper formula value will have been projected:

(4.46) *EΓO* *tu* *eδoka* *lefta*
I.NOMINATIVE him.GEN gave.1SG money.ACC
'It was me that I gave him money.'

4.6.2 Subordinating Conjunctions, Modality/Tense markers

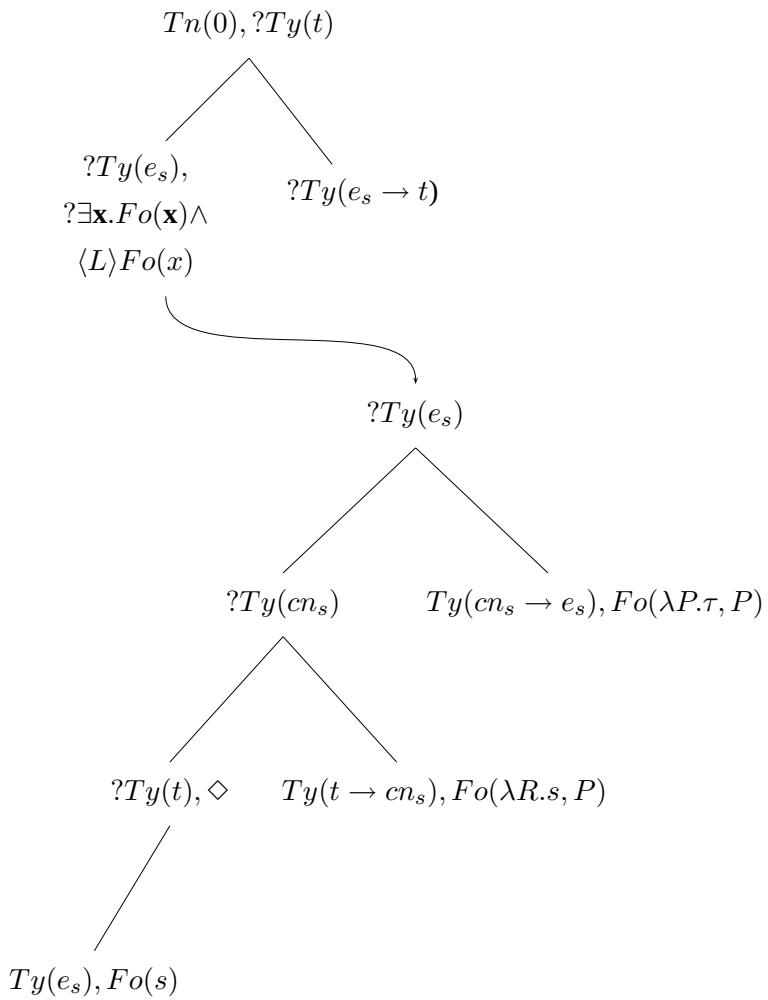
It has already been mentioned that subordinating conjunctions trigger proclisis, the exception being the subordinating conjunctions *epiδi*, *γiati* 'because'. For these specific conjunctions proclisis/enclisis alteration is possible. The crucial question to ask here is what is the general characteristic that links all these conjunctions. In order to answer this, we need to think how subordination is analyzed in DS. Gregoromichelaki (2006) in analyzing conditionals in DS, assumed that subordination includes pairs of LINKed trees sharing a situation argument. For example, in conditional structures, the IF clause is assumed to be linked with the THEN clause via means of a LINK relation. The LINK relation is built via the ACTIONS of the conditional conjunction *if*. The lexical entry for *if*, starting from the type e_s requiring node of the main tree first posits a requirement that the formula value found in this node is shared with the formula value of a LINKed situation argument ($?\exists x.Fo(x) \wedge \langle L \rangle Fo(x)$). Then, the actions of *if* build a LINK relation from the situation argument node to a node which carries a requirement for a type of the same type as the node from which the LINK starts, i.e. e_s:

(4.47) Parsing IF (partial)



Then, *if* further builds the internal nodes of the complex situation argument node, corresponding to the form of quantification involved in conditional sentences, namely universal quantification. The only difference in the structure of the situation argument will be that the internal argument node will be of type *t* instead of type e_s. This type *t* will further project structure that will correspond to the logical form of the IF sentence. The final structure after the actions of the subordinator have been fully induced is shown below:

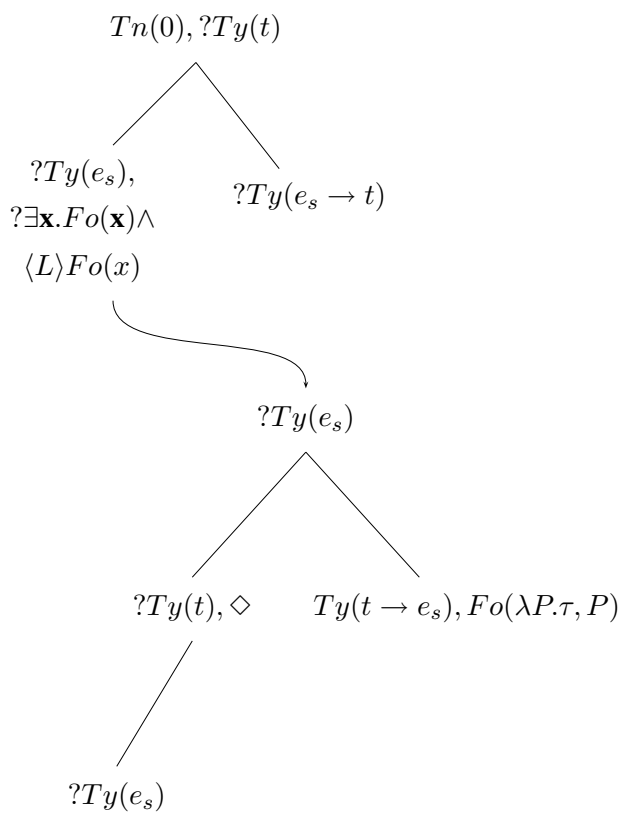
(4.48) Parsing IF (full)



At that point the IF clause itself can be parsed to yield the restrictor for the τ binder. This term, with its universal quantification force, will be copied to the type e_s requiring node of the main tree (the node immediately dominated by the root), satisfying the requirement for a shared term posited in the e$_s$ node ($?\exists\mathbf{x}.Fo(\mathbf{x}) \wedge \langle L\rangle Fo(x)$). By satisfying this requirement, the formula value obtained in parsing the IF clause will act as the context in which the THEN clause is parsed (the proposition expressed by the THEN clause will be true in all situations that the proposition of the IF clause is true). However, Gregoromichelaki's analysis is not fully compatible with recent developments as regards tense and aspect in DS (see Cann, forthcoming; Chatzikyriakidis, forthcoming) presented

in 2.2.2.4 and 3.4.3.1 respectively. In particular, what is rather problematic under Gregoromichelaki's account is the way tense and aspect features are going to be introduced via the entries of the verbs. This is because in subordinate sentences, the subordinators themselves will have constructed the type e_s node, decorating it with a type value before the verb comes into parse. In that respect, the two analyses do not seem to be compatible. However, there seems to be a way to unify the two analyses. The first step in attempting such a unification is to assume that the introduction of the type e_s is not done by the subordinator but rather by the verb of the IF clause, in line with Cann (forthcoming) that take verbs to project this type. The subordinator however will still be assumed to build the situation argument node of the IF clause but will not provide a type for it (only a requirement for a type e_s). This slight modification just proposed will minimally be different with respect to the output tree after the subordinator has been parsed, in that the lowest situation argument node of the LINKed tree (the lowest node in 4.48) will not have a type yet. Instead, a requirement for such a type will be posited. This requirement will be satisfied after the verb comes into parse.

(4.49)

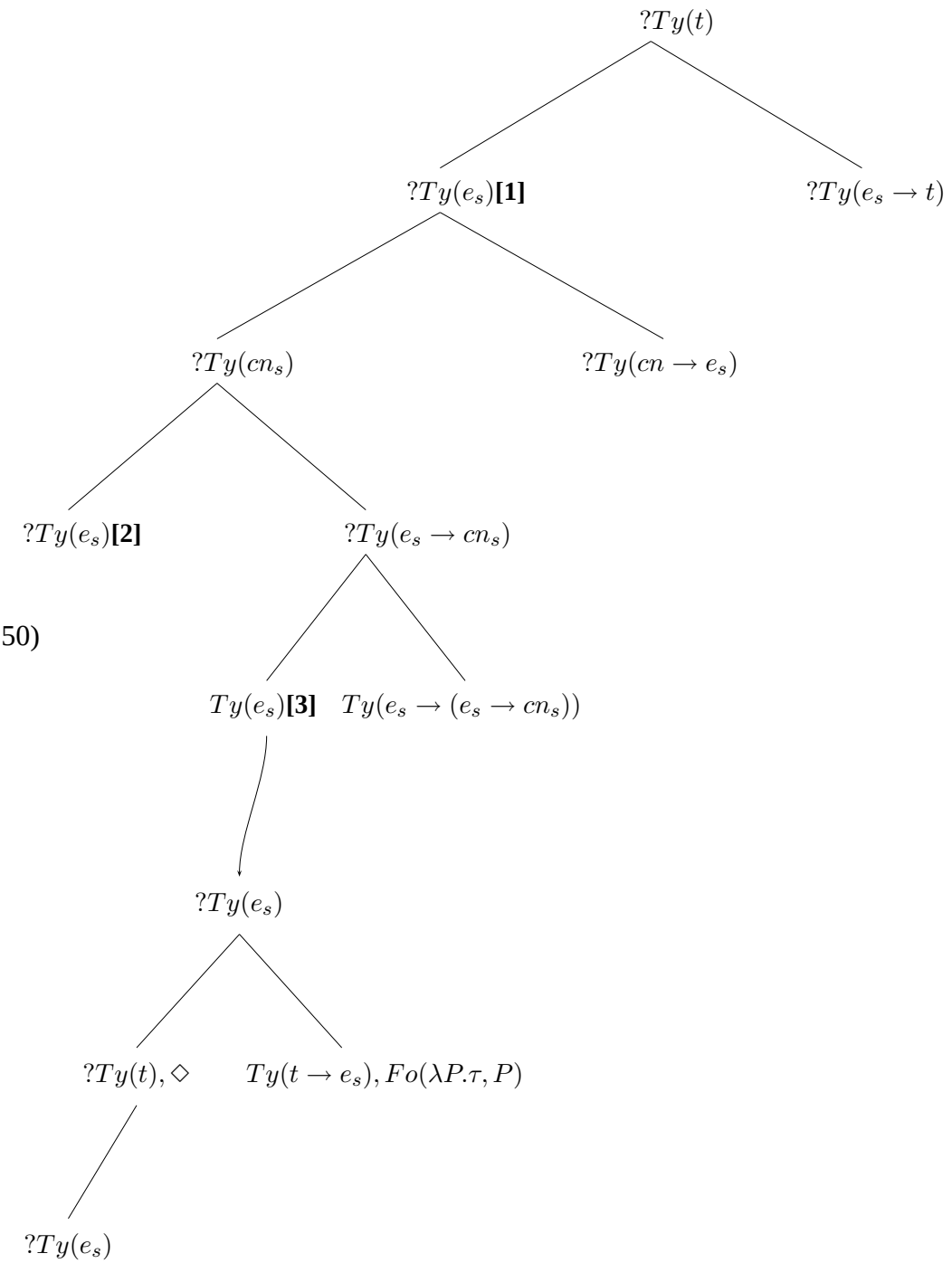


Note that the $t \rightarrow cn_s$ node is changed to $t \rightarrow e_s$ and the two nodes above the $t \rightarrow cn_s$ node omitted. This is because the node $t \rightarrow cn_s$ is not needed anymore given that its function in Gregoromichelaki (2006) is to introduce a freshput variable. Specifically, it "selects and inserts as a Formula value for a node the first available variable that has not appeared earlier in the global tree being constructed" (Gregoromichelaki, 2006: 210). Since such a formula value under the new analysis by Cann (forthcoming) is projected inside the complex situation argument node projected by the verb, such a node is not needed anymore. In that sense, the type value is changed to $t \rightarrow e_s$ where the universal quantification properties associated with conditionals are introduced (in effect moving the formula value found in the omitted functor node $cn_s \rightarrow e_s$ to the $t \rightarrow e_s$ node). However, this modification is not the end of the story. There a number of other things that need to be taken care of. For example, it is not clear under such an analysis how the tense and aspect specifications of THEN clause verb will be projected inside the situation argument node of the initial tree. This is because under Gregoromichelaki's (2006) analysis the situation argument node of

the initial tree and the higher situation argument node of the LINKed are assumed to be identical. Actually, the LINK evaluation rule proposed by Gregoromichelaki (2006) for conditionals does just that, i.e. it copies the formula of the higher situation argument node of the IF clause to the situation argument node of the THEN clause. Given that the verb of the THEN clause will have to project tense and aspect specifications in another situation argument node, the problem that is created is the locus where this situation argument node is going to be projected. A potential solution to this problem is the following: Assume that the two situation argument nodes posited for the construal of IF and THEN clauses are not stipulated in anticipation of their projection to be identical. Then assume complex structure for the situation argument node of the THEN clause in line with Cann (forthcoming) and Chatzikyriakidis (forthcoming). Then parsing of IF will involve a LINK transition from the lowest type e_s[3] node of the THEN clause rather than from the higher one [1]. The effect of parsing *if* is shown below:[9]

[9] Again, as in chapter 3, numbering for the type e_s nodes of the complex situation argument is going to be used.

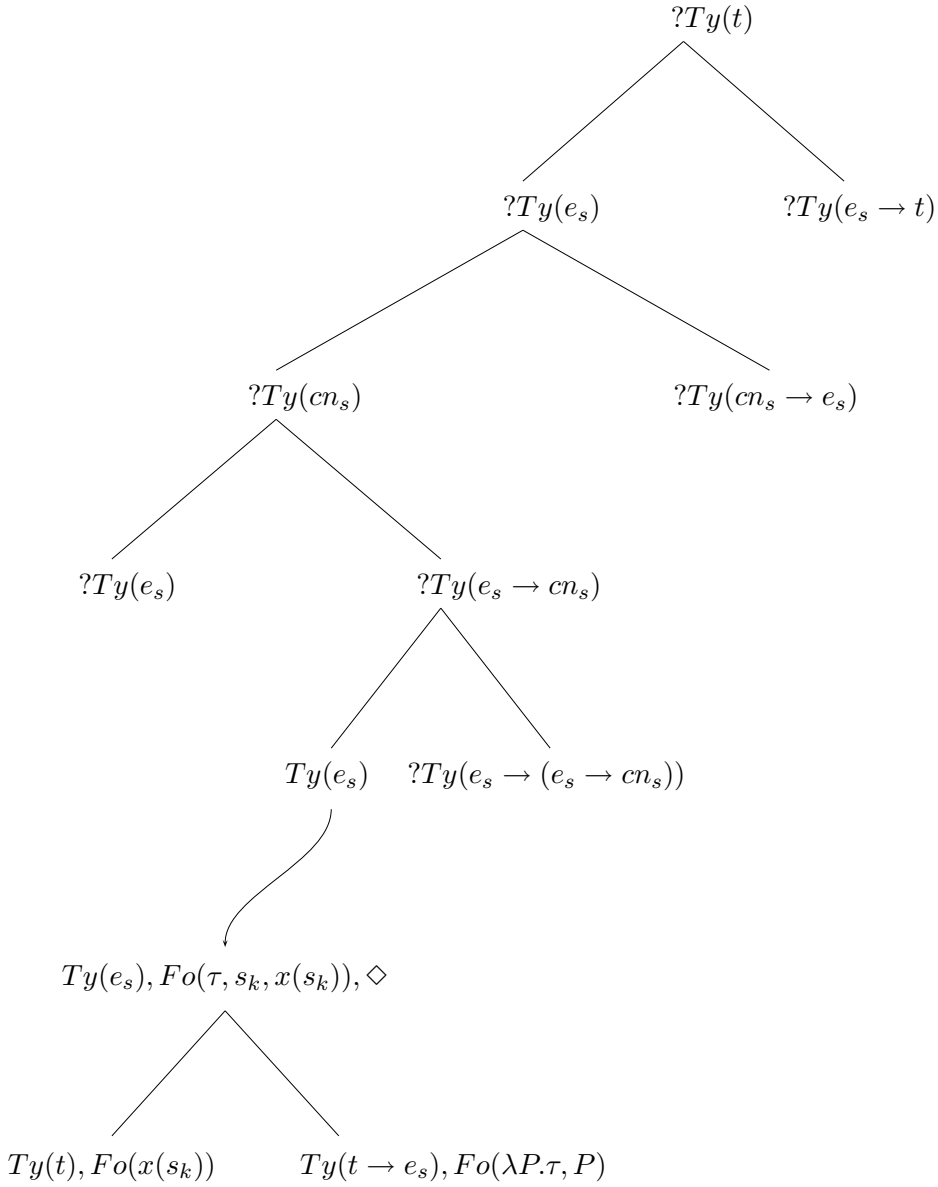
(4.50)



Assuming the IF clause is parsed and a well-formed type *t* proposition is obtained, this obtained proposition combines with the $t \rightarrow e_s$ functor node via functional application and

modus ponens (for *Fo* values and types respectively) and a formula plus a type value are projected in the higher e_s node of the LINKed tree:[10]

(4.51) After parsing the IF clause



Then, I assume that the LINK evaluation rule for subordinate clauses does not involve formula value sharing between the higher situation node of the IF clause and a e_s of the

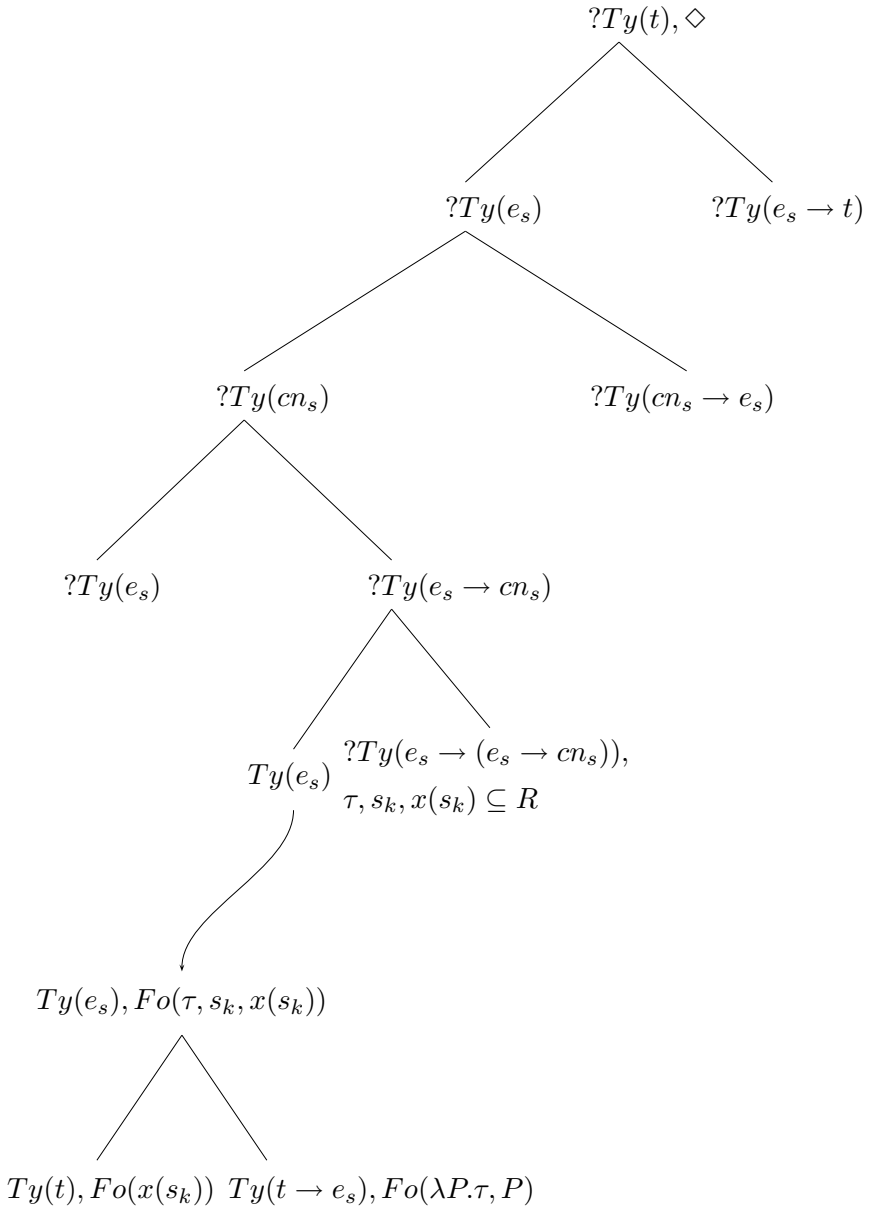
[10]The structure below the type *t* node in the IF clause is omitted for ease of exposition.

THEN clause as in Gregoromichelaki (2006: 214), but identifies the higher e_s node of the IF clause (top node of the LINKed tree) as holding at the reference time **R** of the THEN complex situation argument node. It further leaves the pointer at the top type *t* requiring node of the main tree . In that respect the situation expressed by the THEN clause will be evaluated from the point of view of the IF clause, i.e. given that the IF clause holds. The LINK evaluation rule plus the effect of the rule in tree notation is shown below:

(4.52)

$$\frac{\{\{Tn(X), ?Ty(t), ...\}, \{\langle\uparrow_1\rangle\langle\uparrow_1\rangle\langle\uparrow_0\rangle\langle\uparrow_0\rangle Tn(X)\}, \{\langle L^{-1}\rangle\langle\uparrow_0\rangle\langle\uparrow_1\rangle\langle\uparrow_0\rangle\langle\uparrow_0\rangle Tn(X), Fo(\alpha), \Diamond ...\}...\}}{\{\{Tn(X), ?Ty(t), \Diamond ...\}, \{\langle\uparrow_1\rangle\langle\uparrow_1\rangle\langle\uparrow_0\rangle\langle\uparrow_0\rangle Tn(X), Fo(\alpha) \subseteq R\}, \{\langle L^{-1}\rangle\langle\uparrow_0\rangle\langle\uparrow_1\rangle\langle\uparrow_0\rangle\langle\uparrow_0\rangle Tn(X), Fo(\alpha)\}...\}}$$

(4.53) After LINK evaluation



The proposed solution is a way of unifying the two approaches but needs however to be further checked in detail, since this is just a sketch. In any case, and whatever the proposed unified proposal will be, I believe that the subordinator must not be assumed to project the e_s type value since this type value is assumed to be projected by the verb in Cann (forthcoming). The assumption I am going to make from now on is that conditional,

temporal and cause[11] subordinators build the internal type e_s requiring node of the LINKed tree and leave the pointer at the type *t* requiring node of that same tree as in (4.50).

Returning to clitics, it seems that the relevant generalization capturing proclisis with subordinate conjunctions is the existence of a type e_s requiring node. The proposed trigger is shown below:[12]

(4.54) Situation argument type plus the unfixed node trigger (note that the | symbol stands for inclusive disjunction):

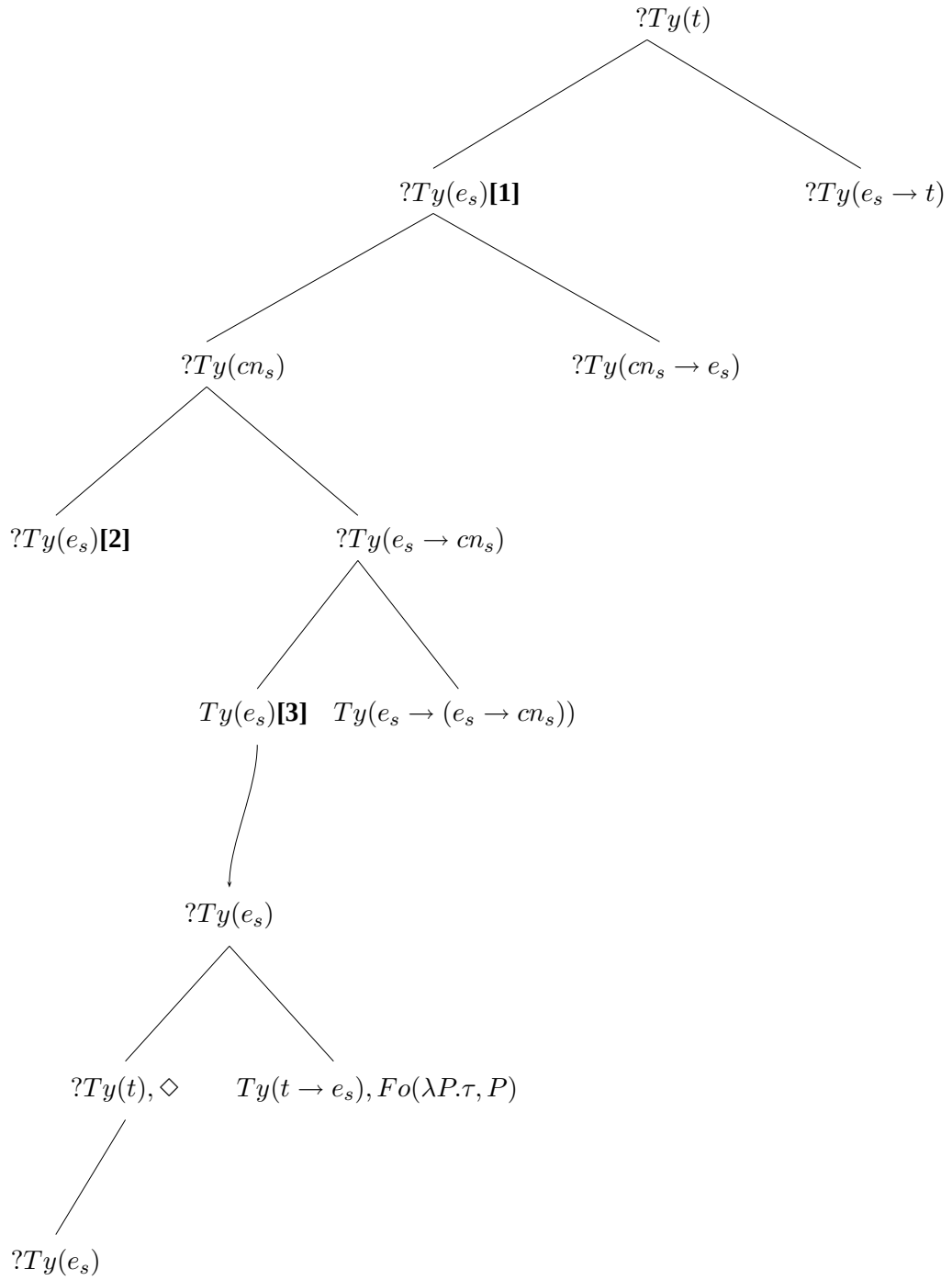
IF	$?Ty(t)$	
THEN	IF	$\langle\downarrow_*\rangle?\exists x.Tn(x))\vert$
		$\langle\downarrow_0\rangle?Ty(e_s)$
	THEN	...
	ELSE	abort
ELSE	abort	

The above trigger will correctly capture proclitic positioning in case a subordinating conjunction has already been parsed by the time the clitic comes into parse. Let us see why. Assuming that the subordinator is parsed, we get, as already said, the following structure:

[11] Variation presented with this subordinator will be dealt with later on.

[12] The way in which a potential overgeneration is caused when both a verb and a subordinate conjunction are present and enclisis instead of proclisis is derived, is going to be discussed later on when all the pieces comprising the CG lexical entry will be put together (4.8).

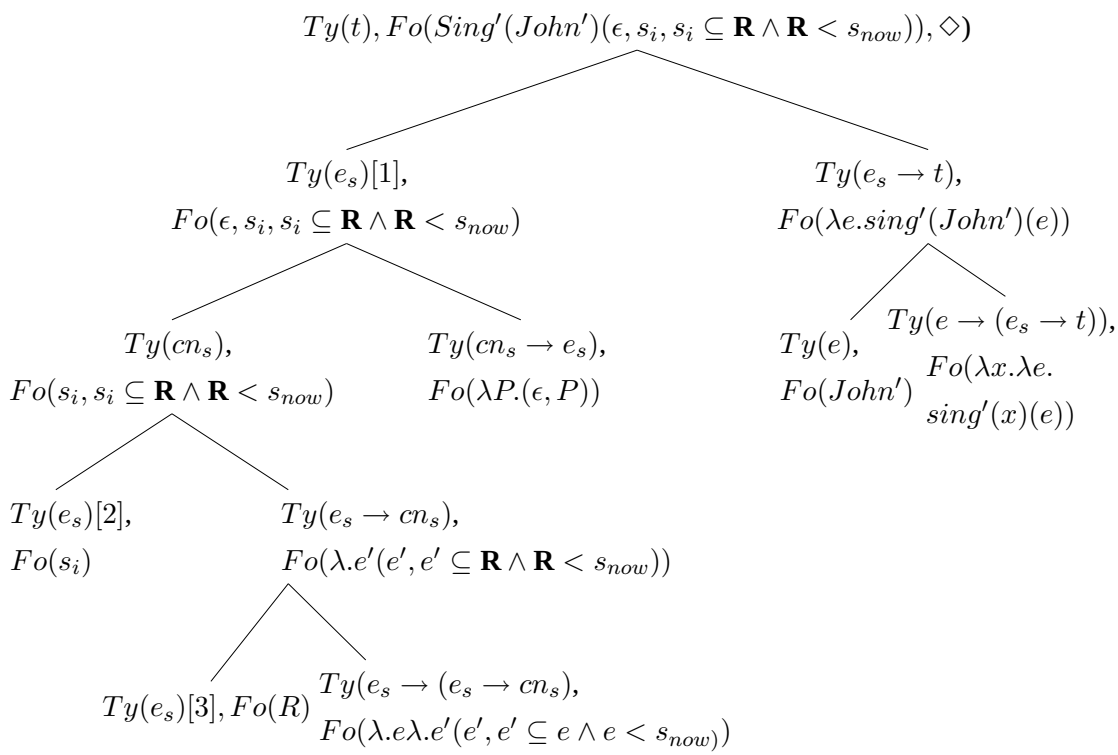
(4.55) Parsing IF



The pointer is at a type *t* requiring node and furthermore a type e_s requirement exists in the argument daughter. In that sense the clitic can be parsed. Proclitic positioning in subordinate clauses is thus correctly captured within this account.

But what about modality/tense markers? Surprisingly the account just sketched in order to capture proclisis for subordinate conjunctions will turn out to be extremely relevant for modality/tense markers as well. Let us explain by reverting to the simplest type of English example: *Mary sang*. As already said, Cann (forthcoming) and Chatzikyriakidis (forthcoming) assume that tense and aspect properties are projected inside the situation argument node. This node, following standard DS assumptions with respect to NP structure, is assumed to involve complex structure. In the repeated example below, illustrating the complete parse of the sentence *John sang*, tense properties are introduced in the lower functor node, specifying that the event described (the singing event) took place in the past (see section 2.2.2.4 for more information):

(4.56) Parsing *John sang*



Tense properties are introduced in the internal functor node ($Ty(e_s \rightarrow (e_s \rightarrow cn_s))$) node). World information, as already discussed in chapter 3, will be introduced in that node as well in case the e_s [2] requiring node of the complex situation argument node contains a formula value encoding both a situation and a world parameter. Given this analysis of tense and aspect, let us think how the lexical entries for modality/tense markers are going to look like. Without going into the exact details of what is the exact contribution of these elements in terms of tense or aspect, we can safely assume that all these elements will project this kind of information in the internal functor node. This means that all these elements will have to build this node as well, given that they appear before the verb and given that the internal functor node will not have been constructed by the time they will come into parse. This last fact has the further consequence that the higher e$_s$ [1] node has to be built as well, decorated with an e$_s$ requirement. The final consequence of all these assumptions is that when any of these elements is parsed, and assuming that the pointer is returned to the initial type *t* requiring node after parsing any of these elements, the clitic can be parsed given the trigger shown in (7.29), i.e. in the presence of a type e_s requirement. As already mentioned, the exact details of the lexical entries of these elements are not going to be fleshed out here, since whatever these details will be, the clitic positioning analysis can effectively stay the same. The relevant part common to all these entries is the construction of the e_s [1] node and its decoration with a type e_s requirement.[13] However, there is a syntactic difference between subjunctive markers *as* and *na* on the one hand and future particle *enna* on the other. Only the latter can occur with any of the subordinating conjunctions:

(4.57) **Otan/an na/as δosi...*
when/if SUBJ give.3SG-PNP
'If/When he should/must give...'

(4.58) *Otan/an enna δosi...*
when/if FUT give.3SG-PNP
'If/When he will give...'

As we have already seen, subordinate conjunctions involve the projection of the type e_s node and its decoration with a type e_s requirement. We can thus assume that the existence

[13]The interested reader is however directed to Giannakidou (2007) for a discussion involving the formal semantics of the subjunctive marker *na* plus the SMG equivalent to *enna* future particle *θa*.

of a e_s requirement or the higher situation node itself in general, is what prevents subjunctive markers but not the future marker *enna* from being parsed. The examples below show two sample lexical entries one for subjunctive markers *na/as* and one for the future marker *enna*. In the first case, the parse proceeds only in case there is no situation argument node, while such restriction is not posited for the future marker *'enna'*. Both entries build the situation argument node, decorate it with a type e_s requirement, and furthermore build the subtree to contain the $?Ty(cn_s)$, $?Ty(e_s \rightarrow cn_s)$, and $?Ty(e_s \rightarrow (e_s \rightarrow cn_s))$ nodes. Aspect/tense information will be projected on the latter node but the details of this information will not be given here (noted as ... in the lexical entry), since it is irrelevant to the content of both the thesis and the specific account given. The lexical entries for *na/as* and *enna* are shown below:

(4.59) Lexical entry for subjunctive markers *na*/*as* (excluding formal detail)

IF $?Ty(t), \langle\downarrow_0\rangle\bot$

THEN $\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(?Ty(e_s));$
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(?Ty(cn_s));$
$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \texttt{put}(?Ty(e_s \rightarrow cn));$
$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \texttt{put}(?Ty(e_s \rightarrow (e_s \rightarrow cn)), ...);$
$gofirst(?Ty(t))$

ELSE abort

(4.60) Lexical entry for future particle *enna* (excluding formal detail)

IF $?Ty(t)$

THEN $\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(?Ty(e_s);)$
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(?Ty(cn));$
$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \texttt{put}(?Ty(e_s \rightarrow cn));$
$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle); \texttt{put}(?Ty(e_s \rightarrow (e_s \rightarrow cn)), ...);$
$gofirst(?Ty(t))$

ELSE abort

Given the above entries, subjunctive markers *na/as* are predicted not to co-occur with subordinate complementizers, since their triggering restriction specifies that nothing must hold at the situation argument node in order for parsing to proceed ($\langle\downarrow_0\rangle\bot$). No such

restriction is found in the entry for the future marker *enna* and thus co-occurrence in that case is possible. What is more, proclisis is predicted to obtain in case any of these elements is parsed. This is because the lexical actions of these elements build the situation argument node and decorate it with a type e_s requirement, a requirement which, as already said, is a proclitic trigger. Thus, such an account correctly predicts proclisis to be the case with subjunctive markers *na/as* as well as the future marker *enna*.

4.6.3 The Complementizer *pu*

The factive complementizer *pu* is another environment where proclisis arises in CG.[14] The examples below exemplify the relevant facts:

(4.61)

a. *Lipume* *pu* *ton* *iδes*
be-sorry.1SG COMP him.CL-ACC saw.2SG

b. **Lipume* *pu* *iδes* *ton*
be-sorry.1SG COMP saw.2SG him.CL-ACC
'I'm sorry that you saw him.'

The rather intriguing fact about CG is that when we move to the non-factive complementizer *oti*, proclisis is no longer the only option. Instead variation is possible. The relevant examples are repeated below:

(4.62)

a. *Ipen* *oti* *eθkiavasen* *to*
said.3SG COMP read.3SG it.CL-ACC

[14]Terzi (1999a: fn 19) mentions that some of the Cypriot speakers consulted indicated that enclisis with *pu* is also possible. I got the same judgments from one of the speakers consulted. All the other speakers indicated that proclisis is the only option. In this thesis, we will assume that clitic placement is strictly proclitic with the factive complementizer *pu*. It will become evident when the analysis for the non-factive complementizer *oti* will be given, how such an account can be expanded to cover variation with *pu*, in case such variation is proven to exist in CG. See the relevant discussion as regards variation with *oti*.

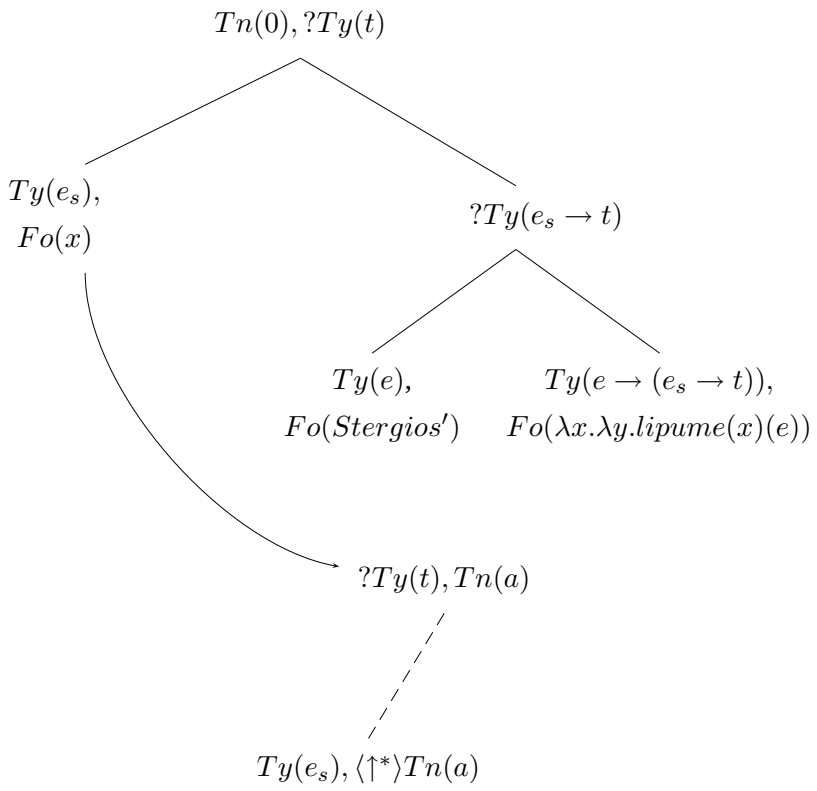
b. *Ipen* *oti* *to* *eθkiavasen*
said.3SG COMP it.CL-ACC read.3SG
'S/He said that s/he read it.'

So, how is proclisis derived in environments where factive complementizer *pu* is present? I will propose two ways of capturing the relevant facts. It will become evident on later on, which of the two is the most plausible. The first proposal follows Roussou (2007) and assumes that the difference between non-factive *oti* and factive *pu* lies in the fact that *pu* is the same lexical item with relative *pu* introducing relative clauses. In that sense, *pu* can bind either a propositional variable (the relative *Wh* reading) or an individual variable (the complementizer reading). Implementing this assumption in DS terms, complementizer *pu* in line with relative *pu* will be parsed on an unfixed node inside a LINKed tree. The LINK relation will start from a type *e* node in the relative reading, while from a type e_s node in the complementizer reading. In the former case, the type *e* node will be provided by the relativized NP, while in the latter case the e_s type node will be provided by parsing the verb subcategorizing for a *pu* complement. A consequence of this is that the unfixed node projected inside the LINKed tree will be of type *e* in the relative clause case, but of type e_s in the complement clause case.[15] A further difference between the Wh/relative construal of *pu* and its complementizer construal, is that in the latter case no requirement of the formula value of the node that the LINK relation starts must be imposed in the LINKed tree.[16] According to that view verbs with *pu* complements will be seen as subcategorizing only for a situation argument and a subject. The verbal complement will then be parsed inside a LINK structure induced from the situation argument node of the main tree. In the tree below the LINK relation is imposed, while furthermore an unfixed node is projected from the type t node of the LINKed tree. In that node *pu* is parsed:

[15]Of course, this is not a problem, since as already said, e_s is a subtype of type *e*. We can further define the ADJUNCTION rules to apply to all type *e* subtypes as well.

[16]Imposing such a requirement would mean that the event denoted by the main sentence is the same event the complement clause expresses.

(4.63) Parsing complementizer *pu*



There is the question of whether the complementizer *pu* induces a type value or just a requirement for that type.[17] Whichever analysis is chosen, the existence of the unfixed node will act as a proclitic trigger. Proclisis under that account is predicted to be the only option with *pu* in all its uses.

Another simpler and more straightforward solution as regards proclisis with *pu*, will use the same reasoning used for deriving proclisis with subordinating conjunctions. The core of such proposal will be that *pu*, in the same sense as subordinating conjunctions, will build the situation argument node decorating it with a requirement for a e_s type. The difference with subordinating conjunctions will lie in the fact that this requirement will be built in the main tree and not in a LINKed tree as it is the case with subordinating conjunctions. Such an analysis will unify the analysis of *pu* with the one given for subordinating conjunctions but will dissociate it from the relative *pu* analysis. Evidence that this account is on the

[17]The other option would be the unfixed node to carry a regular type *e* requirement. Then, what the complementizer does is to update this requirement into a type e_s requirement.

correct track, is the fact that *pu* as well as subordinating conjunctions cannot be followed by any of the subjunctive markers *na* or *as*. With the trigger of these markers being the non-existence of a situation node, parsing of *pu* will immediately block further parsing of any of these two markers according to fact:

(4.64) **Lipame* *pu* *na* *ton* *δis*
be-sorry.1SG COMP SUBJ him.CL-ACC saw.2SG
'I'm sorry that you saw him.'

(4.65) **Lipame* *pu* *as* *ton* *δis*
be-sorry.1SG COMP SUBJ him.CL-ACC saw.2SG
'I'm sorry that you saw him.'

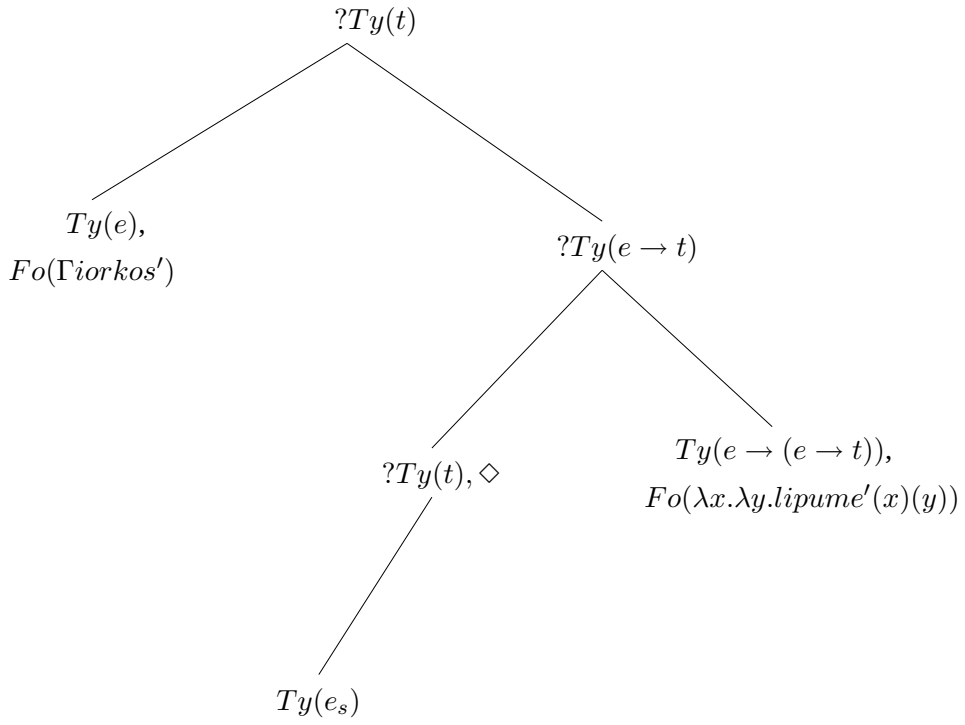
Under this analysis, the actions of the complementizer *pu* will involve the construction of the situation argument node along with its decoration with a requirement for a e_s type:

(4.66) Sample lexical entry for *pu*

IF $?Ty(t), \langle\uparrow\rangle\top$
THEN $\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(?Ty(e_s));$
$...gofirst(?Ty(t))$
ELSE abort

Parsing of *pu* will start from the embedded type *t* argument node of the verb that *pu* is a complement. In that sense, the entry specifies that the type *t* requiring node cannot be the initial node, by positing that something holds above the type *t* requiring node ($?Ty(t), \langle\uparrow\rangle\top$). Assuming that the main verb is parsed first, the effect of parsing *pu* is shown below:

(4.67) After parsing *pu* in *lipume pu ton iδes* ‘I’m sorry that you saw him.’



In the above structure, the clitic can be parsed since there is a situation type requirement in the daughter node.

4.6.4 Negation

The last proclitic environment we are going to look at is environments containing the negation particles *en/min*, ‘not’. As already said, proclisis obtains in the presence of these elements as exemplified below:

(4.68) *En* *ton* *iksero* *(*ton).*
NEG him.CL-ACC know.1SG him.CL-ACC
‘I do not know him.’

(4.69) *Min* *ton* *δis* *(*ton)*
NEG him.CL-ACC see.2SG.PNP him.CL-ACC
‘Do not see him.’

In Bouzouita (2008a,b), proclisis found in negative environments is accounted for by using a negative feature *[+NEG]*. This feature is assumed to be projected by the negation marker in the type *t* requiring node. The lexical entry given by Bouzouita (2008a,b) for Medieval Spanish (MedSp) specifies that the clitic can be parsed in case a *[+NEG]* feature exists in the type *t* requiring node:[18]

(4.70) Lexical for the Medieval Spanish clitic *lo* (taken from Bouzouita, 2008)

IF $?Ty(t)$
THEN IF $[NEG+]$ $\vee$
$(\langle\downarrow_*\rangle Fo(\alpha), ?\exists x.Tn(x))$ $\vee$
$?\exists x.Tns(x)$
THEN make($\langle\downarrow_1\rangle$); go($\langle\downarrow_1\rangle$);
make($\langle\downarrow_0\rangle$); go($\langle\downarrow_0\rangle$);
put($\mathbf{U}, Ty(e), ?\exists x.Fo(x),$
$[\downarrow]\bot, ?\langle\uparrow_0\rangle Ty(e \rightarrow t)$)
ELSE abort
ELSE IF $?Ty(e), \langle\uparrow\rangle\top$
THEN IF $\langle\uparrow_0\uparrow_1^*\rangle^+(?Ty(t), [NEG+], \exists x.Tns(x))$ $\vee$
$\langle\uparrow_0\uparrow_1^*\rangle^+(?Ty(t), \langle\downarrow_*\rangle(Fo(\alpha), ?\exists x.Tn(x)),$
$\exists x.Tns(x))$ $\vee$
$\langle\uparrow_0\uparrow_1^*\rangle^+(?Ty(t), ?\exists x.Tns(x))$
THEN abort
ELSE put($\mathbf{U}, Ty(e), ?\exists x.Fo(x),$
$[\downarrow]\bot, ?\langle\uparrow_0\rangle Ty(e \rightarrow t)$)
ELSE abort

Such a feature is used as a diacritic in Bouzouita's analysis, since a proper analysis of negation in DS is pending (Bouzouita, 2008: 221). The trigger capturing negation is shown below:

[18]Fonts as in Bouzouita, 2008a.

(4.71) Trigger capturing proclisis with negation particles

IF	$?Ty(t)$	
THEN	IF	$[+NEG]$
	THEN	...
	ELSE	abort
ELSE	abort	

If this new proclitic trigger is added, then the lexical entry for clitics in CG will contain three proclitic triggers: a) the unfixed node trigger b) the type e_s requiring node and c) the negation trigger. All three proclitic triggers are shown below:

(4.72) All three proclitic triggers

IF	$?Ty(t)$	
THEN	IF	$\langle\downarrow_*\rangle ?\exists x.Tn(x))\vert$
		$\langle\downarrow_0\rangle ?Ty(e_s)\vert$
		$[+NEG]$
	THEN	...
	ELSE	abort
ELSE	abort	

However, there might be a way to avoid the *[+NEG]* trigger and still be able to capture negation. There are two types of negation in CG, one used in indicative and one in subjunctive or negative imperative environments:

(4.73) *En/*min* *ton* *iδa*
NEG him.CL-ACC saw.1SG
'I did not see him.'

(4.74) *Θelo* *na* *min/*en* *pais*
want.1SG SUBJ NEG go.PNP
'I want you not to go/ I do not want you to go.'

(4.75) *Min* *pais*
NEG go.PNP
'Do not go.'

Furthermore, the negation marker *min* when combined with the Perfective Non Past (PNP) verbal form can give rise to a negative imperative reading. In that sense, sentence (4.75) is also possible without the subjunctive marker *na* as witness the example below:

(4.76) *Na* *min* *pais*
SUBJ NEG go.PNP
'Do not go.'

It seems that the different types of negation interact with modality/tense-aspect in different ways. Giving an analysis of negation in CG might then involve the contribution of a number of restrictions as regards modality or aspect by the two different negation markers that will capture the different semantics in each case as well as the combinatorial properties of the two. Of course, a study of what these restrictions actually are and how to formally encode them is well beyond the scope of this thesis. However, given that modality/tense-aspect information is encoded inside the complex situation argument node, it is quite straightforward to assume that the negation markers will project their restrictions inside this node. Whatever the exact locus of the projection of this information (presumably the lower functor node), negation markers will have to built the higher type e_s node. In that respect, the lexical entry for negation markers will have to project at least a *[+NEG]* feature as well as the type e_s requiring node. The lexical entry is shown below:

(4.77) Lexical entry for negation

IF	$?Ty(t)$
THEN	$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$
	$\texttt{put}(?Ty(e_s)), ...;$
	$gofirst(?Ty(t)), \texttt{put}([+NEG])$
ELSE	abort

Given the above entry, we do not have to refer to the *[+NEG]* feature to capture negation, since proclisis with negation will be captured by the same mechanisms used for modality/tense markers as well as subordinating conjunctions, namely the type e_s requiring trigger. In that sense, the set of proclitic triggers remain two:

(4.78) Proclitic triggers (revised)

IF	$?Ty(t)$	
THEN	IF	$\langle\downarrow_*\rangle ?\exists x.Tn(x))\vert$
		$\langle\downarrow_0\rangle ?Ty(e_s)$
	THEN	...
	ELSE	abort
ELSE	abort	

With negation, I stop the discussion on proclitic environments in CG. The next step is to look at the so-called variation environments, i.e. environments that are associated with both proclisis and enclisis.

4.7 Variation environments

4.7.1 The Case of *oti*

One of the facts that make clitic positioning in CG rather complex, is the existence of environments where variation in clitic positioning is possible. The most striking of these environments is the case of the non-factive complementizer *oti* 'that'. This complementizer, unlike its factive counterpart discussed above, can trigger both proclisis and enclisis. The relevant facts are repeated below:

(4.79)

a. *Ipen* *oti* *eθkiavasen* *to*
said.3SG COMP read.3SG it.CL-ACC

b. *Ipen* *oti* *to* *eθkiavasen*
said.3SG COMP it.CL-ACC read.3SG
'S/He said that s/he read it.'

Proclisis with *oti* is straightforwardly captured assuming an analysis in line with the second approach I proposed for complementizer *pu*, i.e. an analysis where *oti* builds the situation argument node and decorates it with a type e_s requirement:

(4.80) Sample lexical entry for *oti*

IF $?Ty(t), \langle\uparrow^*\rangle\top$

THEN $\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(?Ty(e_s));$
$...gofirst(?Ty(t))$

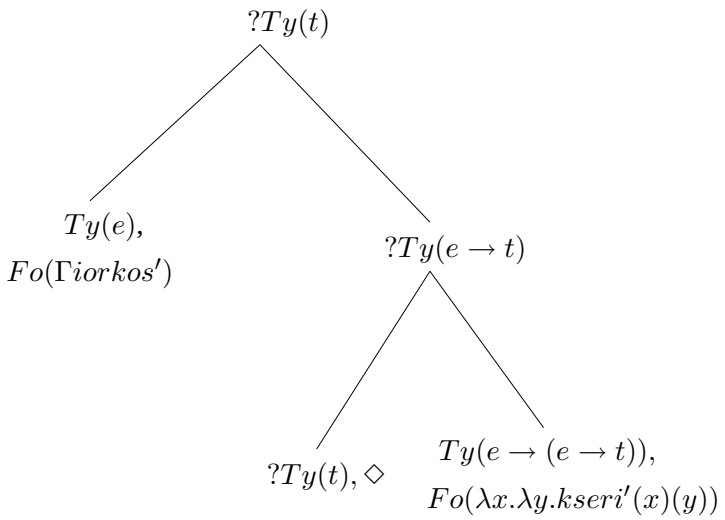
ELSE abort

Such a unifying treatment for both *oti* and *pu* will lead us to choose the second solution posited for *pu* in the previous section. In this approach, the reason for proclisis with *oti/pu* and subordinating complementizers will be the existence of a type e_s requiring node. However, only half the problem is solved, since *oti* can also involve enclitic positioning. Nothing in the entry we have given above will predict this fact. So, what can explain enclitic positioning with *oti*? In order to provide an answer for the latter question, let us look at MMG and Medieval Cypriot Greek (MCG). Clitic positioning with *oti* in MMG is prevalently postverbal, with 30 of the 38 tokens showing postverbal positioning and only 8 of them preverbal (Pappas, 2004). The situation in the Cypriot *chronicles of Makhairas* and *Boustronios* with respect to clitic positioning of *oti* is quite harder to tell. Pappas (2004) does not give a count for *oti* for MCG. A first look at the *chronicle of Makhairas*, reveals a number of cases involving *oti*. However, in their vast majority, these cases involve the use of *oti* as a causal subordinator and not as a non-factive complementizer. However, in both its guises, *oti* can be found with both proclisis and enclisis, with enclisis being prevalent in both cases (5/8 for non-factive *oti* and 36/40 for causal *oti*). I do not know if and how positioning of causal *oti* has influenced positioning with complementizer *oti* or vice versa. What is rather interesting is Pappas' (2004: 36) discussion concerning the classification of *oti* in the category clause initial. Pappas provides two interesting views on the function of *oti*, one given by Mackridge (1985) and one given by Jannaris (1968). Both these writers, argue that *oti* is actually more of a coordinating conjunction rather than a subordinating one. Mackridge states that *oti* shows a pattern of "incomplete subordination" while Jannaris (1968) that in many instances *oti* "corresponds to our modern colon(;)...". The use of *oti* as a coordinating rather than a subordinating conjunction might give us a way out of the variation problem. We can thus assume that *oti*, can be also parsed as a coordinating conjunction. In this scenario, *oti* will build a LINK transition in the same sense coordinating conjunctions do in DS. There are a number of issues with respect to such a transition.

Coordinating conjunctions induce this LINK transition from a type complete to a type requiring node of the same type. However, in the case of *oti* things are different, since no complete type *t* node will exist by the time *oti* will come into parse. This is because a main verb like *ksero* 'know' will subcategorize for an NP subject and a VP object. Therefore, when the complementizer *oti* will come into parse in a sentence like (4.81), no complete type *t* node will exist, since the value for the VP object (the internal type *t* requiring node) will have not yet been provided:[19]

(4.81) *O* *Γiorkos* *kseri* *oti* *kseris* *ton*
the George know.3SG COMP know.2SG him.CL-ACC
'George knows you know him.'

(4.82) Before parsing *oti* in *o Γiorkos kseri oti kseris ton*



The pointer is at the internal type *t* requiring node. The subject has been parsed but no complete type *t* can be given until the embedded clause is parsed. Thus, an analysis where *oti* projects a LINK structure from a type complete node will not work. What we need is an analysis that posits that the LINK relation must start from a type *t* requiring node rather than a type complete one. This node will then be the embedded type *t* requiring node, since the pointer will be at that node when *oti* will come into parse and there is no way of moving the pointer up via recursive applications of COMPLETION, since no type is satisfied in the

[19] Again, the situation nodes are omitted of ease of exposition.

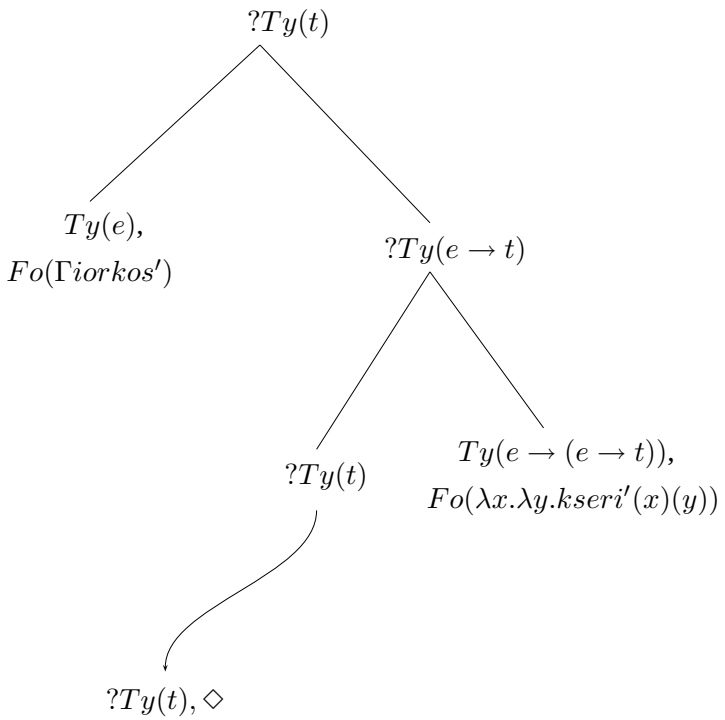
internal type *t* requiring node. In that sense I assume that *oti* builds a LINK transition from the internal type *t* requiring node to a type *t* requiring node as well. The lexical entry for *oti* is shown below (the relevant part being the disjunctive THEN part):

(4.83) Final lexical entry for *oti*

IF	$?Ty(t), \langle\uparrow^*\rangle\top$
THEN	$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(?Ty(e_s))$
	$...gofirst(?Ty(t))\vert$
	$\texttt{make}(\langle L\rangle); \texttt{go}(\langle L\rangle); \texttt{put}(?Ty(t))$
ELSE	abort

Parsing *oti* will produce the following structure:

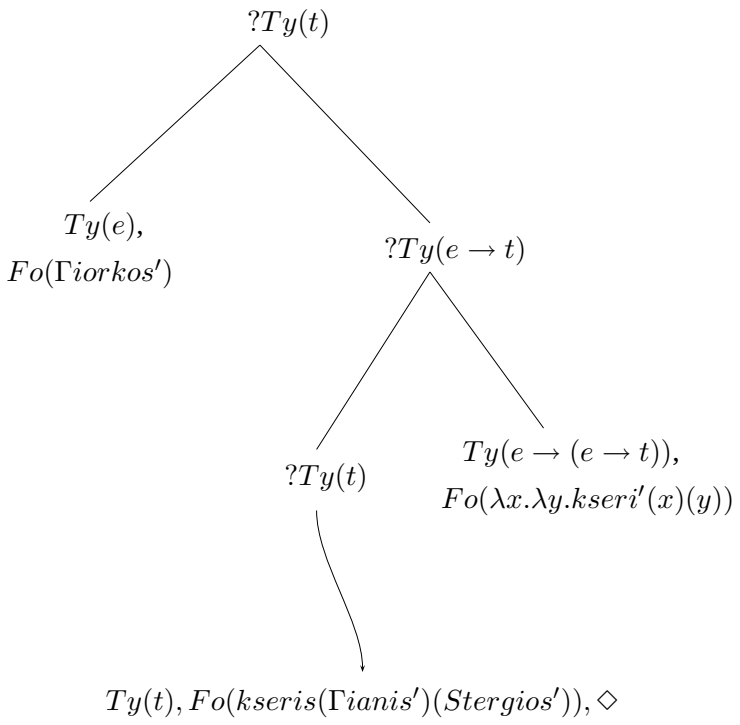
(4.84) After parsing *oti* as a LINK



The pointer is at the type *t* requiring node of the LINKed tree. The embedded clause is parsed, and then the result of parsing the embedded clause is copied to the type *t* requiring

node of the main tree via means of a LINK evaluation rule. The tree below shows the situation after the embedded clause in (4.81) has been parsed:[20]

(4.85) After parsing the embedded clause



At that point, the LINK evaluation rule copies the formula value found in the top node of the LINKed tree and copies it to the node where the LINK starts:

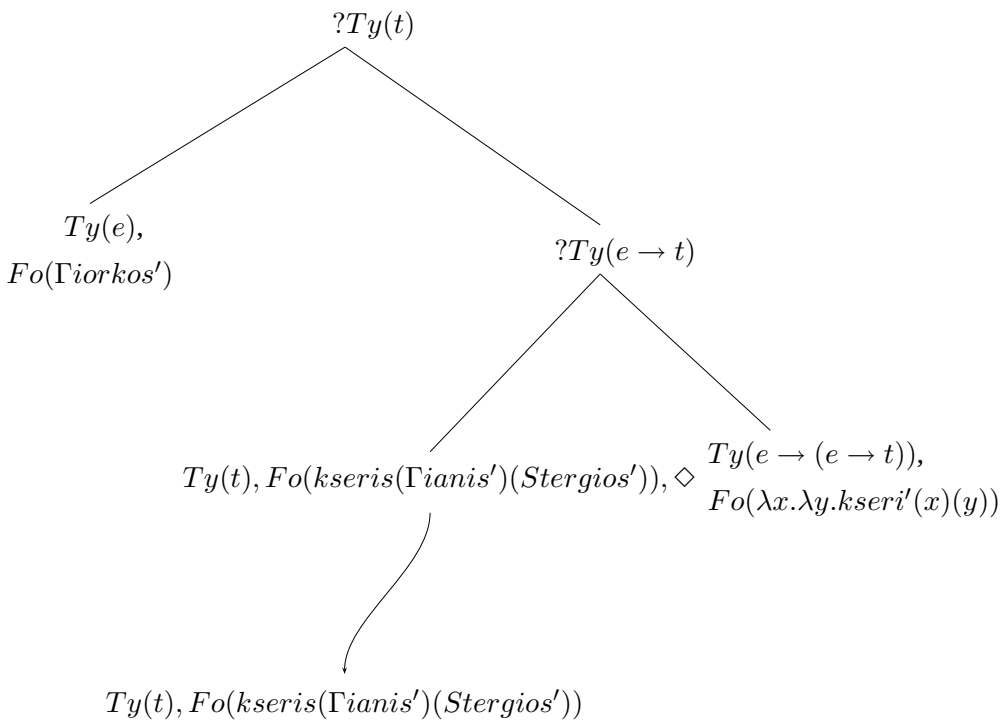
(4.86) LINK evaluation rule for *oti*

$$\frac{\{\{Tn(X), ?Ty(t), ...\}, \{\uparrow^* Tn(X), ?Ty(t), Tn(Y)\}, \{\langle L^{-1}\rangle Tn(Y), Ty(t), Fo(\alpha), \Diamond ...\}...\}}{\{\{Tn(X), ?Ty(t), ...\}, \{\uparrow^* Tn(X), Ty(t), Tn(Y), Fo(\alpha), \Diamond)\}, \{\langle L^{-1}\rangle Tn(Y), Ty(t), Fo(\alpha), ...\}...\}}$$

After LINK evaluation the structure we get is the following:

[20]The subject metavariable is substituted by the value *Stergios'* while the clitic metavariable by the value Γ*ianis*.

(4.87) After LINK evaluation



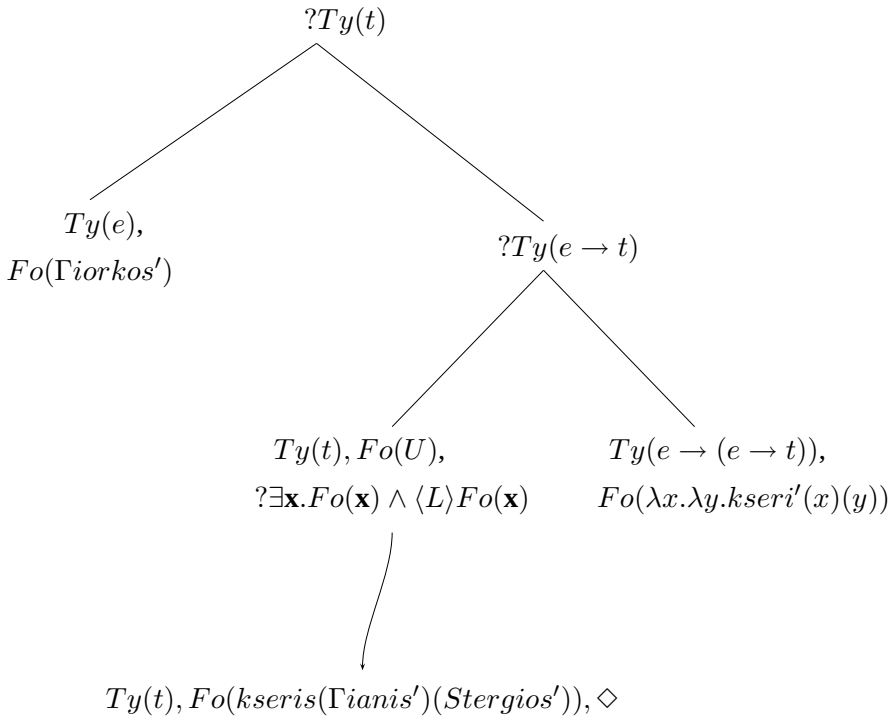
An alternative approach that would avoid the use of a LINK evaluation rule, would be to assume that *oti* decorates the embedded type *t* requiring node with a type *t* value and a formula metavariable. The requirement for the proper formula value will then posit that this proper value must be the same as the one on the top node of the LINKed tree. This alternative lexical entry for *oti* is shown below:

(4.88) Final lexical entry for *oti*

IF	$?Ty(t), \langle\uparrow^*\rangle\top$
THEN	$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle); \texttt{put}(?Ty(e_s))$
	$...gofirst(?Ty(t))\vert$
	$\texttt{put}(Ty(t), Fo(U), ?\exists\mathbf{x}.Fo(\mathbf{x}) \wedge \langle L\rangle Fo(\mathbf{x}))$
	$\texttt{make}(\langle L\rangle); \texttt{go}(\langle L\rangle); \texttt{put}(?Ty(t))$
ELSE	abort

Given this lexical entry, the result of parsing (4.81) using the LINK strategy is shown below:

(4.89) Parsing *o* Γ*iorkos kseri oti kseris ton* 'George knows that you know him.'



The only possible substituent given the situation above is the Fo formula value of the top node of the LINKed tree ($Fo(kseris(\Gamma ianis')(Stergios'))$). Given substitution, we get the correct semantic results without using the LINK evaluation rule.

Treating *oti* both as a regular complementizer and as a coordinating conjunction will predict variation to be possible.[21] In what follows, variation with causal subordinators is examined.

4.7.2 The Case of epiδi

Subordinate clauses introduced with the causal subordinator epiδi (γiati as well) also exhibit variation in CG. The relevant data are repeated below:

[21]For those speakers that accept variation with *pu*, it can be argued that parsing of *pu* as a coordinating conjunction is also possible, possibly in analogy to non-factive *oti*.

(4.90)

a.	*Epiδi*	*enevriases*	*me,*	*en*	*kamno*	*tipota*
	because	made-angry.2SG	me.CL-ACC	NEG	do.1SG	nothing
b.	*Epiδi*	*me*	*enevriases,*	*en*	*kamno*	*tipota*
	because	me.CL-ACC	made-angry.2SG	NEG	do.1SG	nothing

'I'm not doing anything because you pissed me off.'

The variation attested with *epiδi* is problematic, since one would expect epiδi to pattern with the rest of the subordinate conjunctions in terms of clitic positioning and exhibit only proclisis rather than variant positioning. Thus, assuming an analysis similar to the one I have argued for the rest of the subordinating conjunctions, capturing proclisis is the easy case. Under such an analysis, the cause subordinator will project a type e_s requirement that will function as the clitic's proclitic trigger. So far, so good. But, what about enclisis? How is enclisis going to be captured? In order to answer this question, it would be good to look at the properties of the causal subordinator in question or causal subordinators in general. The first thing worth mentioning is that this variant positioning exhibited by *epiδi* has been also attested for the equivalent subordinator in Medieval Spanish (Bouzouita, 2008a). The cause subordinator *ca* 'because' in MedSp shows the same characteristics, exhibiting both proclisis and enclisis (examples and bold in examples from Bouzouita, 2008a):

(4.91)

E	*tu,*	*entra*	*en*	*el*	*archa*	*e*	*todo*	*to*	*casado,*
and	you	enter.2SG	in	the	ark	and	all	your	family
<u>ca</u>	***te***	*vi*	*iusto*	[...]					
because	CL	saw.1SG	just						

'And you and your family, enter in the ark, because I saw you are a just man.'
(*Faz.*: 117)

(4.92)

siempre	*oy*	*dezir*	*que*	*era*	*buen*	*seso*	*atener*	
always	heard.1SG	say	that	was.3SG	good	brain	pay-attention	
se	*omne*	*a*	*las*	*cosas*	*çiertas*	*&*	*non*	*a*
CL	man	to	the	things	certain	and	not	to
ca	*muchas*	*vezes*	*a*	*los*	*que*	*se*	*atienen*	*a*
because	many	times	to	the	that	CL	pay-attention.3PL	to
las	*fuzas*	*contesçe*	***les***	*lo que*	*contesçio*	*a*	*donna*	
the	fantasies	happens.3SG	CL	what	happened.3SG	to	lady	
truana								
Truhana								

‘[...] I always heard that a prudent man pays attention to certainties than to fantasies because what happened to lady Truhana happens often to those who pay attention to fantasies.’ (*Luc.*: VII)

Miriam Bouzouita (2008a: chapter 3), following a generally accepted line of view (Lapeza 1978; Menendez Pidal 1980), notes that sentences introduced with *ca* can function as either subordinate or coordinate sentences. The reason for this structural duality can be traced back to Latin where two different kinds of causal relations existed, one pattering with coordinate and one with subordinate structures. The same structural duality is found in Ancient Greek (AG, Tzartzanos, 1940). It is worth mentioning that the variation environments attested with *epiδi* in CG are also attested in MCG, so variation is not an innovation of the modern CG system. A quick check in the *chronicle of Makhairas* returned four instances of clitics appearing in *epiδi* clauses. Half of them exhibited proclisis and the other half enclisis, so variation with *epiδi* was already at stake in MCG. This kind of evidence will allow us to pursue an analysis of *epiδi* as a coordinating conjunction in the same sense we did for the *oti* complementizer.[22] What I propose is that *epiδi*, just like *oti* can be parsed as both a subordinating and a coordinating conjunction. In the first case, parsing *epiδi* will give us a type e_s requirement, and thus proclisis will obtain. In the second case, *epiδi* will

[22]Another way to explain the variant positioning exhibited with causal subordinates is to attribute it to variant positioning found with the complementizer *oti*. Given that *oti* could be either a non-factive complementizer or a subordinate of cause conjunction, variation exhibited by non-factive *oti* could spread to causal *oti* and from causal *oti* to subordinates of cause in general. However, no data are available to vindicate such a claim and as such at the moment it cannot be anything more than a claim.

be parsed as a coordinating conjunction where no such a requirement will exist. Actually, the analysis of *epiδi* as a coordinating conjunction will be much closer to the DS analyses of classic coordinating conjunctions like *and* or *but* than the analysis provided for *oti*. This is because *oti* was assumed to build a LINK relation from a type *t* requiring node to another type *t* requiring node or, according to the alternative analysis provided, the LINK relation was created from a type complete node but with no complete formula value (metavariable). This is unlike the analyses of coordinate conjunctions like *and* or *but* where the LINK relation is built from a type and formula complete node into a type *t* requiring node. In the case of *epiδi* however, one can maintain an analysis where the causal conjunction builds a LINK relation from a type and formula complete node to a type *t* requiring one. Let us see how this will be done. Say we want to parse a subordinate of cause structure involving a clitic in enclitic position followed by the main clause:

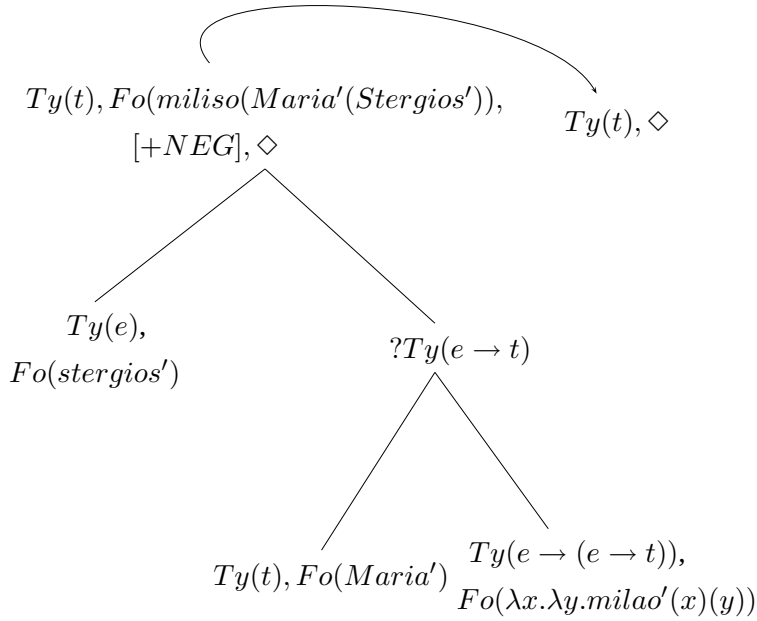
(4.93)

En	*su*	*milao,*	*epiδi*	*enevriases*	*me*
NEG	you.CL-GEN	talk	because	made-angry.2SG	me.CL-ACC

'Because you made me angry, I'm not talking to you.'

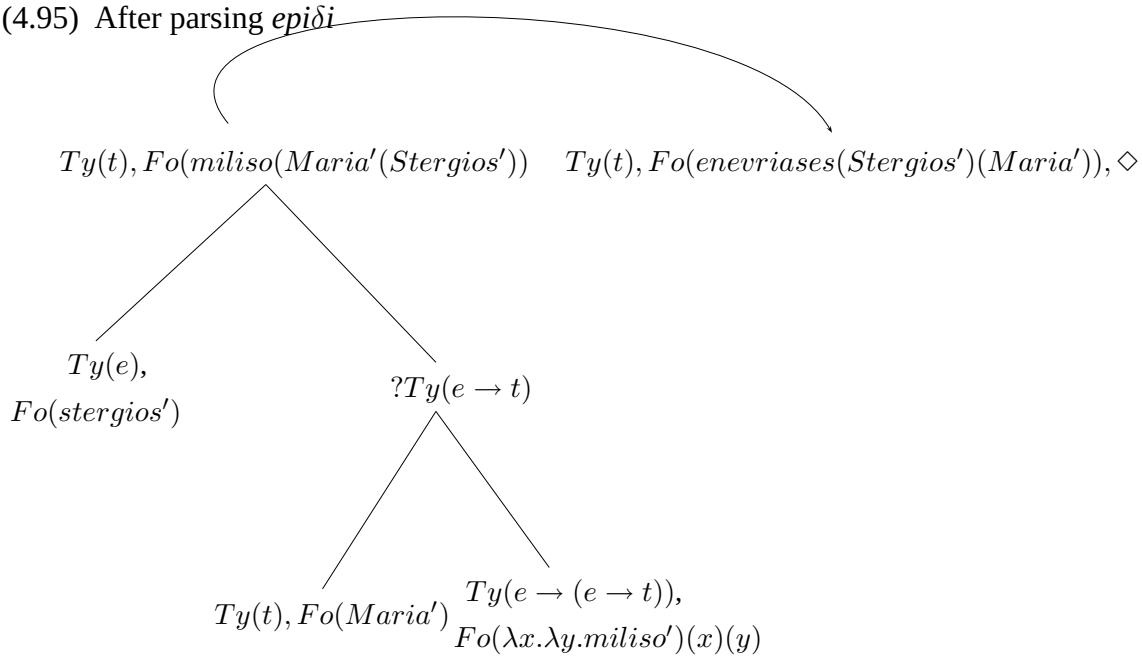
First we parse the main clause obtaining a type *t* and a formula value in the initial node. At that point epiδi comes into parse and creates a LINK relation from the type *t* complete node to a type *t* requiring node and leaves the pointer at the latter node:

(4.94) After parsing *epiδi*



The pointer is left at the type *t* requiring node of the LINKed tree. No proclitic trigger exists and thus parsing of the clitic is impossible. On the other hand, the verb can be parsed since its triggering point ($?Ty(t)$) is satisfied, and now parsing of the clitic is possible giving rise to enclitic positioning:

(4.95) After parsing *epiδi*



At that point the LINK evaluation rule for causal subordinates applies. This LINK evaluation rule will be minimally different to the one assumed for coordinate conjunctions like *and* or *but* in that it links the two sentences via a causal relation:

(4.96) LINK evaluation rule for *epiδi*

$$\frac{\{\{Tn(X), Ty(t), Fo(\alpha), ...\}, \{\langle L^{-1}\rangle Tn(X), Ty(t), Fo(\beta), \Diamond ...\}...\}}{\{\{Tn(X), Ty(t), Fo(\alpha) \leftarrow Fo(\beta), \Diamond, ...\}, \{\langle L^{-1}\rangle Tn(X), Ty(t), Fo(\beta), ...\}...\}}$$

The $\leftarrow$ symbol is taken to encode the causal relation.[23] Formally $Fo(\alpha) \leftarrow Fo(\beta)$ reads as: $Fo(\alpha)$ has a cause in case $Fo(\beta)$ is true. In simpler terms, $Fo(\alpha)$ is the cause of $Fo(\beta)$.[24] The analysis of causal subordinates in line with coordinating conjunctions will give us enclisis than proclisis. Given that subordinates of cause can be also parsed as regular subordinates, variation with *epiδi* is predicted to be possible.[25]

To recapitulate, proclisis with epiδi will be captured in the same sense as in the case of subordinate conjunctions, namely because of the projection of an e_s requirement by the subordinator. I further assume that *epiδi* can be also parsed as a coordinating conjunction. In this last case no situation argument is projected and thus enclisis is the only option in the presence of a clitic. Given this, variation with epiδi is correctly captured within such an account.

4.7.3 Complex Conjunctions/marker with *tze*

Now, I will take a look at one of the most puzzling positioning restrictions with respect to CG clitics. This restriction involves a number of complex subordinators and the complex negation marker 'en tze'. All these complex subordinators/markers are consisted of two parts. The first part is a marker or subordinate conjunction, while the second part is the coordinate conjunction *tze*. The rather intriguing fact about all these complex elements is that they fail to trigger proclisis unlike the elements that comprise the first part of the complex in isolation do. The relevant data are shown below:

[23] Taken from Litschitz (1998).

[24] The formal details of what a causal relation is are not going to be discussed here. The interested reader is directed to Lewis (1986) for a discussion of causal relations and Liftschitz (1998), Shafer (1998) for causal logic models using situation calculus and event trees respectively.

[25] The same reasoning applies to variation found with the causal subordinate *γiati*, 'why'.

(4.97)

a. *En tze iδa ton*
NEG and saw.1SG him.CL-ACC

b. **En tze ton iδa*
NEG and him.CL-ACC saw.1SG
'I did not see him.'

a. *An tze iδa ton...*
if and saw.CL-ACC him.CL-ACC...

b. **An tze ton iδa*
if and him.CL-ACC saw.CL-ACC...
'Even though I saw him...'

Both negation *en* and the conditional conjunction *ean*, as discussed earlier, categorically trigger proclisis. What is even more puzzling is that *en tze* is used to denote emphatic negation in CG. Thus in this case, two elements that in general trigger proclisis in CG, i.e. a negation marker and an emphatic element, fail to trigger proclisis. The culprit behind this behavior seems to be the coordinating conjunction *tze*, which in a way cancels triggering of preverbal positioning. Indeed, this is the analysis already pursued in Agouraki (2001), where *en tze* is argued to involve conjunction of two CPs. The negation marker occupies the head of the CP1 phrase, while conjunction *tze* occupies the SpecCP2, the verb occupying $CP2_0$. It is very hard to account for enclisis in these constructions without resorting to stipulations. The problem is that coordinating conjunctions in general do not trigger enclisis but rather do not constitute a trigger for proclisis. In the cases where a coordinating conjunction is present, enclisis obtains in the absence of any proclitic triggers and not due to the presence of the coordinating conjunction. Furthermore, the coordinating conjunction can be part of sentences exhibiting proclisis in case any of the proclitic triggers is present in the sentence:

(4.98) *Tuto en aliθja tze en ton aδiko pu alakse γnomi*
this is truth and NEG him.CL-ACC blame COMP changed.3SG opinion
'This is true and I do not blame him for changing his mind.'

(4.99) *Men* *lalis* *pelares* *tze* *enna* *me* *anagasis* *na* *mino*
NEG talk.2SG rubbish and FUT me.CL-ACC force SUBJ stay.1SG
spiti
home
'Do not talk rubbish and force me to stay at home.'

In giving an analysis of *en tze*, let us first look at the elements comprising the complex negation particle *en tze* and see what the existing DS analyses are, if any, with respect to each of these elements separately. The negation marker has already received an analysis in this chapter. According to this analysis the negation marker builds the situation argument node and then returns to the type *t* requiring node where it further projects a *[+NEG]* feature on that node. On the other hand, coordinating conjunctions, as already discussed in chapter 2, have traditionally been characterized in DS as LINKed structures, the mainstream assumption being that they induce a LINK relation from a type complete node to a node of the same type (Cann et al., 2005). The reasoning behind such treatment is that coordinating conjunctions start a new domain, be it sentential, verbal or nominal. However, no actions are induced in that same domain by coordinating conjunctions in contrast to subordinate conjunctions which further introduce structure in that new LINKed domain. In that sense, a clitic following a coordinating conjunction is a sentence initial clitic, since it is the first element inducing structure in the domain. The reason that proclisis cannot obtain with *en tze,* and enclisis is only possible might very well be attributed to this latter fact, i.e. that *tze* actually starts a new domain different than the one *en* is parsed. In that sense, the situation argument node projected by the negation marker is parsed in a different domain, given that *tze* will induce a LINK transition from the tree where negation is parsed to another tree structure. The first thing we need to look at in formalizing this assumption, is the node from which the LINK transition associated with *tze* is going to be created. When *tze* comes into parse in a sentence like (4.96.a), the only word already parsed will be the negation operator *en*:

(4.100) After parsing *en*

$$?Ty(t), [+NEG], \diamond$$

$$?Ty(e_s)$$

The pointer returns to the type *t* requiring node. The problem is that now, no node is type complete. In that sense, the assumption that coordinating conjunctions project a LINK relation from a type complete node to a node requiring the same type cannot be maintained in this case. However, it may seem possible to assume that *tze* is in fact inducing a LINK relation from a type *t* requiring node node to another type *t* requiring node. The other way to look at the problem is to assume that these complex elements are parsed as one element and not separately. Such an assumption is not implausible at all, given that nothing can intervene between the two elements in all these constructions and furthermore, while the two elements can appear on their own, the interpretation that they give rise to in combination is not a compositional accumulation of the content of the entries of each of the separate elements. For example, *an tze* 'even though', cannot be captured assuming an analysis where the two elements are treated as separate entries, since in this case what we get is the actions induced by the conditional *an* plus the actions of the coordinating conjunction *tze*, which by no means capture the semantics of concession. Given these facts, it is quite natural to assume one lexical entry for the whole complex for elements like these. Given an analysis for *en tze* as involving one lexical entry, we can assume that *en tze* first marks the type *t* requiring node with the *[+NEG]* feature and further creates a LINK transition to a type *t* requiring node. This new LINKed domain will be the domain in which the rest of the sentence is going to be parsed. What I am assuming is that the sentence is parsed within the context of a negative specification, within the context of negation. In that sense, such an analysis is quite close to the DS analysis of HTLD or relative clause structures. Left dislocated arguments in HTLD constructions or the relativized elements in relative clauses can be seen as setting the context in which the rest of the clause in HTLD structures or the relative clause in relative clause structures are going to be parsed. In the same sense, *en tze* can be seen as providing the context (a negative context) in which the

sentence is to be parsed. The lexical entry for *en tze* is shown below:

(4.101) Lexical entry for *en tze*

IF	$?Ty(t)$
THEN	$\texttt{put}([+NEG]); \texttt{make}(\langle L \rangle);$
	$\texttt{go}(\langle L \rangle); \texttt{put}(?Ty(t))$
ELSE	abort

The result in tree notation is shown below:

(4.102) After parsing *en tze*

$$?Ty(t), [+NEG] \qquad ?Ty(t)$$

Notice that no type e_s requiring node is assumed to be projected by *en tze*. This is done for reasons of simplicity only, since it does not make any difference in the analysis presented, given that the type e_s requiring node will be projected in the domain where the LINK starts, i.e. in a tree different to the one where the clitic is going to be parsed as we will see. This type e$_s$ requiring node cannot constitute a proclitic trigger, as the clitic will be parsed in a different domain, namely the domain of the LINKed tree. After parsing *en tze* the pointer is at the type *t* requiring node of the LINKed tree. Let us say we want to parse *en tze iδa ton* 'I did not see him'. *En tze* is parsed first giving us the structure just discussed. Then, the rest of the sentence is parsed in the LINKed type *t* requiring domain. Assuming that all metavariables (subject and object) have been substituted by values from the context, we end up with the following structure:

(4.103) After parsing *en tze iδa ton*

$$?Ty(t), [+NEG] \qquad Ty(t), Fo(i\delta a(\Gamma iani)(Stergios'))\Diamond$$

At this point, a LINK evaluation rule will take place. This LINK evaluation rule will copy the formula and type of the LINKed tree to the tree where where the LINK starts:

(4.104) LINK evaluation rule for *en tze*

$$\frac{\{?Ty(t), +[NEG], \langle L\rangle Ty(t)\}, \{Fo(x), Ty(t), \langle L^{-1}\rangle Ty(t)\Diamond, ...\}}{\{Fo(x), Ty(t), [+NEG], \Diamond\}, \{Fo(x), Ty(t)\}}$$

After LINK evaluation applies the structure we get is the following:

(4.105) After parsing *en tze iδa ton* 'I did not see him'

$$Ty(t), Fo(i\delta a(\Gamma iani)(Stergios')), [+NEG] \qquad Ty(t), Fo(i\delta a(\Gamma iani)(Stergios'))$$

An alternative approach that would avoid making use of a LINK evaluation rule specific to *en tze* would assume that *en tze* projects a type *t* value and a formula metavariable in the initial type *t* requiring node. Then, the requirement for a proper formula value to be found has a further restriction that needs this formula value to be identical to the one found on the top of the LINKed tree. This alternative lexical entry for *en tze*, similar to the alternative proposed for *oti*, is shown below:

(4.106) Lexical entry for *en tze* (alternative)

IF $?Ty(t)$

THEN $\texttt{put}([+NEG], Ty(t), Fo(U), ?\exists\mathbf{x}.Fo(\mathbf{x}) \wedge \langle L\rangle Fo(\mathbf{x}));$
$\texttt{make}(\langle L\rangle); \texttt{go}(\langle L\rangle); \texttt{put}(?Ty(t))$

ELSE abort

Given the above lexical entry, no LINK evaluation rule is needed, since the formula metavariable can only be substituted by the value in the LINKed tree. Both lexical entries will give us the correct results.

The core of the proposed analysis is that the complex negation particle *en tze* links two type *t* requiring nodes. The node where the LINK starts, carries the *[+NEG]* feature. The pointer is at the other type *t* requiring node. It is at that domain where the rest of the sentence is parsed. The sentence is parsed given a negative context, i.e. the complex negation marker sets a negative context against which the sentence is going to be parsed. This treatment of complex negation is an effective way to get out of the enclisis problem. Proclisis

is not possible with *en tze* since the clitic is parsed in a domain where no proclitic triggers exist (the clitic will be parsed on the LINKed tree). In that respect, the only possibility with *en tze* is enclisis. The same reasoning can be argued to apply for the other complex markers/conjunctions like *ean tze* 'even though'. *Ean tze*, in the same sense as *en tze*, will be parsed as one complex element linking a type *t* requiring node with another node. A LINK evaluation rule will then do the appropriate copying, so that the semantics of concession are captured. There are a number of details that need to be taken care of for the case of *ean tze*, since concessive clauses just like subordinate clauses will already be linked with the consequent clause via a LINK relation. I will not flesh out the exact details of how the *ean tze* conjunction (or other complex conjunctions with *tze*) works. What is rather crucial, is that there is a way to account for these cases where enclisis unexpectedly arises fairly easy within DS. With this said, it is now time to look at how the complete entries for CG will look like.

4.8 Gluing the Pieces Together

In the previous section, I have given an account of CG clitic placement. This account proceeded in a stepwise manner, solving the problems one by one. What I have not done yet though, is putting these pieces together to form the full lexical entry for 3rd person accusative clitics in CG. This is what I will discuss in this section. We have already defined our triggering point for the clitic in question, i.e. a type *t* requiring node. We first add the proclitic triggers. The entry now has the format as shown below:

(4.107) The proclitic triggers

IF	$?Ty(t)$	
THEN	IF	$\langle\downarrow_*\rangle ?\exists x.Tn(x)\vert$
		$\langle\downarrow_0\rangle ?Ty(e_s)\vert$
	THEN	...
	ELSE	abort
ELSE	abort	

We now add the enclitic trigger:[26]

(4.108) The full set of triggers

IF	$?Ty(t)$	
THEN	IF	$\langle\downarrow_*\rangle ?\exists x.Tn(x)\vert$
		$\langle\downarrow_0\rangle ?Ty(e_s)\vert$
		$\langle\downarrow_1^+\rangle Ty(x)$
	THEN	...
	ELSE	abort
ELSE	abort	

The next step is to add the actual structure projected by the clitic. Starting with 3rd person accusative clitics, these actions will assumed to be identical to the ones posited for SMG and GSG 3rd person accusative clitics. Thus, 3rd person accusative clitics will build the direct object node and decorate the same node with a type value and a formula metavariable. The full entry is shown below:

[26]The order is arbitrary. We could well off start with the enclitic trigger.

(4.109) The full set of triggers plus actions

IF $?Ty(t)$

THEN IF $\langle\downarrow_*\rangle ?\exists x.Tn(x)|$
$\langle\downarrow_0\rangle ?Ty(e_s)|$
$\langle\downarrow_1^+\rangle Ty(x)$

THEN **(**$\texttt{make}((\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle))$**);**
$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle)$**;**
$\texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle)$;
$\texttt{put}(Ty(e), Fo(U_x), ?\exists\mathbf{x}.Fo(\mathbf{x})$;
$gofirst(?Ty(t)))$

ELSE abort

ELSE abort

The next thing we are going to look at is whether any overgeneration is caused via the various interactions of the triggers proposed. Note that such a check was not necessary for SMG and GSG because the two triggers posited there are mutually exclusive, in the sense that if $Mood(Imp)$ exists then the first trigger that specifies that a verb must not be present is obligatorily not true and vice versa. However, in the case of the triggers proposed here things are not quite the same. It is possible that more than one of these triggers will be present when the clitic comes into parse. Thus, we will have to see, what our entry predicts in that case. The first case will involve the situation where both an e_s requirement and an unfixed node will exist in the tree structure. Such structures are commonly found in CG and can involve a number of different structures, e.g. a focused object+negation+clitic, a subordinating conjunction+focused object+clitic, a focused adverb+the future particle+clitic and a number of similar constructions. In these cases, both triggers are satisfied, so the interpretation of the disjunction must be inclusive. An exclusive interpretation of disjunction will undergenerate, since it will rule out examples like the one below:

(4.110) *SIMERA* *enna* *tu* *δoso* *to* *vilvio*
today FUT him.CL-GEN give the.CL-ACC book.ACC
'This is the book I'm going to give him today.'

In the above example both an unfixed node and a type e_s requiring daughter node will exist when the clitic comes into parse. Assuming an exclusive interpetation of the disjunction between the unfixed node trigger and the type e_s requiring node trigger, examples like the above would be ruled out contrary to fact. However, things are different when we look at the interaction of the enclitic trigger with the proclitic ones. An inclusive interpretation of disjunction in that case will overgenerate, since it will predict that cases where both a proclitic and the enclitic trigger are present and the clitic appears postverbally should be grammatical, contrary to fact. For instance, examples like the ones shown below will be predicted to be grammatical given an inclusive interpretation of disjunction between the proclitic and the enclitic trigger:

(4.111) **En iksero ton*
NEG know him.CL-ACC
'I do not know him.'

(4.112) **TUTO TO VIVLIO eδosa tu*
this.ACC the.ACC book.ACC gave.1SG him.CL-GEN
'This is the book I gave him.'

A straightforward way out of this overgeneration is to posit that the disjunction between the proclitic triggers and the enclitic one is to be interpreted exclusively rather than inclusively. This small modification in the entry will treat the overgeneration caused. We will encode this exclusive interpretation of disjunction using the || symbol:[27]

[27]This is basically the XOR operator.

(4.113) Updated lexical entry

IF $?Ty(t)$

THEN IF $\langle\downarrow_*\rangle ?\exists x.Tn(x)|$
$\langle\downarrow_0\rangle ?Ty(e_s)||$
$\langle\downarrow_1^+\rangle Ty(x)$

THEN **(**$\texttt{make}((\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle));$**)**
$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle)$
$\texttt{put}(Ty(e), Fo(U_x), ?\exists \mathbf{x}.Fo(\mathbf{x});$
$gofirst(?Ty(t)))$

ELSE abort

ELSE abort

The updated lexical entry is rather minimal considering the complexity of the phenomena itself. The reason for this minimality is the two major generalized triggers capturing proclisis. These two triggers manage to capture the complexity of proclitic positioning by generalizing over the elements that trigger proclisis and grouping them into two major categories, those that project unfixed nodes and those that project an e_s requirement after being parsed. I will argue that these generalizations provide a straightforward explanation of how the clitic systems in MG dialects have evolved from their respective Medieval Greek counterparts. I postpone this discussion until chapter 7, where a first sketch of a unitary diachronic account of the dialects under consideration will be provided.

4.9 Clitic clustering and ordering in CG

Clitic ordering inside clitic clusters in CG is identical to clitic cluster ordering in GSG, i.e. DAT-ACC order is obligatory in all environments. The examples below illustrate the relevant claim:

(4.114)

a.	*Eδoke*	*tu*	*to*
	gave.3SG	him.CL-GEN	it.CL-ACC

b. **Eδoke* *to* *tu*
gave.3SG it.CL-ACC him.CL-GEN
'S/He/It gave it to him.'

(4.115)

a. *Do* *tu* *to*
give.IMP him.CL-GEN it.CL-ACC

b. **Do* *to* *tu*
gave.IMP it.CL-ACC him.CL-GEN
'Give it to him.'

(4.116)

a. *En* *tu* *to* *edoke*
NEG him.CL-GEN it.CL-ACC gave.3SG

b. **En* *to* *tu* *edoke*
NEG it.CL-ACC him.CL-GEN gave.3SG
'S/He/It did not give it to him.'

(4.117)

a. *Pjos* *tu* *to* *edoke?*
Who.NOM him.CL-GEN it.CL-ACC gave.3SG

b. **Pjos* *to* *tu* *edoke?*
Who.NOM it.CL-ACC him.CL-GEN gave.3SG
'Who gave it to him?'

(4.118)

a. *Otan* *tu* *to* *edoke...*
when him.CL-GEN it.CL-ACC gave.3SG

b. **Otan* *to* *tu* *edoke...*
otan it.CL-ACC him.CL-GEN gave.3SG
'When s/he/it gave it to him...'

Capturing ordering within clitic clusters in CG, will require the same mechanisms that capture clitic cluster ordering in GSG. In that sense, we just add the restriction $[\downarrow_1^+]?\exists\mathbf{x}.Tn(\mathbf{x})$ as a further restriction in the proclitic environments for both genitive and accusative clitics and the restriction $[\downarrow_1^+][\downarrow_0]?Ty(x)$ as a further restriction in the enclitic environment for genitive clitics only. By doing that, ordering in CG is predicted to be strictly DAT-ACC.[28] Furthermore, such a restriction will also disallow *V NP cl* structures to be generated. Given that an NP parsed after a verb will be parsed on a fixed node, the restriction $[\downarrow_1^+][\downarrow_0]?Ty(x)$ will have to abort given that one fixed argument will be present in the tree. The updated version of 3rd person accusative clitics in CG is shown below:

(4.119) Updated lexical entry for 3rd person accusative clitics (ordering included)

IF $?Ty(t)$

THEN IF $\langle\downarrow_*\rangle?\exists x.Tn(x), [\downarrow_1^+]?\exists\mathbf{x}.Tn(\mathbf{x})|$
$\langle\downarrow_0\rangle?Ty(e_s), [\downarrow_1^+]?\exists\mathbf{x}.Tn(\mathbf{x})||$
$\langle\downarrow_1^+\rangle Ty(x)$

THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle)$**)**;
$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle)$;
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle)$;
$\texttt{put}(Ty(e), Fo(U_x), ?\exists\mathbf{x}.Fo(\mathbf{x}))$;
$gofirst(?Ty(t))$

ELSE abort

ELSE abort

[28]The illicit clitic clusters that give rise to PCC violations will be discussed in chapter 6.

(4.120) Updated lexical entry for genitive clitics in CG (ordering included)

IF $?Ty(t), Tn(a)$

THEN IF $\langle\downarrow_*\rangle ?\exists x.Tn(x), [\downarrow_1^+]?\exists \mathbf{x}.Tn(\mathbf{x})|$
$\langle\downarrow_0\rangle ?Ty(e_s), [\downarrow_1^+]?\exists \mathbf{x}.Tn(\mathbf{x})||$
$\langle\downarrow_1^+\rangle Ty(x), [\downarrow_1^+][\downarrow_0]?Ty(x)$

THEN **(**$\mathtt{make}(\langle\downarrow_1\rangle); \mathtt{go}(\langle\downarrow_1\rangle)$**);**
$\mathtt{make}(\langle\downarrow_1^+\rangle); \mathtt{go}(\langle\downarrow_1^+\rangle)$**;**
$\mathtt{make}(\langle\downarrow_0\rangle); \mathtt{go}(\langle\downarrow_0\rangle)$;
$\mathtt{put}(\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a)$;
$Ty(e), Fo(U_x), ?\exists \mathbf{x}.Fo(\mathbf{x})$;
$?\exists \mathbf{x}.Tn(\mathbf{x}), gofirst(?Ty(t)))$

ELSE abort

ELSE abort

With these last entries, I conclude the discussion on the CG clitic system.

4.10 Conclusions

In this chapter I looked at the clitic system of CG. Firstly, I have shown that the CG positioning system is more complex than assumed in the literature. This added complexity of CG system lies in the existence of variation environments, a fact noted in the literature only by Revithiadou (2006) and only as regards the factive complementizer *oti*. My data concur with Revithiadou as regards variation with *oti*, while further show that variation is also attested with the the subordinate of cause conjunctions epiδi and γati. I first argued that all the analyses proposed for CG positioning (Terzi, 1999b; Agouraki, 2001; Condoravdi & Kiparsky, 2002; Revithiadou, 2006) are inadequate to capture positioning in CG. Then, I provided an account of CG positioning by arguing that the apparent complexity with proclitic positioning can be derived assuming two generalized proclitic triggers. These two proclitic triggers were shown to be adequate enough to capture the complexity of proclisis in CG. Following Bouzouita (2008a,b), it was argued that the presence of an unfixed node consitutes a proclitic trigger. In that sense, a group of elements parsed on unfixed nodes can give rise to proclisis according to fact (focused and Wh elements). Building

on assumptions by Gregoromichelaki (2006) and Cann (forthcoming), according to which every sentence contains an obligatory situation argument, I argued that the second proclitic trigger consists in the existence of a type e_s requiring node. Modifying Gregoromichelaki's (2006) analysis, according to which subordinate conjunctions project the type e_s and not the requirement for the same type, to fit the recent assumptions as regards tense and aspect presented in 2.2.2.4 and 3.4.3.2, I assumed that subordinating conjunctions project a requirement for such a type and not the actual type itself. This e_s requiring node will then neatly capture proclisis with modality/tense markers, given that these elements will project tense/aspect information inside the complex situation argument node. Such a trigger captures proclisis with negation, in effect avoiding the use of a third proclitic trigger that will specifically refer to negation (like e.g. the [+NEG] feature used in Bouzouita, 2008a,b). Enclisis on the other hand is captured assuming a generalized trigger which allows the clitic to be parsed in case any verbal type has been parsed first. These three triggers are adequate enough to capture the complexity of the CG positioning system. Variation in positioning with *oti* and *epiδi* was captured assuming that these elements can be also parsed as coordinating conjunctions. Ordering inside the clusters was explained using the same mechanism used in deriving strict DAT-ACC ordering in GSG.

In conclusion, the CG clitic system can be seen under our account as a system in which three generalized parsing triggers (or strategies) give rise to a highly complex positioning system. We will see in chapter 7, how these generalized triggers emerged from KG and through MCG by the syntactic encoding of pragmatic preferences via a process of routinization (Pickering and Garrod, 2004; Bouzouita, 2008a,b). In the following chapter, I will examine the clitic system of the last dialect under consideration, to wit Pontic Greek (PG).

Chapter 5

Pontic Greek

5.1 Some Historical and Sociological Notes

The term Pontus refers to the area in North-East Asia minor, bordering with the Black Sea to its north. Greeks first settled there as early as the 7th century BC, and continuously lived there until the population exchange of 1922-3. I will use the term Pontic Greek (PG) to refer to the dialect as it is spoken today in Modern day Greece, since a form of Pontic, Romeyka Pontic (Sitaridou, 2010a) is still spoken today in some villages of Trabzon and the Of area. The roots of PG, like all other Modern Greek dialects (with the exception of Tsakonika), can be traced back to Koine Greek, i.e. the form of Greek used during the Hellenistic years. However, as Mackridge (1987) notes, PG came to be markedly distinct to the other Greek dialects probably due to the Seljuk invasion of the 11th century which split the Pontus area from the other areas of the Byzantine empire.[1] The PG spoken in Asia Minor was not a uniform dialect but included several sub-dialects. A number of classifications have been proposed by the years (see Papadopoulos, 1995; Triantafyllidis, 1981 and Nicholas, 1995). However, all these sub-dialects tend to neutralize or have already done so in favour of a PG koine. The population exchange brought 1.5 milion Pontic Greeks into mainland Greece, mainly in the regions of Macedonia and Thrace. From that point the contact between PG and SMG begins. The fact that the PG dialect was treated with hostility by the official

[1] PG along with Cappadocian Greek comprise the Asia Minor Greek dialectal group. For more information about Cappadocian see Dawkins (1916) and Janse (2006).

Greek state of the time as well as the need to maintain a distinct cultural idenity led to the creation of a PG Koiné neutralizing sub-dialectal idiosyncrasies. Moreover, Pontic Greeks from different areas of Pontus ended up living together in mainland Greece. This last fact combined with the tendency of PG people to prefer marriages with "their own kind", played an important role in the development of a koine form as Chatzisavvidis (1995) and Tompaidis (1996) note.[2] The data I am going to use in this thesis come from this form of koiné Pontic. I will draw these data from my own personal database of PG. This database is the result of a fieldwork visit undertaken in the villages of Nea Nikomideia and Palaia Likogianni in the preferecture of Imathia in the central area of the Macedonian region in Greece, as well as the village of Mikroklisoura in the preferecture of Grevena in the Nortwestern area of the Macedonia region in Greece. The fieldwork took part in discontinuous time intervals during July/August 2009. Data taken from other sources will be cited accordingly.

5.2 The Data

Before the PG data are presented, it is worth looking at the clitic forms comprising the PG clitic system. PG exhibits considerable differences in terms of clitic morphology compared to the other three dialects we have discussed. One of the major differences is that PG clitics are syncretic across the board (even 3rd person clitics). The system is comprised of accusative clitics only, which can function as both direct and indirect objects. Furthermore, PG exhibits a number of alternative morphological forms corresponding to a single clitic. The table below illustrates the morphological forms one can find in PG (adapted from Drettas, 1997):[3]

[2] See Chatzisavvidis (1995) for a detailed discussion on the sociolinguistics of PG.

[3] The fact that some of the forms below are exactly the same as strong pronouns has led Michelioudakis & Sitaridou (2010) to claim that Romeyka Pontic (RP) forms like *aton(a)* 'him' are not clitics but rather weak pronouns. We will see however later on in this chapter, that at least for some cases and at least for PG such an assumption cannot be maintained.

(5.1) Pontic Greek clitic forms

	1st	2nd	$3^{rd} masculine$	$3^{rd} feminine$	$3^{rd} neuter$
Singular	eme(n)	ese(n)	aton	aten	at(o)
	em	es		atine(n)	a(t)
	me(n)	se(n)	æton	æten	æto
	m	s		ætine(n)	æ(t)
Plural	ema(s)	esa(s)	ats	ats	ata
	emasen(e)	esasen(e)	atsen(e)	atsen(e)	a
	ma(s)	sa(s)	æts	æts	æta
	masen(e)	sasen(e)	ætsen(e)	ætsen(e)	æ

The PG clitic system is unique among the clitic systems found in the rest of the MG dialects in that positioning restrictions are not only defined solely with respect to the verb but furthermore the type of the verbal element involved does not affect clitic positioning. In that respect, clitic positioning in PG is always postverbal, no matter the form of the verb or the nature of any preceding element:

(5.2)

a. *Ekser* *aton*
know.1SG him.CL

b. **Aton* *eksero*
him.CL know.1SG
'I know him.'

(5.3)

a. *Δos* *me* *avuto* *to* *vivlio*
give.2SG-IMP me.CL this the book

b. **Me* *δos* *avuto* *to* *vivlio*
him.CL give.2SG-IMP this the book
'Give me this book.'

(5.4)

a. *Ki* *kser* *aton*
NEG know.1SG him.ACC

b. **Ki* *aton* *eksero*
NEG him.ACC know.1SG
'I do not know him.'

(5.5)

a. *Pios* *entok* *aton?*
who hit.1SG him.CL

b. **Pios* *aton* *entoke(n)?*
who him.CL hit.1SG
'Who hit him?'

(5.6)

a. *Esi* *k* *epices* *ekino* *do* *ipa* *se*
you.NOM NEG did.2SG this what told.NOM you.CL

b. **Esi* *k* *epices* *ekino* *do* *se* *ipa*
you.NOM NEG did.2SG this what you.CL told.1SG
'You did not do what I told you' [Papadopoulos 1955: 172 for [a], judgment for [b] my data]

(5.7)

a. *Ipe* *me* *na* *leγo* *se* *pola*
told.3SG me.CL SUBJ say you.CL many

b. **Ipe* *me* *na* *se* *leγo* *pola*
told.3SG me.CL SUBJ you.CL say many
'S/He told me to tell you a lot (of things)' [Papadopoulos 1955: 172 for [a], judgment for [b] my data]

The PG system seems, in comparison to the other three examined systems, to be the least complicated one in terms of positioning, the only restriction being the presence of a verbal element. In case such a verbal element exists, the clitic appears to its immediate right. Notwithstanding this, the PG clitic system exhibits a number of peculiarities that are not found in any other MG dialect. The most striking one, is the unavailability of clitic clusters comprised of two 3rd person clitics. In other words, 3rd person clitic clusters are illicit in PG as the examples below illustrate:

(5.8)

a. **Eδek aton ato/a*
gave.3SG him.CL it.CL
'S/He/It gave it to him.'

b. **Eδek ats ato/a*
gave.3SG them.CL it.CL
'S/He/It gave it to them.'

However, clitic clusters of a 1st/2nd person plus a 3rd person clitic are licit in PG:

(5.9)

a. *Eδek m ato/a*
gave.3SG me.CL it.CL
'S/He/It gave it to me.'

b. *Eδek s ato/a*
gave.3SG you.CL it.CL
'S/He/It gave it you.'

Things get more complicated, since in cluster environments only one out of the four possible forms for 1st/2nd person clitics are possible (forms *m/s*):

(5.10)

a. *Eδeke m/*eme(n)/*em/*me(n) ato/a*
gave.3SG me.CL it.CL
'S/He/It gave it to me.'

b. *Eδek* *s/*ese(n)/*es/*se(n)* *ato/a*
gave.3SG you.CL it.CL
'S/He/It gave it you.'

3rd person clitic forms beginning with *æ* are also impossible in clusters:

(5.11)

a. **Eδeke* *m* *æ(t)/æ*
gave.3SG me.CL it.CL
'S/He/It gave it to me.'

b. **Eδek* *s* *æ(t)/æ*
gave.3SG you.CL it.CL
'S/He/It gave it you.'

However, the reason that the *æ* forms cannot appear in clusters seems to be of a phonological nature rather than a syntactic one. As Condoravdi and Kiparsky (2002) note, *æ* forms are the result of regular phonological processes of the PG system. According to them, the form *æ* is derived via vowel contraction (synalepha), where the ending *i* vowel of the verb meets the beginning *a* vowel of the 3rd person clitic. The result is the *æ* sound:

(5.12) Pontic Greek vowel contraction rule

/i,e+a/ → æ (Condoravdi & Kiparsky 2002: 20)

Vowel contraction is a common phenomenon in PG and, as Condoravdi & Kiparky (2002) note, it is also a postlexical phonological process applying across word boundaries as well. Thus, the *æ* forms depend on the phonological context and cannot be considered autonomous forms.[4] However, the *m/s* forms do not depend on the phonological context but rather on whether these appear in clusters or not.

One last property of PG clitics that will not be dealt with in this thesis and furthermore dissociates PG from the rest of MG dialects is the fact that PG clitics can take wide scope over coordination. In 5.13 the clitic appearing in the coordinated clause gets scope over

[4]Note that such a fact argues for a word status for PG clitics as Condoravdi & Kiparsky (2002: 39) note.

coordination. Note that this is optional and a construction where the clitic is repeated in both constructions, i.e. the only option in any other MG dialect, is also grammatical as shown in 5.14:

(5.13) *Ekops* *ke* *emaiγreps* *ato*
cut.1SG CONJ cooked.1SG it.CL
'I cut it and cooked it.'

(5.14) *Ekops* *ato* *ke* *emaiγreps* *ato*
cut.1SG it.CL CONJ cooked.1SG it.CL
'I cut it and cooked it.'

As already said, the above data are not going to be dealt with in this thesis but will constitute a subject for further research, since an account of elliptical structures needs to be first taken into consideration before such an attempt is being made. But even if I was ready to provide such an account, what one further needs is the PG ellipsis data, which at the moment I do not have.

5.3 Existing Approaches

5.3.1 Condoravdi and Kiparsky, 2002

Even though a large literature exists on PG, only a very small part of it deals with PG syntax on a theoretical level. The only paper I know of that tries to give an account of the PG positioning restrictions is Condoravdi and Kiparsky (2002).[5] Condoravdi and Kiparsky, as already discussed in 3.2.1.4, propose a general analysis of clitics for all dialects of MG by classifying these different clitic systems into three major categories (see chapter 3). Under this analysis, PG clitics fall into type B. Type B clitics according to Condoravdi and Kiparky (2002) are assumed to be X_0 clitics, syntactically adjoined to a lexical head. Given their status as X_0 elements, they are head adjoined to V_0 and not to a phrasal projection as it is argued for CG type systems. Condoravdi and Kiparsky (2002) argue, contra Drettas

[5]There are a number of other works that do discuss clitics in PG but are either descriptive only (Drettas, 1997) or deal with phenomena associated with PG clitics but do not give an account of PG clitics (Tsakali, 2008).

(1997), that PG clitics are not agreement suffixes. Their argument is based on evidence from perfect tense constructions as well as coordinating conjunctions. In perfect tense constructions, PG clitics attach to the infinitive rather than to the auxiliary. The argument here is that if agreement affixes attach only to finite verbal forms, then attachment of PG clitics to the infinitive rather than the auxiliary remains unexplained:[6]

(5.15) *An ihame ndosne se, ihes maθine to maθema s*
if had.1PL beat.INF you.CL had.2SG learn.INF the lesson yours
'If we had beaten you, you have learned your lesson.' [Condoravdi and Kiparsky, 2002 citing Papadopoulos, 1955: 174]

Furthermore, coordinated verbs can share a single clitic in PG, a fact that, according to Condoravdi and Kiparsky (2002), should not be the case if PG clitics were indeed affixes:

(5.16) *Ekops ke emaiγreps ato*
cut.1SG CONJ cooked.1SG it.CL
'I cut it and cooked it.'

(5.17) *Ekinos pa eprostaksen na luzne ke plin aten*
he PA ordered SUBJ shampoo CONJ wash her.CL
'He ordered them to shampoo him and wash him.' [Condoravdi and Kiparsky, 2002 citing Fostiropoulou, 1938: 190]

The fact that word-level phonological processes apply between the verb and the clitic is also an argument used by Condoravdi & Kiparsky, in order to prove that PG clitics are in fact full words.

Condoravdi and Kiparsky's discussion concentrates on the status of clitics. Assuming that one has to take a decision on the affix-word status of PG, Condoravdi and Kiparsky's

[6]Note that the example below contains the infinitival forms *ndosne* beat and ma*θ*ine 'learn'. It is an open question whether the infinitive is still used in modern day PG. Papadopoulos (1955: 87) claims that the infinitive was still active in PG, while the dialect of Of exhibited a form of inflected infinitives as well. Tombaidis (1988) on the other hand, disputes such a claim and argues that the infinitive is not found in PG. I have not done any research on whether the infinitive exists in modern day PG and if this is the case in which contexts. It is worth mentioning however, that infinitival constructions do exist in Romeyka Pontic (RP, Sitaridou, 2010b), i.e. the form of PG by Turkish Muslims in some villages of modern day Trabzon and Of region in Turkey. Sitaridou (2010b) also reports the existence of inflected infinitives in the RP variety of the Of region. Furthermore, I will not get into a discussion of whether PG has true perfect constructions or not (see Tompaidis, 1988 for a discussion). The part of the analysis relevant to infinitives is based on data by Papadopoulos (1955).

arguments are quite convincing (especially the phonological ones). However, the argument that affixes attach only to finite forms (and thus that PG clitics cannot be suffixes since clitics in perfect tense constructions attach to the infinitive rather than the auxiliary) has been already disputed for type C dialects (SMG type dialects) in section 3.2.1.4. Assuming that Condoravdi & Kiparsky are right and clitics in PG are indeed syntactic X_0 elements (words) and not suffixes, their analysis is adequate enough to capture the positioning restrictions in PG. However, no clitic system is analyzed in depth in Condoravdi & Kiparsky (2002) and as such a number of issues remain unresolved for PG. For instance, clitic clusters are not discussed. Given this state of affairs, such an analysis fairs well as regards PG clitic positioning but does not deal with clitic clusters. My goal is to present an alternative, more complete account of the PG clitic system, giving a dynamic account of both clitic positioning and clustering.

5.4 A DS Analysis

The PG clitic positioning system, as we have already seen, is rather easy to be described compared to the other clitic systems we have looked at. The only restriction on clitic placement is that the clitic must follow any verbal element, whether finite or non-finite. Encoding of this positioning restriction in DS terms can be done extremely straightforward. Actually, the enclitic trigger used for CG is what we need in order to account for PG clitic positioning. Given this trigger, the clitic can only be parsed in case a verbal type exists in one of the functor nodes:

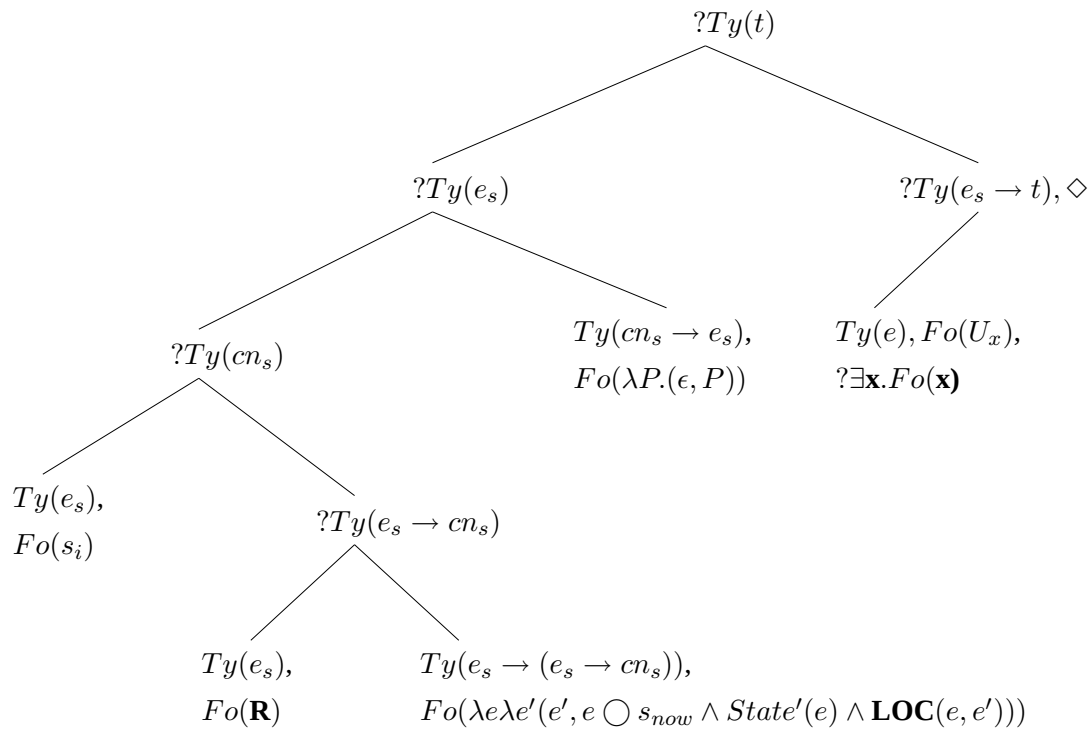
(5.18) Trigger ensuring enclisis

IF	$?Ty(t)$	
THEN	IF	$\langle\downarrow_1^+\rangle Ty(x)$
	THEN	...
	ELSE	abort
ELSE	abort	

The above trigger will be true in case any type of verb has been parsed. Such an analysis will furthermore predict that clitics in PG will attach to the infinitive rather than

the auxiliary in examples like (5.15). Assuming an analysis of PG auxiliaries in the same vein as the analysis given for SMG *exo* 'have' in section 3.4.3, clitics in PG cannot be parsed after the auxiliary has already been parsed. This is because the lexical entry for the auxiliary will leave the pointer at the type $?Ty(e_s \rightarrow t)$ t requiring node and as such the clitic's initial trigger (*?Ty(t)*) will not be satisfied. But even in case pointer was left at the type *t* requiring node via means of the lexical entry of the auxiliary, parsing of the clitic would still be impossible. This is because in parsing an auxiliary no verbal type is projected. Therefore, the embedded trigger of PG clitics ($\langle\downarrow_1^+\rangle Ty(x)$) will not be satisfied. The structure after parsing an auxiliary is shown below:

(5.19) The effect of parsing an auxiliary



A verbal type will be provided only when the infinitive is parsed in the example given. Thus, the clitic can be parsed only after the infinitive has already done so. The trigger I have provided is thus adequate enough to capture the strict enclitic properties of the PG clitic system. The next step is to look at the actual actions projected by clitics in PG. Since

PG clitics are syncretized across the board, I argue that all PG clitics will project locally unfixed nodes. Thus, all types of clitics in PG will have the general entry shown below:

(5.20) Lexical entry for PG clitics

IF $?Ty(t, Tn(a))$
THEN IF $\langle\downarrow_1^+\rangle Ty(x)$
THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1^+\rangle); \texttt{go}(\langle\downarrow_1^+\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a));$
$\texttt{put}(Ty(e), Fo(U_x), ?\exists\mathbf{x}.Fo(\mathbf{x});)$
$\texttt{put}(?\exists\mathbf{x}.Tn(\mathbf{x}), gofirst(?Ty(t)))$
ELSE abort
ELSE abort

The above lexical entry correctly accounts for the positioning facts, while furthermore capturing the fact that PG clitics are structurally underspecified, in the sense that they can occupy more than one structural position in the tree structure (direct and indirect object positions). This last fact is effectively captured by assuming that PG clitics project locally unfixed nodes.

5.4.1 Clitic Clusters

As we have seen in the presentation of the data, PG is rather idiosyncratic when it comes to clitic clusters, given that clusters of 3rd person clitics are illicit. The only licit clitic clusters in PG are combinations of a 1st/2nd plus a 3rd person clitic. However, even in these cases the situation is quite complex, since out of the 4 possible forms for 1st/2nd person clitics, only one can be used in clitic clusters, i.e. the phonologically reduced *m/s* forms. All the other forms are unacceptable. The relevant examples are repeated below:

(5.21)

Edeke	*m*	*ato/*	**eme(n)*	*ato/*	**em*	*ato/*	**me*	*ato*
gave.3SG	me.CL	it.CL	me.CL	it.CL	me.CL	it.CL	me.CL	it.CL

'S/He/It gave it to me.'

(5.22) *Edeke* *s* *ato/* **ese(n)* *ato/* **es* *ato/* **se* *ato*
gave.3SG you.CL it.CL you.CL it.CL you.CL it.CL you.CL it.CL
'S/He/It gave it to you.'

The lexical entries we provided for clitics in PG provide a natural explanation for the lack of 3rd person clitic clusters. Remember that PG clitics were assumed to project locally unfixed nodes. Assuming that two clitics have been parsed, two locally unfixed nodes will be projected. But then, these two nodes will collapse into one by means of treenode identity and thus the parse will never be successful. In more detail, let us say we want to parse the illicit cluster *aton a* 'him it'. The result of parsing the two clitics will be a situation where two locally unfixed nodes are present in the tree structure:

(5.23) The result of parsing two PG clitics

?Ty(t),Tn(a)

$Fo(U_{male'})$,	$Fo(V_{neut'})$,
$Ty(e)$,	$Ty(e)$,
$?\exists \mathbf{x} Tn(\mathbf{x})$,	$?\exists \mathbf{x} Tn(\mathbf{x})$,
$\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a)$	$\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a)$

However, in the above structure both unfixed nodes bear the same underspecified treenode modality, i.e. $\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a)$, and thus they are indistinguishable in terms of address. This treenode identity has the effect of collapsing the two unfixed nodes into one. The situation we end up with is one where there is one unfixed node carrying incompatible formula metavariable presuppositions. The parse can never be successful, since update of one of the metavariables will result in incompatibility with the other. The structure after the two unfixed nodes have collapsed into one is shown below:

(5.24) The two nodes collapse into one

?Ty(t),Tn(a)

$$Fo(V_{neut'}), Fo(V_{male'}),$$
$$Ty(e), ?\exists \mathbf{x} Tn(\mathbf{x}),$$
$$\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a)$$

Within this line of reasoning the unavailability of 3rd person clitic clusters follows from a very general tree-logic constraint of the system, namely the "no more than one unfixed node at a time" constraint. A rather welcomed result of this account is that the same account will be used in order to explain the Person Case Constraint (PCC). Actually, the proposed account for PG will be a strong argument favoring an analysis of the PCC as a hard-wired logic constraint rather than a restriction involving person feature interaction (see chapter 6).

The proposed account predicts that no clitic clusters should be possible in general in PG, contrary to the facts. The fact that 1st/2nd plus 3rd person clitic combinations like *em ato/es ato* 'me it/you it' or *emen ato/esen ato* 'me it/you it' are not possible is easily explainable given the "one unfixed node at a time constraint". However, such an approach cannot account for the licit clusters *m ato/ s ato* 'me it/you it', where a 1st/2nd person clitic forms a cluster with a 3rd person clitic. By taking a closer look at how the forms *m/s* actually behave in PG, a straightforward solution emerges. We have already seen that *m/s* are the only possible 1st/2nd forms in clitic clusters. This is not the end of the story however. The forms *m/s* can appear only inside clitic clusters and as such are impossible in single clitic constructions:

(5.25) **Entoke m*
gave.3SG me.CL
'S/He/It hit me.'

(5.26) **Edeke m avuto to vivlio*
gave.3SG me.CL this.ACC the.ACC book.ACC
'S/He/It gave me this book.'

One might claim that the reduced forms *m/s* in cluster constructions like *m ato/s ato* are the result of "ekthlipsis", in which case the final vowel of a word disappears in the presence of the beginning vowel of the next word. In the case we are examining, the final vowel of the form *me*, i.e. *e*, disappears in the presence of the beginning vowel *a* of the 3rd person form *ato*, giving rise to the form *m ato*. However, if "ekthlipsis" was at play here, we would expect the same phenomenon to occur in (5.26), where the reduced form *m* is followed by a word beginning with the same vowel the form *ato* begins with. However, (5.26) is ungrammatical. Given this, the reason for the use of different forms might be syntactic after all. Given the lexical entries for clitics in PG, there is a wholly natural explanation for this phenomenon. As we have already seen, clitics in PG are taken to project locally unfixed nodes. Under this approach no cluster should be possible in PG, since more than one unfixed node with the same underspecified address will be present in that case. This is the crucial point in the account provided. The three forms used in single clitic constructions project locally unfixed nodes and thus cannot combine with any other clitic form. Now, given that *m/s* can appear only in clitic clusters, where they can be only interpreted as indirect objects, it seems natural to pursue an analysis along the following lines: the licit clitic clusters are parsed as one single lexical entry, in which the contribution of the *m/s* forms is to identify themselves with the indirect object node and further project a type value and a formula metavariable in that node. Then, the second clitic in the cluster can actually build a locally unfixed node. The construction becomes totally unproblematic, since it is not subject to the "no more than one unfixed node at a time" constraint. The lexical entry I propose for licit clitic clusters in PG is shown below:

(5.27) Lexical entry for PG clitic clusters

IF $?Ty(t), Tn(a)$

THEN IF $\langle\downarrow_1^+\rangle Ty(x)$

THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$**)**

$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle)$

$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$

$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$

$\texttt{put}(Ty(e), Fo(U_{Sp'/Hr'}), ?\exists\mathbf{x}.Fo(\mathbf{x})); gofirst(?Ty(t));$

($\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$**)**

$\texttt{make}(\langle\downarrow_1^+\rangle); \texttt{go}(\langle\downarrow_1^+\rangle);$

$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$

$\texttt{put}(\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a));$

$\texttt{put}(Ty(e), Fo(V_{x'}), ?\exists\mathbf{x}.Fo(\mathbf{x}), \texttt{go}(\langle\uparrow_0\rangle\langle\uparrow_1\rangle));$

$gofirst(?Ty(t))$

ELSE abort

ELSE abort

We can further assume that the second clitic in these clusters also builds fixed structure, since fixing the first clitic in the indirect object node will leave only one possibility of update to the second one. Assuming that the second clitic in the cluster projects fixed structure will give us the same results. Thus, the two entries are pretty much equivalent:

(5.28) Alternative entry for PG clitic clusters

IF $?Ty(t)$

THEN IF $<\downarrow_1^+> Ty(x)$

THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle)$**);**

$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle)$**;**

$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle)$;

$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle)$;

$\texttt{put}(Ty(e), Fo(U_{Sp'/Hear'}), ?\exists\mathbf{x}.Fo(\mathbf{x})$;

$\texttt{go}(\langle\uparrow_0\rangle\langle\uparrow_1\rangle)); \texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle)$**;**

$\texttt{put}(Ty(e), Fo(V_x), ?\exists\mathbf{x}.Fo(\mathbf{x}))$;

$gofirst(?Ty(t))$

ELSE abort

ELSE abort

However, the question pending is why similar clusters did not develop for 3rd person clitics: why are clusters comprised of two 3rd person clitics, with one of the two clitics in a reduced form (say *a* 'it'), not possible? Why is a cluster of the form *aton a* 'him it' not possible? Looking at the distributional properties of the reduced forms of 3rd person clitics, one notices a major difference compared to the reduced 1st/2nd person clitic forms. Unlike the reduced forms *m/s* that cannot appear on their own, i.e. in single clitic constructions, the reduced 3rd person form *a* can appear on its own in single clitic constructions as either a direct or an indirect object:

(5.29) *Ehasen* *a*
lost3SG it.CL
'S/He/It lost it.'

(5.30) *Eδeken* *a* *kat*
gave3SG it.CL something
'S/He/It gave it (e.g. the child) something.'

Given the above facts, these reduced forms have their own lexical entry according to which they project locally unfixed nodes. In that sense, a clitic cluster comprised of two

3rd person clitics cannot be formed, even in case one of the two clitics exhibits the reduced form *a*.[7]

5.4.2 A Repair Strategy?

Before I conclude the discussion on PG clusters, it is essential to note a construction used in PG to express 3rd person clitic clusters. A number of the speakers I consulted, when asked for a translation of SMG sentences involving 3rd person clitic clusters, translated the sentences in PG using a construction that involves the use of the direct object clitic plus a locative, the exact status of which (whether it is a clitic or not) remains unknown to me:

(5.31) *Eδek ato ki*
gave it there
'S/He/It gave it to him/her

One speaker used the exact structure when asked to translate from a SMG sentence which contained a cluster of a 2nd person plus a 3rd person clitic:

(5.32) *Eδek ato ki*
gave it there
'S/He/It gave it to you.' [One speaker]

Unfortunately, I do not have enough data to decide whether *ki* is a clitic and as such forms a clitic cluster with the preceding 3rd person clitic or not. In case *ki* is a clitic no problem is created for the analysis I have provided, since *ki* is not underspecified in these structures but is always interpreted as the indirect object. In that sense, *ki* cannot be assumed to project a locally unfixed node. On the contrary, it will be assumed to project fixed structure or possibly a LINK structure on the assumption that it is the same element with the homophonous locative adverb *ki* 'there'.[8] In that respect, constructions like the one in (5.32) are predicted to be grammatical according to the account provided. The lexical entry for *ki* is shown below:

[7] What is rather intriguing under such an analysis, is that the most morphologically impoverished form is the one that is more specified in terms of syntactic information, an assumption that is contra what Cardinaletti & Starke (1999) claim.

[8] Another plausible explanation suggested to me by Sitaridou (p.c) is that *ki* is a genderless shortened form of the deictic pronoun *ekinos/kinos* 'that one'.

(5.33) Lexical entry for *ki*

IF $?Ty(t)$

THEN IF $\langle\downarrow_1^+\rangle Ty(x)$

THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$**)**

$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$

$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$

$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$

$\texttt{put}(Ty(e), Fo(U_x), ?\exists\mathbf{x}.Fo(\mathbf{x}));$

$\texttt{put}(?\exists\mathbf{x}.Tn(\mathbf{x}), gofirst(?Ty(t)))$

ELSE abort

ELSE abort

Where x $\in$ {Male, Female, Neuter}[9]

As already said, the data available do not suffice to give a fuller analysis of *ki*. However, at first look it does not seem to be problematic within the account sketched. On the contrary it seems to vindicate it, since *ki* is used exceptionally in environments where two 3rd person clitics would occur in SMG, in order to side-step the "one unfixed node at a time constraint". Given that *ki* has a fixed interpretation as an indirect object, it can be encoded either as projecting fixed structure or as a LINK structure, in which case a *ki* will be assumed to be the same lexical entry for the homophonous locative adverb. Both encodings will avoid the "one unfixed node at a time constraint". However, as already said, further research is needed in order to see the exact nature of *ki* in these constructions.

5.4.3 Are PG Clusters Problematic for Current Linguistic Theory?

The PG clitic system can be described as being regular as regards clitic positioning compared to the highly complicated CG positioning system or even the less complicated but still not that regular as the PG system, SMG positioning system (see the chapters on CG and SMG, 4 and 3 respectively). However, even though the PG positioning system is quite

[9]For the speaker that translated the 2nd person plus 3rd person clitic with a locative the set must contain the additional member 'Hearer'.

straightforward to describe (and possibly give a syntactic analysis), it is not that straightforward, as we have already seen, with respect to clitic clustering. PG clusters are unique among MG dialects (at least to my knowledge) in that clusters of two 3rd person clitics are illicit. Such a fact is not only idiosyncratic across MG dialects but seems to have no counterpart in other clitic languages such as e.g. the Romance languages. Person restrictions are quite common in clitic languages (e.g. the PCC) but a restriction of that kind, namely a ban on two 3rd person clitics, is specific to PG. So, the question is whether this idiosyncrasy of the PG clitic system poses any problem to current linguistic theory and if so, how serious this problem is. Well, on minimalist grounds, at least as regards analyses in the line of Anagnostopoulou (2003, 2005), Adger & Harbour (2007), Nevins (2007), Rezac (2003, 2008a) and Michelioudakis (2009), the unavailability of 3rd person clusters within a system that does allow clusters is highly problematic, since under these analyses 3rd person clitic clusters are exactly those clusters that are allowed to be generated (along with clusters of 1st/2nd dative plus a 3rd person accusative). For example in Anagnostopoulou (2003, 2005), Rezac (2003), Adger & Harbour (2007), the PCC is explained as a person checking failure. The core idea (with variant formalizations in the accounts mentioned) is that the legibility of a clitic cluster relies on whether there is more than one person feature or not. Assuming more than one person feature, the PCC obtains. For example in Anagnostopoulou (2003), clitic clusters check their features against one functional head bearing one number and one person feature that can only be checked once. Assuming that datives and 1st/2nd person accusative clitics are always specified for person, the PCC effects are predicted. However, in a cluster comprised of two 3rd person clitics no PCC obtains, since the person feature in this case is just one (on the assumption that only the dative and not the accusative will bear a person feature. See next chapter for more details on person based accounts of the PCC).[10] Within alternative theories of generative grammar, notably HPSG and LFG, the unavailability of 3rd person clitic clusters has to be defined as a pure lexical phenomenon where the unavailable clusters just do not exist in the lexicon. Within HPSG (Miller & Sag, 1997; Monachesi, 1993, 1998a,b, 1999), as already discussed in section 3.3.3, clitics are treated as affixes. In that sense, a complex entry is given for both the clitic

[10]Similar problems are caused in the other accounts as well, i.e. Nevins (2007), Rezac (2008a) and Michelioudakis (2009).

and its host (the verb). Given the different clitic forms and clitic clustering possibilities, such analyses have to enumerate all the possible verb-clitic and all verb-clitic cluster combinations, which obviously does not provide us with any explanation as regards the PCC or any other person restriction. In that sense, the analysis in terms of tree-growth presented fairs better to any minimalist or HPSG analysis in that it naturally predicts clusters of two 3rd person clitics to be illicit by a general restriction on underspecification, a hard-wired treegrowth constraint.

5.4.4 Romeyka Pontic - Michelioudakis & Sitaridou, 2010

In Romeyka Pontic (RP), as reported in Michelioudakis & Sitaridou (2010) clitic clusters are not possible at all. The construction in (5.34) is not considered a clitic cluster by Michelioudakis & Sitaridou (2010), since the two elements can be separated by a full NP as shown in (5.35):

(5.34) *Eδiksane* *me* *aton(a)*
Showed.3PL me him
'They showed him to me.'

(5.35) *EδotΣen-eme* *o* *Mehmet* *aton(a)*
gave.3SG-me.CL the Mehmet.NOM him/it
'Mehmet gave me him/it.'

Michelioudakis & Sitaridou (2010) conclude that *aton(a)* 'him' in (5.34) is not a clitic. That might well be true but this is not the end of the story. RP also exhibits a reduced clitic form for 3rd person pronouns:[11]

(5.36) *O* *Mehmetis* *emenan* *eδotΣen* *æ*
The Mehmet.NOM me gave.3SG it.CL
'Mehmet gave it to me.'

[11]The exact gender and number values of this reduced form are not specified in Michelioudakis & Sitaridou (2010). In the examples presented, the form *æ* is a the 3rd person neuter singular clitic. However, in Özkan (2010) the same clitic form is reported as the 3rd person masculine singular clitic. In PG on the other hand, the reduced form *a* is the 3rd person neuter clitic in both numbers.

(5.37) **O Mehmetis eδotΣe m(e) æ*(unattested)
The Mehmet gave.3SG me.CL it.CL
'Mehmet gave it to me.'

Given this reduced form, it is not at all clear that the absence of 3rd person clitics is the reason for the unavailability of clitic clusters, since the reduced form *æ* can be safely assumed to be a clitic (at least as regards its morphology and positioning). The question remains, why RP does not have clitic clusters? The explanation is straightforward given the account already proposed for PG clitics. Assuming that clitics in RP will project locally unfixed nodes, clitic clusters are predicted to be illicit. Surprisingly, Michelioudakis & Sitaridou (2010) report constructions involving two clitics (a cluster) or possibly one clitic and a weak or strong pronoun that are PCC violations. These constructions are straightforwardly accounted via the same account I have proposed for PG clitics and the PCC but for the moment I am going to postpone the discussion on the PCC and these constructions in RP until the next chapter.

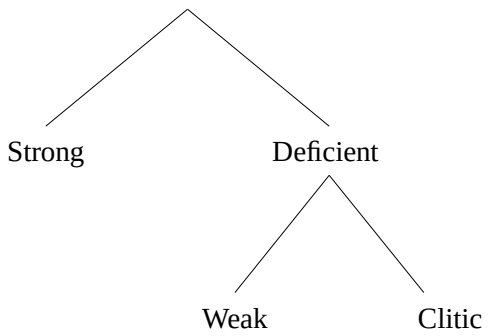
In conclusion, comparing the clitic systems of PG and RP one finds major similarities but a major difference also. The positioning system and the morphological syncretism are similar in both dialects. In that sense, the analysis given for PG clitics will be extremely relevant for RP clitics as well. Analyzing RP clitics as projecting locally unfixed nodes, one predicts that no clitic clusters should be possible, in accordance to the facts. However, PG makes use of 1st/2nd plus 3rd person clitic clusters by using a special form for 1st/2nd person clitics, i.e. the form *m/s*. As I have shown in this chapter, these highly reduced morphological forms cannot appear on their own but only in clusters, while they are always interpreted as indirect objects. Thus, the forms *m/s* are not underspecified anymore but can be seen as projecting fixed structure (i.e. fixing their position in the indirect object node). Given this last assumption, the "no more than one unfixed at a time" constraint is not violated anymore and such clusters are predicted to be licit.[12]

[12] A far more general question would be why clitic clusters have managed to emerge in PG but not in RP. This question however, is beyond the scope of this thesis and as such will not be dealt with.

5.4.5 PG Clitics: Are These Really Clitics?

We have seen that Michelioudakis & Sitaridou (2010) have argued, on the basis of data like (5.35), that *aton(a)* is not a clitic in RP. Whether the same element and similar 3rd person clitics like *aten/ato* 'her/it') is not a clitic but rather a weak pronoun in PG assuming Cardinaletti & Starke's (1999) tripartite classification of pronouns is something that needs to be elucidated:

(5.38) Cardinaletti & Starke's (1999) pronoun classification



There are a number of diagnostics that could potentially lead someone to claim that forms like *aton(a)* are not clitics, like for example the fact that the form *aton(a)* is of the same form with the equivalent strong pronoun, a fact that according to Cardinaletti & Starke would point towards a weak pronoun status for elements like *aton(a)*. However, it seems that there is a safe way to prove that at least in some cases *aton(a)* is a clitic rather than a weak pronoun in PG. We have seen that PG exhibits reduced weak forms for pronouns in all persons, notably 1st/2nd person *m/s* 'me/you' and 3rd person neuter *a* 'it/these'. These elements can be safely assumed to be clitics, since they are phonologically reduced and their positioning is strictly immediately postverbal. What is more, the forms *m/s* can only appear in clitic clusters and are not possible on their own:

(5.39) *Eδeken* *a* *kat*
gave.3SG it.CL something
'S/He/It gave it (e.g. the child) something.'

(5.40) **Eδeken* *kat* *a*
gave.3SG something it.CL

'S/He/It gave it (e.g. the child) something.'

(5.41) *Eδeke* *m* *a*
gave.3SG me.CL it.CL
'S/He/It gave me it.'

(5.42) **Entoke* *m*
gave.3SG youCL it.CL
'S/He/It hit me.'

(5.43) **Eδeke* *m* *avuto* *to* *vivlio*
gave.3SG me.CL this.ACC the.ACC book.ACC
'S/He/It gave me this book.'

The fact that these reduced forms can form clitic clusters where no element can intervene is also pointing towards their clitic status:

(5.44) **Eδeke* *m* *o* *Γiorikas* *a*
gave.3SG me.CL the George it.CL
'George gave me it.'

Given all these facts, is quite easy to see that at least in some cases forms like *aton(a)* 'him' (as well as forms like *ato* 'it' and *ate(n)* 'her') are also clitics. This is because the reduced forms *m/s* can also combine with them in PG. Given that *m/s* can only appear in clitic clusters, then forms *aton(a)/ato/ate(n)* are also clitics in these cases :

(5.45) *Eδeke* *m* *aton(a)/ate(n)/ato*
gave.3SG me.CL him/it/her.CL
'S/He/It gave me him/her/it.'

(5.46) **Eδeke* *m* *o* *Γiorikas* *aton(a)/ate(n)/ato*
gave.3SG me.CL the George him/it/her.CL
'S/He/It gave me him/her/it.'

The above facts prove that at least in cases like (5.45) *aton(a)/ato/ate(n)* should be treated as clitics. More data are needed in order to decide whether PG exhibits only strong and clitic pronouns or strong, weak and clitic pronouns. I will not pursue this issue here. It

is worth mentioning, that even in RP, where no clitic clusters are reported, reduced forms like *m* 'me' or *a* 'it' are very difficult to be accommodated under a weak pronoun analysis, given their phonological reduction and strict enclitic positioning. Furthermore, it is not clear to me whether the tripartite distinction proposed by Cardinaletti & Starke (1999) has any theoretical consequences in giving a DS analysis. For example, it is not clear that an element classified as a weak pronoun under Cardinaletti & Starke (1999) will receive a different analysis than an element classified as clitic. I will not get into a discussion on the issue but it is worth mentioning that even if RP is proven to involve only weak pronouns and not clitics (such a claim is not sustainable for PG as discussed above), the account proposed will most probably stay the same.

5.5 Conclusions

In this chapter the PG clitic positioning system was discussed. The strict enclitic nature of the PG positioning system was neatly captured by assuming a trigger which referred to the presence of a type value in any of the functor nodes. Furthermore, novel data as regards clitic clusters in PG were presented. More specifically, it was shown that PG lacks 3rd person clitic clusters but does allow clusters of a 1st/2nd plus a 3rd person clitic. Given that PG is syncretic across the board, PG clitics were assumed to project locally unfixed nodes. This last assumption straightforwardly explains the reason why clusters of two 3rd person clitics are illicit, since assuming that both 3rd person clitics will project locally unfixed nodes, such constructions are out via the "one unfixed node at a time" constraint. The licit clusters are not a problem either, given that the clitic forms standing for 1st/2nd person clitics in these clusters are special forms that cannot appear on their own but only in clusters and they are always interpreted as indirect objects. Thus, these clitic forms are not structurally underspecified, since their position in the tree structure is fixed. Therefore, these specific clitic forms are assumed to project fixed structure and as such constructions involving these special forms plus a 3rd person clitic are not subject to the "one unfixed node at a time" constraint. This analysis turns out to be extremely relevant for RP as well, since it does provide an explanation for why no clitic clusters are possible in RP. Assuming that clitics will project locally unfixed nodes in RP, this is expected. Lastly, it was shown

that Michelioudakis & Sitaridou's (2010) claim that RP does not have clitics but rather weak pronouns cannot extend to PG, at least for the clitic clusters presented.

Chapter 6

General Clitic Phenomena Found in all Dialects

6.1 The Person Case Constraint

The PCC is a clitic co-occurrence restriction, which states that a dative clitic cannot co-occur with a 1st/2nd person accusative clitic. The restriction is found across a remarkable number of both related and unrelated to each other languages, spanning from Romance and Greek to Kiowa and Basque (see Rezac, 2008b for the Basque data and Adger & Harbour, 2007 for the Kiowa data). The examples below from Spanish and SMG exemplify the restriction:

(6.1) **Le me ha dado*
it.CL-DAT me.CL has given
'S/he has given me to him.' [Spanish]

(6.2) **Mu se exi δosei*
me.CL-DAT you.CL-ACC has given
'He/She/It has given you to me.' [SMG]

The above restriction is referred to in the literature as the strong PCC version. Another weaker version of the PCC has been claimed to exist in some varieties of Catalan, Italian

and Spanish (see Bonet, 2007; Bianchi, 2006 and Cuervo, 2002 respectively).[1] Under this looser version of the restriction, the ban is not against datives in general but only against 3rd person datives. In that sense, the weak PCC version precludes clusters of a 3rd person dative plus a 1st/2nd person accusative clitic but however allows combinations of a 1st/2nd person dative plus a 1st/2nd person accusative:[2]

(6.3) *Te m' ha recomanat la Mireia*
you.CL me.CL has recommended the Mireia
'Mireia has recommended me to you/you to me.' [Catalan-Bonet, 2008]

(6.4) *Lui mi ti presento*
he me.CL you.CL introduces
'He introduces me to you/ you to me.' [Some varieties of Italian]

(6.5) **Gli mi ha dato*
them.CL-DAT me.CL-ACC has given
'He/She has given me to them.' [Some varieties of Italian]

In SMG, only the strong version of the constraint is attested and the equivalent SMG clitic sequences to the ones shown above, are all ungrammatical. The same facts hold for GSG and CG:

(6.6) **Mu se eδose*
me.CL.DAT you.CL.ACC gave
'S/He/It gave you to me .'[SMG]

(6.7) **Eδoke mu se*
gave me.CL.DAT you.CL.ACC
'S/He/It gave you to me .'[CG]

(6.8) **Mu se edike*
me.CL.DAT you.CL.ACC gave
'S/He/It gave you to me. '[GSG]

[1]See also Chatzikyriakidis & Kempson 2009 for argumentation that the weak PCC version is not a robust constraint in these varieties and as such should not be attributed to a general syntactic property in these cases.

[2]There is also another version of the constraint exhibited in Romanian in which case sequences of a dative clitic and a 1st person accusative are licit while sequences of a dative plus a 2nd person accusative are ungrammatical. Furthermore, no PCC restrictions arise with postverbal singular clitics but do however arise with postverbal plural clitics. See Savescu (2007, 2009) and Nevins & Savescu (2008) for the relevant data.

PG, on the other hand, is idiosyncratic as regards the person restrictions. PG seems to allow (at least for some speakers) sequences of two 1st/2nd person clitics. PG is thus is subject only to the strong PCC constraint. It is worth noting that sequences of 1st/2nd person clitics are also reported by Michelioudakis & Sitaridou (2010) for RP:

(6.9) *Eδikse/eneδikse* *m* *esen*
showed.3SG me.CL you.CL
'(S)he/It showed you to me.' [PG - Some speakers]

(6.10) *Eδiksane* *m* *ese*
showed.3PL me.CL you.CL
'They showed you to me.' [Sürmene variety of RP, Michelioudakis & Sitaridou, 2010]

The PCC seems rather puzzling at least at first sight and seems to resist a principled syntactic explanation. Adding to the elusiveness of such constructions, it is worth mentioning that the restriction cannot be derived on semantic grounds (in the sense that such constructions are semantically rare), since in every language exhibiting the constraint, there is an equivalent syntactic construction to express the semantics of PCC constructions, a fact noted in the literature as a repair (Bonet, 2007; Rezac, 2008a, among others). In the examples shown below, taken from SMG and French, the dative clitic has been substituted by a preposition plus the strong form of the pronoun. The result in both cases is grammatical:

(6.11) *Me* *sistisan* *se* *sena*
me.CL-ACC introduced to you.ACC
'They introduced you to me.' [SMG]

(6.12) *Je* *t'* *ai* *presenté* *à* *lui*
I you.CL have introduced to him
'I introduced you to him.' [French]

The literature on the PCC is extensive and ranges from purely functional approaches like Haspelmath (2004), in which the PCC is taken to result from infrequent usage of such constructions, to purely morphological accounts arguing for a separate level of morphology like Bonet (1991, 1994) and Heap (2005), or to purely minimalist syntactic accounts, in

which the PCC is argued to derive from general mechanisms of the Probe/Agree system (Bejar & Rezac, 2003; Anagnostopoulou, 2003, 2005; Rezac, 2008a,b; Adger and Harbour 2007; Nevins, 2007; Michelioudakis, 2009 among others). In what follows a number of the most recent prominent analyses with respect to the PCC are reviewed.

6.1.1 Minimalist Accounts of the PCC

Most of the minimalist accounts that have been proposed in the last few years are in agreement that the PCC is the result of the inability of a given clitic to check its features against a given head (Bejar & Rezac, 2003; Anagnostopoulou 2003, 2005; Rezac, 2008; Adger & Harbour, 2007; Michelioudakis, 2009 among others). The details of the analysis are different in each case but the core of the analysis relies crucially on feature checking. In Anagnostopoulou (2003, 2005) an agree relation is established between clitics and a functional head F. The PCC under this account arises when the functional head F valuing the clitics' features cannot check any given feature more than once. In more detail, the functional head F carries a person and a number feature that can be checked, as already mentioned, just once. Anagnostopoulou then assumes that dative clitics carry a person feature (even though for 3rd person dative clitics this feature is specified as [-]) but not a number feature. 1st/2nd accusative clitics carry both a number and a person feature while 3rd person accusative clitics only a number feature. All these assumptions put together will give us the PCC. For example, assuming a dative clitic has checked its person feature against F, 1st/2nd person accusative clitics cannot check their person feature anymore. Given that dative clitics are always specified for person, the only possible combinations in the presence of a dative clitic are those combinations consisting of a dative plus a 3rd person accusative clitic. In Bejar & Rezac, 2003 on the other hand, the PCC arises due to dative intervention rather than the inability of a head checking the same feature twice. In this account, an agree relation must license a 1st/2nd person accusative clitic. In PCC combinations no such relation can be established due to the intervening dative. 3rd person accusative clitics do not participate in this relation. Similarly, Adger & Harbour (2007) argue that the participant features of 1st/2nd person accusative clitics are responsible for the PCC. There, 1st/2nd person accusative clitics, are licensed inside the applicative head (Appl) by virtue of being

direct objects (VP being the complement of Appl). SpecAppl bears participant features due to animacy restictions associated with indirect objects. The generalization proposed by Adger and Harbour is as follows:

(6.13) **Adger and Harbour's Generalization**

> *The requirements which a functional head requires its specifier to bear cannot be used as probes in the head's complement domain* [Adger & Harbour, 2007: 21]

In the presence of an Appl head, i.e. in ditransitive constructions, a 1st/2nd person accusative clitic is excluded, since assuming that the specifier of Appl will always bear participant features, the participant features of these clitic forms will remain unchecked. Since feature checking of direct object clitics is done against the Appl head, a 1st/2nd person accusative clitic will not be able to check its features given (6.13). This is because these features cannot be used as probes in the complement domain of Appl, given that SpecAppl already bears participant features. This checking inability will then give us the PCC.[3]

Taking a closer look at all the aforementioned analyses, one notices that their common starting point and actual basis is the decision one makes with respect to which clitics have which features. Such a decision is crucial to these accounts, as the actual analysis derives from such a decision. For example, assuming that 3rd person accusative clitics do bear person features (independently disputed in the minimalist literature, Bianchi 2006; Nevins 2007; Michelioudakis, 2009), no clitic sequences will ever be possible under Bejar & Rezac, 2003, Anagnostopoulou (2003, 2005), Adger & Harbour (2007) or Rezac (2008a). The question is whether these features really exist, and if they do, what is the motivation behind the proposed specifications. For example, there seems to be no principled reason of why 3rd person dative clitics should be characterized as bearing a [-person] feature, whereas on the other hand 3rd person accusative clitics do not bear any such feature at all (Anagnostopoulou 2003, 2005) other than the fact of matching the distributional idiosyncrasy with some supposedly general mechanism. It is rather unclear what the difference is between a [-person] feature, i.e. the negative specification of a feature, and the

[3] A similar approach is put forth in Michelioudakis (2009), where the PCC arises due to the inability of a dative to check the participant feature of a probe in the presence of a closer 1st/2nd person goal.

total absence of such a feature. The same critique can be applied to Adger and Harbour's (2007) analysis in which 3rd person dative clitics bear participant features (even though unspecified), whereas 3rd person accusative clitics do not. Furthermore, Anagnosopoulou (2003, 2005) has to simply define in a rather post-hoc manner that number features of dative clitics are not checked. She argues that dative heads are defective and as such their number feature is not accessible for checking purposes. This claim is built on data drawn from participial agreement languages, where dative clitics, contrary to accusative clitics, fail to trigger this form of agreement. However, such evidence is not decisive as the pattern is only partial (pointed out in Bonet 2007: footnote 14), with Catalan participial agreement being possible with 3rd person accusative clitics but not with 1st/2nd accusative clitics.

In addition, not all languages confirm some of the specific claims made so and would require additional stipulation. Adger and Harbour (2007) assume that indirect objects always bear participant features. However, at least for languages like SMG, this restriction does not hold, since a number of constructions with an inanimate indirect object are perfectly grammatical. There is a tendency for indirect objects to be animate, but this fact remains not more than a tendency; it is certainly not a categorical restriction:

(6.14) *Tis* *eδose* *mia* *klotsia* *(tis* *kareklas)*
her.CL-GEN gave a kick.ACC the.GEN chair.GEN
'He kicked the chair.'

(6.15) *Tis* *eδose* *mia* *klotsia* *(tis* *Marias)*
her.CL-GEN gave a kick the.GEN Mary.GEN
'He kicked Mary.'

So again, the feature assignment, though as a methodology taken to be an interface between syntax/morphology and interpretation, is in fact not grounded either in semantic or morphophonological motivation. Examples where a 3rd person dative clitic refers to an inanimate dative introduced in previous discourse are also problematic. These constructions should be ungrammatical according to Adger and Harbour (2007) since inanimate objects do not bear participant features: [4]

[4]The same problem has been reported by Bonet (2007: 25) for Catalan in discussing the same analysis.

(6.16) *A: Pos eγine etsi to vivlio? B: Tu eδosa mia klotsia kata*
A: how happened that the book B: it.CL-GEN gave a kick by
laθos
mistake
A: 'Why is the book like that?' B: 'I kicked it by mistake.'

(6.17) *A: Pos ton xalases ton ipologisti sto grafio? B: Apla,*
A: how it.CL-ACC damaged the computer in-the office B: simply
tu estila enan io
it.CL-GEN sent a virus
A: 'How did you manage to destroy the computer in the office?' B: 'I just sent it a virus.'

In this connection, I believe that even though a number of ditransitive constructions involving inanimates are somewhat degraded in terms of grammaticality compared to the ones involving animate NPs, a generalization banning inanimates from ditransitive constructions is on the wrong track. The exact details of these constructions remain unresolved, since the data are not clear cut to guarantee a definite answer. For example, there are a number of constructions involving inanimate indirect objects that are, if not ungrammatical, at least questionable as to their grammaticality. Substituting the genitive marked NP with the accusative NP headed by preposition *se* 'to', the sentence becomes grammatical. The peculiar thing is that a genitive clitic can be used to refer back to that PP construction:

(6.18) *A: ???Δose mia efkeria tis dimokratias B: Θa tis*
A: give a chance the.GEN democracy.GEN B: FUT her.CL-GEN
δoso
give.1SG
A: 'Give democracy a chance.' B: 'I will.'

(6.19) *A: Δose mia efkeria sti dimokratia B: Θa tis δoso*
A: give a chance to-the democracy B: FUT her.CL-GEN give-I
A: 'Give democracy a chance.' B: 'I will.'

Further research is needed in order to understand what is the exact correlation between animacy and double object constructions. However, it is clear to me that a generalization grounded in some semantic property such as participanthood, however abstractly construed,

like the one given by Adger and Harbour (2007), cannot be sustained. In that respect, at least for SMG, the assumption that all indirect objects are interpreted as animates is dubious.[5]

Anagnostopoulou (2005) discusses the weak version of the PCC and argues that it derives via multiple rather than single agree against a functional head. However, in order to maintain such an account, an additional constraint has to be stipulated in order to exclude the *me-lui* cases. The proposed generalization bans multiple agree just in case the elements entering this agree relation have contradictory values, with the explanation lying in 3rd person dative clitics not being be able to co-occur with 1st/2nd person accusative clitics given that the former are specified as [-person] while the latter as [+person]. This is openly a stipulation, one that does the job and so observationally adequate, one might say:

(6.20) **A condition on multiple agree**

> *Multiple agree can take place only under non - conflicting feature specifications of the agreeing elements* [Anagnostopoulou 2005: 221]

However, the above condition does not have much to say about why such *me-lui* constructions are banned. And it is unclear what exactly counts as a conflicting feature specification. 1st/2nd person cannot have the same person specifications: they will both be specified as +person but additionally they must have something to distinguish between 1st and 2nd person. If this is the case, then it is very hard to see why this further specification as 1st or 2nd person does not consitute a conflicting specification. But if it does, then how is it that in some languages, other combinations are possible? The sense of seeking a principled explanation for why such restrictions emerge on a broad cross-linguistic basis (albeit with variants) is close to having become irrevocably lost. We appear to be searching for a feature stock that, in the event, gets extended language by language in an essentially post-hoc fashion.

[5]There is an additional unresolved problem in Adger and Harbour (2007), i.e. how to derive the checking of participant features of a *v* head only in the presence of a direct object. See Michelioudakis (2009) for more details.

6.1.1.1 Nevins, 2007

The account proposed by Nevins (2007), despite being a feature based account as well, departs in one major respect with respect to which clitics have which features. This is because Nevins argues that 3rd person accusative clitics have in fact person features, specified as [-Auth,-Part] (Author, Participant).[6] The crucial point Nevins' in analysis lies in the notion of contrastiveness of a feature, which is defined as follows:

(6.21) *An instance of a feature F is contrastive within a set of other features S if both values of F may occur in S* (Nevins, 2007: 289)

1st person clitics are specified as [+Auth,+Part], whereas 2nd person clitics as [-Auth,+Part]. A multiple agree account in the sense of Anagnostopoulou (2005) is given, in which the core idea lies in the intervention effects caused when an element bearing a feature α intervenes between the probe and the features it is looking for. The idea is that different PCC effects arise depending on the number of excluding values by the relativization domain as set with respect to the probe controlling the domain. Given a certain parametrization of a certain domain, the following two conditions must hold (I present the informal versions of these conditions only):

(6.22) **Contiguous Agree**: *There can be no interveners betwen P and x that are not in the domain of relativization that includes x.*

(6.23) **Matched values**: *All elements within the domain of relativization must contain the same value.*

Nevins first discusses the weak version of the PCC arguing that in that case the search is relativized to marked values of [Participant]. In that case, the only illicit cases are the ones where an element specified as [-part] intervenes, i.e. a 3rd person dative. The strong version of the PCC on the other hand is derived by assuming that the probe looks for contrastive [Author] values. It is essential to note that an [Author] value, whatever its specification, will not be contrastive in the presence of a [-Part] feature since the only binary realization

[6]The assumption that 3rd person accusative clitics carry a person feature is also made in Bianchi (2006) and Michelioudakis (2009).

possible given [-Part] is [-Part,-Author], [-Part,+Auth] being logically impossible. Thus, a 3rd person dative will count as an intervener. Combinations of a 1st person with a 2nd person clitic or vice versa will be out via (6.23) since the two clitics will bear conflicting contrastive values [+Auth,+Part] [-Auth,+Part]. However, the account as proposed overgenerates, allowing illicit clitic clusters to be generated within a language exhibiting the strong version of the PCC. First of all, clusters comprised of two identical 1st/2nd person clitics cannot be ruled out. Clusters like 'me me' in Spanish or 'mu me' in SMG are impossible to rule out given Nevin's account, since both clitics will be specified [+Participant, +Author]. Someone might argue that such examples must be ruled out independently via the binding theory or alternatively by a phonological rule that disallows two identical clitics to form clusters. The first claim, even though well to the point, is problematic for SMG. This is because in SMG no principle B violation occurs in a construction where both a 1st person clitic and a 1st person pronoun are both co-indexed and within the same domain. They are only licensed if one of these is external in some way to be structurally defined:

(6.24) *Mu* *edose* *emena*
me.CL-GEN gave.3SG me
'She/He/It gave me myself.'

(6.25) *Me* *edose* *se* *mena*
me.CL-ACC gave.3SG to me
'She/He/It gave me myself.'

Given the above data, it seems that at least for SMG, combinations of the same clitic form must be dealt within the PCC and not binding theory. But even if such constructions are to be dealt within binding theory and as such are irrelevant to a PCC account, the account would still overgenerate. This is because the account also precludes clusters of two 1st person clitics, one of them being a singular and the other a plural clitic, e.g. *me nos* or *te vos* in Spanish, *'mas me'* in SMG.

6.1.1.2 Ormazabal & Romero, 2007

An alternative albeit related account to the account mentioned so far, is that of Ormazabal and Romero (2007). Under this account, the PCC is dissociated into two distinct phenomena. The first is a universal tendency for object agreement to display sensitivity to animacy, while the second is a restriction on multiple object agreement. Ormazabal and Romero (2007) propose the following two generalizations:

(6.26) **Object animacy realization** *Object relations, in contrast to subject and applied object relations, are sensitive to animacy* [Ormazabal & Romero, 2007: 335]

(6.27) **Object agreement constraint**

If the verbal complex encodes verbal agreement, no other argument can be licensed through verbal agreement [Ormazabal & Romero, 2007: 336]

The above two generalizations are argued to capture the complexity the PCC exhibits; and one of the Ormazabal and Romero (2007) predictions is that in clitic languages where the PCC is active, it should not apply to ethical datives. This is true for many languages (French, Spanish to name a few, see Jouitteau & Rezac, 2008 and Ormazabal & Romero (2007) respectively, but not for SMG:

(6.28) *Pai ke mu to δinei sto Γiorγo anti na*
goes and me.CL-GEN it.CL-ACC gives to George instead SUBJ
to δosi se mena
itCL-ACC give to me
'He gives it to George instead of giving it to me (and I'm angry with that or I'm expressing irony with respect to such an action).'

(6.29) **Pai ke mu se dinei sto Γiorγo anti na*
goes and me.CL-GEN youCL-ACC gives to George instead SUBJ
se dosei se mena
you.CL-ACC give to me
'He gives you to George instead of giving you to me (and I'm angry with that or I'm expressing irony with respect to such an action).'

A further problem with the Ormazabal and Romero (2007) account is their characterization of the 3rd person Spanish clitic *lo*. Since their second generalization does not allow two objects to agree with the verb, they argue for a non-agreement, determiner-like analysis for *lo* to explain the grammaticality of sentences involving two argument clitics, where the accusative clitic is *lo*. They use a number of examples to establish that *lo* does not agree with the verb. Ormazabal and Romero (2007), based on data involving clitic doubling constructions, argue that *lo* is indeed a different kind of clitic from dative or 1st/2nd person accusative clitics: the latter can only double in particular environments, and when they do they must be interpreted as [+Specific]. However, the account does not extend successfully to SMG. The reason is that the equivalent SMG clitic for *lo* can double at least the same phrases as do 1st/2nd person accusative clitics, hence presumably agreeing with these in the relevant sense. The first two sentences below (6.30 and 6.31) are ungrammatical in Spanish, yet grammatical in SMG:

(6.30) *To* *iδame* *to* *spiti*
it$_{cl-acc}$ saw the house
'We saw the house.'

(6.31) *Tus* *iδame* *merikus* *sto* *maγazi*
them.CL-ACC saw some in shop
'We saw some of them in the shop.'

(6.32) *Mas* *fadazome* *merikus* *sti* *filaki*
us.CL-ACC imagine some in prison
'I imagine some of us in prison.'

(6.33) *Θa* *sas* *δo* *merikus* *avrio*
FUT you.CL-ACCsee some tomorrow
'I will see some of you tomorrow.'

Furthermore, the specificity explanation for such cases is inapplicable, since even though some have argued that such a restriction holds for SMG doubling constructions (Iatridou, 1995; Anagnostopoulou, 1997) there are plenty of sentences involving a doubled bare quantifier or a doubled non-specific indefinite which do not exhibit or optionally exhibit a specificity effect:

(6.34) *Polus anθropus den tus enδiaferi*
many people NEG them.CL-ACC care
'Many people do not care.' (specific or non - specific).

(6.35) *Mia kokini bluza ti θelo afton ton kero*
one red blouse her.CL-ACC want this the time
'I need a red blouse at this time of the year.' (specific or non - specific).'

(6.36) *Merika pota ta pino apopse*
some.ACC drinks.ACC them.CL-ACC drink.1SG tonight
'I would have some drinks tonight.' (specific or non - specific).

(6.37) *Mia bluza θa tin agoraza*
one blouse FUT her.CL-ACC bought
'I would buy a red blouse.' (non specific only)

(6.38) *Opioδipote traγuδi tu to sfiriksis, θa to peksi*
whichever song him.CL-GEN it.CL-ACC whistle, FUT it.CL-ACC play
'He will play any song you will whistle to him.'

Similar observations are made by Kallulli (2000) arguing that CLLD does not encode specificity in SMG and Albanian. The above examples from SMG suggest that a determiner-like analysis of 3rd person accusative clitics loses its empirical support and at least for SMG cannot maintained, leaving the Ormazabal and Romero account unable to capture the PCC effects as displayed at least for SMG.

This is where I will stop the discussion as regards the various analyses that have been given directing the reader to Bonet 1991, 1994; Cuervo, 2002; Heap 2005; Bianchi, 2006; among others for more analyses on the PCC. In what follows, I will try to provide a DS analysis of the PCC phenomenon based on Kempson & Cann (2007) and Chatzikyriakidis & Kempson (2009), arguing that the phenomenon is purely syntactic (concurring in that respect with all the analyses presented but disagreeing with e.g Bonet 1991, 1994; Heap 2005) and actually derives from a very general, hard wired treegrowth constraint. As things will turn out, this restriction will be a restriction on the nature of underspecification of structure in general.

6.2 A DS analysis

6.2.1 The Martins' Observation and Latin Scrambling

In Martins (2002), a very interesting observation is made. It is argued that clitic positioning can be seen as a reflection of word-order patterns of an earlier system. For example, clitic clustering in the Romance languages reflects the word order patterns of the earlier Latin system. Kempson & Cann (2007) as well as Chatzikyriakidis & Kempson (2009), concurring with the Martins observation, tried to provide a formal account of this claim, by arguing that clitics in Romance can be seen as calcified processing strategies used in the earlier Latin system. According to such an account, the same processing strategies used in Latin scrambling are those governing the clitic systems of Romance languages. Latin is assumed to make use of both unfixed and locally unfixed nodes as well as LINK structures. The crucial point in these accounts is the actions induced by parsing different kinds of case marking NPs. It is argued that case marking can be used in the following three senses: a) as constructive case b) as output filter case and c) as underspecified case. These three forms of case along with the use of the LINK strategy are argued by Kempson & Cann (2008) and Chatzikyriakidis & Kempson (2009) to be the four parsing strategies responsible for the Latin scrambling system.

Constructive case refers to a situation where case marking provides unambiguous information as regards the NP's structural position in the tree structure. Constructive case then fixes the unfixed node's address the NP is parsed on (if parsed on an unfixed node), by updating the underspecified relation of the unfixed node (the underspecified treenode address) to a fixed relation (fixed treenode address) provided. For example assume we want to parse *Lesbiam Catullus amavit* 'Catullus loved Lesbia'. The NP *Lesbiam* is parsed on a locally unfixed node. The case marking of *Lesbiam*, unambiguously signalling an accusative direct object, further updates the unfixed node's address by simply providing it with a fixed treenode address (010), i.e. it identifies the NP with the direct object position:

(6.39) Parsing *Lesbiam* in *Lesbian Catullus amavit*

Local*Adjunction		*Lesbiam*		Constructive case
$Tn(0), ?Ty(t)$	$\mapsto$	$Tn(0), ?Ty(t)$	$\mapsto$	$Tn(0), ?Ty(t), \Diamond$
$\vdots$		$\vdots$		$\vert$
$\langle\uparrow_1^*\rangle Tn(0)$		$\langle\uparrow_1^*\rangle Tn(0),$		$\langle\uparrow_1\rangle Tn(0),$
$\vert$		$?Ty(e \rightarrow t)$		$?Ty(e \rightarrow t)$
$?Ty(e)$		$Ty(e), Lesbia'$		$\vert$
$\langle\uparrow_0\rangle\langle\uparrow_1^*\rangle Tn(0), \Diamond$		$?\langle\uparrow_0\rangle Ty(e \rightarrow t)$		$\langle\uparrow_0\rangle\langle\uparrow_1\rangle Tn(0), Lesbia'$
		$\langle\uparrow_0\rangle\langle\uparrow_1^*\rangle Tn(0)\Diamond$		$?\langle\uparrow_0\rangle Ty(e \rightarrow t)$

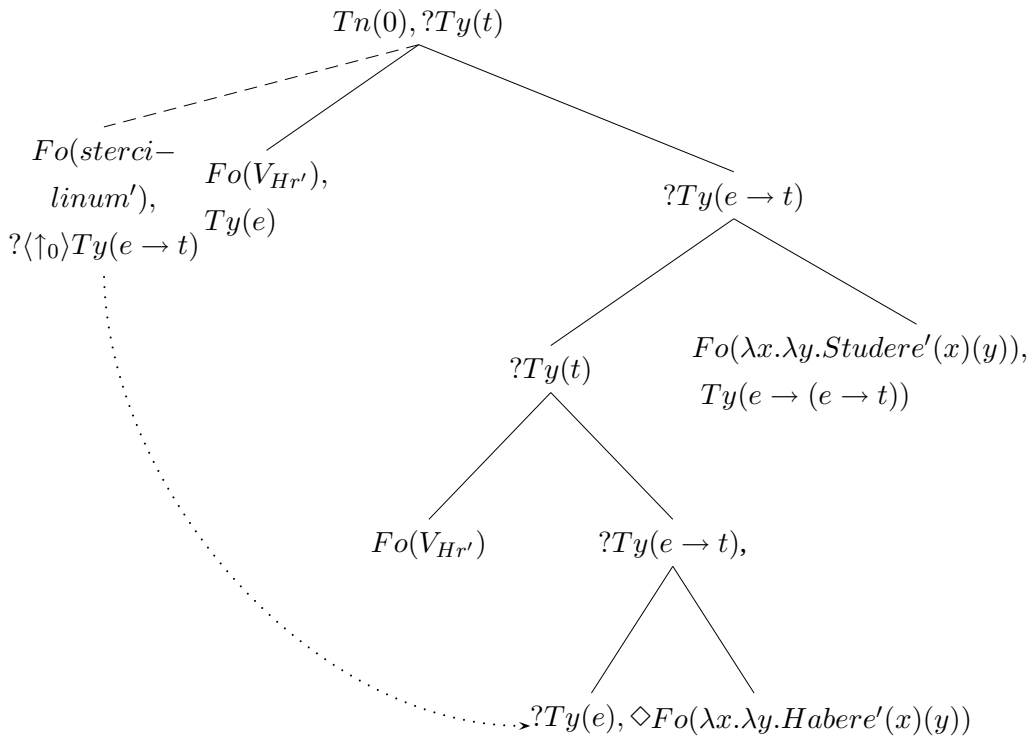
Case can also act as an output filter. In this case, case restricts the potential fixing sites of the unfixed node but does not however fix the treenode relation itself. This is definitive of case and illustrated by the effect of case-decorated left-peripheral constructions. As (6.41) illustrates, the left dislocated *stercilinum* is parsed as decorating an unfixed node. It projects the specification $?\langle\uparrow_0\rangle Ty(e \rightarrow t)$ along with type and formula values in the unfixed node. This specification will ensure that the NP must end up on an argumental node (0 node) immediately dominated by a node carrying a predicate type ($e \rightarrow t$). What this means is that the NP must be parsed as a direct object, no matter how deeply embedded that node dominated by a predicate type will turn out to be:

(6.40)

Stercilinum	*magnum*	*stude*	*ut*	*habeas*
dunghill.ACC	big	ensure.2SG	that	have.2SG

'Ensure that you have a large dunghill.'

(6.41) Parsing *stercilinum magnum stude ut habeas*



The last type of case is underspecified case. This type of case is unable to fix or even restrict the potential fixing sites of an unfixed node. It does not do anything with respect to the fixing of the unfixed node's address or with respect to restricting the potential fixing sites of the same unfixed node. A good example of this type of case would involve Latin neuter nouns of the 4th declension. These types of nouns are highly syncretized in their singular forms and exhibit the same morphological forms for singular accusative, dative and ablative.[7] Given this syncretism, and assuming that all three forms have a single lexical entry, this lexical entry should encode this underspecification via means of syncretism. In order to do that, we just assume that this type of case does not provide any update information with respect to treenode addresses or output node filters in the single lexical entry for all three different forms of case marking. In that respect, assuming that *cornu* is parsed on a locally unfixed node, the result we get is shown below:

[7]In other words the template apporach to such Latin forms distinguish them as discrete homonyms is simply false, a misleading if familiar pedagogical heuristic.

(6.42) Parsing *cornu*

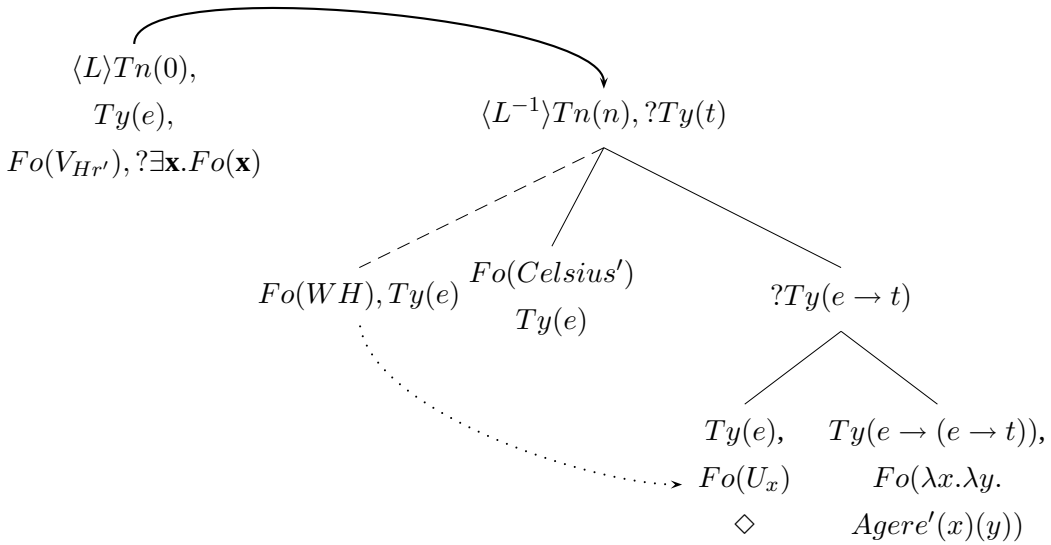
Local*Adjunction		*Cornu*
$Tn(0), ?Ty(t)$	$\mapsto$	$Tn(0), ?Ty(t)$
$\vdots$		$\vdots$
$\langle\uparrow_1^*\rangle Tn(0)$		$\langle\uparrow_1^*\rangle Tn(0),$ $?Ty(e \rightarrow t)$
$\vert$		$/$
$?Ty(e),$ $\langle\uparrow_0\rangle\langle\uparrow_1^*\rangle Tn(0), \Diamond$		$Ty(e), Fo(cornu')$ $\langle\uparrow_0\rangle\langle\uparrow_1^*\rangle Tn(0), \Diamond$

In the above example, parsing *cornu* does not provide us with any further information as regards the unfixed node's address or as regards the potential fixing sites of than unfixed node. This fact will enable *cornu* to be parsed in more than one structural position according to fact.

The fourth parsing strategy used in Latin scrambling does not involve the projection of unfixed nodes but rather pairs of separate trees, i.e. LINKed trees. The LINK strategy, as we have already seen in chapters 2 and 4, has a number of different uses in DS. It is used to encode relative clauses, topicalized subjects, as well as subordinate and coordinate structures. In the case of topicalized subjects, the topicalized subject is parsed on a partial tree containing only a type *e* node which is LINKed to a type *t* requiring tree. The rule of TOPIC STRUCTURE REQUIREMENT then projects a requirement for a copy of the formula of the topic to be found in the main tree. This move is indeed required, since assuming that no requirement would be posited, the topic would be impossible to be constructed as an argument of the proposition expressed by the main tree contrary to fact. However, there is a case in which parsing an NP in a type *e* node LINKed with a type *t* requiring node is possible, without however the need of projecting a requirement for a value of the type *e* formula to be found in the main tree. This case has been traditionally characterized as an ethical dative and involves an NP being interpreted very loosely with respect to the proposition expressed by the main sentence. The NP is not parsed as an argument of the proposition but rather expresses a number of weak relations with respect to this proposition, ranging from affinity to anger and irony. It is this weak relation that Kempson & Cann (2008) and Chatzikyriakidis & Kempson (2009) encode as a LINK relation with no requirement of the

NP's formula value in the main tree:

(6.43) *quid* *tibi* *Celsius* *agit?*
what you.DAT Celsius does
'How, pray, is Celsius?'



In the above structure, the ethical dative is parsed on a LINK structure pretty much like HTLD elements. The difference however lies in the fact that no requirement for a copy of the term projected by the ethical dative-marked expression is posited.

6.2.2 The PCC as a Restriction on Underspecification

Cann & Kempson (2008) and Chatzikyriakidis & Kempson (2009), as already mentioned, argue that clitics in Romance languages can be seen as encoding one of the aforementioned four strategies existing in the Latin scrambling system. Non-syncretized clitics, like for example 3rd accusative clitic *lo* in Spanish, following Bouzouita (2008a,b), are assumed to match the constructive case strategy. In that respect 3rd person accusative clitics are assumed to project fixed structure (the same was assumed in this thesis for the same type of clitics in SMG, GSG and CG).[8] On the other hand, datives and 1st/2nd person clitics are assumed to be structurally underspecified on independent grounds. 1st/2nd person accusative

[8]Note that if one wants to be accurate, 3rd person clitics must first project a locally unfixed node and then fix it immediately if an exact correlation between constructive case and 3rd person clitics is to be drawn. The

clitics by virtue of being case syncretized whereas non syncretized dative clitic forms like *le* by virtue of the dative being structurally underspecified as regards its construal (argumental, benefactive-malefactive, possessor, ethical datives among possibly more functions). As they argue, dative and 1st/2nd person clitics are assumed to project a locally unfixed node without even any output filter, matching the underspecified case in the Latin system. The proposed specifications for Spanish clitics *me/te*, *le* and *lo* in Spanish taken from Chatzikyriakidis & Kempson (2009) are shown below:[9]

(6.44) Lexical entry for *me/te*

IF $?Ty(t), Tn(a)$
THEN IF $[\downarrow_1^+]?Ty(x)|$
$Mood(Imp)$
THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$ **)**
$\texttt{make}(\langle\downarrow_1^+\rangle); \texttt{go}(\langle\downarrow_1^+\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a);$
$Ty(e), Fo(U_{Sp'/Hr'}), ?\exists\mathbf{x}.Fo(\mathbf{x})$
$?\exists\mathbf{x}.Tn(\mathbf{x})); gofirst(?Ty(t))$
ELSE abort
ELSE abort

lexical entries provided for *lo* in Chatzikyriakidis & Kempson (2009) omit the intermediate step of projecting the locally unfixed node. This is what I have assumed in this thesis as well (see the lexical entries for 3rd person accusative clitics in chapters 3 and 4).

[9] I have changed the Kleene star (*) modality used in Chatzikyriakidis & Kempson (2009) to the Kleene plus (+) one for the sake of consistency with the rest of the thesis. Also note that the higher situation are not used in Chatzikyriakidis & Kempson (2009).

(6.45) Lexical entry for *le*

IF $?Ty(t), Tn(a)$

THEN IF $[\downarrow_1^+]?Ty(x)|$
$Mood(Imp)$

THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1^+\rangle); \texttt{go}(\langle\downarrow_1^+\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a);$
$Ty(e), Fo(U_x), ?\exists\mathbf{x}.Fo(\mathbf{x})$
$?\exists\mathbf{x}.Tn(\mathbf{x})); gofirst(?Ty(t))$

ELSE abort

ELSE abort

(6.46) Lexical entry for *lo*

IF $?Ty(t)$

THEN IF $[\downarrow_1^+]?Ty(x)|$
$Mood(Imp)$

THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle)$
$\texttt{put}(Ty(e), Fo(U_x), ?\exists\mathbf{x}.Fo(\mathbf{x}));$
$gofirst(?Ty(t))$

ELSE abort

ELSE abort

Given the lexical entries above, the PCC follows directly from the "no more than one unfixed node at a time" constraint. In parsing *me te*, we end up with one locally unfixed node (since the two locally unfixed nodes will collapse into being the same node) carrying both a metavariable with a *Sp'* restriction as well as a metavariable with a *Hr'* restriction. Update of one of the two metavariables will not update the other and vice versa. The process of parsing *me te* is shown below:

(6.47) After parsing *me* in *me te*

$$Tn(a), ...?Ty(t), \diamond$$

$$\langle\uparrow_1^+\rangle Tn(a),\text{?Ty(x)}$$

$$\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a),$$
$$Ty(e), Fo(U_{Sp'})$$
$$?\exists\mathbf{x}.Fo(\mathbf{x}), ?\exists\mathbf{x}.Tn(\mathbf{x}).$$

(6.48) After parsing *te* in *me te*:

$$Tn(a), ...?Ty(t), \diamond$$

$$\langle\uparrow_1^+\rangle Tn(a),\text{?Ty(x)}$$

$$\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a),$$
$$Ty(e), Fo(U_{Sp'}), Fo(V_{Hr'})$$
$$?\exists\mathbf{x}.Fo(\mathbf{x}), ?\exists\mathbf{x}.Tn(\mathbf{x})$$

It is worth noting that such an account will not exclude cases where one of the clitics of a PCC construction is a strong pronoun. This is because, under standard DS assumptions (Kempson et al., 2001; Cann et al., 2005 among others), strong pronouns will involve a type *e* trigger. In that sense, strong pronouns will be parsed in a fixed treenode position in case the verb has already been parsed or on a regular unfixed node in case they are preverbal.[10] However, in the latter case the "no more than one unfixed node at a time" constraint is not violated in the presence of a clitic parsed on a locally unfixed node, because the two unfixed nodes will have distinct treenode addresses, i.e. one being on a regular

[10]One might ask why the strong pronoun cannot be parsed on a locally unfixed node. There are two reasons for this. The first one is that the rule of LOCAL *ADJUNCTION is not assumed to be a general computational rule for Romance languages (see Bouzouita, 2008 for MedSp). Furthermore, even if LOCAL *ADJUNCTION was assumed to be a general rule of the system of these languages, projecting strong pronouns on such nodes would undergenerate, since strong pronouns can be always interpreted outside their local domain.

unfixed node specified as $\langle\uparrow^*\rangle Tn(a)$ whereas the other on a locally unfixed node specified as $\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a)$. In that sense the restriction will be relevant to nodes with the same underspecified addresses and not to underspecified nodes but with different underspecified addresses.

The account proposed by Cann & Kempson (2008) and Chatzikyriakidis & Kempson (2009) is indeed an intriguing account, since the PCC derives from a very general principle of the tree-growth system. However, Greek seems to be problematic to such an account since no syncretism is found in 1st/2nd person clitics, at least in the singular. The solution to this problem proposed by Chatzikyriakidis & Kempson (2009) (see also Chatzikyriakidis, 2009b) is based on the assumption that 1st/2nd person accusative clitics are also underspecified but singular 1st/2nd person accusative clitics further carry an output filter that dictates the fixing site of the unfixed node as a direct object:

(6.49) Lexical entry for 1st/2nd person accusative clitics in SMG

IF $?Ty(t), Tn(a)$
THEN IF $[\downarrow_1^+]?Ty(x)|$
$Mood(Imp)$
THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1^+\rangle); \texttt{go}(\langle\downarrow_1^+\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a);$
$Ty(e), Fo(U_{Sp'/Hr'}), ?\exists\mathbf{x}.Fo(\mathbf{x})$
$?\exists\mathbf{x}.Tn(\mathbf{x}), ?\langle\uparrow_0\rangle(Ty(e \rightarrow t)));gofirst(?Ty(t))$
ELSE abort
ELSE abort

The fact that of the clitics of SMG that are associated with inducing an unfixed node, it is only 1st/2nd person accusative clitics that contain a filter on output determining some case specification, though a stipulation, is buttressed by the full paradigm of 1st/2nd person clitics. Though singular 1st/2nd person clitics are non-syncretic, their plural counterparts are syncretized with respect to case (*mas*.1PL, *sas*2PL), so the non-syncretism which we

have taken to be definitive of an output-filter specification is in contrast to 3rd person clitics that are non-syncretic across the board. If a unitary analysis of 1st/2nd person clitics in Greek is to be provided, there are then two choices, either to encode plural clitics as projecting fixed nodes, or to encode singular clitics as projecting unfixed nodes despite their non-syncretic forms. The first option is clearly on the wrong track, since it will predict that plural 1st/2nd person clitics are not structurally underspecified. On the other hand, the second option can be naturally encoded given the analysis just proposed, i.e. assuming that 1st/2nd person accusative clitics, even though unfixed, can be defined as projecting a case-filter on output while nevertheless not incrementally fixing the structural relation. Adger & Harbour (2007) also explain syncretism in the Greek case by referring to the non-syncretic plural 1st/2nd person forms, albeit with different argumentation (see Adger and Harbour (2007) for the relevant argumentation).

Vindication that at least the strong PCC effects are grounded in a processing constraint comes from Pontic Greek (PG), a dialect which as already discussed in chapter 5, disallows any combinations of 3rd person clitics:

(6.50)

a. **Edeke aton ato/a*
gave.3SG him.CL it.CL
'He gave it to him.'

b. **Edeke ats ato/a*
gave.3SG him.CL it.CL
'He gave it to them.'

PG is syncretic across the board in the sense that all clitics can appear as either direct or indirect objects. Assuming an analysis of clitics in PG where these project locally unfixed nodes, the above facts are straightforwardly captured. Recall that the illicit 3rd person combinations would be a major challenge to feature oriented explanations (Bejar & Rezac, 2003; Anagnostopoulou, 2003, 2005; Rezac, 2008a; Nevins, 2007; Adger and Harbour, 2007; Michelioudakis, 2009), since it is very hard to see how such constructions would be precluded in clitic systems where a person/participant/author feature is taken to be the culprit behind person restrictions.

The analysis proposed in Chatzikyriakidis and Kempson (2009) captures the PCC in Greek using the same mechanisms as in syncretized languages like Spanish or Italian. Furthermore, the analysis has the advantage of capturing the PG facts where sequences of 3rd person clitics are altogether banned.[11]

6.2.3 The PCC in the other Dialects

GSG and CG follow the pattern found in SMG as regards the PCC. Both dialects, like SMG, follow the strong version of the PCC:

(6.51) **Eδoke* *tu* *me*
gave.3SG him.CL-GEN me.CL-GEN
'She/He/It gave me to him.' [CG]

(6.52) **Tu* *me* *edika*
him.CL-GEN me.CL-GEN gave.3SG
'She/He/It gave me to him.' [GSG]

(6.53) **Eδoke* *mu* *se*
gave.3SG me.CL-GEN you.CL-ACC
'She/He/It gave you to me.' [CG]

(6.54) **Mu* *se* *edika*
me.CL-GEN you.CL-ACC gave.3SG
'She/He/It gave you to me.' [GSG]

(6.55) **Eδoke* *su* *me*
gave.3SG you.CL-GEN me.CL-ACC
'She/He/It gave me to you.' [CG]

(6.56) **Su* *me* *edika*
you.CL-GEN me.CL-ACC gave.3SG
'She/He/It gave me to you.' [GSG]

[11] In a similar vein, data from Gascon further vindicate this analysis. Gascon is case syncretic with 3rd person clitics, except that in clitic clusters the clitic standing for the direct object must change to neuter (Nicol, 2005), since two identical clitics are not allowed (as the system proposed would predict). Furthermore, the neuter clitic is marked for accusative and, given the DS perspective, being not syncretic, it would naturally be modellable as inducing the construction of a fixed node.

However, PG seems to exhibit the weaker version of the constraint at least for some speakers. In that sense, clusters of a 1st plus a 2nd person clitic or vice versa are licit in PG as witness the examples below:

(6.57) *Eδikse/eδeknise* *m* *ese(n)*
showed.3SG me.CL you.CL
'S/He/It showed you to me.'

Note that the interpretation in the above examples is always DAT-ACC. The ordering ACC-DAT can be only produced with one clitic and one strong pronoun headed by the preposition 'to':

(6.58) *Eδikse/eδeknise* *me* *s* *ese(n)*
showed.3SG me.CL to you
'S/He/It showed you to me.'

(6.59) *Eδikse/eδeknise* *se* *s* *eme(n)*
showed.3SG you.CL to me.CL
'S/He/It showed you to me.'

There is also the possibility of DAT-ACC ordering with the second element being a stressed strong pronoun:

(6.60) *Eδikse/eδeknise* *me* *ESE(N)*
showed.3SG me.CL you
'It was you that she/he/it showed to me.'

(6.61) *Eδikse/eδeknise* *se* *EME(N)*
showed.3SG you.CL me
'It was me that she/he/it showed to you.'

In that respect, the only example involving clitic clusters is (6.57). So, how are cases like this going to fit within the account proposed? The answer is extremely straightforward and involves the same reasoning used in chapter 5 in order to explain clusters of a 1st/2nd plus a 3rd person clitic. Remember that in cluster constructions only the forms *m/s* can be used. Furthermore, these forms cannot stand on their own in single clitic constructions.

Thus, *m/s* are always interpreted as indirect objects. But if this is true, then they are not underspecified anymore, and as such do not project a locally unfixed node. The consequence is that a cluster containing one of these forms *m/s* will not be subject to the "one unfixed node at a time constraint", since only one locally unfixed node will exist after parsing a cluster like the one in (6.57). Clusters like these are parsed as involving one single entry, where the *m/s* forms first fix their position in the indirect object node and the other clitic projects an unfixed node:

(6.62) Lexical entry for *m ese(n)* 'to me you'

IF $?Ty(t), Tn(a)$
THEN IF $<\downarrow_1^+> Ty(x)$
THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$
$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(Ty(e), Fo(U_{Sp'}), ?\exists\mathbf{x}.Fo(\mathbf{x})); gofirst(?Ty(t));$
($\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1^+\rangle); \texttt{go}(\langle\downarrow_1^+\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(Ty(e), Fo(V_{Hr'}), ?\exists\mathbf{x}.Fo(\mathbf{x}), \langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a));$
$gofirst(?Ty(t))$
ELSE abort
ELSE abort

Thus, the peculiar characteristics of PG, i.e. the fact that it allows clusters of a 1st plus a 2nd person clitics and vice versa, while disallowing 3rd person clitics altogether, receives a straightforward explanation under the account proposed.

It is worth noting that similar data have been preported for RP in Michelioudakis & Sitaridou (2010):

(6.63) *Eδiksane* *m* *ese*
showed.3PL me.CL you.CL
'They showed you to me.' [Sürmene variety of RP, Michelioudakis & Sitaridou, 2010]

Note that in the above example the reduced form *m* is used. The form *m* is the form that is only used in clusters in PG. I do not know whether this fact holds for RP as well. In case however this is true and *m* is indeed used in RP only in cluster constructions, then two important conclusions are derived: RP has clitic clusters contrary to what has been claimed by Michelioudakis & Sitaridou (2010) and the explanation of constructions like (6.63) is effectively the same with the one provided for the equivalent PG facts. More data are definitely needed to further elucidate the situation as regards clitic clustering in RP. To be more specific, we need clear evidence to show whether RP has clitic clusters and if yes, what kind of clitic clusters these are. On the other hand, more data from PG as regards the weak PCC constructions are also needed in order to determine whether such clusters are generalized for every speaker or vary from speaker to speaker as my preliminary data show.

6.2.4 Ethical Datives

The so-called ethical datives (EDs) have long provided a challenge to both descrptive and theoretical linguistics (see e.g. van Hoecke, 1996; Sitaridou 1998; Michelioudakis 2007; Franco & Huidboro, 2008; Jouitteau & Rezac, 2008). It is quite hard to define what the term 'ethical dative' means. Roughly, EDs are non-argumental datives usually expressed by clitics in clitic languages. The meaning of these datives is discourse oriented in most cases and refers to the the speaker's or hearer's involvement in the event expressed by the sentence in which the ED appears. The semantic relation of EDs with respect to the proposition expressed by the sentence is rather loose and can take a number of interpretations ranging from malefactive/benefactive to the expression of anger, irony, sadness, disappointment. In SMG, EDs are expressed by dative clitics and can be found with all types of active and middle verbs. A number of examples are shown below:

(6.64) *Mu* *to* *arostisan* *to* *peδi*
me.CL-GEN it.CL-ACC made-sick.3PL the child
'They made the child sick (and I'm affected negatively).'

(6.65) *Tora* *pu* *mas* *efiγe* *i* *atiχia,* *ola* *tha* *pane* *kala*
now that us.CL-GEN went the bad-luck all FUT go good
'Everything will be ok, since bad luck has gone away (and this is to our advantage).'

(6.66) *O iliθios pai ke mu to δini sto Γiorγo anti*
the idiot goes and me.CL-GEN it.CL-ACC gives to-the George instead
na to δosi se mena
SUBJ it.CL-ACC give to me.ACC
'Instead of giving it to me, the idiot is giving it to George (and I'm reinforcing the claim he is an idiot'.

The range of interpretations an ethical dative can get is wide and depends largely on the context. For example in (6.66), the contribution of the ED is to reinforce the speakers' belief that the person giving something to George is an idiot. This is a case where the semantic contribution of the ED is extremely loose with respect to the proposition as compared to examples like (6.64) and (6.65) where the semantic contribution of the ethical is far more clear (even though still loose). A fact often noted in the literature is that EDs escape the PCC (Albizu, 1997; Jouitteau & Rezac, 2008 among others):

(6.67) *Demain je (me) vous (me) emmène en vacances*
tomorrow I me.CL you.CL me.CL take in vacations
'Tomorrow I will take you on vacation.'[French - Jouitteau & Rezac, 2008]

(6.68) *Me li van dir que havia suspès l' examen*
me.CL him.CL-DAT said that had failed the exam
'They told him (on me) that he had failed the exam.'[Catalan - Albizu, 1997]

However, SMG provides the exception to this fact, since the PCC is active with EDs as well:[12]

(6.69) *Mu se arostisan to peδi*
me.CL-GEN you.CL-ACC made-sick.3PL the child
'They made the child sick (and I'm affected negatively).'

(6.70) **O iliθios pai ke mu se δini sto Γiorγo anti*
the idiot goes and me.CL-GEN you.CL-ACC gives to-the George instead
na se δosi se mena
SUBJ you.CL-ACC give to me.ACC
'Instead of giving it to me, the idiot is giving it to George'.

[12]The same facts hold for CG. PG lacks ethical datives altogether, while for GSG, unfortunately there are no relevant data available. The discussion on ethical datives will concentrate on SMG.

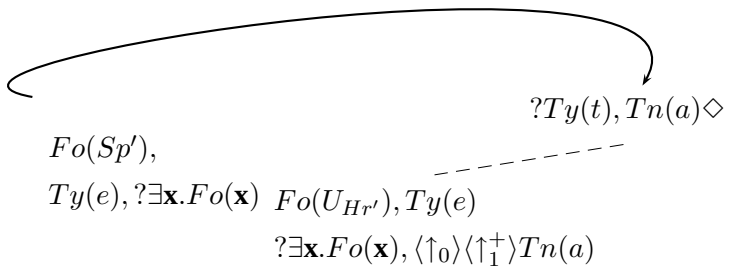
The standard DS analysis as regards EDs assumes the development of homonymy for the dative, i.e. an alternative strategy available for dative clitics which parses the clitic on a LINKed node. The analysis is based on the fact that EDs are non-argumental and furthermore their semantic contribution is rather weak or in a way peripheral to the main clause. Kempson & Chatzikyriakidis (2009) analyze ethical datives in languages like Spanish or Italian as being parsed on a LINKed node. In that sense, the lexical entries for dative clitics in these languages are assumed to have developed a second parsing strategy, which builds the LINKed node and decorates it with a type *e* value standing for the ED. No requirement for a copy of the term of the ED is posited in the main tree, and as such the ethical dative is interpreted as non-argumental. The lexical entry for datives proposed by Kempson & Chatzikyriakidis (2009) are shown below:

(6.71) Lexical entry for datives in Italian/Spanish

IF	$?Ty(t)$	
THEN	IF	$[\downarrow_1^+]?\exists x.Tn(x)\vert$
		$Mood(Imp)$
	THEN	**(**$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$**)**
		$\texttt{make}(\langle\downarrow_1^+\rangle);\texttt{go}(\langle\downarrow_1^+\rangle);$
		$\texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle);$
		$\texttt{put}(\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle?Ty(t);$
		$Ty(e), Fo(U_x), ?\exists\mathbf{x}.Fo(\mathbf{x})$
		$?\exists\mathbf{x}.Tn(\mathbf{x}));gofirst(?Ty(t))$
	ELSE	$\texttt{make}(\langle L\rangle);\texttt{go}(\langle L\rangle);$
		$\texttt{put}(Ty(e), Fo(U), ?\exists x.Fo(x);\texttt{go}(\langle L^{-1}\rangle)$
ELSE	abort	

Parsing of an ethical dative on a LINK structure predicts that the PCC should not be active in these constructions, since the "no more than one unfixed node at a time" constraint is not operative in these cases. Parsing an illicit PCC combination like *me te* in Spanish, is predicted to be grammatical assuming one of the clitics is interpreted as an ED according to fact:

(6.72) Parsing *me te* with *me* being an ethical dative



It is obvious that such an analysis will produce the wrong results for SMG, since as already said, ethical datives do NOT escape the PCC in SMG. If ethical datives in SMG are not analyzed as LINKs, then how can they be analyzed? Abstracting away from framework dependent assumptions, the question can be rephrased as follows: If ethical datives are not non-argumental, then what they are? What I am going to argue is that ethical datives are actually arguments, even though optional ones, of the verb in SMG. The solution I am going to propose is based on Marten's (2002) concept of optional arguments. Under this approach, verbs are always underspecified with respect to their type. Verbs subcategorize for all their obligatory arguments but their type is underspecified and can host additional optional arguments if necessary. For example, a monotransitive verb will be specified as Ty($e^* \rightarrow$ (e $\rightarrow$ (e $\rightarrow$ t))) instead of Ty(e $\rightarrow$ (e $\rightarrow$ t)). Assuming such an account, EDs in Greek will be argued to be parsed as optional arguments rather than non-arguments.[13] By analyzing EDs as optional arguments, the lexical entry for dative clitics in SMG can stay unaltered, since EDs will be dealt with in the same sense argumental datives are. Such an analysis will work. The question however is whether there is any sort of evidence for such a treatment of SMG EDs. The first thing that should be noted is that such a proposal is not novel for SMG. Michelioudakis (2007) considers EDs (contra almost everyone) to be non-core arguments that are introduced by a higher applicative head assigning a sentient interpretation and bearing no propositional role. The evidence for such treatment of EDs comes from doubling constuctions. Doubling of an ED with a strong pronoun is licit, at

[13]See Marten (2002) for the formal details of such a proposal.

least in some cases. It is not that usual or natural as argumental clitic doubling but in general it seems to be acceptable as witness the examples below:

(6.73)

a. *Emena mu arostise to peδi*
me.GEN me.CL-GEN made-sick the child

b. *Mu arostise to peδi emena*
me.CL-GEN got-sick the child me.GEN
'S/He made my child sick (and this is bad for me).'

This fact has already been noted by Michelioudakis (2007). However, Michelioudakis (2007) claims that doubling of an ED is only possible in case the strong pronoun is topicalized. This last argument is made in order to claim that strong pronouns are not merged in vp internal position, but are rather base-generated in a D-Linked position. Given that, they can participate in dislocation constructions but not in clitic doubling constructions in the sense of Anagnostopoulou (2006). Michelioudakis (2007) provides the following examples:

(6.74)

a. *(Emena), #δen mu troi o Janis kreas, #(emena)*
me.GEN NEG me.CL-GEN eats the John meat me.GEN

b. *?*Δen mu troi emena kreas o Janis*
NEG me.CL-GEN eats me.GEN meat the John

c. **Δen mu troi o Janis emena kreas*
NEG me.CL-GEN eats the John me.GEN meat
'To my disappointment, John does not eat meat.'

However my intuitions as well as those of my informants strongly disagree with the above judgments. All the above examples are grammatical (given the right intonation) according to my intuitions. The point to prove here is not that strong pronouns in these constructions are merged in vp internal position (since this is rather irrelevant to a DS account) but that ED doubling is more widespread than what Michelioudakis' (2007) assumes. Even

the examples involving clitic doubling (and not dislocation) of a 3rd person ethical dative by a full NP are to me at least as grammatical as the examples involving dislocation contra Michelioudakis (2007: 14). Michelioudakis gives the following judgments with respect to doubling of a 3rd person ethical dative by a full NP (cf. the differences in judgement between dislocation and doubling (CLLD and CD) constructions in Michelioudakis, i.e. % against *):

(6.75)

a. *%Tis Marias, δen tis meletai o Γianis arketa*
the Mary.GEN NEG her.CL-GEN studies the John enough

b. **Δen tis meletai o Γianis tis Marias arketa*
NEG her.CL-GEN studies the John the Mary.GEN enough

c. *%Δen tis meletai o Γianis arketa tis Marias*
NEG her.CL-GEN studies the John enough the Mary.GEN

d. **O Γianis δen tis meletai tis Marias (arketa)*
the John NEG her.CL-GEN studies the Mary.GEN enough

e. **Δen tis meletai (arketa) tis Marias (arketa) o Γianis(arketa)*
NEG her.CL-GEN studies enough the Mary.GEN enough the John enough
'To Mary's disappointment, John does not study enough.'

My judgments as well as those of my informants are uniform across all these sentences and again given the right intonation and context all the above examples can be grammatical. Additional examples that do not conform to the grammaticality judgments posited by Michelioudakis are shown below:

(6.76) *Δe mu katestrepsan emena tipota afu su ipa*
NEG me.CL-GEN destroyed.3PL me nothing because you.CL-GEN told.1SG
'I told you so, they didn't destroy anything on me.'

(6.77) *Siγa ti su kanane esena, mu xoun katastrepsei*
slow what you.CL-GEN did.3PL you me.CL-GEN have.3PL destroyed
emena praγmata
me things
'It is not a big deal what they did to you, you must see how many things they have destroyed on me.' [free translation]

(6.78) *Entometaksi, mas peθeni emas i γiaγia ke δen*
meanwhile us.CL dies us the grandmother.NOM and NEG
kserume ti na kanume
know.3PL what SUBJ do.3PL
'In the meanwhile, grandmother dies and we do not know what to do.'

What is more, even the claim that EDs should only be clitic pronouns can be challenged for SMG. A number of examples involving a strong pronoun acting as an ED without the ethical dative clitic can be found in SMG. Here are some:

(6.79)

a. *Emena/EMENA troi poli to pedi tu antra mu kai den ksero ti*
me.GEN eats very the child the husband my and not know what
na kano
SUBJ do
'To my disadvantage, my husband's child eats a lot and I do not know what to do.'

b. *Emena/EMENA to edose sto Γiorγo o ilithios*
me.GEN it.CL-ACC gave to-the George the idiot
'The idiot gave it to George (and I'm angry or I'm reinforcing that the giver is an idiot).'

The above is quite natural with both a contrastive focus and a non-focus interpretation of the strong pronoun. If only the focus interpretation of ethical datives were possible with strong pronouns and no doubling, one could claim that this is due to the inability of clitics to receive a focus interpretation. However, non-focus interpretations, as we have seen above, are also possible. The situation is extremely complicated and indeed Michelioudakis' (2007: 15) claim that strong pronouns are not possible without the ethical clitic is true in a number of instances. However, the fact that constructions like the above exist

remains a fact. Furthermore, the claim that only strong pronouns in left or right dislocation can double an ethical clitic is also true in some instances. For example, in ditransitive constructions where the ethical clitic appears in a clitic cluster along with the clitic standing for the theme, doubling of the ethical clitic is only possible when the strong pronoun is left or right dislocated:

(6.80)

a.	*Emena*	*mu*	*to*	*eδose*	*sto*	*Γiorγo*	*o*	*ilithios*
	me.GEN	me.CL-GEN	it.CL-ACC	gave	to-the	George	the	idiot

b.	*?Mu*	*to*	*eδose*	*sto*	*Γiorγo*	*o*	*ilithios*	*emena*
	me.CL-GEN	it.CL-ACC	gave	to-the	George	the	idiot	me.GEN

c.	**Mu*	*to*	*eδose*	*emena*	*sto*	*Γiorγo*	*o*	*ilithios*
	me.CL-GEN	it.CL-ACC	gave	to-the	George	me.GEN	the	idiot

'The idiot gave it to George (and I'm angry or I'm reinforcing that the giver is an idiot).'

The situation is far from clear, since a number of ethical constructions rely on contextual or intonational information for grammaticality, a fact already noted by Michelioudakis (2007). The discussion is vast and the number of parameters that need to be taken into consideration is far too high to be discussed in this thesis. The claim of this thesis is that EDS in SMG can be analyzed as optional arguments, and as such can be given a unitary analysis along with argumental clitics. In that sense, I concur with Michelioudakis (2007), although disagreeing in the way such a fact is formalized, and further disagreeing with pretty much everyone else. There might be a functional explanation for why SMG did not develop a distinct strategy for EDs in the sense Spanish or Italian clitics did by developing a second strategy of parsing dative clitics as LINK structures. Looking at the distribution of ethical datives in SMG we see that from the three types of ethical datives identified for Spanish in Franco & Huidboro (2008), only one type is possible in SMG. The fact that EDs are not that widespread might suggest that linguistic input was not enough in order for a separate strategy to be developed for EDs in SMG, contrary to what happened in Spanish. However, such a claim remains speculative and needs to be backed by further empirical evidence

that I do not have at the moment. However, it is something that might point towards an explanation of the reason ethical datives in SMG are behaving as arguments in a number of instances.

Ending our discussion on the behaviour of EDs ethical datives under the PCC, it should be noted that there is a construction involving EDs that the analysis proposed cannot capture in this present form. The construction in question involves a sequence of two dative clitics, of which one is an ED. In general, sequences of two dative clitics are banned in SMG, even when one of them is an ED. However, there are a number of examples as reported in Michelioudakis (2007) involving an ethical plus an argumental dative which are at least grammatical for some speakers. Sitaridou (1998: 7) rightly points out that sequences of two dative clitics are not strictly ruled out in SMG:

(6.81)

a. *%Na* *mu* *tis* *δosis* *pola* *filia*
SUBJ me.CL-GEN her.CL-GEN give.2SG many kisses
'Give her many kisses, on behalf of me.'

b. *%Na* *min* *mu* *tis* *aγorasis* *tipota*
SUBJ NEG me.CL-GEN her.CL-GEN buy,2SG nothing
'Do not buy her anything (I strictly order you).'

My analysis will rule out the above examples since both dative clitics will be assumed to project locally unfixed nodes, and as such should be subject to the PCC given the "one unfixed node at a time" constraint. What should be noted however, besides the fact that these examples are not generally accepted, is that the form of the 3rd person argumental dative clitic plays a role in the grammaticality of the above examples. My intuitions as well as those of most of my informants accept both the examples presented above, but however do not accept the same examples in case the 3rd person feminine clitic is substituted with the 3rd person masculine one. What is worse, the first example involving a pure ditransitive verb is judged to be more ungrammatical than the second example, in which a verb that has a benefactive clitic is involved:

(6.82)

a. **Na mu tu δosis pola filia*
SUBJ me.CL-GEN him.CL-GEN give.2SG many kisses
'Give him many kisses, on behalf of me.'

b. *???Na min mu tu aγorasis tipota*
SUBJ NEG me.CL-GEN him.CL-GEN buy.2SG nothing
'Do not buy him anything (I strictly order you).'

Under the light of these data, I am reluctant to declare the proposed analysis inadequate. However, the fact that a number of examples are left unaccounted for remains. At the moment, there is no good explanation for that under the proposed account. It must be noted however, that declaring the account proposed inadequate cannot be based on the present data alone but further needs to take more solid data into consideration to decide whether such an account is on the right track or not. In that sense, I leave this issue as a subject of further research.

With this last remark, I stop the discussion on ethical datives and their behaviour under the PCC, directing the interested reader to Sitaridou (1998) and Michelioudakis (2007) for extensive and illuminating discussions on the issue.

6.3 Doubling

The term clitic doubling is ambiguous in the linguistic literature. It has both a general and a specific meaning. The general meaning refers to constructions where an argumental clitic doubles a full NP. However, there is a more specific interpretation of the term clitic doubling, which specifically refers to one kind of such construction, i.e. the construction where the doubled NP follows the clitic. I will use the term clitic doubling (CD) to refer to that specific construction and the term Doubling to refer to the phenomenon of doubling in general.

Doubling comes into two basic guises. The first one involves dislocation of the doubled NP to the left periphery:

(6.83) *Ton* *Γiorγo* *ton* *ksero*
the.ACC George.ACC him.CL-ACC know
'I know George'. [SMG]

In the second guise, already coined as CD, the doubled NP follows the clitic and appears postverbally:

(6.84) *Ton* *ksero* *ton* *Γiorγo*
him.CL-ACC know the.ACC George.ACC
'I know George'. [SMG]

Doubled NPs in SMG violate Kayne's generalization (1975) according to which a doubled NP must be always headed by a case assigning preposition. This is indeed what we find in languages like Spanish, where the doubled NP is headed by the preposition *a*:

(6.85) *Lo* *vi* **(a)* *tu* *papá*
him.CL-ACC saw to your father
'I saw your father.'

SMG behaves in the exact opposite way, i.e. ungrammaticality obtains if the doubled NP is headed by a preposition:

(6.86)

a. **Ston* *Γiorγo* *ton* *ksero*
to-the.ACC George.ACC him.CL-ACC know

b. **Ton* *ksero* *ston* *Γiorγo*
him.CL-ACC know to-the.ACC George.ACC
'I know George'.

In what follows, I'm going to discuss and provide an analysis of CLLD and CD in SMG first and then move on to the implications of such an analysis for CG and PG.

6.3.1 Clitic Left Dislocation and Clitic Doubling in SMG

Clitic Left Dislocation (CLLD) is an instance of left dislocation whereby a left dislocated NP is doubled by a clitic. CLLD should not be confused with Hanging Topic Left Dislocation (HTLD), a construction also involving a left dislocated NP and a doubling clitic. There are a number of diagnostics distinguishing between the two. The first diagnostic involves case connectivity. CLLD requires case connectivity between the clitic and the doubled NP, HTLD does not:

(6.87) *Ton* *Γiorγo* *ton* *ksero*
the.ACC George.ACC him.ACC know
'I know George.' [CLLD]

(6.88) *O* *Γiorγos,* *ton* *ksero*
the.NOM George.NOM him.ACC know
'I know George.' [HTLD]

Notice that the HTLD example above has a comma after the dislocated NP indicating an intonational break. Such an intonational break is needed for HTLD but is not neccessary in CLLD:

(6.89) *Ton* *Γiorγo(,)* *ton* *ksero*
the.ACC George.ACC him.ACC know
'I know George.' [CLLD]

(6.90) *O* *Γiorγos*(,)* *ton* *ksero*
the.NOM George.NOM him.ACC know
'I know George.' [HTLD]

Furthermore, HTLD, contrary to CLLD, is insensitive to islands (example adapted slightly modified from Anagnostopoulou, 1999):

(6.91) *O* *Γiorγos,* *xtes* *γnorisa* *tin* *γineka* *pu*
the.NOM George.NOM yesterday met the.ACC woman.ACC that
ton *pantreftike* *ke* *aporo* *pos* *tin* *andexi*
him.ACC married and wonder how him.CL-ACC stands

(6.92) **Ton* *Γiorγo,* *xtes* *γnorisa* *tin* *γineka* *pu* *ton*
the.ACC George.ACC yesterday met the.ACC woman.ACC that him.ACC
pantreftike *ke* *aporo* *pos* *tin* *andexi*
married and wonder how him.CL-ACC stands
'George, yesterday I met the woman who married him and I wonder how she can stand him.'

It is worth mentioning however that CLLD is not sensitive to complement clause islands, both factive and non-factive:

(6.93) *Ton* *Γiorγo* *ksero* *oti* *ton* *xtipises*
the.ACC George.ACC know.1SG that him.CL-ACC hit.2SG
'I know that you hit George.' [non-factive complement]

(6.94) *Ton* *Γiorγo* *lipame* *pu* *ton* *xtipises*
the.ACC George.ACC be-sorry.1SG that him.CL-ACC hit.2SG
'I'm sorry that you hit George.' [factive complement]

One last criterion distinguishing between the two structures is embedding. HTLD is a root phenomenon only, while CLLD can be found in embedded contexts as well:

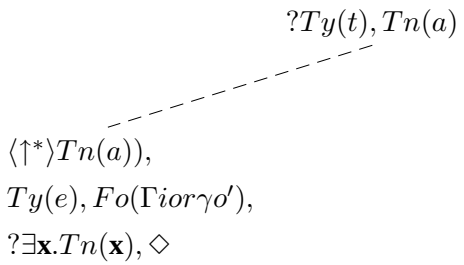
(6.95) *Mu* *ipe* *oti* *ton* *Γiorγo* *ton* *xtipise*
me.CL-GEN said.3SG that the.ACC George.ACC him.CL-ACC hit.3SG

(6.96) **Mu* *ipe* *oti* *o* *Γiorγos,* *ton* *xtipise*
me.CL-GEN said.3SG that the.NOM George.NOM him.CL-ACC hit.3SG
'S/He/It told me that s/he/it hit George.'

CLLD has received considerable attention in the linguistic literature and a number of analyses have been proposed by the years within the generative tradition. The core debate with respect to CLLD focuses on whether CLLD involves base generation of the NP in the left periphery (Cinque 1990; Iatridou 1995; Anagnostopoulou 1997 among others) or movement of the NP from a canonical argument position to the left periphery (Cinque 1977; Kayne 1994; Philippaki & Spyropoulos 2002 among others). The exact details of these analyses are not going to be discussed here. The interested reader is directed to Cinque (1977, 1990), Iatridou (1995), Anagnostopoulou (1997) and Philippaki et al. (2002) among others for a detailed discussion.

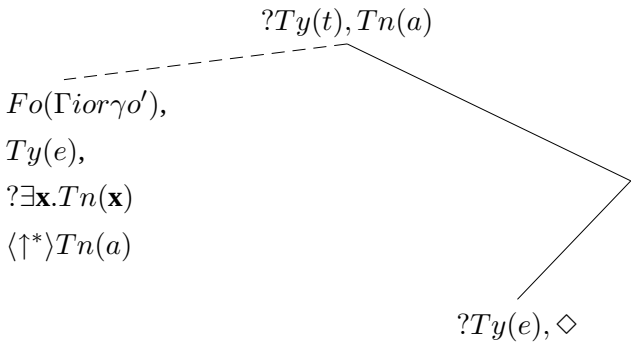
The analysis I am going to propose is based on existing DS analyses of both CLLD and CD (Cann et al. 2005; Gregoromichelaki 2010). In discussing CLLD, Cann et al. (2005) argue that there are two ways in which this can be analyzed. According to the first, the left dislocated NP is parsed on an unfixed node. In a CLLD case like (6.87) the rule of *ADJUNCTION applies first. Then, the full NP is parsed on that unfixed node:

(6.97) Parsing *ton Γiorγo* 'the.ACC George.CL-ACC' in *ton Γiorγo ton ksero* 'I know George' on an unfixed node:



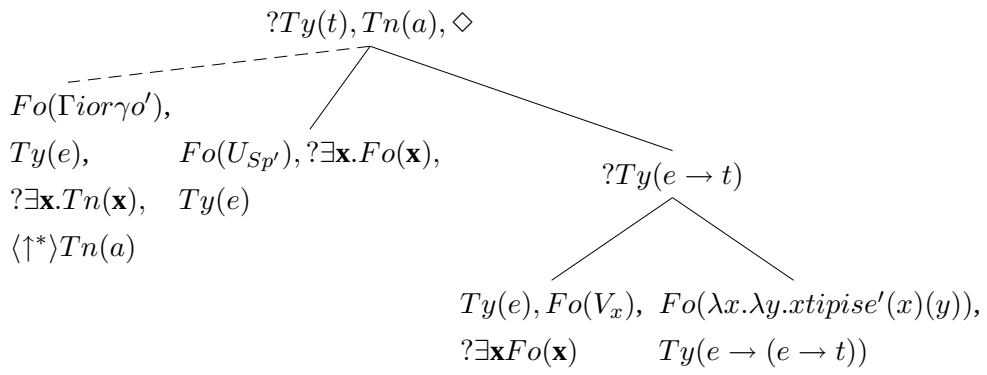
Then the accusative clitic comes into parse, building and decorating the direct object node with a type value and a formula metavariable:

(6.98) Parsing *ton* 'him.CL-ACC' in *ton Γiorγo ton ksero* 'I know George'



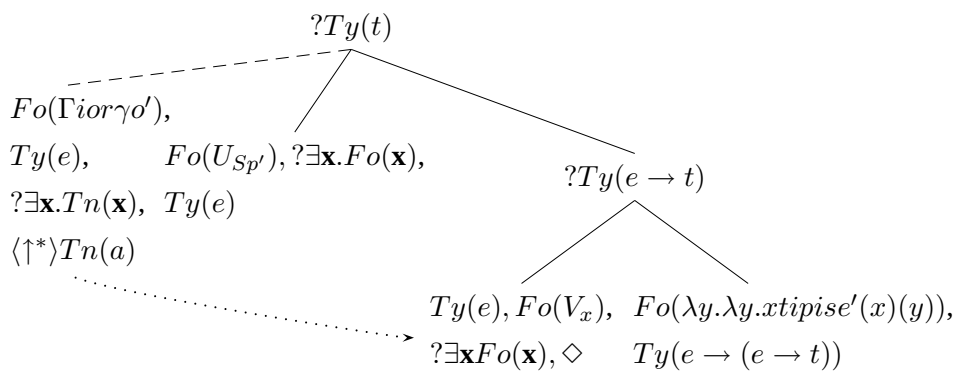
The next step is parsing of the verb. The resulting structure after the verb has been parsed is shown below:

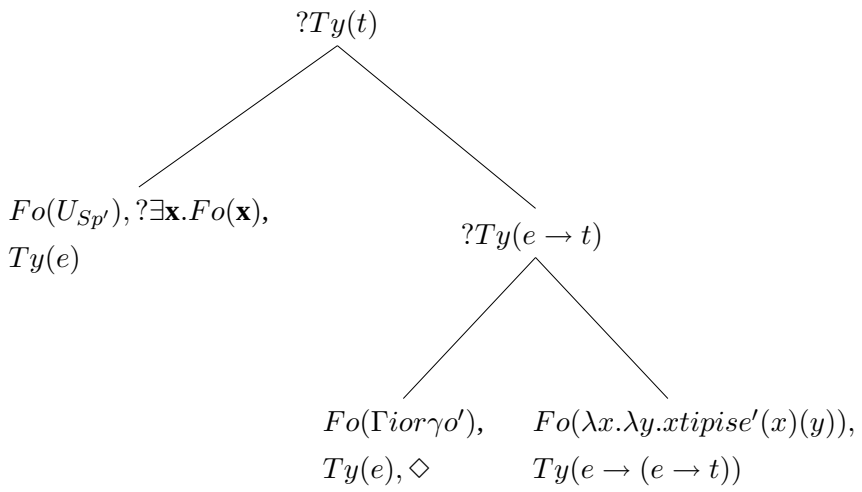
(6.99) Parsing the verb



The pointer can move down to the object node via ANTICIPATION since an unsatisfied requirement exists in that node ($?\exists \mathbf{x} Fo(\mathbf{x})$). At that point, MERGE can apply, unifying the unfixed node with the object node:

(6.100) Before and after MERGE of the unfixed node

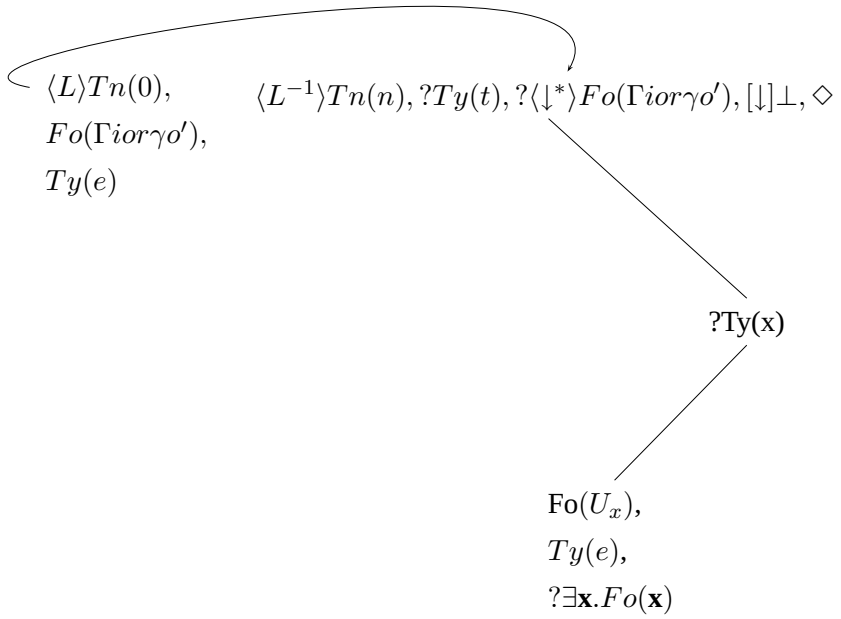




From that point on, the parse can be completed following application of the relevant rules.

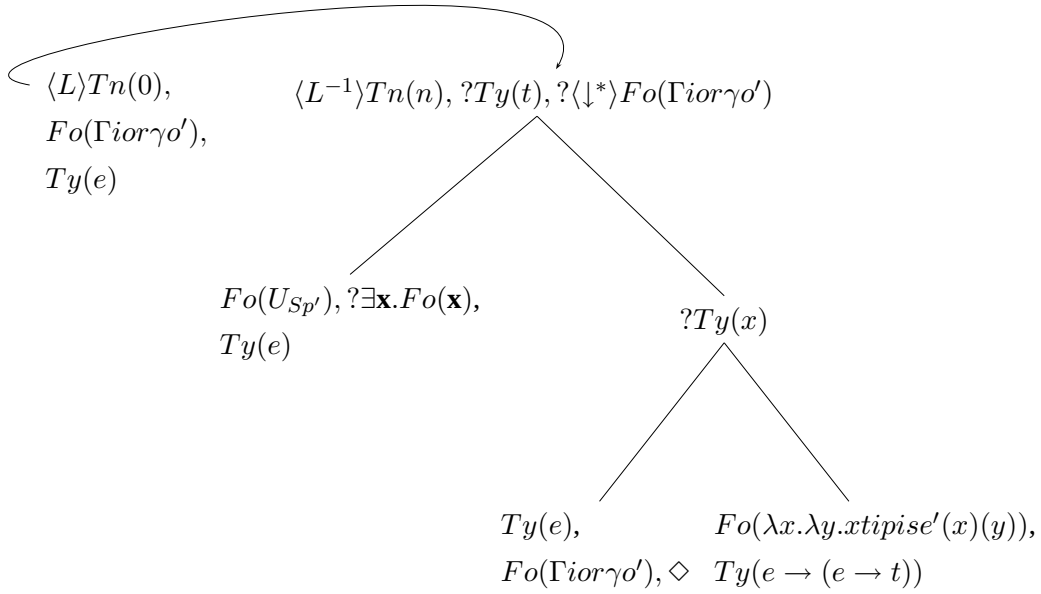
The second way of analyzing CLLD proposed by Cann et al. (2005) assumes that CLLD structures make use of the same strategy used in HTLD constructions, namely the LINK strategy. The only difference between the two structures is the choice of the modal operator in each case. HTLD, being insensitive to relative clause islands, makes use of the $\langle D \rangle$ operator in defining the tree space in which the copy of the dislocated NP, i.e. the clitic pronoun, must be found. On the other hand, CLLD structures use the kleene star (*) operator, effectively excluding cases where the clitic appears inside a relative clause (since relative clauses will extend over a separate tree structure not captured by (*)). In the following tree, the dislocated NP has been parsed on a LINK structure while a requirement for a copy of its *Fo* value is posited in the LINKed tree. This formula value must be found somewhere below the LINKed tree (*) but not anywhere outside this tree (as the *D* operator would predict). The clitic is parsed in the LINKed tree and projects a type value and a formula metavariable:

(6.101) Parsing *ton* Γ*iorγo ton* in ‘ton Γiorγo ton xtipise’ with LINK



The verb is parsed providing the rest of the propositional structure. The next step is substitution of the metavariable projected by the clitic by the formula value appearing in the structure where the LINK begins (the dislocated NP):

(6.102) Parsing *xtipise* in ‘ton Γiorγo ton xtipise’



The question that arises is which of the two structures is relevant for SMG CLLD. Cann et al. (2005) argue that the unfixed node analysis is suited for languages where quantifiers are allowed to appear in CLLD constructions. On the other hand, the LINK structure analysis is suited for cases where only definites and quantifiers with specific interpretations are possible in CLLD. The reasoning used is based on the analysis of scope. As we saw in our introduction to DS, quantified NPs need scope evaluation within an individual tree. Cann et al. (2005) however argue that indefinites can extend scope across independent domains. Within this line of reasoning, Cann et al. (2005) analyze direct object CLLD in Spanish as involving the LINK strategy, but indirect object CLLD where quantified NPs are freely allowed, as involving the unfixed node strategy. However, things are not that simple. Given the analysis of indefinites presented in Cann et al. (2005), it is impossible even for indefinites parsed on a LINK structure to be part of a scopal statement. This is because the sole definition of the scopal requirement *?Sc(x)* precludes the case where a quantified NP (any kind of quantified NP) is parsed on a LINK structure:

(6.103) $?SC(x) =_{def} ?\langle\uparrow_0\rangle\langle\uparrow_1^*\rangle\exists y.Scope(y < x) \vee Scope(x < y)$

Under the above definition, the quantified phrase must participate in the most local scopal statement above. The modalities used to capture this locality $?\langle\uparrow_0\rangle\langle\uparrow_1^*\rangle$ will not capture the case where the quantified NP is parsed on a LINK structure, since in that case the $\langle L\rangle$ relation will be needed in order to reach the type *t* requiring node of the LINKed tree where scope is going to be calculated. The structure below shows a case where an indefinite has been parsed on a LINK structure (omitting the complex NP structure):

(6.104)

$\langle L\rangle Tn(0)$
$Fo(\epsilon, student', (x)), ?Sc(x),$
$Ty(e)$

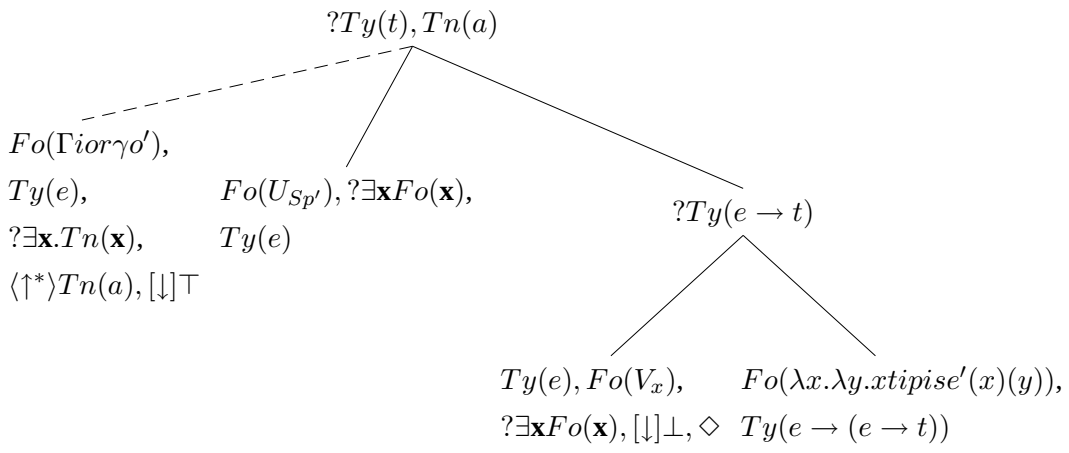
$Tn(0), ?Ty(t), ?\langle D\rangle Fo(\epsilon, student', (x)), Scope(S_i)$

In the above structure, the scopal requirement *?Sc(x)* can never be satisfied since by definition it cannot reach the node where scope is calculated, the reason being that the modalities used in the definition of *?Sc(x)* do not capture the case where a LINK relation

needs to be traversed. What is worse, the predicate *?Sc(x)* is assumed in both Kempson et al. (2001) and Cann et al. (2005) to be projected from the lexical entry of common nouns and not determiners (Kempson et al. 2001: 239; Cann et al. 2005: 107). Such a fact will predict that definites are also impossible on LINK structures given the definition of *?Sc(x)*. In that respect, even the HTLD account will have to be modified in case we want left dislocated elements introduced as topics to participate in the scopal statement. It seems in that respect that an analysis of CLLD in SMG as involving an unfixed node in the left periphery that will later on will unify with the fixed clitic node is more plausible. At the moment, the only possible way to get scope interactions is within a tree that contains a type *t* node. Such a fact, as already said, is impossible in cases where the dislocated NP is parsed on a LINK structure. An additional bonus of analyzing CLLD as involving an unfixed node, is that unavailability of doubling the NP with a strong pronoun can be easily explained assuming that strong pronouns unlike clitics have a bottom restriction ($[\downarrow]\bot$), in effect a terminal node restriction that will preclude any development below the node the strong pronoun is parsed (Gregoromichelaki 2010: ftn 27). Assuming that every NP involves complex structure, unification (via MERGE) of the NP with the strong pronoun in a sentence like (6.105) would be impossible, due to the bottom restriction (note that complex structure of the NP, as already said, is denoted by the statement $[\downarrow]\top$):

(6.105) **Ton Γiorγo ksero afton*
the George know.1SG him.STRONG
'I know George.'

(6.106) After parsing the strong pronoun *afton* in ‘ton Γiorγo ksero afton’



At this point, unification cannot take place since the statements $[\downarrow]\top$ and $[\downarrow]\bot$ are contradictory.

Moving now to CD, we should firstly distinguish between CD and Clitic Right Dislocation (CRD). Even though SMG exhibits right dislocation of both subjects and objects, CD seems to be a different construction. First of all, there is a different intonational pattern involved in the two cases. CD does not need an intonational break, while such a break is essential for CRD constructions:[14]

(6.107) *Ton* *ksero* *ton* *Γiorγo*
him.CL-ACC know.1SG the George
‘I know George.’ [CD]

(6.108) *Ton* *ksero,#* *ton* *Γiorγo*
him.CL-ACC know.1SG the George
‘I know him, George.’ [CRD]

Perhaps, the most crucial feature that distinguishes CD from CRD is the fact that in CD constructions the doubled NP does not have to be the rightmost element of the clause. On the contrary, in CRD constructions the doubled NP is always the rightmost element of the clause:

[14]The intonational evidence is summarized in Anagnostopoulou (1999) and the interested reader is referred there for more details.

(6.109) *Ton* *kseri* *ton* *Γiorγo* *o* *Γianis*
him.CL-ACC know.3SG the George the John
'John knows George.' [CD]

(6.110) **Ton* *kseri,#* *ton* *Γiorγo* *o* *Γianis*
him.CL-ACC know.3SG the George the John
'John knows George' [CRD]

Another piece of evidence that supports the claim that the two structures are different is the fact that there are languages that do allow CRD but not CD. French is one such example (Jaeggli, 1986):

(6.111) *Je* *l'* *ai* *vu#* *l'* *assassin*
I him.CL-ACC have seen the murderer
'I saw him, the murderer.'

(6.112) **Je* *l'* *ai* *vu* *l'* *assassin*
I him.CL-ACC have seen the murdere
'I saw the murderer.'

DS has a natural mechanism to account for right dislocation effects. This mechanism involves the introduction of a rule, the rule of RECAPITULATION. This rule can be seen as the right periphery analogue of the rules used for HTLD constructions, namely TOPIC STRUCTURE INTRODUCTION and TOPIC STRUCTURE REQUIREMENT. The difference between the two rules is that the rule of RECAPITULATION builds a LINK transition from a type *t* complete node, i.e. a complete proposition, to a type *e* requiring node by further positing that the formula value of the LINKed tree must be shared with some subterm of the main tree. The rule is shown below:

(6.113) RECAPITULATION

$$\frac{\{...\{Tn(0),...,?Ty(t),Fo(\phi),...\}\{\uparrow_* Tn(0),Tn(n),Ty(e),Fo(\alpha),...\Diamond\}...\}}{\{...\{Tn(0),...,?Ty(t),Fo(\phi),...\}\{\uparrow_* Tn(0),Tn(n),Ty(e),Fo(\alpha),...\}...\}}$$

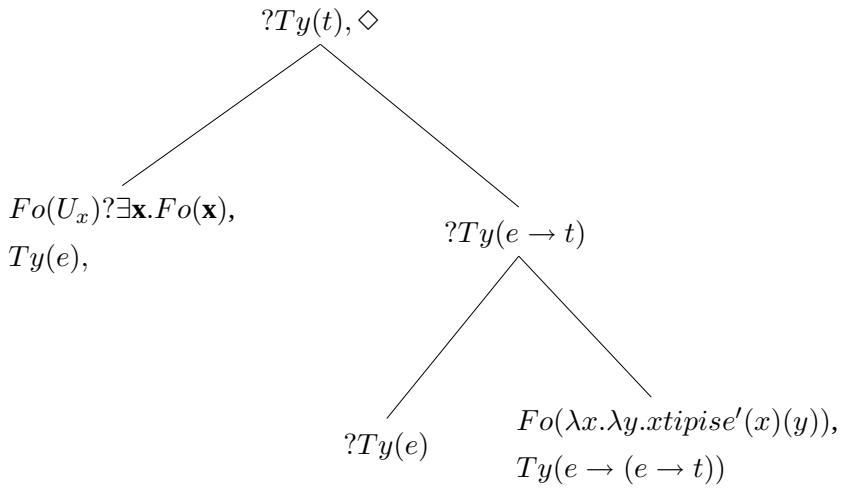
$$\{\langle L^{-1}\rangle Tn(0),?Ty(e),?Fo(\alpha),\Diamond\}$$

The above rule provides a straightforward explanation to right dislocated constructions like the one shown below:

(6.114) *Xtipise* *ton* *Γiorγo* # *o* *Γianis*
hit the George # the John
'He hit George, John.'

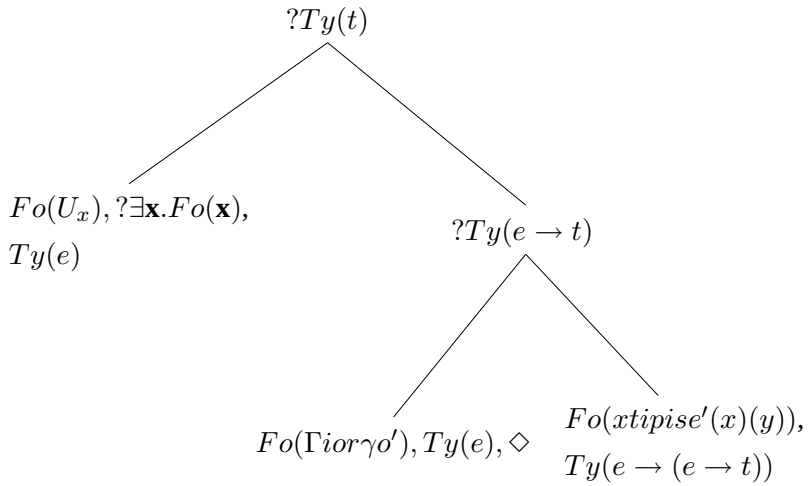
The above construction contains a right dislocated subject, i.e. *o Γianis* 'the John'. The first step in parsing such a sentence is parsing of the verb. The verb projects the whole propositional template as well as a type value and a formula metavariable in the subject node:

(6.115) After parsing the verb



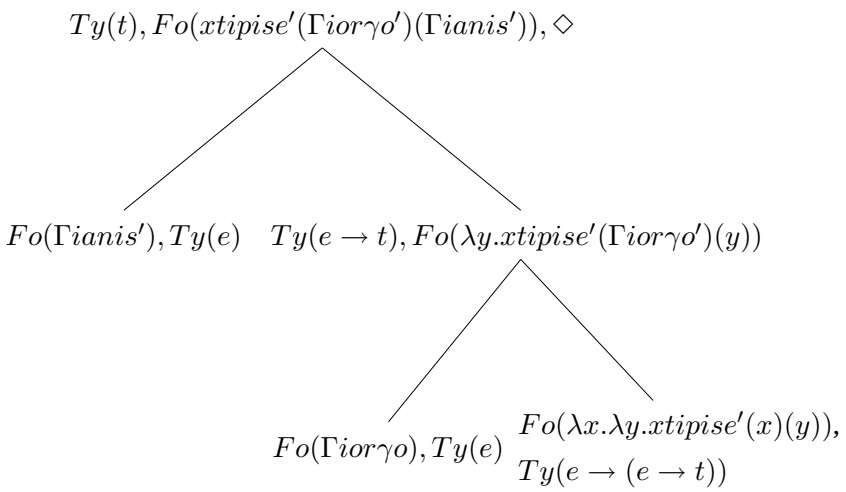
The pointer is left at the type *t* requiring node and the object comes into parse. The triggering point of the object will be a type *e* requiring node. Such a node is the object node but the pointer has to be moved down there. This is done by applying the rule of ANTICIPATION twice. The pointer reaches the direct object node and now the object can be parsed:

(6.116) After parsing the object



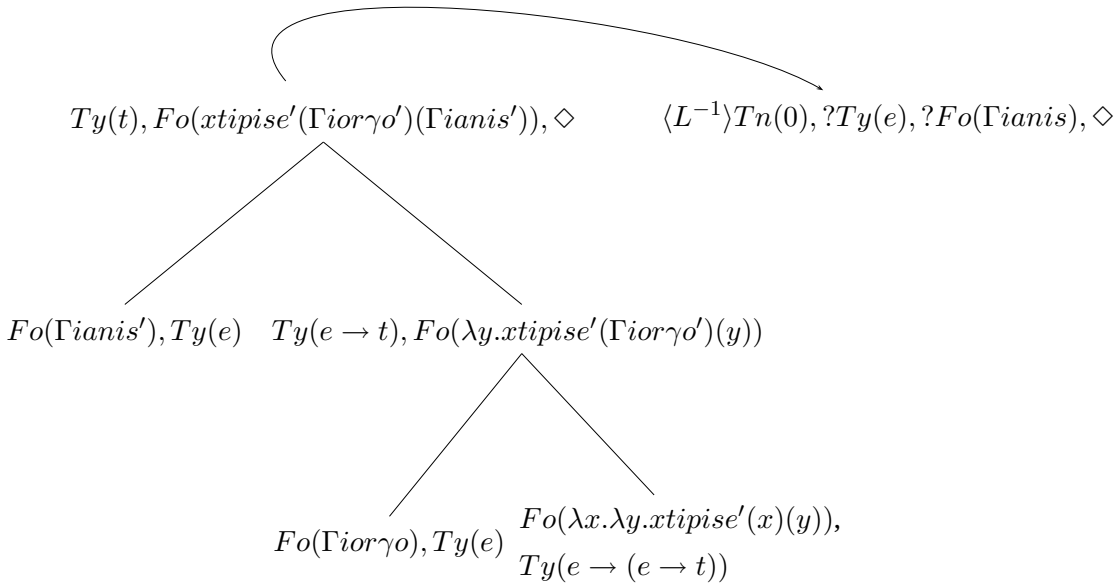
Now, assuming that the subject metavariable gets updated by the value Γ*ianis* 'John', we get a well-formed parse applying ELIMINATION:

(6.117) The result



Notice that parse of the sentence could stop here, since a well formed parse has been established. However, in case of a dislocated argument the rule of RECAPITULATION can apply, which will create a LINK transition from the type complete *t* node to a type *e* requiring node. It will further posit that a formula value of type *e* found somewhere in the main tree must be the formula value of the node of the LINKed tree. The structure after RECAPITULATION has applied is shown below:

(6.118) After RECAPITULATION has applied



At that point the right dislocated subject can be parsed in the LINKed tree. Now, The same reasoning can also be used to account for right dislocated structures in which the right dislocated element is coreferential with the clitic. In these cases, we first parse the sentence as a regular sentence involving a clitic but we do not parse the doubled NP. We obtain a well-formed parse and then the rule of RECAPITULATION is applied parsing the right dislocated object in the LINKed tree. The rule of RECAPITULATION will thus work for cases of CRD as well.

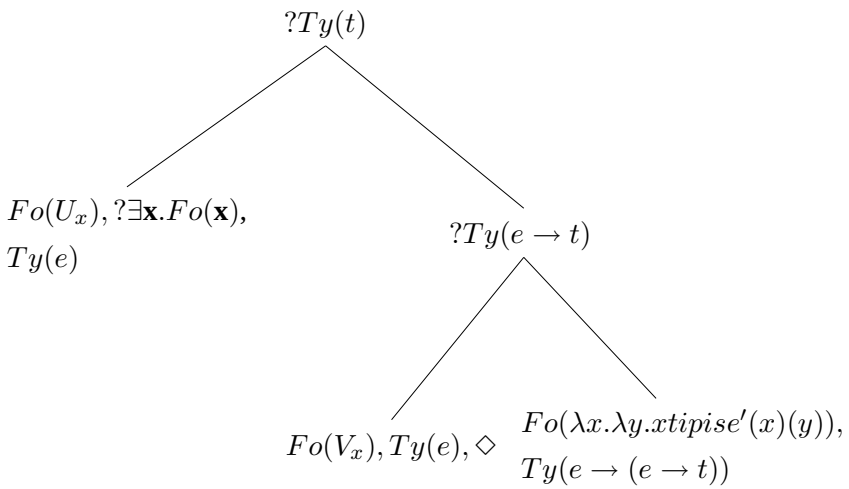
However, in SMG there are instances of genuine CD in which the doubled NP is not the last element in the clause, while no intonational break before the doubled NP exists. Such an example is repeated below:

(6.119) *Ton* *kseri* *ton* Γ*ior*γ*o* *o* Γ*ianis*
him.CL-ACC know.3SG the George the John
'John knows George.' [CD]

Gregoromichelaki (2010) treats the above sentences in the same sense as CLLD structures, namely as involving MERGE, i.e. unification, of the clitic node with an unfixed node hosting the doubled NP. However, there are a number of differences in parsing the two

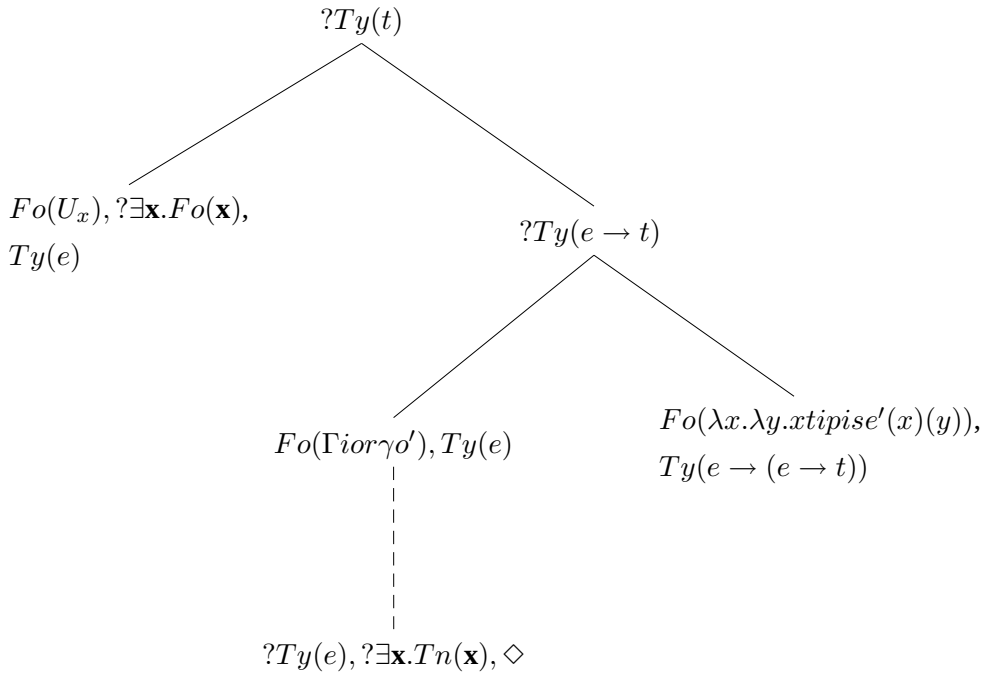
constructions (CLLD and CD) even though the core of the proposal stays the same. In a sentence like (6.119), there is no unfixed node to provide the referent for the metavariable provided by the clitic. In that respect, Gregoromichelaki (2010) argues that the referent for the pronoun must be salient enough in the discourse in order to be updated via the context and not by the natural language string itself. But let us see the construction process step by step. The clitic is parsed first, projecting the direct object node and decorating the same node with a type value and a formula metavariable. Next the verb comes into parse, projecting the rest of the propositional template along with a type value and a formula metavariable in the subject node:

(6.120) After parsing *ton kseri* in *ton kseri ton* Γ*ior*γ*o o* Γ*ianis*



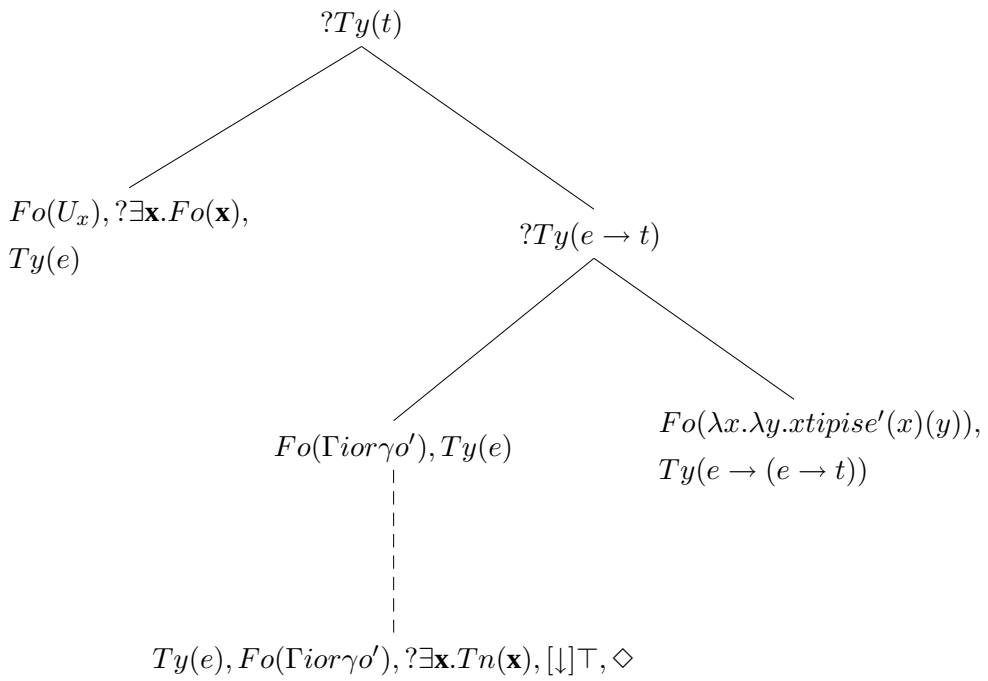
The pointer is back at the top node. ANTICIPATION will enable the pointer to move to the direct object node again. By SUBSTITUTION the formula metavariable gets a proper formula value from the context, say Γ*ior*γ*o*. Then, the rule of LATE *ADJUNCTION applies and projects an unfixed type *e* requiring node from the type complete *e* node (remember that LATE *ADJUNCTION needs a type complete node to apply). The result is shown below:

(6.121) Applying LATE *ADJUNCTION



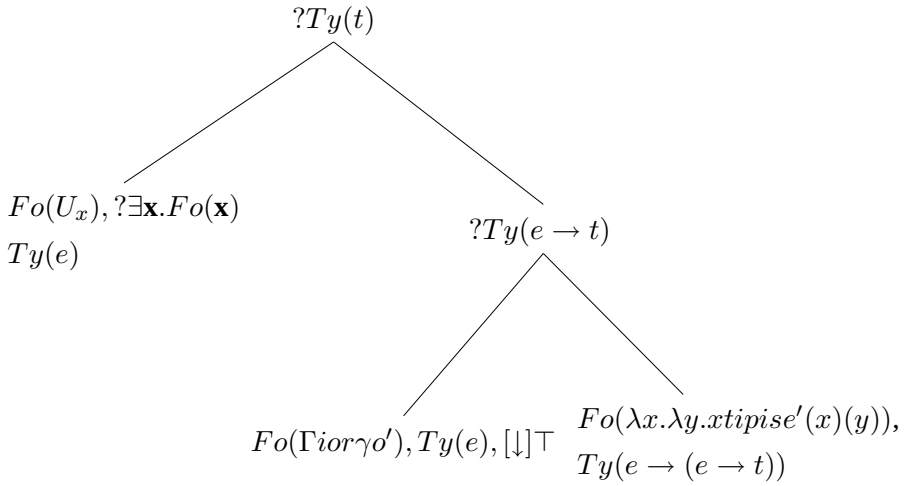
The full NP *ton* Γ*ior*γ*o* is then parsed on the unfixed node projected via LATE *ADJUNCTION:

(6.122) After parsing *ton* Γ*iorγo* in *ton kseri ton* Γ*iorγo o* Γ*ianis*



The next step is unifying the unfixed node with the direct object node. Notice that in order for MERGE (unification) to take place the two nodes must carry compatible decorations. In that sense, if the formula value of the fixed node does not match the formula value of the unfixed node, such unification is not possible. MERGE is possible given the state of affairs in the above tree. Applying MERGE will give us the following structure:

(6.123) After MERGE



Then the subject, *o Γianis*, can be parsed using LATE *ADJUNCTION or RECAPITULATION. Both options are possible and will lead to a well-formed parse. In conclusion, the analysis of CD provided by Gregoromichelaki (2010) and based on Cann et al. (2005) assumes CLLD and CD to involve the same mechanism of unification (MERGE). However, given the dynamic nature of the DS framework, there is a difference between the two constructions. In CLLD constructions, the doubled NP is the element introducing the referent in context. In that sense, given a left-to right parsing perspective, the doubled NP will have already introduced a referent when the clitic comes into parse in CLLD constructions and thus can act as a referent for the clitic. In CD constructions on the other hand, the doubled NP comes as a kind of confirmation for an already chosen referent for the clitic's metavariable. This fact according to Gregoromichelaki (2010), is what gives rise to the different discourse properties of the two constructions. In CD constructions, and given that the metavariable of the clitic must get a formula value from context, since the full NP has not yet been parsed and as such has not yet provided a referent, the referent for the clitic must be salient enough to guarantee substitution of the metavariable. This can actually provide us with an explanation of the well-known claim that a number of bare quantifiers seem not to be tolerated in CD constructions but are fine in CLLD constructions (Iatridou, 1990):

(6.124)

a. *Tria provlimata mono o Kostas ta elise*
three problems only the kostas them.CL solved

b. *Mono o Kostas (*ta) elise tria provlimata*
only the kostas them.CL solved three problems
'Only Kostas solved three problems.'

Iatridou claims that certain noun NP classes are not tolerated in CD but are fine in CLLD but does not specify which these NP classes are. In the same vein, Anagnostopoulou (1999), proposes that CD in general is subject to a referentiality constraint. Then, parametric variation exhibited in different languages as regards which semantic noun classes are allowed to be CDed in each case is subject to a referential constraint which can be different

from language to language. The core idea is that referentiality is a scalar notion and as such different elements on that scale can encode a referential feature in some languages and not in some others. The referentiality scale, adapted from Anagnostopoulou & Giannakidou (1995) is shown below:

(6.125) Anagnostopoulou & Giannakidou's (1995) referentiality scale

referential indefinites > partitives > weak definites > novel definites > proper names and definite descriptions > definites > demonstratives > anaphoric pronouns

The proposal in Anagnostopoulou (1999) is that the relevant property encoded in Sportiche's (1993) ClP projection is referentiality. Parametric variation in different languages is captured by assuming that each clitic language can choose to overtly mark referentiality on different types of NPs within the scale. It must be noted that points in the scale are in a *superset-subset* relation and as such, if a language chooses to mark a given element for referentiality, then it also marks the elements in the scale that are lower than this element in the scale.

However, the issue is more complicated and the distinctions in grammaticality judgments between the two constructions are not than fine-grained as presented in Iatridou (1990) or Anagnostopoulou (1999). For example, I and a number of my informants judge sentences containing clitic doubled bare quantifiers at least acceptable or in some cases fully grammatical:

(6.126) *(?)* *Mono* *o* *Kostas* *ta* *elise* *tria* *provlimata*
only the kostas them.CL solved three problems
'Only Kostas solved three problems.'

There are a number of other examples that should not be grammatical according to what Iatridou (1990), Anagnostopoulou (1999) among others claim. However, such examples are not rare. The crucial factor seems to lie in how salient the referent of the clitic is in the context. Here are some examples, either constructed or taken by the internet (see Appendix B for more examples):

(6.127) *Ine* *oli* *tus* *poli* *kala* *peδja.* *Tus* *ksero* *merikus*
are all them very good guys them.CL know some
'They are all very nice guys. I know some of them.

(6.128) *Ine poli aksii i musiki tis, tus ksero kapius*
are very worthy the musicians POSS them.CL some
'The band's musicians are very worthy, I know some of them.'
[http://www.musicheaven.gr/html/modules.php?name=News& file=article& sid=1358 as retrieved on the 28th of April, 2010, 00:00:55 GMT]

Even non-specific indefinites seem to be tolerated in clitic doubled position in some cases (more examples in Appendix B):

(6.129) *Telika tora pu to skeftome malon exis δikio. Tin xriazome mia kaθaristria*
lastly now that it think maybe have right her.CL-ACC need one cleaning-lady
'Coming to think of it, you are maybe right. I need a cleaning lady.'

(6.130) *To xtipao ena saduitsaki tora*
it.CL-ACC hit one sandwich now
'I could/would eat a sandwich now.'

Partitives are also fine in CD environments (again see Appendix B for more examples):

(6.131) *I taksi tou Vageli exi poli kala peδia. Sto leo γati ta ksero kapia apo afta*
the class of Vaggelis has very good children you-it say because them.CL-ACC know.1SG some.ACC of them
'The children in Vaggelis' class are very nice. I say that because I know some of them.'

The examples show that the referentiality scale proposed by Anagnostopoulou (1999) does not act as a strict constraint but rather as a preference. In that sense, it would be overgenerating to posit an analysis which allows some semantic classes to be CDed but disallows some others (even though it is quite straightforward to do so in DS). The reason that CD seems to be more restricted than CLLD in terms of which NPs can be doubled seems to involve incrementality. I think that Gregoromichelaki (2010) is absolutely right in saying that the different discourse properties of the two different structures derive from the different dynamics of each construction. I think that the reason CD seems to have a

more restricted distribution compared to CLLD must be attributed to that same reason and not to any inherent syntactic constraint on CD. In CD constructions, the referent has not been introduced yet by the time the clitic comes into parse, citing Gregoromichelaki: "In ClD, under this analysis, the doubled DP can only come later as a confirmation of a referent already selected. Unlike what happens in CLLD, the doubled DP in ClD cannot be solely responsible for introducing the referent in the context, as such a referent has already been utilized for the earlier resolution of the clitic" (Gregoromichelaki, 2010: 44, Cld stands for Clitic Doubling). Given this line of reasoning, one should expect that less referential NPs would be harder to access than more referential ones. However, there is no commitment as to whether some NP classes will be banned or not. This seems to me to be the correct way to go, since, as we have seen above the assumption that certain semantic classes of NPs cannot be CDed is disputed from the examples provided. Thus, it would seem that the differences between CLLD and CD, at least for SMG, should not be attributed to a some kind of a referentiality constraint relevant for CD but not CLLD, but rather to a preference induced by the mere dynamics of a time-linear, lef-to-right system.

The analysis proposed by Gregoromichelaki has a number of additional bonuses. The most striking one is that it explains Cechetto's (1999) observation that CD constructions are not subject to a principle C violation, an observation that has not been given considerable attention in the literature. CD constructions involve an R-expression which is bound by the clitic, something that should be out, at least according to standard binding theory:

(6.132) *Ton*$_i$ *ksero ton* Γ*ior*γ*o*$_i$
him.CL-ACC know the George
'I know George'

The answer to this puzzling behaviour receives a straightforward explanation given the DS analysis given. I do not want to go into the specifics of how binding theory is formulated in DS (see Cann et al., 2005) but the core idea for principle C is that the R-expression is not coreferential with a co-argument. However, these restrictions apply at the time the R-expression comes into parse and are not calculated at the end of the parse. In that sense, no principle C violation occurs given the CD analysis proposed by Gregoromichelaki, since the R-expression (doubled NP) is parsed on an unfixed node and as such it is not a co-argument with the clitic at the time of its introduction. The fact that it will be a co-argument

later on when the unfixed node will unify with the object node is totally irrelevant, since as already said, the binding restrictions apply in a time-linear, dynamic manner and not statically at the end of the parse.

The account proposed by Gregoromichelaki gets the doubling facts as regards SMG right, while it furthermore captures a number of important side-effects as regards doubling (e.g. the Cechetto observation). The next step is to see how a doubling account can be put forth for the rest of the dialects examined in this thesis.

6.3.2 Some Thoughts on Doubling in Cypriot and Pontic Greek

6.3.2.1 Cypriot Greek

CG exhibits both CLLD and CD as witness the examples below:

(6.133) *Ton* *Γiorγo* *ksero* *ton*
the George.ACC know.1SG him.CL-ACC
'I know George.'

(6.134) *Ksero* *ton* *ton* *Γiorγo*
know.1SG him.CL-ACC the George.ACC
'I know George.'

CG exhibits genuine CD as the examples below, where the subject (focussed or unfocussed) follows the doubled NP, show:

(6.135) *Efae* *tin* *tin* *turta* *o* *Γianis/* *O* *ΓIANIS*
ate.3SG it.CL-ACC the cake.ACC the John.NOM the John.NOM
'John ate the cake.'

(6.136) *Ton* *Γiorγo* *kseri* *ton* *o* *Γianis/* *O* *ΓIANIS*
the George.ACC know.3SG him.CL-ACC the John.NOM the John.NOM
'John knows George'

CG allows indefinites or bare quantifiers in CLLD or even in CD constructions:

(6.137) *Merikus* *ksero* *tus/* *Ksero* *tus* *merikus*
some know them.CL-ACC know them.CL-ACC some
'I know some of them.'

(6.138) *Tora pu to skeftome, san na eshis dikeo telika. Θelo*
now that it think, like SUBJ have right in-the-end want.1SG
tin mian kaθaristria/ mian kaθaristria θelo tin
her.CL-ACC a cleaning-lady a cleaning-lady want.1SG her.CL-ACC
'Coming to think of it, you are right. I need a cleaning lady.'

What is rather problematic in CG, given the analysis for CLLD and CD for SMG plus the positioning analysis of CG given in chapter 4, is that CLLD fails to trigger proclisis in CG. CLLD is associated with enclisis in CG, something that is unexpected assuming that CLLD involves parsing of the doubled NP on an unfixed node. Given that the presence of an unfixed node is a trigger for proclisis, one would expect proclisis to obtain in CLLD constructions contrary to fact. Given this state of affairs, either the unfixed node trigger needs to be re-examined or the CLLD analysis as involving parsing of the doubled NP on an unfixed node is not relevant for CG. The first solution involves getting rid of the unfixed node trigger as a proclitic trigger in general. However, this move is highly implausible and highly uneconomical, since by getting rid of this restriction, we get rid of a proclitic generalization that captures proclisis with a wide range of elements (Wh elements, subjects/adverbs/PPs/temporal expressions). Then the only way to capture proclisis with all these elements is by separately listing them, a move that does not seem to be plausible at all. Therefore, I believe that such a thought needs to be abandoned. But if such a solution is not to be pursued, how is CG CLLD analyzed? A solution that would work is to assume that doubled NPs are not parsed on unfixed nodes but rather as involving a LINK structure. This solution seems more plausible than the first one, but however needs evidence to support it. It seems that such evidence indeed exists for CG. This evidence comes from focussed object OV structures. CG does not allow focussed OV structures but rather uses clefts to denote object focus (examples from Tsiplakou et al., 2007):

(6.139) **TES KUNNES efaen o Tassos*
the nuts.ACC ate.3SG the Tassos.NOM
'Tassos ate THE NUTS.'

(6.140) *en TES KUNNES pu efaen o Tassos*
is the nuts.ACC that ate.3SG the Tasos.NOM
'It is the nuts that Tasos ate.'

Given that focussed objects in OV structures are assumed to be parsed on an unfixed node, it seems that CG does not allow object NPs to be parsed on an unfixed node. However, deictic object NPs are possible in OV position as shown below:

(6.141) *TUTO* *TO* *VIVLIO* *tu* *eδoka*
this the book him.CL-GEN gave.1SG
'This is the book I gave him.'

In that respect, we should ban object NPs from being parsed on an unfixed node except in the case of deictic NPs. For the moment, the lexical entry I am going to propose disallows accusative NPs to be parsed on an unfixed node but does not however exclude deictic NPs. The lexical entry is shown below:[15]

(6.142) Entry for accusative NPs

IF	$?Ty(e), \langle\uparrow^*\rangle\langle\downarrow\rangle\top$
THEN	...
ELSE	abort

The trigger says that if the pointer is at a type *e* requiring node and if following the modality $\langle\uparrow^*\rangle\langle\downarrow\rangle$ there is structure, then the NP can be parsed. Such a situation will not be true in case an unfixed is projected via *ADJUNCTION. By definition the rule of *ADJUNCTION must apply only when there is no other induced structure in the tree. The restriction I have posited describes an underspecified situation where at least some structure exists besides the initial type *t* requiring node. This restriction will predict accusative NPs to be impossible on unfixed nodes projected by *ADJUNCTION but will however allow such NPs to be parsed on an unfixed node projected by LATE *ADJUNCTION. This is because by definition the rule of LATE *ADJUNCTION will apply to a node with a fixed treenode address other than the initial node. In that sense, at least one node below the initial node will exist, thus the restriction $\langle\uparrow^*\rangle\langle\downarrow\rangle\top$ will be satisfied. Such an entry will preclude OV structures in CG. But, if accusative NPs cannot be parsed on unfixed nodes projected by *ADJUNCTION, then CLLD does not involve unfixed nodes in CG also. The

[15]The lexical entry, as already said, does not exclude deictics NPs. This issue is not going to be pursued here.

situation points towards an analysis of CG CLLD as involving a LINK structure. Let us see if such a solution is feasible. We have already seen that LINK structures have already been used in DS for doubling constructions (HTLD and CLLD, Cann et al. 2005; Gregoromichelaki, 2010). In both constructions the doubled NP is parsed on a type *e* node which is LINKed to a type *t* requiring node. A requirement for a copy of the NP's formula to be found somewhere in the main tree is posited. In HTLD, this requirement is encoded using the *D* operator while for CLLD using the Kleene star operator (*). This is done, as we have already said, in order to capture the different island properties associated with the two constructions. The problem that we encounter with CLLD and HTLD constructions in general, given the analysis of scope in DS, is that the scopal statement *?Sc(x)* is defined in such a way that it does not allow NPs parsed in a separate tree to participate in the scopal statement of the LINKed tree. Since quantifiers are possible in CLLD and since we want these quantifiers to participate in scopal statements, we should first update the definition of *?Sc(x)* to allow NPs that have been parsed in a separate tree to participate in the scopal statement of the LINKed tree. The *?Sc(x)* definition given in Kempson et al. (2001) and Cann et al. (2005) repeated below allows NPs to participate in the scopal statement of a type *t* node only if these NPs are within the local domain of this type *t* node:

(6.143) $?SC(x) =_{def} ?\langle\uparrow_0\rangle\langle\uparrow_1^*\rangle\exists y.Scope(y < x) \vee Scope(x < y)$

What we want is to include cases where an NP is parsed in a tree LINKed to the tree the scopal statement is calculated. In order to do that, we need to add the additional operator $\langle L\rangle$ in the address where the scopal statement is to be found. We introduce this operator as a disjunction to the existing modality. The new statement is shown below:

(6.144) $?SC(x) =_{def} ?(\langle\uparrow_0\rangle\langle\uparrow_1^*\rangle \vee \langle L\rangle)\exists y.Scope(y < x) \vee Scope(x < y)$

Given this updated version of *?Sc(x)*, NPs parsed in a tree LINKed with the main tree can participate in the scopal statement. In that sense, it is possible to define CLLD in CG as involving a LINK structure transition rather than an unfixed node. The tree below shows the result of parsing the doubled NP in (6.145) as a LINK structure. The *?Sc(x)* can now be satisfied since the node where the scope statement is found can also be the LINKed node, i.e. the type *t* requiring node. Assuming that all type *t* nodes have a scopal statement, the doubled NP can effectively enter into the scopal evaluation:

(6.145) *Enan* *maθiti* *kseri* *ton* *o* *Γianis*
a student.ACC knows.3SG him.CL-ACC the John.NOM
'John knows a student.'

(6.146) Parsing *enan maθiti* in *enan maθiti kseri ton o Γianis*

$\langle L \rangle Tn(0),$
$Fo(\epsilon, x, ma\theta itis'(x)), [\downarrow]\top, ?Sc(x),$
$Ty(e)$

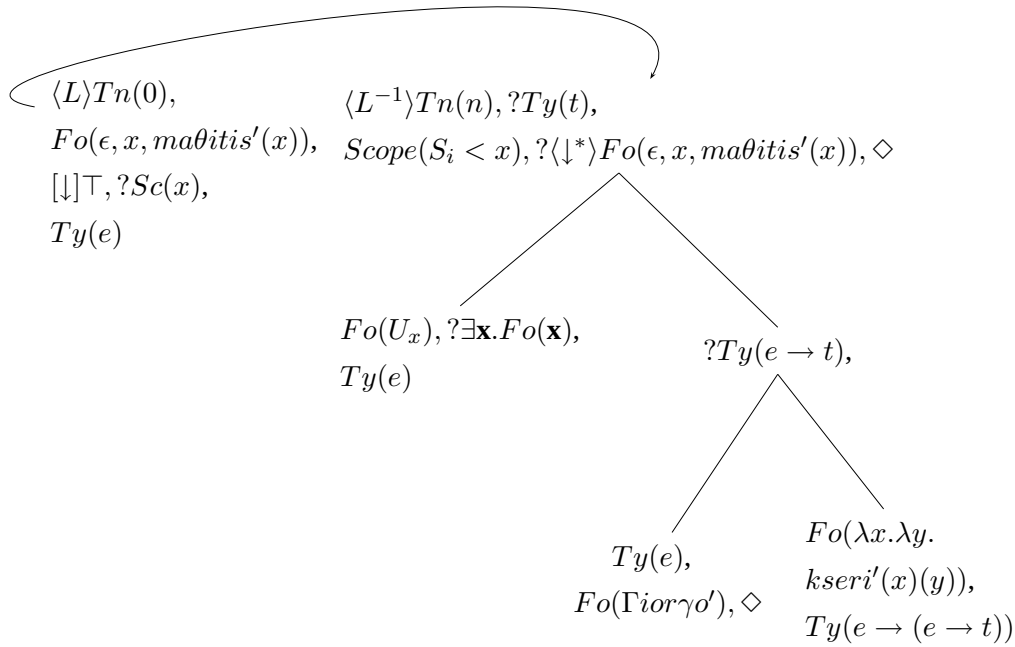
$Tn(0), ?Ty(t), Scope(S_i < x),$
$?\langle \downarrow^* \rangle Fo(\epsilon, x, ma\theta itis'(x)), \diamond$

In the above structure the variable x of the indefinite NP participates in the scopal statement in the main tree. The details of how scope is calculated when multiple quantifiers are present are not going to be discussed here. However, the interested reader is directed to Kempson et al. (2001), Cann et al. (2005) for an introduction to scopal calculation in DS, and most importantly Gregoromichelaki (2010) for an analysis of scope in SMG CD constructions. The important thing at the moment is that NPs parsed in a separate tree domain as in above can participate in scopal statements. Such a move is necessary, since quantified NPs in CLLD constructions interact scopally with other quantified phrases in CG:

(6.147) *Ena* *vivlio* *tu* *Chomsky* *eθkjavasan* *to* *poli*
a.ACC book.ACC of Chomsky read it.CL-ACC many.NOM
maθites
students.NOM
'Many students read a book by Chomsky.' [$\exists < Many \vee Many < \exists$]

Returning to the actual example in (6.145), the verb comes into parse projecting the rest of the propositional structure. Then the clitic is parsed, projecting its information in the direct object node. The formula metavariable of the clitic is then substituted by the formula value of the doubled NP, satisfying the requirement for a shared term to be found between the two trees:

(6.148) Parsing *ton* in ‘enan maθiti kseri ton o Γiorγos’



Given this account, proclisis is predicted to be impossible in CLLD structures according to fact. This is because the CLLDed NP will be on a LINK structure rather than an unfixed node. In that respect, no proclitic trigger will exist, and as such the clitic will be unable to get parsed before the verb. However, there is one last pending problem in giving a LINK analysis of CLLD and this involves case connectivity. The analysis proposed cannot capture the fact that in CLLD structures the doubled NP must be matched in case with the clitic. Such a problem is easily solved however if one assumes that the TOPIC STRUCTURE REQUIREMENT rule for CLLD structures will also involve a case filter that will ensure case connectivity. The revised TOPIC STRUCTURE REQUIREMENT rule for CLLD structures is shown below. Note furthermore that the rule is defined in case structure exists before the ?Ty(t) requiring node. This last thing will predict CLLD to be possible in embedded contexts as well, contrary to what holds for HTLD:

(6.149) TOPIC STRUCTURE REQUIREMENT for CLLD

$$\frac{\{...\{Tn(0), ?Ty(t)\}, \{\langle L\rangle Tn(0), Fo(a), Ty(e), \diamond\}}{\{\{Tn(0), ?Ty(t), ?\langle\downarrow^*\rangle(Fo(a) \wedge ?\langle\uparrow_0\rangle Ty(e \rightarrow t)), \diamond\}\}, \{\langle L\rangle Tn(0), Fo(a), Ty(e)\}}$$

Where D $\in \{\downarrow_0, \downarrow_1, \downarrow^*, L\}$

In the above rule, besides the introduction of a requirement for a copy of the doubled NP to be found somewhere in the LINKed tree, a further case filter must be true in the node where the *Fo* value of the doubled NP will be found. This case filter ($?\langle\uparrow_0\rangle Ty(e \rightarrow t)$) identifies the node where the copy of the doubled NP will be found with the direct object position.[16] Given this modification, the problem of case connectivity is solved and a LINK account of CLLD is thus feasible. CD in CG involves the same process as in SMG, namely application of LATE *ADJUNCTION for the parse of the doubled NP.

Recapitulating, I assume that CG CLLD involves the use of a LINK structure rather than an unfixed node. The definition of *?Sc(x)* was modified in order to allow doubled NPs parsed as a LINK structure to appear in the scopal statement of the main tree. This move will predict that CLLD should not induce proclisis according to fact. On the other hand, CD in CG is assumed to make use of the exact same mechanisms used in SMG CD. It is an open question as to whether SMG CLLD also involves a LINK structure but this is something that I will not pursue in this thesis. In principle, both of the analyses sketched will give us equivalent results.

6.3.2.2 Pontic Greek

The situation as regards doubling in PG is quite different from what we found in SMG or CG. In a nutshell, PG allows CLLD structures but disallows CD. Doubling constructions where the full NP follows the clitic are considered to be instances of CRD rather than CD. Tsakali (2007) claims that PG exhibits CRD and not CD on the basis of examples like the ones shown below:

(6.150) *Ton Γiorγo ekser aton*
the George know.3SG him.CL
'I know George.'

(6.151) *Efae aten (o Γianis), ti supa*
ate.3SG her.CL the John the soup
'John ate the soup.'

[16] Similar considerations will apply for indirect object doubling, assuming that necessary changes in the form of the case filter.

Tsakali (2007) does not give the examples showing that examples like (6.151) are indeed cases of CRD and not CD. The data I have confirm Tsakali's (2007) claim. A doubling construction where the doubled NP follows the clitic and a focused subject or in general a subject of any sort comes at the end are not grammatical in PG:

(6.152) **Efaen aten ti supa o Iorδanis/O IOR∆ANIS*
ate.3SG her.CL the soup.ACC the Jordan.NOM
'Jordan ate the soup.'

Quantifiers can be doubled in CLLD constructions in PG as witness the examples below:

(6.153) *Ena kokino bluza xriaskom ato (ena kokino bluza)*
a red blouze need it a red blouze
'I need a red blouze.'

(6.154) *Ena δjo peδja ekser ata (ena δjo peδja)*
one two children know them.CL one two children
'I know some of the guys.'

(6.155) *Ults ekser ats (ults)*
all know them.CL all
'S/He knows everyone.'

In approaching doubling in PG, we should keep in mind that genuine CD is not an option. CD, as we have already seen, is accounted for by using LATE *ADJUNCTION projected from the direct object node after the clitic is already parsed. The doubled NP is then parsed on that unfixed node. MERGE applies and unifies the two nodes. Cann et al. (2005) argue that the reason pronouns in English or Germanic (SMG strong pronouns as well) cannot be doubled is because these pronouns retain a bottom restriction in their lexical entry that disallows any structure to be projected below the node they project their decorations. Assuming that an NP will involve complex structure, MERGE of a pronoun with a full NP will be impossible given the bottom restriction of the strong pronoun. Then, Cann et al. (2005) argue that some pronominals (like clitics in languages like SMG) have lost their bottom restriction and as such can MERGE with a full NP that has complex

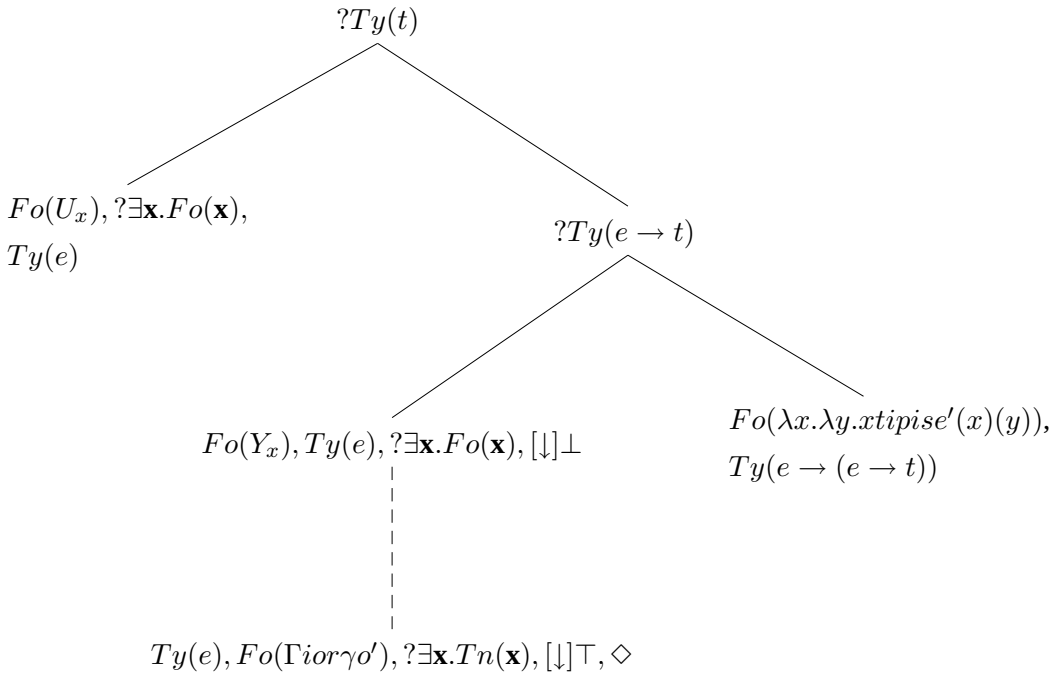
structure. However, clitics in some languages have retained this bottom restriction and as such MERGE of a clitic with a full NP is impossible. The same assumption is made by Gregoromichelaki (2010), as we have already seen, for SMG strong pronouns. Following this line of reasoning, I will argue that PG clitics retain a bottom restriction that disallows them from unifying (via MERGE) with a full NP. The updated lexical entry for PG clitics is shown below:

(6.156) Updated entry for PG clitics (bottom restriction added)

IF $?Ty(t)$
THEN IF $\langle\downarrow_1^+\rangle Ty(x)$
THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1^+\rangle); \texttt{go}(\langle\downarrow_1^+\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(\langle\uparrow_0\rangle\langle\uparrow_1^+\rangle Tn(a);$
$Ty(e), Fo(U_x), ?\exists\mathbf{x}.Fo(\mathbf{x});$
$?\exists\mathbf{x}.Tn(\mathbf{x}), [\downarrow]\bot); gofirst(?Ty(t))$
ELSE abort
ELSE abort

Under such an entry unification of the clitic node with an unfixed node where a full NP has been parsed on, is impossible. CD is then predicted to be impossible for PG, since the full NP on the unfixed node projected via LATE *ADJUNCTION cannot MERGE with the object node below:

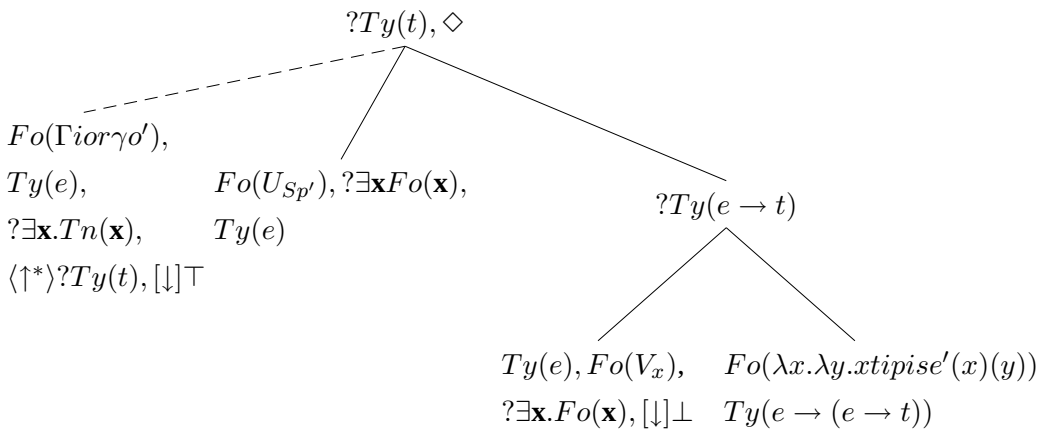
(6.157) Parsing ton Γiorγo in *ekser aton ton* Γ*iorγo* 'S/He/It knows George'[17]



MERGE between the object node and the adjuncted node is impossible given the conflicting statements $[\downarrow]\top$ and $[\downarrow]\bot$ found in the adjuncted and the object node respectively. Another consequence of this assumption is that CLLD cannot be captured using the unfixed node mechanism for the same reason, i.e. MERGE between the unfixed node and the object node will be impossible:

[17]The locally unfixed node projected from the clitic is assumed to have found its position in the tree structure presented.

(6.158) Parsing *ton* Γ*iorγo ekser aton* with an unfixed node

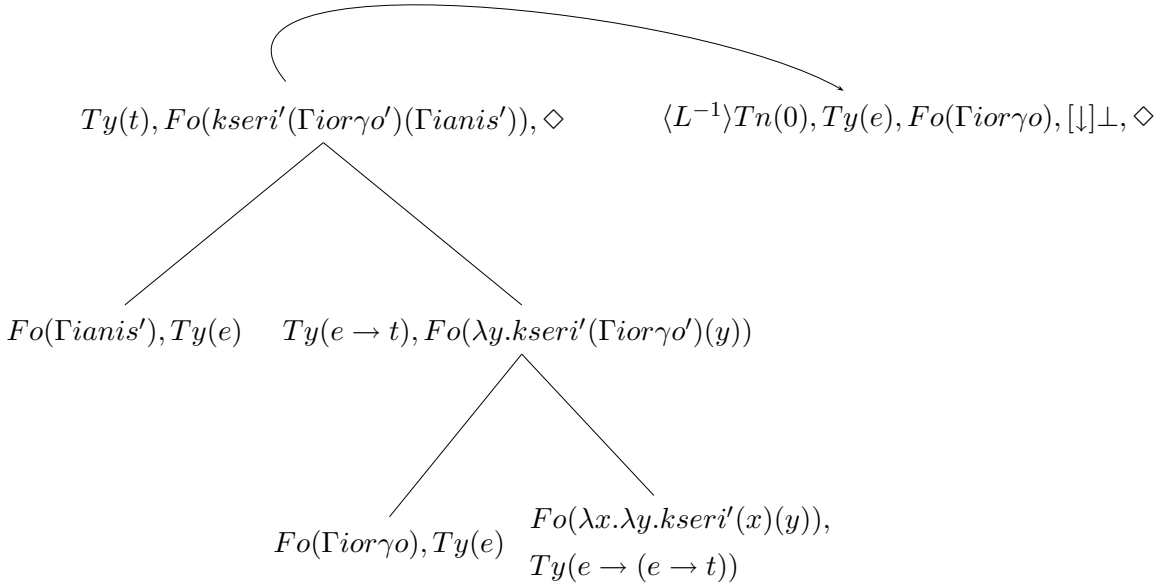


Given that MERGE of an unfixed node with the node where the clitic is parsed is impossible, then the prediction is that PG doubling involves LINK structures. CLLD will then be captured using the same analysis I have used in the case of CG. On the other hand, CRD will be captured using RECAPITULATION. In both cases LINK structures will be used. The analysis will correctly predict that genuine CD cases will be out in PG. Examples like the one shown in (6.159) will be impossible to parse in PG:

(6.159) **Kser aton to* Γ*iorγo, o* Γ*ianis/O* Γ*IANIS*
know.1SG him.CL the George the John
'John knows George.'

Let us see why. The clitic will project a bottom restriction so the only way for the object to be parsed will be via RECAPITULATION. However, parsing of the object with RECAPITULATION we are left with the subject at the end. The parse has finished (we got a well-formed type *t* formula value) and we are left with the subject.

(6.160) Parsing *to* Γ*ior*γ*o* in *kser aton to* Γ*ior*γ*o o* Γ*ianis/O* Γ*IANIS* with RECAPITULATION



The parse is complete and the pointer returns back to the type *t* node. There is nothing additional to do. RECAPITULATION cannot apply again and no outstanding requirements exist in the tree. The problem is that we are left with the dislocated subject which we have not parsed and there is no way to parse it. In that sense, sentences like (6.159) are predicted not to be possible given our account. The account predicts that in cases of doubling where the doubled NP follows the clitic, the doubled NP should be the last argument to get parsed, i.e. it predicts that PG does not have genuine CD but CRD according to fact. However, CRD like CLLD structures involve case connectivity between the clitic and the doubled NP. Thus, we will have to also account for the fact that the CRDed doubled NP in CRD structures is case-connected with the clitic. Remember that one of the uses of case in DS is to identify a node with a given structural position. For example, accusative NPs in SMG will involve a case filter that will identify them as accusative, i.e. $?\langle\uparrow_0\rangle Ty(e \rightarrow t)$. Assuming parsing of the same NP in a LINKed type *e* node after RECAPITULATION has applied, such a filter will never be true, since the LINKed tree will be the sole node in the tree and as such will not have a mother. However, given the case connectivity of the doubled NP in the LINKed tree, we can assume that the case filter of the doubled NP refers back to the clitic and requires

that the clitic must obey the case filter $?\langle\uparrow_0\rangle Ty(e \rightarrow t)$. It is extremely straightforward to assume a case filter that will capture both the regular cases and the case where the accusative NP will be parsed on a LINKed tree projected by RECAPITULATION. We just have to add the additional operator $\langle L^{-1*} \downarrow^*\rangle$ and our problem is solved. The new case filter for an accusative NP will be of the form $\langle L^{-1*} \downarrow^*\rangle ?\langle\uparrow_0\rangle Ty(e \rightarrow t)$. The relations $\langle L^{-1*} \downarrow^*\rangle$ can be potentially empty, in which case the case filter refers to the tree where the NP is parsed in (the regular cases). In case the relation $\langle L^{-1*} \downarrow^*\rangle$ is not empty, the case filter refers to the tree found after traversing the inverse LINK relation, capturing the fact that the doubled NP in a CRDed example should be case connected with the clitic.

Summarizing, I provided an account of PG doubling arguing that clitics in PG have retained their bottom restriction which disallows MERGE of an unfixed node with the object node. Doubling is then taken to involve LINK structures in PG. CLLD is captured using the same LINK mechanism we have used for CG CLLD. On the other hand, the fact that there is no genuine CD in PG but only CRD is straightforwardly explained, since MERGE of the full NP that has been parsed on an unfixed node (LATE *ADJUNCTION) is impossible.

6.4 Conclusions

In this chapter, the PCC and doubling in the dialects under consideration were discussed. It was argued that the analysis proposed by Cann & Kempson (2007) and Chatzikyriakidis & Kempson (2009) was adequate to capture the PCC in SMG. Under that account, the PCC is seen as a restriction on underspecification, a hard-wired logic constraint on tree-growth, where no more than one unfixed node of the same type can exist in a given tree.[18] Assuming that dative clitics as well as 1st/2nd person clitics project locally unfixed nodes, while 3rd person accusative clitics on the other hand do not but rather project fixed structure, the PCC facts as regards SMG are neatly derived. Such an analysis has the bonus of unifying the PCC with the unavailability of 3rd person clitic clusters in PG, by attributing them to the

[18]But see Chatzikyriakidis 2009a for an account where the PCC is taken to arise from the interactions of triggering points of different clitics and as such is assumed to be highly language dependant and not a general constraint found across languages.

exact same phenomenon, i.e. the "no more than one unfixed node at a time" constraint. The fact that PG as well as RP, seem, at least for some speakers, to allow combinations of a 1st plus a 2nd person clitic (and vice versa) are straightforwardly explained using the same reasoning used to explain the availability of clusters of a 1st/2nd person plus a 3rd person accusative clitic in PG. In these constructions, the reduced clitic forms (*m/s*) are not underspecified, since they are always interpreted as indirect objects, and are otherwise impossible in one single constructions. In that sense, these clitic forms can only appear in clitic clusters and when they do they are always interpreted as indirect objects. Hence, clitic clusters involving the forms *m/s* are licit since *m/s* project fixed structure rather than a locally unfixed node, thus the "no more than one unfixed node at a time" constraint is not operative in these cases.

Building on assumptions by Cann at al. (2005) and Gregoromichelaki (2010), an account of doubling of SMG, CG and PG was proposed. It was shown that Gregoromichelaki's (2010) account as regards CLLD in SMG cannot be maintained for CG for independent reasons, and a LINK analysis was proposed for CG CLLD by modifying the definition of *?Sc(x)* to allow NPs on a LINK structure to participate in the scopal statement of the main tree and by further imposing a case filter ensuring case connectivity in the requirement of the copy. CD in both SMG and CG was captured assuming that the doubled NP is parsed on an unfixed node projected via LATE *ADJUNCTION that merges (unifies) with the clitic node later on. Such a unification is possible, since clitics in SMG and CG are assumed to have lost their bottom restriction, and as such can unify with a node that has additional complex structure (full NPs). The unavailability of CD in PG on the other hand, is attributed to the fact that PG clitics retain a bottom restriction and as such cannot unify with the unfixed node where the doubled NP is parsed. In that sense, the only possible structure for PG is CRD and not CD. Furthermore, the account proposed by Gregoromichelaki (2010) as regards CD in SMG and adopted in this thesis for both SMG and CG, is further backed by evidence showing that the alleged impossibility of bare quantifiers or non-specific indefinites to appear in CD constructions, is more of a pragmatic preference rather than a strict syntactic constraint. The explanation provided assumes that the reason CD is more restricted than CLLD as regards the range of NP types participating in these constructions, derives from the mere definition of incrementality. In CLLD constructions a

referent for the clitic's metavariable is already salient in the discourse when the clitic comes into parse (given that it is provided by the doubled NP), whereas such an explicit referent is not present in the tree structure when the clitic comes into parse in CD constructions (given that the doubled NP has not been parse yet). Thus, there is no need to use a referentiality scale in the sense of Anagnostopoulou & Giannakidou (1995) and Anagnostopoulou (1999) or any other hard-wired syntactic constraint in order to derive something that seems to be no more than a preference.

Chapter 7

The Diachronic Development: A First Sketch

7.1 Introductory Remarks

The four dialects presented in this thesis have their medieval dialectal counterparts and do not derive from a common medieval language. In what follows, I will provide a first sketch of how the medieval clitic systems of the dialects presented in this thesis influenced the way the respective modern systems have developed. In doing that, I will first try to show the way in which a pragmatically governed clitic system such as the one I will argue is found in Koine Greek (KG),[1] ended up losing its pragmatic basis (Bouzouita, 2008a,b) and transformed the pragmatic reasoning underlying the system into syntactic/parsing constraints via means of routinization (Pickering and Garrod, 2004; Bouzouita, 2008a,b). Unfortunately, the data we have from KG cannot make fine-grained distinctions between different dialects of KG. In that sense, I will have to assume that all medieval dialects examined derive pretty much from a uniform KG clitic system.

[1]The term Koine Greek refers to post-classical Greek, roughly from 300 BC to 300 AD. Furthermore, I will use the term clitics for KG weak object pronouns. Someone might argue that KG did not have clitic pronouns but just weak pronouns. The decision on whether these elements are clitics or weak pronouns are irrelevant to the account that is going to be proposed. In that sense, I will use the term clitics for reasons of consistency.

7.2 The Koine Greek Clitic System: Competing Grammars?

The Koine Greek clitic system presents an interesting case of an evolving system in which constraints of an earlier system seem to be replaced by constraints of a newer system. Janse (1993) argues that the KG clitic system of the New Testament can be adequately described assuming that: a) clitics in KG regularly follow their governing hosts, i.e. the verb in this case and b) a version of Wackernagel's law that is however not specified is responsible for proclisis.[2] The law of Wackernagel applies exceptionally according to Janse (1993) and depends on the nature of the element in first position. The elements that trigger this application of the Wackernagel law are interrogative elements, the negation marker *mē*, demonstrative/ strong pronouns and subordinating conjunctions (examples from Janse, 1993):[3]

(7.1)

a. *Tis* *mu* *ēpsato*
who.NOM me.CL-GEN touched
'Who touched me?'

b. *Egō* *se* *edoksasa* *epi* *tēs* *gēs*
I you.CL-ACC glorify in the earth
'I glorified you in the earth.'

c. *Oudeis* *se* *katekrinen*
none you.CL-ACC blamed
'No one blamed you.'

d. *Mē* *mi* *kopus* *parehe*
NEG me.CL-DAT trouble give
'Do not give me trouble.'

[2]The law of Wackernagel refers to a restriction according to which clitic elements appear in second position. There are two different versions of second position, strict second position, i.e. following the first element of the clause and XP second position, i.e. following the first XP (see Schütze, 1994 and Progovac, 1996 for Serbo-Croatian 2P position clitics, Garrett, 1996 for Hittite 2P clitics and Taylor, 1996 for AG 2P clitics among others. See also Janse, 1994, 2000).

[3]The *ē* and *ō* symbols represent the long AG vowels η and ω respectively.

Janse (1993: 99) then argues that the relevant formulation of Wackernagel law for KG is the one where clitics follow either the first word of the phrase or the first phrase, the first word being a subordinating conjunction or a word in focus. Given Janse's description, one would expect postverbal clitics not to be possible with wh-elements or subordinating conjunctions. However, such a fact is not true as shown by Pappas (2006). The data from the *Oxyrynchus Papyri* (Oxy) he presents shown in (7.2), depict a situation where enclisis with wh-elements and complementizers is not only found, but, furthermore, in the case of complementizers, postverbal positioning is more common than proclitic positioning (37/54 tokens). Postverbal positioning involving a wh-element is also common in the *Oxyrynchus Papyri* (16/34 tokens):

(7.2) Clitic positioning in the Oxyrynchus Papyri (vols. 1-56)

Environment	Preverbal	Postverbal
Clause initial	4	231
Infinitival complement	12	0
Adverbs	34	14
NP-object	21	15
NP-Subject	10	13
PP	13	37
Complementizers	17	37
Wh-expressions	18	16

Pappas' data pose a major challenge for all the existing accounts of the KG clitic system, notably Janse (1993) and Taylor (1995, 2002). Taylor (2002) proposed an account of KG clitics which crucially relies on the following three assumptions: a) KG are enclitics, b) they attach to the left edge of a VP and c) prosodic Inversion (PI) plus Φ restructuring are operative. As Pappas (2006: 322) notes Taylor's (2002) account would predict complementizers, negative markers and wh-expressions to be always associated with preverbal positioning, since these will belong to the phonological phrase of the verb. Hence, the clitic does not undergo PI because a host will be available to its right. In Taylor (1995: 124), it is argued that clitics not only need a host to their right, but furthermore, this host must be

a stressed element. Then, it is assumed that complementizers in KG are in a process of becoming unstressed elements. This process, as claimed by Taylor (1995), is not yet complete and as such complementizers in KG are sometimes stressed and sometimes unstressed, thus the variation in positioning. This argument is not used in Taylor (2002), so Pappas' (2006) claim is right given that it relies on Taylor (2002). But even if, for the sake of the argument, Taylor (1995) is right and indeed some complementizers in KG are stressed and some are not, the account is not unproblematic. Assuming that negation markers and wh-elements are inherently stressed elements, Taylor's account would predict preverbal positioning in these cases. However, as shown by Pappas (2006) variation in positioning is found in both environments. In that sense, enclisis is also possible in these environments:

(7.3) *takhu* *erkhei* *hina* *idōmen* *se*
quickly come so-that see.SUBJ you.CL-ACC
'Come quickly, so that we may see you.' [Pappas, 2006: 323]

(7.4) *kan* *egō* *mē* *graphō* *soi*
even-if I NEG write you.CL-DAT
'Even if I do not write to you.' [Pappas, 2006: 323]

The data presented in Pappas (2006) are also highly problematic for the account given by Condoravdi and Kiparky (2002: 26). Condoravdi & Kiparsky (2002) propose that prior to the emergence of the Σ phrase, focused elements and negation would have been fronted to SpecCP.[4] Subsequent movement of the verb in C would predict clitics to appear postverbally in these cases, contradicted by Pappas' findings. The clause structure assumed for KG and MPG by Condoravdi & Kiparsky (2001) is shown below:

(7.5) Clause structure assumed for MPG (Condoravdi & Kiparksy, 2002)

$[_{CP}[_{C'}[_{C}\mathrm{V}_j[_{TP}\mathbf{Cl}[_{TP}[_{T}\ \mathrm{V}_j][_{VP}\ \mathrm{t}_j]]]]]]]$

Furthermore, the same account predicts complementizers to be associated with preverbal positioning only, since assuming that complementizers will appear in C, there is no slot for the verb to move to anymore. This latter claim, as already discussed, is also falsified by

[4]The Σ phrase hosts Wh elements, negation/tense markers and complementizers.

Pappas' findings. Given this state of affairs, Pappas (2006: 324) notes that it is unlikely that the KG clitic system can be captured within one grammatical system and argues that the variation exhibited in KG provides evidence that different subsystems of the grammar are in fact competing. He argues that evidence for competing grammars in KG come from a number of transitions in the grammar of Greek during the Graeco-Roman period, notably the transition from verb final to verb medial word order (Taylor, 1994), from infinitival to finite complementation (Joseph, 1983) and the emergence of new functional categories via clause restructuring (Philippaki & Spyropoulos, 2004). Direct evidence for such a competition, according to Pappas, can be found in existing examples of scribal errors and corrections. Pappas notes two striking examples of this sort. In the first example, the clitic is written twice, i.e. in both preverbal and postverbal position, a fact that according to Pappas (2006: 324) indicates "uncertainty as to its correct position":

(7.6) *ean* *moi* *paradois* *moi* *tous* *anthrōpous*
if me.CL-DAT give me.CL-DAT the people
'If you deliver the people to me.' [Pappas, 2006 apud Oxy: 2981, 2nd cent]

In the second example the clitic is put in both the traditional Wackenagel position as well as in a position preceding the verb before it is erased and the verb is written above it:[5]

(7.7) *Hoti* *egō* *gar* *autos* *soi* *anelipōs* *[[soi]]* *\ō/*
Because I PART self you.CL-DAT constantly you.CL-DAT write
'Because I myself constantly write to you.' [Pappas, 2006 apud Oxy: 2980, 2nd cent;]

Pappas then presents an analysis of the KG clitic system that involves three subsystems. The first subsystem captures the traditional Wackernagel system (no formulation is given by Pappas). The second subsystem is the system described in Taylor (2002). Pappas proposes that clitics in that system are X_{max} enclitics adjoined to IP. They are able to appear before the verb as long as there is no phonological boundary Φ to their left. The clausal structure assumed for this subsystem is the one proposed by Philippaki & Spyropoulos (2004):

(7.8) $[_{CP}[_{MP}[_{NegP}[_{IP}\mathbf{Cl}[_{I'}[_{I}\ V_j][_{VP}\ t_j]]]]]]]$

[5] [[X]] indicates erasure, \X/ indicates insertion from above).

As can be seen from the above structure, the clitic moves out of its thematic position and it is adjoined to the left of the IP. In the presence of a phonological boundary to the left, PI applies and enclitic positioning obtains. In the presence of wh-elements, complementizers and negation, the clitics appear preverbally since all these elements appear in the same phonological domain as the clitic and the verb. Thus, in these cases no phonological boundary is present to the left of the clitic, thus no application of PI. The cases that are not covered by these two subsystems, namely postverbal positioning in the presence of complementizers, wh-elements and negation is captured via a third subsystem. This third subsystem is what Condoravdi & Kiparsky (2002) assume for PG. According to that subsystem, clitics are X_0 enclitics that attach to the right of the verb. One of the things worth mentioning here, already noted by Pappas (2006: 325), is that such competing systems will have to admit that a certain degree of ambiguity is present in the system, since as Pappas notes "speakers would not have been able to distinguish postverbal clitics that are the result of PI from those that belong to the innovative subsystem" (Pappas, 2006: 325). Pappas' analysis in terms of competing grammars is a rather intriguing and interesting one. However, I will argue that a better alternative to such a proposal is possible.

7.2.1 Competing Grammars or a Pragmatically Governed System?

As is already evident from our previous discussion, the KG clitic system presents a complicated situation that seems to resist any grammatical generalization. This complexity led Pappas (2006) to assume that such a system cannot be captured within one syntactic account. In that sense, Pappas presents an analysis where the KG clitic system is derived by making use of three different grammatical subsystems. These subsystems are different grammars in competition in the sense of Kroch (1989). The problem with such an account however lies in the time span that this competition extends over. It is relatively easy to accept that once a new grammatical feature is introduced in the grammar, it does not immediately replace the old feature but for some time both of the features are present in the grammar. Assuming three competing subsystems with respect to clitic placement, one would expect one of them to prevail in some future stage of the language. However, such a fact is not borne out by the data Pappas himself presents (Pappas, 2004; Pappas, 2006).

Pappas (2004) notes that no generalization can be given for the MMG clitic system, and as such none of the proposed subsystems proposed for KG would work for MMG. Such a thing is not a problem in itself, since one can very easily assume that new subsystems arose and different competing grammars are now at play in MMG. What is really hard to understand is how a system like MPG is going to be accommodated under a competing grammar hypothesis. The data from MPG are extremely interesting, since they do differ from both the MCG and the MMG data in crucial aspects as shown in Pappas (2006). The difference lies in the fact that the number of proclitic environments are far less in MPG than in the other two medieval varieties. The data from MPG are repeated below:

(7.9) Clitic placement in MPG (adapted from Pappas, 2006: 316)

Environment	Preverbal	Postverbal	Date
Clause initial	0	19	13th, 14th, 15th
Fronted constituent	0	10	13th
Temporal expression	0	3	13th, 14th, 15th
ouk	0	1	13th
kath$\bar{o}$s	0	2	13th
epei	0	1	13th
m$\bar{e}$p$\bar{o}$s	0	1	13th
hina	14	0	13th, 14th, 15th
as	1	0	13th
me	1	0	13th
os	1	0	13th
Subject	1	0	13th
Wh-expression	7	19	13th, 14th, 15th

It is very hard to see how a system like the one shown in the above table will be captured by one or more of the subsystems proposed in Pappas (2006). The nature of the problem seems to be the fact that the MPG system as presented above seems to resist generalizations, even partial ones. The major tendency appears to be enclitic positioning. The only environment that seems to categorically trigger proclisis is the one involving the complementizer *hina*. Fronted constituents as well as temporal expressions that constitute proclitic

environments in other Medieval Greek systems are not associated with proclisis in MPG. Furthermore, a number of subordinate conjunctions do not trigger proclisis but enclisis (*kathōs* 'as' *epei* 'because', *mēpōs* 'that' (dubitative)). Even wh-expressions that are the proclitic environments par excellence in MMG and MCG, exhibit variation in MPG with enclisis being more frequent than proclisis (19/26 tokens). It is very difficult to imagine a minimalist analysis of the above facts. For example, let us take the case of wh-elements, which, as already said, present variation in positioning. Assuming the clausal structure in (7.8), and the assumptions made by Taylor (2002), one would expect proclisis with wh-elements. However, enclisis is not only possible in MPG but it seems to be the preferable option.[6] At that point, an analysis assuming competing grammars can further use another subsystem in the same sense Pappas (2006) following Condoravdi & Kiparky's analysis for PG clitics did for KG. This subsystem will assume clitics to be X_0 enclitics. This will solve the problem for wh-expressions. However, what cannot be accounted for is categorical proclitic positioning with *hina*. Assuming these two subsystems, and no matter where *hina* is assumed to be instantiated in the clausal structure (C_0 or M_0), it is impossible to get the facts right. The two subsystems will predict variation with *hina* as well, contrary to fact. As far as I can see, there is no viable explanation for the transition from KG to MPG at least within GB/Minimalist grounds. One last problem for the analysis proposed by Pappas (2006) is the existence cases in which the clitic appears sentence initially. Although rather rare (4/231 in the Oxyrynchus Papyri), these constructions exist and cannot be captured by any of the three proposed subsystems. The old Wackernagel system did not tolerate clitics in first position by definition, the system based on PI will predict PI in these cases since no host for the clitic will be available and the system based on Condoravdi & Kiparsky (2002) assumes clitics to be enclitics to the verb. Thus, no system will predict sentence initial clitics to be available in KG, contrary to fact.

If an analysis based on competing grammars cannot provide us with an explanation of the transition from KG to Medieval Greek, then is there an alternative account that will fare better with respect to this transition? As I have already shown in chapter 3 and this chapter (table 7.2), a number of elements that are categorically restricted to either proclisis

[6]Even though it might not be that wise to posit preference statements based on such short samples. In any case, the fact is that enclisis is possible with Wh-elements in MPG.

or enclisis in MMG show variant positioning in KG. Table 7.2 is repeated below:

(7.10) Clitic positioning in the *Oxyrynchus Papyri* (vols. 1-56)

Environment	Preverbal	Postverbal
Clause initial	4	231
Infinitival complement	12	0
Adverbs	34	14
NP-object	21	15
NP-Subject	10	13
PP	13	37
Complementizers	17	37
Wh-expressions	18	16

An alternative way of looking at the above data is to assume that the various syntactic constraints proposed in order to capture clitic positioning in KG or the different grammar subsystems exhibiting different syntactic constraints (as in Pappas, 2006) are actually not there. Then, following Bouzouita (2008a, 2008b) building on Pickering & Garrod (2004), it can be argued that the clitic positioning system of KG is not governed by syntactic constraints but rather by underlying pragmatic preferences. Within this line of reasoning, I argue that the complicated picture arising from the KG data is the result of a number of positioning tendencies depending on a pragmatically governed clitic system. Assuming such a pragmatically governed system, one can get a surprisingly neat and economical account of the KG clitic facts, by making the at first sight extravagant assumption that no restrictions exist in the entry for KG clitics at all! This may seem quite exaggerated as a statement but by taking a closer look, one will realize that such an assumption is not that extravagant in the end. Before the consequences of such a lexical entry are discussed, let us first see how such an entry is to be formulated. Encoding no restrictions in the entry means that the only thing we need is just an initial trigger. This trigger would again assumed to be a type *t* requiring node. Given that this trigger is satisfied, then parsing of the clitic can proceed. The lexical entry for an accusative marked clitic in KG is shown below:[7]

[7]Ordering and clustering in KG or the medieval dialects is not going to be discussed. The entries that will be given will refer to accusative clitics and will assume fixed structure associated with these clitics. This

(7.11) Lexical entry for an accusative clitic in KG

IF $?Ty(t)$

THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$**)**

$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$

$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$

$\texttt{put}(Ty(e), Fo(U_x);$

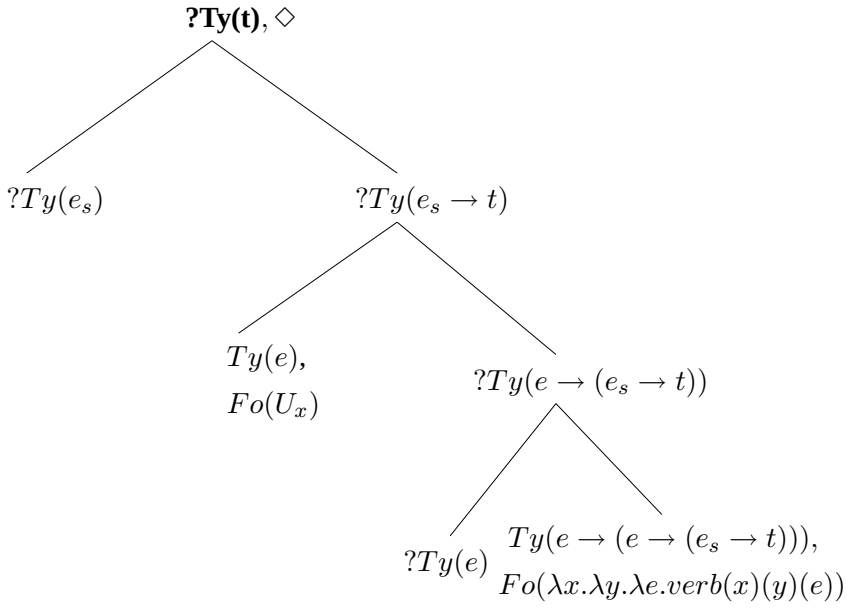
$?\exists\mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$

ELSE abort

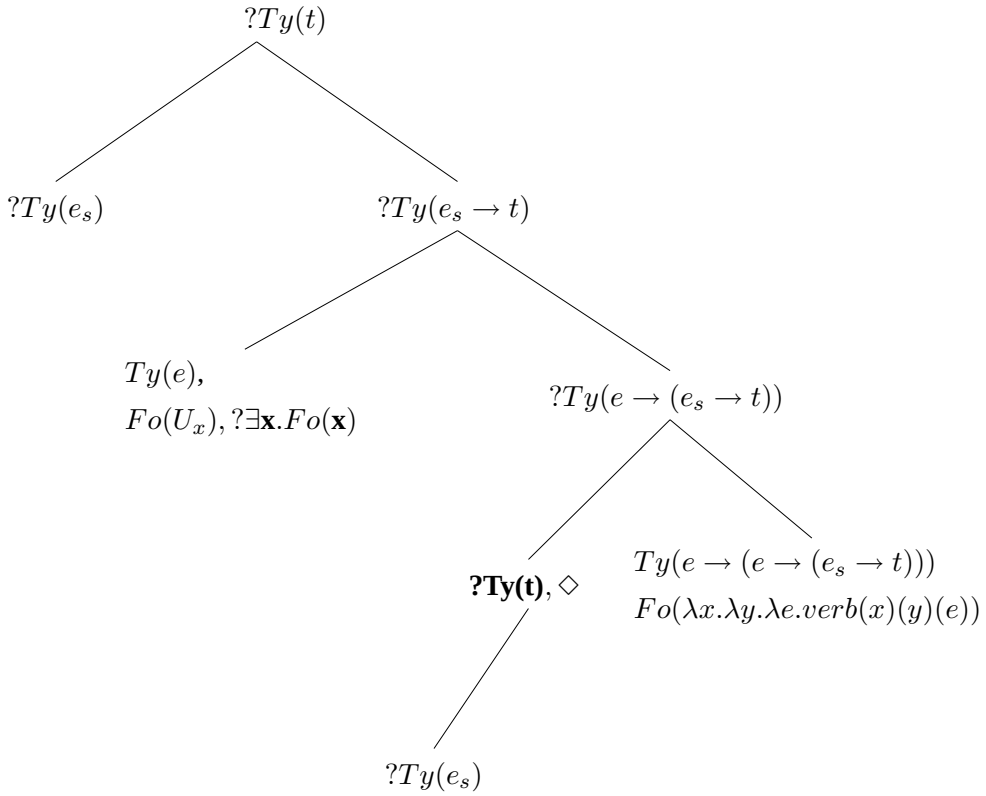
But let us see what the above entry will predict. I begin with the enclitic cases. In these cases a verb has already been parsed when the clitic comes into parse. Following the assumptions made in this thesis for verbs in the modern dialects examined, I assume that the pointer in KG is left at the type *t* requiring node as well when the verb is parsed. Parsing of the clitic is possible after a verb has already done so. In case a complementizer is parsed (these includes both complementizers and subordinate conjunctions), as discussed in chapter 4, the pointer is again at the type *t* requiring node, so the clitic can be parsed, giving rise to proclisis in environments where a complementizer or a subordinating conjunction exists. However, in the presence of a complementizer/subordinating conjunction the clitic can be also parsed after the verb. Variation is thus correctly predicted. The same is true for negation marker *mē* as well as focused elements, i.e. objects, subjects, PPs and adverbs. Assuming parsing of these elements on an unfixed node and given the fact that these will project a type value in the unfixed node, COMPLETION applies moving the pointer to the type *t* requiring node, making parsing of the clitic possible. The examples below exemplify the structure we get after parsing a transitive verb, a complementizer, a subordinating conjunction and a focussed constituent parsed on an unfixed node:

is done for illustration purposes and does not imply that all accusative are fully specified in KG and the medieval dialects. The data needed to decide on the issue are not available and as such this issue will not be dealt with here. The relevant part in these entries is positioning rather than clitic ordering or structural underspecification in KG or the medieval dialects.

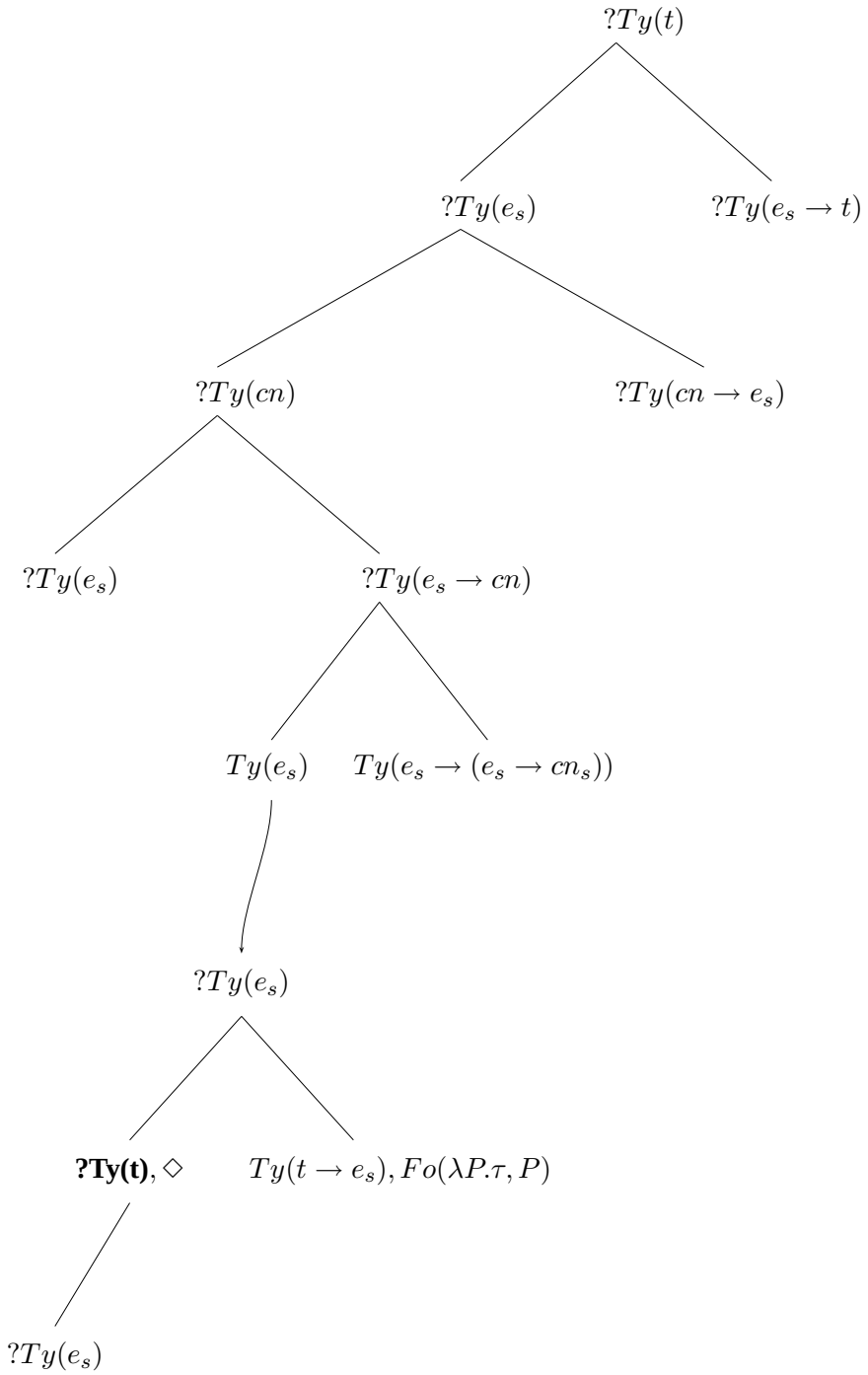
(7.12) Parsing a transitive verb



(7.13) After parsing a complementizer (embedded clause)



(7.14) Parsing a subordinate conjunction



(7.15) Parsing a focused constituent on an unfixed node and application of COMPLETION

$Tn(n)$, **?Ty(t)**, $\diamond$

$\langle\uparrow *\rangle Tn(n)$

$Ty(x)$

$Fo(x')?\exists\mathbf{x}.Tn(\mathbf{x})$

Note that the entry as it is will also predict sentence initial clitics to be possible according to the data in the *Oxyrynchus Papyri*. However, given that no sentence initial clitics are found in the other KG texts such as the *New Testament* or *The Shepherd of Hermas*, someone might argue that indeed a restriction on sentence initial clitics exists. In case this is true, a slight modification in the entry will ban first position clitics. This modification involves a further condition which will state that something must hold below the type *t* requiring node in order for the clitic to get parsed:

(7.16) Lexical entry for an accusative clitic in KG excluding sentence initial clitics

IF $?Ty(t), \langle\downarrow^{+}\rangle\ \top$

THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$**)**

$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$

$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$

$\texttt{put}(Ty(e), Fo(U_x);$

$?\exists\mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$

ELSE abort

Under the account proposed, the apparent syntactic restrictions are the result of pragmatic preferences rather than syntactic constraints. These pragmatic preferences can be relevance-driven, a claim already made by Bouzouita (2008a: chapter 6) for Latin weak pronouns. According to Bouzouita, an explanation for the Latin weak pronoun system can be given assuming a processing effort minimization account in the sense of Cann & Kempson (2008). Cann & Kempson (2008) argue that clitics being anaphoric expressions, need to be identified by a substituent in the context. It is then argued that parsers require

the search space for such substituents to be as small as possible in order for cognitive effort to be minimized under general relevance theoretic considerations (Sperber & Wilson, 1986). Anaphoric expressions must then appear as early in the clause as possible, a strategy aiming to minimize search space. In the case of Latin, weak pronouns tend to appear immediately after an emergent propositional domain has been established. For example the rule of *ADJUNCTION associated with focused constituents in both Bouzouita (2008a) and in the account presented here can only apply at the beginning of a clause, main or embedded, since the rule takes effect only if no daughter nodes exist below the node that the unfixed node is projected (see this thesis: chapter 2).[8] Such a signal for an emerging propositional domain is provided by complementizers or subordinate conjunctions as well. All these elements project the situation argument node and decorate it with a requirement for a type e_s, signaling a new propositional domain (Bouzouita, 2008a,b uses a tense requirement since the older DS version with no situation nodes is assumed). Negation has similar properties, since it also projects a situation argument requirement (negation is captured using the feature *[+NEG]* as a proclitic trigger in Bouzouita, 2008a,b). However, unlike Latin, in which no sentence initial clitics were possible, KG as we have already said seems to allow sentence initial clitics, although extremely rare. In that sense no Tobler-Mussafia like restriction (first position restriction) needs to be encoded in the entry for KG clitics.

Recapitulating, I argue that the apparent complexity of the KG system that has been claimed to refuse systematization within one single grammatical system (Pappas, 2006) can be neatly explained assuming that the putative syntactic constraints are no more than pragmatic preferences driven by cognitive effort minimization. In that sense, I have proposed a lexical entry for KG clitics free of any syntactic restrictions. Such an entry will

[8]Note that given the newest assumptions in DS where a situation argument is explicitly encoded in the tree structure, the fact that *ADJUNCTION can apply only when no other node is present below, will predict that no unfixed node should be possible in case a situation argument exists. This has the consequence that an unfixed node cannot be used after Modality/Tense/Negation markers and subordinating conjunctions. For Modality/Tense/Negation markers such a fact is true. As regards subordinating conjunctions, there are no concluding data where both a subordinating conjunction and an element parsed on an unfixed node exist, and the proclisis obtained can be safely attributed to the element parsed on the unfixed node. If these cases exist, then a minimal modification in the *ADJUNCTION rule that will allow the *ADJUNCTION rule in the presence of the immediately argument dominated node will allow such structures. Given that we do not have these data, the rule of *ADJUNCTION is kept as it is.

correctly predict the KG facts. Of course, such an entry cannot predict the exact percentages in which varation can occur and cannot make any predictions on how strong these pragmatic preferences are.

Given the proposed pragmatically governed account of KG, it is now time to look at the transition from such a system to the systems of the medieval dialects of Pontic, Cypriot and Standard Modern Greek.

7.2.2 The Transition from Koine to Medieval Greek

The next step in sketching a diachronic account of clitics in the dialectal systems we are discussing, is to look at the transition from KG to the respective medieval dialects of the dialects under consideration. These medieval dialects exhibit a number of similarities but a number of subtle differences with respect to clitic positioning as well. As Pappas (2006) has shown, the picture emerging from the systems of MPG, MCG and MMG is one of a cline towards generalization of proclisis, with MPG leaning the most towards enclisis while MMG leaning the most towards proclisis.[9] In particular, the proclitic environments in MMG are more than the respective proclitic environments in MCG and the proclitic environments in MCG are more than the ones found in MPG. The tables below taken from Pappas (2004, 2006) illustrate this claim:

[9]Unfortunately, I do not have the data as regards the medieval dialect of Southern Italy. In that sense, the transition from the medieval dialect of Southern Italy to GSG will not be discussed in this thesis, leaving this issue as a subject of future research. However, see Minas (1994) for a discussion on the grammar of medieval Southern Italian Greek.

(7.17) Clitic positioning in MMG (adapted from Pappas, 2004)

Environment	Preverbal	Postverbal
Clause initial	59	719
Coordinating conjunction	58	681
oti	8	30
δioti	4	0
Reduplicated object	39	79
Wh-elements	439	5
Negation	431	3
na,ina,as	1525	4
ean,an,pos	324	2
Object,PP,non-temporal adverb	898	90
Subject	334	130
Temporal expression	86	63
Imperatives	25	305

(7.18) Clitic positioning in the Cypriot Chronicles (vols. 1-56, adapted from Pappas, 2004)

Environment	Preverbal	Postverbal
Clause initial	0	208
Reduplicated object	0	15
Function word	101	3
Fronted constituent	1	14
Subject	0	29
Gerund	0	19
Imperative	0	4

(7.19) Clitic placement in MPG (adapted from Pappas, 2006:316)

Environment	Preverbal	Postverbal	Date
Clause initial	0	19	13th, 14th, 15th
Fronted constituent	0	10	13th
Temporal expression	0	3	13th, 14th, 15th
ouk	0	1	13th
kath ō s	0	2	13th
epei	0	1	13th
m ē p ō s	0	1	13th
hina	14	0	13th, 14th, 15th
as	1	0	13th
m ē	1	0	13th
ō s	1	0	13th
Subject	1	1	13th
Wh-expression	7	19	13th, 14th, 15th

The first thing one notes when comparing the medieval dialects to KG is that a number of environments that are unstable in the KG system, in the sense that they exhibit variation, tend to become fixed by choosing one of the two possibilities, enclisis or proclisis. This is for example the case for wh-elements, the negation marker *mē* and subordinate conjunctions in MMG and MCG (note that in the MCG table all these elements are grouped as a common category, i.e. function word). MPG, on the other hand, does not follow MMG and MCG and presents a more conservative picture with proclisis being categorical only with the complementizer *hina* 'in order to'. There are instances of categorical preverbal placement with some complementizers and categorical enclitic placement with others but the number of tokens is too small to be conclusive (e.g. complementizer *ōs* associated with enclisis has only one token). Sentence initial clitics are not found in MCG and MPG but are found in MMG. In what follows, I will argue in line with Bouzouita (2008a,b) that the transition from KG to the medieval dialects can be summarized as follows: the pragmatic reasoning governing the KG clitic system was progressively lost and gave rise to stricter

positioning systems where pragmatic preferences were lexically encoded as syntactic restrictions. This whole process of syntactic constraints substituting pragmatic preferences can be explained as a routinization process in the sense of Pickering & Garrod (2004). Routinization is in itself a very common concept in the psycholinguistic literature (Garrod & Pickering, 2004; Pickering & Garrod, 2004; Ruh et al., 2005 among others) and can be seen as a process establishing a number of routines that reduce cognitive load, in effect parsing shortcuts. According to Pickering & Garrod (2005: 89), a routine is "an expression that is "fixed" to a relatively great extent". Routinization in the psycholinguistic literature is something like long-term alignment. Short-term alignment can occur between dialogue participants within the time span of only a few sentences. Alignment can range from low-level lexical alignment to pure syntactic alignment (see Garrod & Doherty, 1994; Branigan et al., 2005). Once these alignment patterns get repeated, they are stored in the memory as routines. This long-term alignment is what we call routinization. Bouzouita (2008a,b) using this concept of routinization, argues that syntactic change is in fact driven by routinization.[10] In our case, the repetition of a number of clitic positioning patterns that had an underlying pragmatic basis, got routinized and became a kind of a processing shortcut. Once this had happened, the pragmatic basis was lost and subsequently replaced by the encoding of these routinizations as syntactic constraints inside the lexicon.

The following step is to show how the routinization hypothesis is going to be transformed into a syntactic account of diachronic change. In particular, we need to provide a formal account of the transition from KG to each individual medieval system. Starting with the MMG system, we notice striking similarities to the system found in modern day CG. To start with, we notice that the elements that are assumed to be parsed as unfixed nodes trigger proclisis, in the same way as in CG. This category of elements includes wh-elements, fronted objects, PPs and adverbs as well as subjects. Notice that objects, PPs, adverbs as well as subjects may all involve alternative parsing strategies, e.g. subjects and adverbs can be parsed on a LINK structure in topic and non-focussed environments respectively. In that respect, the variation in those cases is expected, since alternative parsing strategies are always available. Wh-elements on the other hand, are standarly assumed to involve

[10] As Bouzouita (2008: chapter 6: 314) aptly notes the concept of routinization has intriguing similarities with a number of other terms (automatization, syntactization) proposed in the historical linguistics literature. See Bouzouita (2008: chapter 6) for discussion and references.

only one parsing strategy, that of the unfixed node, and as such they are expected to exhibit categorical proclisis, a fact which is not far from the truth (only 4/444 instances of enclisis with wh-elements are found in Pappas, 2004). The next category of elements that comprised a set in our account of CG clitics, was the set of elements assumed to be inducing a requirement for a type e_s in the daughter node of a type *t* requiring node. This category includes subordinating conjunctions, subjunctive markers *na/as*, future marker *enna* as well as negation particles *den* and *min*.[11] These elements seem to behave accordingly in MMG, since proclisis is near categorical and in some cases categorical (e.g. 1436/1439 instances of proclisis for the subjunctive marker *na*, 237/238 instances of proclisis for temporal/comparative conjunctions and 24/24 instances of proclisis with conjunction *hina* 'in order to'). Similarly negation exhibits near categorical proclisis (431/434 tokens). What seems to be slightly different is the behaviour of MMG with respect to sentence initial clitics. There are a number of instances of sentence initial clitics (59/778). Proclisis is also found with coordinating conjunctions (58/739 tokens). Assuming a similar entry for the one given for CG will undergenerate in these cases. A solution to this problem would be encoding an additional trigger in the entry for MMG that will allow sentence initial clitics. This trigger would be actually the opposite of what we have used for KG in order to exclude the same kind of clitics. Instead of using the $\top$ operator, we will instead use the $\bot$ operator, positing a restriction that nothing must hold below the current node.[12] The proposed entry is shown below:

[11]Note that predominant enclisis with the archaic form of negation *ou* is not going to be dealt with in this thesis. It should be kept in mind that a full analysis of the clitic system of MMG will have to take care of this fact as well.

[12]The other way to look at sentence initial clitics within such a system is to assume that first position clitics are actually the result of overgeneralization of proclisis in the system. The claim is that the first traits of the SMG system where proclisis is the norm are found in these texts. Given the number of proclitic environments in MMG, such a claim is not implausible at all. It is very easy to imagine a situation where proclisis is generalized to all environments due to the number and frequency of proclitic triggers. The same reasoning can be used for coordinating conjunctions. It might in that case argued that categorical proclisis with subordinating conjunctions gave rise to overgeneralizing proclisis to all conjunctions. See the discussion on the transition from MMG to SMG, this chapter.

(7.20) Lexical entry for clitics in MMG

IF $?Ty(t)$
THEN IF $\langle\downarrow_*\rangle ?\exists x.Tn(x)|$
$\langle\downarrow_0\rangle ?Ty(e_s)|$
$[\downarrow^+]\bot||$
$\langle\downarrow_1^+\rangle Ty(x)$
THEN **(**$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(Ty(e), Fo(U_x), ?\exists \mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$
ELSE abort
ELSE abort

The lexical entry below shows the transition from KG to MMG:

(7.21) The transition from KG to MMG

KG clitics

IF $?Ty(t)$
THEN **(**$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(Ty(e), Fo(U_x);$
$?\exists \mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)));$
ELSE Abort

$\mapsto$

MMG clitics

IF $?Ty(t)$
THEN IF $\langle\downarrow_*\rangle ?\exists x.Tn(x)|$
$\langle\downarrow_0\rangle ?Ty(e_s)|$
$[\downarrow^+]\bot||$
$\langle\downarrow_1^+\rangle Ty(x)$
THEN **(**$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(Ty(e), Fo(U_x);$
$?\exists \mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$
ELSE abort
ELSE abort

Comparing the lexical entries of KG and MMG, we see the loss of a pragmatically governed system (no syntactic restrictions) and the encoding of these restrictions syntactically.

Moving on to the case of MCG, things are somewhat different, since proclisis is encountered in less environments. One of the things that one would expect given a system

where proclisis is not widespread and given the hypothesis that sentence initial clitics in MMG are the result of overgeneralization of proclisis due to the great number of proclitic environments, is that sentence initial clitics should not be found in MCG, something which is true. The difference between MCG and MMG is that most of the environments associated with the unfixed node strategy do not induce proclisis but rather enclisis. Fronted constituents are near categorically associated with enclisis (1/14) while the same holds for subjects (29/29). Hence, the only elements that are parsed on an unfixed node and trigger proclisis in MCG are the set of wh-elements. In that sense, it can be claimed that the unfixed node strategy in MCG is not yet generalized as a proclitic trigger. The encoding of the entry for MCG will thus refer to the unfixed node strategy but only as applied to Wh-elements. It is then the generalization of the parsing strategy of wh-elements, i.e. the unfixed node strategy, that will give rise to the spread of proclisis to constrastive focus environments in CG. On the other hand, a type e_s trigger will be present in MCG in the same sense it is present in MMG or CG. In specific, the function word category (see 7.18) that includes subordinating conjunctions, modality/tense markers, negation particles and wh-elements is nearly categorical proclitic (101/104 tokens). All these elements, except wh-elements, are assumed to be captured via the type e_s requirement trigger. The lexical entry for MCG will thus involve the following triggers: One capturing enclitic positioning, the e$_s$ requirement trigger, and a restricted unfixed node trigger applied only to wh-elements. The lexical entry is shown below:

(7.22) The transition from KG to MCG

KG clitics

IF $?Ty(t)$

THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$**)**

$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$

$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$

$\texttt{put}(Ty(e), Fo(U_x);$

$?\exists\mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$

ELSE abort

$\mapsto$

MCG clitics

IF $?Ty(t)$

THEN IF $\langle\downarrow^*\rangle?\exists\mathbf{x}.Tn(\mathbf{x})$

$\wedge Fo(WH)|$

$\langle\downarrow_0\rangle?Ty(e_s)||$

$\langle\downarrow_1^+\rangle Ty(x)$

THEN **(**$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$**)**

$\texttt{make}(\langle\downarrow_1\rangle); \texttt{go}(\langle\downarrow_1\rangle);$

$\texttt{make}(\langle\downarrow_0\rangle); \texttt{go}(\langle\downarrow_0\rangle);$

$\texttt{put}(Ty(e), Fo(U_x);$

$?\exists\mathbf{x}.Fo(\mathbf{x});$

$gofirst(?Ty(t)))$

ELSE abort

ELSE abort

Lastly, turning to MPG, one sees that this dialect is the most conservative of the three in the sense that spread of proclisis is strikingly more limited than in the other two. The only environment that is categorically related to enclisis is complementizer *hina*, 'in order to'. There are other instances of categoricality, e.g. enclisis with temporal and dubitative conjunctions *epei* and *mēpōs* respectively, but the number of tokens is too small to derive a safe conclusion. Our account will have to rely on the data available, even though these, as said, cannot be decisive with respect to a number of environments. The situation as presented by the data can be described framework-independently as a situation where fronting does not play a role in clitic positioning (10/10 cases where a fronted constituent is present are associated with enclisis), and proclisis is not generalized as regards subordinating conjunctions and wh-expressions. Under a DS perspective, the generalized proclitic triggers, i.e. the type e_s and the unfixed node trigger, cannot be maintained for MPG. It seems that in MPG, the lexical entry for clitics had to specify each of the elements that are associated with proclisis separately. The fact that proclisis did not spread in MPG had the effect of overloading the lexicon with information regarding every single element exhibiting proclisis. The latter , as I will argue, was the basic reason PG obtained a strictly enclitic rather than a proclitic system. The entry I am going to present for MPG will thus involve actual

reference to each separate proclitic item. In what follows I will use diacritics for the encoding of *hina* and the various subordinating conjunctions. However, it should be kept in mind that a proper analysis will have to get rid of these diacritics. For the moment, imagine these diacritics to stand for a subset of the actions associated with parsing these elements. Since I cannot really go into the actual actions these elements induce, i.e. give a proper syntactic analysis for all these elements, diacritics will be used. Furthermore, the *Fo(WH)* trigger we have used for MCG will have to be used after an exclusive disjunction (||). This will predict that variation is possible with wh-expressions. Note that the situation gets worse in terms of complexity if we add a proclitic trigger specifically for subjects, assuming that subjects give rise to proclisis when on an unfixed node. Such a trigger will refer to the existence of a nominative case filter ($\langle\downarrow^*\rangle?\langle\uparrow_0\rangle Ty(t)$). The lexical entry for MPG is shown below:

(7.23) Lexical entry for MPG

IF		$?Ty(t)$
THEN	IF	$+HINA\|$
		$+\bar{\text{O}}S\|$
		$AS\|$
		$[+NEG]\|$
		$\langle\downarrow^*\rangle?\langle\uparrow_0\rangle Ty(t)\|\|$
		$\langle\downarrow_1^+\rangle Ty(x)\|$
		$\langle\downarrow^*\rangle Fo(WH)$
	THEN	**(**$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$**)**
		$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$
		$\texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle);$
		$\texttt{put}(Ty(e), Fo(U_x);$
		$?\exists\mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$
	ELSE	abort
ELSE	abort	

The lexical entry is extremely complex but this can be justified as MPG seems to resist a simpler analysis. It is extremely difficult for example to distinguish between the subordinate conjunctions causing enclisis and those that cause proclisis or the variation observed

with wh-expressions. Regardless the choice of the framework used, one will have to posit a number of stipulations to capture the facts. I will argue that lexical complexity is what prevented proclisis to spread, producing a strict enclitic system in later stages of the language.

7.2.3 From Medieval to Modern Greek

The next step in giving a diachronic account of clitic placement in the dialects presented is to see how the medieval ancestors of these have developed into today's systems. Let us start with SMG. As already seen in chapter 3, the clitic system of SMG is exclusively verb centered, with proclisis being generalized to all finite verbs while enclisis being retained only for non-finite verbal forms, i.e. imperatives and gerunds. The system of MMG, on the other hand, is a system where given the absence of proclitic environments, enclisis obtains. In that sense, the form of the verb does not seem to play any role in clitic positioning in MMG. The question that arises is how such a non-verb centered system like the one exhibited by MMG gave rise to the respective SMG system? In our discussion regarding the clitic systems of the medieval dialects of SMG, CG and PG we have seen that MMG exhibited the biggest number of proclitic environments. In that sense, proclisis was more generalized in MMG than in MCG or MPG. This informal observation might be the key in explaining the transition from one system to the other. Note that in the case of MMG, a number of sentence initial clitics can be also found. The number of tokens is not neglectable and might point out that a process towards a verb centered system is already taking place in MMG (59/778). Further evidence might come from the number of proclitic tokens with coordinating conjunctions (58/739). The question is what is the trigger and the nature of such a process. Bouzouita (2008a) in discussing the transition from Medieval Spanish (MedSp) to Rennaisance Spanish (ReSp) and from Resp to Modern Spanish (ModSp), argues that both routinization and re-analysis are at play. The argument is that once the pragmatic basis of the Latin system got lost, pragmatic preferences were encoded in the lexicon of MedSp. The loss of the pragmatic basis, led to a speaker/hearer mismatch, since the existence of alternative parsing strategies for parsing one and the same lexical element could easily lead to a situation where the parsing strategy used by the speaker and the one parsed by the hearer do not match (e.g. an unfixed node used by the speaker is analyzed as

a LINK from the hearer). However, unlike SMG that exhibits a similar development to the ModSp clitic system, the systems of CG and PG are completely different and as such can be used as a testing ground for the "routinization driving syntactic change" hypothesis.

I argue that the high number of proclitic environments in MMG is the actual cause of proclisis generalization in SMG. The shift towards a generalized proclitic system seems to be already at play in MMG since a number of cases of sentence initial clitics are found. Let us take a look again at the lexical entry given for MMG:

(7.24) Lexical entry for clitics in MMG

IF $?Ty(t)$
THEN IF $\langle\downarrow_*\rangle ?\exists x.Tn(x)|$
$\langle\downarrow_0\rangle ?Ty(e_s)|$
$[\downarrow^+]\bot||$
$\langle\downarrow_1^+\rangle Ty(x)$
THEN **(**$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle)$**);**
$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle)$;
$\texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle)$;
$\texttt{put}(Ty(e), Fo(U_x), ?\exists\mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$
ELSE abort
ELSE abort

Notice that the above entry also has a condition capturing first position clitics. I do not know if such condition is really part of the MMG system or is the result of a process of change towards generalized proclisis being already at play. Let us assume for the moment that it is part of the actual system. Notice that the proclitic triggers are separated from the enclitic one by means of an exclusive disjunction. Justification of such a move was given in discussing CG in chapter 4. The exclusive disjunction ensures that the enclitic trigger can be satisfied only in case none of the other proclitic triggers are. There is no such disjunction between the proclitic triggers, meaning that the clitic can be parsed after one or more of the proclitic triggers are present (the disjunction in between the proclitic triggers has an inclusive interpretation). So how did proclitic generalization take place? One of the reasons already noted by Bouzouita (2008a) is the availability of different processing strategies for

parsing one and the same syntactic element. For example, a subject can be parsed either on an already introduced unfixed node or as a LINK structure. Assuming a parsing/hearer mismatch where the hearer analyzes an unfixed node as being a LINK, the hearer associates proclitic positioning with the LINK strategy as well. The direct consequence of that is that other structures involving LINK structures that were not associated with proclisis, notably coordinating conjunctions, can now be associated with proclisis:

(7.25) Speaker's structure

?Ty(t)

$\langle\uparrow^*\rangle ?Ty(t)$
$Ty(e), Fo(x')$
$?\exists\mathbf{x}.Tn(\mathbf{x}), \diamond$

(7.26) Hearer's reanalysis

$\langle L\rangle Tn(0)$ $\langle L^{-1}\rangle, Tn(n), ?Ty(t), ?\langle\downarrow_*\rangle, Fo(x'), \diamond$
$Fo(x')$
$Ty(e)$

Given a generalization of the LINK strategy, the number of elements causing proclisis is now vast, since all preverbal arguments, all conjunctions and all modality/tense markers are now associated with proclisis. Spread of proclisis had the consequence that the embedded proclitic triggers were not discernible anymore. Within this context, the next step involves generalization of the initial *?Ty(t)* trigger by collapsing all the embedded proclitic triggers and developing a general proclitic trigger which ensures that no verb has been parsed. Given that almost all elements appearing before the verb could cause proclisis given the parser/hearer mismatch described, the development of a general trigger that ensures that no verb has been parsed and the collapse of all proclitic triggers in favor of this more general proclitic trigger is not implausible at all. Thus, the first step towards the verb centered system of SMG is proclisis generalization:

(7.27) The transition from MMG to SMG: Step 1: Proclisis spread

IF $?Ty(t)$

THEN IF $[\downarrow_1^+]?Ty(x)||$
$\langle\downarrow_1^+\rangle Ty(x)$

THEN **(**$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(Ty(e), Fo(U_x), ?\exists\mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$

ELSE abort

ELSE abort

Note that the above entry contains two conditions that are contradictory. The first one ($[\downarrow_1^+]?Ty(x)$), states that if no verb is present, then the clitic can be parsed, while the other states that if a verb is present then you can parse the clitic ($<\downarrow_1^+> Ty(x)$). Such a formulation would allow the clitic to be freely associated with both proclisis and enclisis for any type of verb, and as such will not be able to distinguish anything at all, given that it will freely allow everything. However, generalization of proclisis, with the exception of focussed constituents, involved only finite forms and as such the ability of non-finite contexts[13] to retain enclitic dominance was due to the fact that most of the elements triggering proclisis were incompatible with non-finite forms (wh elements, complementizers, subordinating conjunctions, modal and tense particles are impossible in non-finite environments). In that sense, the enclitic trigger even though retained was modified to apply only for non-finite elements (imperatives and gerunds in SMG and GSG):

[13]The term non-finite is used in a pre-theoretical sense here and does not carry any assumptions that the term usually carries, e.g. imperatives bearing no tense. I will use the term non-finite to refer to imperatives and gerunds.

(7.28) The transition from MMG to SMG: step 2: Enclitic trigger modification

IF $?Ty(t)$
THEN IF $[\downarrow_1^+]?Ty(x)||$
$+NON-FINITE$
THEN **(**$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(Ty(e), Fo(U_x), ?\exists\mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$
ELSE abort
ELSE abort

Given the complementarity of the two embedded IF conditions, the exclusive disjunction is not relevant anymore, since a situation where both a verbal element and a non-finite verbal form have been parsed is logically impossible:

(7.29) The transition from MMG to SMG: step 3: Exclusive disjunction drop

IF $?Ty(t)$
THEN IF $[\downarrow_1^+]?Ty(x)|$
$+NON-FINITE$
($\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(Ty(e), Fo(U), ?\exists\mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$
ELSE abort
ELSE abort

Moving on to the case of CG, things are different, since the respective medieval system of Cypriot is different from MMG, in the sense that the number of proclitic environments are considerably fewer. Remember that in MCG, the unfixed node strategy is not yet generalized as a proclitic trigger, and as such, fronted constituents and subjects do not induce proclisis. However, Wh-elements do trigger proclisis. In that sense, the trigger we need is one specifically referring to the presence of a wh-element. On the contrary,

the situation argument node trigger is active, since subordinating conjunctions and modality/tense/negation markers trigger proclisis almost categorically. No sentence initial clitics are possible. The lexical entry proposed for MCG is shown below:

(7.30) Lexical entry for MCG

IF	$?Ty(t)$	
THEN	IF	$\langle\downarrow_*\rangle ?\exists \mathbf{x}.Tn(\mathbf{x}) \wedge Fo(WH)\vert$
		$\langle\downarrow_0\rangle ?Ty(e_s)\vert\vert$
		$<\downarrow_1^+> Ty(x)$
	THEN	**(**$\mathtt{make}(\langle\downarrow_1\rangle); \mathtt{go}(\langle\downarrow_1\rangle);$**)**
		$\mathtt{make}(\langle\downarrow_1\rangle); \mathtt{go}(\langle\downarrow_1\rangle);$
		$\mathtt{make}(\langle\downarrow_0\rangle); \mathtt{go}(\langle\downarrow_0\rangle);$
		$\mathtt{put}(Ty(e), Fo(U_x), ?\exists \mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$
	ELSE	abort
ELSE	abort	

The transition from MCG to CG involves the emergence of a number of new proclitic environments, fronted objects/subjects, fronted adverbs, fronted PPs/temporal expressions. Abstracting away from the number of these environments, we can assume that the new proclitic environments is actually one, i.e. a fronted element at the left periphery. One of the strategies associated with fronted elements in DS, is the use of unfixed nodes, i.e. the projection of nodes that have not yet been identified structurally in the tree structure. Now, the question is how this rise in proclitic environments took place. Remember that we have used a trigger which encodes the presence of a WH metavariable in order to capture proclisis with WH elements. Remember that WH elements are assumed to be parsed as unfixed nodes, i.e. they use one of the strategies that fronted constituents use. In that sense, I argue that hearers generalized the case-specific trigger $\langle\downarrow_*\rangle ?\exists.xTn(x) \wedge Fo(WH)$ into a generalized trigger encoding just the presence of an unfixed node by getting rid the second part of the conjunction referring to the Wh-feature ($Fo(WH)$):

(7.31) The transition from MCG to CG

MCG				CG		
IF		$?Ty(t)$		IF		$?Ty(t)$
THEN	IF	$\langle\downarrow_*\rangle?\exists\mathbf{x}.Tn(\mathbf{x}) \wedge Fo(WH)\vert$		THEN	IF	$\langle\downarrow_*\rangle?\exists\mathbf{x}.Tn(\mathbf{x})\vert$
		$\langle\downarrow_0\rangle?Ty(e_s)\vert$				$\langle\downarrow_0\rangle?Ty(e_s)\vert$
		$\langle\downarrow_1^+\rangle Ty(x)$				$\langle\downarrow_1^+\rangle Ty(x)$
	THEN	**(**$\mathtt{make}(\langle\downarrow_1\rangle); \mathtt{go}(\langle\downarrow_1\rangle)$**);**			THEN	**(**$\mathtt{make}(\langle\downarrow_1\rangle); \mathtt{go}(\langle\downarrow_1\rangle);$ **)**
		$\mathtt{make}(\langle\downarrow_1\rangle); \mathtt{go}(\langle\downarrow_1\rangle);$				$\mathtt{make}(\langle\downarrow_1\rangle); \mathtt{go}(\langle\downarrow_1\rangle);$
		$\mathtt{make}(\langle\downarrow_0\rangle); \mathtt{go}(\langle\downarrow_0\rangle);$				$\mathtt{make}(\langle\downarrow_0\rangle); \mathtt{go}(\langle\downarrow_0\rangle);$
		$\mathtt{put}(Ty(e), Fo(U_x);$				$\mathtt{put}(Ty(e), Fo(U_x);$
		$?\exists\mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$				$?\exists\mathbf{x}.Fo(\mathbf{x}));$
						$gofirst(?Ty(t))$
ELSE		abort		ELSE		abort
ELSE	abort			ELSE	abort	

Then, the transition from MCG to CG involves proclisis spread via generalization of the unfixed node trigger. It is very tempting to compare the CG system with that of MMG, since their similarities are striking, i.e. CG is almost identical in terms of clitic positioning to MMG. However, proclisis in MMG seems to be more generalized than in CG. The crucial difference seems to be that first position clitics are possible in MMG, while absent in CG. The same goes for proclisis after coordinating conjunctions. The difference between the MCG and the CG clitic system is the number of proclitic environments in these two different stages of the dialect, with CG having developed a number of additional proclitic environments via generalization of the unfixed node trigger, a trigger absent in MCG (see 7.33). Now given that the MMG clitic system developed into a system where proclisis was generalized in all finite environments, while enclisis was kept only for non-finite environments, and given the striking similarities between CG and MMG, it would be tempting to propose that CG too is in a process of change towards a clitic system similar to SMG. This is what analogical reasoning would lead us to expect. However, the account as sketched is not deterministic in the sense that there is nothing in the system that would guarantee such an outcome. There are no concrete numbers of proclitic environments or proclitic environment frequencies that would guarantee a change from the CG to a SMG-like system. In any

case, what is rather interesting is the fact that instances of a shift towards an SMG-like system, i.e. a system that allows first position clitics given that proclisis is generalized across the board in finite environments, have already been reported in the sociolinguistic literature as a result of code-switching between CG and SMG (Tsiplakou, 2009). If what I argue is in any way correct, then unexpected first position proclisis of the kind reported in Tsiplakou (2009) might also have an explanation in terms of a more general tendency of CG towards proclisis. Of course, I do not claim that code-switching is not at play in these cases, but rather that code-switching effects might be further enhanced by a general tendency towards proclisis generalization in CG.

The last transition we need to look at is the transition from MPG to PG. Remember that MPG had the fewer proclitic environments compared to the other two medieval systems. However, this fact did not lead to a simpler lexical entry but to a considerably more complex one. This paradox is easily explained given that no generalization across linguistic elements is present in PG. For example, some subordinating conjunctions are associated with proclisis while others with enclisis or variant positioning. The effect of this situation is the emergence of a lexical entry that has to separately list all the proclitic environments, given that no proclitic generalizations can be made. In that sense, even though the proclitic environments are fewer in MPG, the proclitic triggers will be more than the ones in MMG and MCG, since every element inducing proclisis is listed separately as already shown in 7.23 and repeated below:

(7.32) Lexical entry for MPG

IF		$?Ty(t)$
THEN	IF	$+HINA\|$
		$+AS\|$
		$+\bar{\text{O}}S\|$
		$[+NEG]\|$
		$\langle\downarrow^*\rangle?\langle\uparrow_0\rangle Ty(t)\|\|$
		$\langle\downarrow_1^+\rangle Ty(x)\|$
		$\langle\downarrow^*\rangle Fo(WH)$
	THEN	$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$
		$\textbf{(}\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$
		$\texttt{make}(\langle\downarrow_0\rangle);\textbf{)}\texttt{go}(\langle\downarrow_0\rangle);$
		$\texttt{put}(Ty(e), Fo(U_x)$
		$?\exists\mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$
	ELSE	abort
ELSE		abort

The lexical entry is extremely complex and shows in a very transparent way that it is extremely difficult to find generalizations within such a system, a fact already noted by Pappas (2006). The only generalized trigger is the trigger ensuring enclisis ($\langle\downarrow_1^+\rangle Ty(x)$). Given this, and also the fact that the proclitic environments were considerably fewer than those found in the other two medieval dialects, I will argue that the reason for the transition towards a strict enclitic system in PG is totally dependent on these two previous facts. Generalization of proclisis via generalization of parsing strategies as proclitic triggers (e.g. a generalized unfixed or type e_s node proclitic trigger as in the case of MMG) was not possible given that proclisis did not spread across categories in MPG. This last fact in combination with the complexity of the lexical entry shown above and the fact that the enclitic trigger is the only generalized trigger (spanning across all verbal types), gives us a straightforward explanation of why the PG clitic system turned out to be strictly enclitic. Since proclisis failed to be generalized, or even get spread across linguistic categories or parsing strategies, and given the complexity of the entry presented above, PG clitics followed the opposite route and generalized the most general trigger of the entry shown above, i.e.

the enclitic trigger. In that sense, the transition from MPG to PG involves the drop of all proclitic triggers.[14] The final lexical entry contains just one trigger:

(7.33) The Transition from MPG to PG

MPG

IF $?Ty(t)$
THEN IF $+HINA|$
$+as|$
$+\bar{\text{O}}S|$
$[+NEG]|$
$\langle\downarrow^*\rangle?\langle\uparrow_0\rangle Ty(t)||$
$\langle\downarrow_1^+\rangle Ty(x)|$
$\langle\downarrow^*\rangle Fo(WH)$
THEN **(**$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(Ty(e), Fo(U_x)$
$?\exists\mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$
ELSE abort
ELSE abort

$\mapsto$

PG

IF $?Ty(t)$
THEN IF $\langle\downarrow_1^+\rangle Ty(x)|$
THEN **(**$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$**)**
$\texttt{make}(\langle\downarrow_1\rangle);\texttt{go}(\langle\downarrow_1\rangle);$
$\texttt{make}(\langle\downarrow_0\rangle);\texttt{go}(\langle\downarrow_0\rangle);$
$\texttt{put}(Ty(e), Fo(U_x)$
$?\exists\mathbf{x}.Fo(\mathbf{x}), gofirst(?Ty(t)))$
ELSE abort
ELSE abort

7.3 Conclusions

In this chapter I presented a first sketch of how a diachronic account of the dialects under consideration can be put forth using Dynamic Syntax. I argued that the transition from KG to the respective medieval dialects involves a process of routinization (in the sense of Pickering & Garrod, 2004 and applied to syntax in the sense of Bouzouita, 2008a,b) whereby an earlier pragmatic governed system (the KG clitic system) lost its underlying pragmatic basis and encoded pragmatic preferences as syntactic restrictions. The difference in the positioning systems of the modern forms of the dialects is then explained by looking at the

[14] It should be said that in order to buttress such a claim, more must be said about the complexity of lexical entries and moreover about the potential loss of a number of computational rules, an objection already raised by Marten (p.c). Unfortunately, such an issue cannot be disccussed here for reasons of space.

positioning systems of their respective medieval ancestors. SMG derived from a system which already started to generalize proclisis in medieval times. Proclisis was generalized via parsing strategy generalization, i.e. a given parsing strategy used for parsing a number of elements ended up being a proclitic trigger. However, the fact that different parsing strategies were possible in parsing one and the same element (e.g. parsing fronted elements as unfixed nodes or LINK structures), gave rise to parsing/hearer mismatches, and had the result of further generalizing proclisis to more environments. The proclitic environments ended up being so many that was hard to distinguish proclitic triggers anymore. All the proclitic triggers collapsed in favour of one general proclitic trigger. On the other hand, the fact that most of the elements triggering proclisis were impossible in non-finite contexts had the result of enclisis being retained in non-finite environments. The emerging SMG system kept the old enclitic trigger of MMG by modifying it to apply to non-finite verbal forms only. The situation is different in CG. CG derived from a system, i.e. MCG, in which proclitic environments are fewer than those found in the modern system or the medieval ancestor of SMG (MMG). Looking at MCG, one sees that the unfixed node strategy is not yet generalized. In that sense, the unfixed node strategy is specified to apply only for WH elements. Then, the transition to MCG involves generalization of the unfixed strategy to apply to all elements that are parsed as unfixed nodes (subjects, objects, PPs, temporal expressions, adverbs). Lastly, PG derives from a system were proclitic environments are far fewer than the ones found in MCG or MMG. In MPG no general parsing trigger as a proclitic trigger is found, since the elements inducing proclisis cannot be generalized into one class. Hence, each of the elements must be encoded separately in the lexical entry for the clitic. The fact that proclitic environments were few in conjunction with the complexity of the lexical entry and the fact that only the enclitic trigger had a generalized flavor, gave rise to a system where all proclitic triggers were dropped and only the enclitic trigger was retained in PG.

Chapter 8

Conclusions

8.1 Conclusion Summary

8.1.1 General Notes as Regards Clitics

Clitics under the account proposed, and all DS accounts on clitics (Cann et al., 2005; Bouzouita, 2003, 2008a,b; Chatzikyriakidis, 2006, 2009a,b,c; Gregoromichelaki, 2010), are treated as ordinary pronouns in that they are assumed to project a type value and a formula metavariable that needs to be updated from the context. However, what is exceptional as regards clitics is that these create their own structure, and, in so doing, they have become more complex than just being a pronoun occurring in a certain position. Their triggering point is specified to be a different type requiring node than the type they induce. This is totally different to what ordinary pronouns are assumed to do in DS, since pronouns, similarly to regular NPs, have as their trigger a type requiring node that is the type requirement of the actual type they project (*?Ty(e*)). Thus, the lexical entry for a pronoun has as its triggering point the node in which type and formula information is going to be provided (*?Ty(e)*). However, in the case of clitics, this triggering point is a type *t* requiring node. Starting from this node, clitics create their own structure in which they project type and formula information (by building a fixed node or an unfixed one). This behaviour of clitics is independently needed given the locality restrictions associated with clitics. Given these restrictions, a type *e* requiring trigger for clitics will predict that locality violations in the

case of clitics should be fine contrary to fact (given parsing of the clitic on a node introduced by *ADJUNCTION). This behavior, the fact that clitics induce their own structure, is one of the differences between clitics and pronouns and follows the natural observation as regards positioning of pronouns on the one hand and clitics on the other, the former being freer compared to the latter.

8.1.2 Clitic Positioning and Ordering

Clitic positioning in the dialects under consideration was discussed in this thesis under both a synchronic and a diachronic perspective. Given the synchronic perspective, positioning restrictions were argued to derive from a number of triggers found in the entries of clitics in each case. In the case of SMG and GSG, two such triggers were posited, one accounting for proclisis and one for enclisis. As regards CG, it was argued that the complexity of the positioning system with respect to the number of elements triggering proclisis can be reduced to two parsing triggers. These two parsing triggers are two generalized parsing strategies that are used in order for the elements inducing proclisis to be parsed. In that sense, the complexity of the CG system can be reduced to the encoding of two parsing strategies as proclitic triggers in the entry for clitics. These two strategies are the use of an unfixed node and the projection of the situation argument node and its decoration by a type e_s requirement. Fronted constituents and Wh elements, following standard DS assumptions are parsed (or can be parsed in the case of fronted constituents) on an unfixed node, while subordinating conjunctions and all sort of modality/tense/markers introduce a type e_s requiring node. Thus, the complexity of the CG system is captured using two lexical triggers plus a third trigger used for the regular enclitic cases. This third trigger encodes the presence of a verbal type and is activated only in case none of the other proclitic triggers are operative. On the other hand, the PG positioning system was neatly captured assuming a generalized enclitic trigger that allows the clitic to get parsed in case a verbal type of any sort is present in the tree. The development of these triggers and the reason for the differences in clitic positioning found in the dialects under consideration receive a straightforward explanation under the diachronic account sketched in chapter 8. This diachronic account stems from KG and adopts, as its point of departure, the challenge that

the KG clitic system apparently resists any formalization within one grammatical system (Pappas, 2006). Pappas argues indeed that it requires a competing grammars account. In DS terms, the observed facts find a straightforward formalization by making the at first sight rather extravagant claim that the lexical entry for KG clitics does not involve any positioning restrictions over and above that of the initial type t requiring trigger. This very simple lexical entry was shown not to be extravagant at all since such an entry will give us the KG clitic positioning facts. The reason for this simplicity is the assumption that the positioning restrictions of KG clitics are not syntactic restrictions but rather pragmatic preferences. It is the routinization (in the sense of Pickering & Garrod, 2004) of these pragmatic preferences, with subsequent loss of an underlying pragmatic basis (Bouzouita, 2008a,b) that led to the encoding of these preferences as syntactic restrictions in the respective medieval systems. This encoding of syntactic restrictions is not uniform across dialects, and therefore different systems arose in the medieval dialects. MMG developed a system where two parsing strategies were encoded as triggers for parsing the clitic, the unfixed node strategy and the type e_s requirement strategy. These two strategies, combined with a generalized enclitic trigger will roughly give us the system of MMG.[1] On the other hand, the MCG system was not able to generalize both these parsing strategies but used the unfixed node trigger specifically applied to Wh elements. Then, the transition from MCG to CG involves the generalization of the unfixed node strategy to all elements that can be parsed on unfixed nodes. The interesting thing about such an account is that though the CG system looks more complex than the MCG one, in the sense that the number of environments triggering proclisis is more in CG, the lexical entry for clitics in CG is simpler than that of MCG. The reason for this is that CG makes use of two generalized parsing strategies as proclitic triggers whereas MCG makes use of one generalized parsing strategy, plus a second strategy which is however case specific (applies only to Wh elements). This case specific parsing strategy must be further specified as to apply only for Wh elements. Such an account shows that the CG is only superficially more complex than MCG, given that the entry for CG clitics is less complex than the respective MCG one. The transition from

[1] I say roughly because a number of idiosyncrasies of the MMG system like the fact that negation *ou* is associated with enclisis contrary to what expected were not dealt with in this thesis, since the diachronic account proposed was a first attempt to provide such an account, and thus cannot be seen as full diachronic account of the development of the clitic systems in the dialects under consideration.

MCG to CG involves the removal of one case specific trigger and its replacement by a more general trigger. A strong vindication of the account just sketched is the development of the PG clitic system via MPG. Remember, that MPG had fewer proclitic triggers than either MMG or MCG. Again, this did not lead to a simpler lexical entry but to a considerably more complex entry, given that no general triggers capturing a group of elements could be given. In that respect, every proclitic environment in MPG had to be separately listed, leading to a highly disjunctive lexical entry. The fact that the only trigger which was not case specific was the enclitic trigger, and given the highly disjunctive nature of the entry for MPG led to the drop of all proclitic triggers and the generalization of enclisis in all cases.

Diachronic change under this account is driven by both routinization and the development of generalized parsing triggers. These parsing triggers can be thought of as parsing shortcuts, facilitators of the parsing process and hence, lexical entries increasingly seeking to minimize the number of them (which was the case for MPG for example). These two factors can give us a straightforward explanation of why the dialects under consideration followed different routes and ended up with these different clitic positioning systems (in the way just sketched). The different positioning systems of the modern dialects are due to their different medieval counterparts. MMG involved a system where proclisis was generalized to elements parsed via two parsing strategies, the unfixed node and the type e_s strategy. However, MMG shows signs of unexpected proclitic placement, e.g. first position clitics or proclisis with subordinating conjunctions. This fact points towards a system which was already starting to move towards the SMG system. Such a transition can be seen as an increase in the number of proclitic environments due to speaker/hearer mismatches. For example, given that fronted constituents can be also parsed as LINK structures, a possible parsing of a proclitic case involving an unfixed node would be by analyzing the unfixed node as a LINK. This misanalysis made proclisis available with all the elements that could be parsed as LINK structures, i.e. subordinating conjunctions, subjects/objects/adverbs of any sort. Then, the environments triggering proclisis were so many that were not discernible anymore. Thus, a generalized proclitic trigger (basically the trigger SMG has) emerged. The CG clitic system is also the result of proclisis generalization, but to a lesser extent. The difference between MCG and CG is the removal of a case specific trigger in

favor of a more general one. Then, this more general trigger increased the number of elements triggering proclisis, since all the elements that were using the parsing strategy in question in order to get parsed were now associated with proclisis, which is no different than what happened to MMG. The fact than no general proclitic trigger was present in MPG, led in PG following the opposite route, i.e. that of enclisis generalization. The drop of all the case specific proclitic triggers gave rise to the simple (in terms of positioning) clitic system of modern day PG.

8.1.3 Clitic Climbing and Restructuring

One other important issue dealt with in this thesis was CC. CC is only operative in GSG and only for a restricted number of verbs of the so-called restructuring class. The account proposed in this thesis brings climbing inducing verbs on a par with auxiliary verbs and argues that the reason for CC should be attributed to the same reason clitics appear before the auxiliary and not before or after the participle in languages like SMG. The account assumes that climbing inducing verbs do not project any verbal type: rather, they project their semantics inside the complex situation argument node in the same sense auxiliaries are assumed to do in Cann (forthcoming). The account proposed is a straightforward formalization of the functional status assumed for restructuring verbs by Cinque (2001, 2006) and Cardinaletti & Shlonsky (2004). The account is monoclausal, since no verbal type is projected by the restructuring verb and the infinitive provides the type value within the same domain the climbing inducing verb has been parsed. Another consequence of such an account, is that CC constructions are not control structures anymore, since the subject of both the infinitive and the restructuring verb are identical under the proposed account (with the auxiliary commonly inducing it as a skeletal structure). Such a prediction is to be welcomed, since as Cinque (2006) argues, CC constructions are not control constructions. Such an account is directly extendable to optional CC languages where more complicated restructuring phenomena are attested, e.g. unavailability of negating an infinitive when restructuring has taken place or the impossibility of using the same adverb twice in a CC environment:

(8.1)

a. *Voglio* *non* *vederlo*
want.1SG NEG see.INF-IT
'I want not to see it.'

b. **Lo* *voglio* *non* *vedere*
it.CL want.1SG NEG see.INF
'I don't want to see it.'

(8.2)

a. *Maria* *vorrebbe* *già* *averlo* *già* *lasciato*
Maria would already have.INF-IT already left
'Maria would already want to have already left him.' [Cinque, 2006]

b. **Maria* *lo* *vorrebbe* *già* *aver* *già* *lasciato*
Maria it.CL would already have.INF already left
'Maria would already want to have already left him.' [Cinque, 2006]

The account proposed straightforwardly captures these two facts, as already argued in Chatzikyriakidis (2009c, forthcoming). In particular, assuming a trigger for negation that aborts in case any functor nodes exist (a trigger independently needed to capture the fact that negation always precedes verbs of any type in languages like Italian), it is immediately predicted that negation will never follow a climbing inducing verb in case restructuring has taken place. Then, unavailability of negation in CC environments boils down to a simple word order phenomenon. The lexical entry for negation will have to have a trigger that will abort in case any type of verb has been parsed (including auxiliaries). Given the treatment of restructuring verbs as auxiliary-like, the same entry will predict that negation cannot follow a restructuring verb. Structures where the same adverb is repeated twice are also straightforwardly accounted, since whatever the analysis one presents for adverbs, such examples will be ruled out using the same reasoning one would use to rule out cases where two adverbs appear within a simple clause like the one shown below:

(8.3) **Maria* *già* *ha* *già* *comprato* *una* *macchina*
Maria already has already bought a car
'Maria has already bought a car.' [Italian]

Given that the account proposed assumes monoclausality, cases where an adverb is repeated twice should be attributed to a ban on repeating the same adverb within a single clause, and as such are unproblematic.

8.1.4 The PCC and Person Restrictions in General

One further important outcome of this thesis is the vindication of the PCC account proposed by Cann & Kempson (2008) and Chatzikyriakidis & Kempson (2009). As we have already seen, Cann & Kempson (2008) and Chatzikyriakidis & Kempson (2009) claimed that the PCC is due to a hard-wired constraint on tree-growth. In this thesis, strong evidence from PG that such an account is on the correct track was given. The problematic PG data, according to which no 3rd person clitic clusters are possible, receive a straightforward explanation once we assume that PG clitics project locally unfixed nodes, given their syncretic nature. Such an account then predicts clusters of 3rd person clitics not to be possible in PG and further vindicates a feature-free account within the lines of Cann & Kempson (2008) and Chatzikyriakidis & Kempson (2009). This is a particularly robust confirmation, as there is no natural explanation of these phenomena from minimalist assumptions (see section 6.1). Furthermore, the licit PG clusters comprised of a 1st/2nd person plus a 3rd person clitic also receive a straightforward explanation given the idiosyncratic properties of the 1st/2nd person clitic forms used in these cases. These forms, *m/s*, cannot appear on their own but only in clusters and when appearing in clusters, they are always associated with the indirect object node. Thus, these clitics are not underspecified but rather receive a fixed position in the tree. Therefore, they do not project locally unfixed nodes but rather fixed structure, in which case the "no more than one unfixed node at a time" constraint is not active and clusters involving a 1st/2nd person plus a 3rd person clitic are predicted to be licit in PG. The same reasoning applies to weak PCC constructions in PG (possibly in RP as well). It is worth noting, that in RP no clitic clusters are reported to be possible (Michelioudakis & Sitaridou, 2010), which is exactly what our account would predict given that RP is also syncretic across the board. Ethical datives in SMG were also discussed and were shown to behave more like arguments rather than adjuncts in SMG. Following Michelioudakis (2007), I have argued that ethical datives in SMG should be seen as

optional arguments. Such an assumption will predict that ethical datives are not immune to the PCC in SMG, since these are subject to the "no more than one unfixed node at a time constraint", according to fact. I have shown that the ethical dative doubling data Michelioudakis (2007) provides in order to prove that ethical datives behave more like arguments are indeed correct and have furthermore argued that ethical dative doubling is even freer than what Michelioudakis assumes.

8.1.5 Clitic Doubling

An account of doubling in SMG, CG and PG was also given. Following Gregoromichelaki (2010) it was argued that the differences between CLLD and CD structures as regards the range of NPs that can participate in these constructions is straightforwardly accounted for given the dynamics of the system. In that respect, the fact that CD is more restricted than CLLD as regards the classes of NPs that can participate in CLLD but not in CD, derives from the very basic fact that in CLLD, a referent is already introduced (by the doubled NP) when the clitic comes into parse and as such update of the clitic's metavariable can be done easily by MERGE of the doubled NP node with the clitic node, whereas in CD no referent has been introduced yet when the clitic comes into parse. Thus, in the latter case, update of the metavariable can be done purely by the context and as such the substituent value must be salient enough in the discourse. This very basic fact of word order which the left-to-right dynamic nature of DS is able to reflect directly, will give us the more restrictive distribution of CD compared to CLLD. No hard-wired constraints based on a referentiality scale in the sense of Anagnostopoulou & Giannakidou (1995) need to be asssumed, since, as shown in this thesis, these are not hard-wired constraints but rather derive from the anaphoricity effect just mentioned. A number of examples were provided that show that a number of noun classes that in the literature are reported not to be able to participate in CD construnctions, can actually participate in these constructions (see also Appendix B for more examples). CLLD in CG was argued to involve a LINK structure rather than an unfixed node. Such a move is based on the fact that CG does not use OV focus constructions and as such was argued that preposed objects cannot be parsed on an unfixed node. But, if objects cannot be parsed on unfixed nodes, then CLLD cannot be analysed in CG in

terms of the contruction of an unfixed node. This account is further confirmed by the fact that enclisis rather than proclisis is the case in CLLD constructions. Assuming a lexical entry for clitics in the style given in chapter 4, such a distribution is unexpected given that the presence of an unfixed node constitutes one of the proclitic triggers in CG. On this account, this is seen as specialisation of the process of unfixed node construction to a lexically induced trigger. With this shift, a modification of the HTLD rule was proposed that allowed the NP parsed on a LINK structure to participate in the scopal statement of the LINKed tree, while furthermore changing the modality of the copy requirement from D to *. This last move, captures the fact that CLLD is sensitive to relative clause and adverb islands but however not sensitive to complement clause islands.

8.2 Further Research

8.2.1 Scope over Coordination

A number of issues as regards dialect specific or general clitic related phenomena were not dealt with in this thesis. The first of these concerns a general characteristic of the behaviour of clitics, namely the fact that clitics cannot take wide scope over coordination. Thus, in a coordinating construction, the clitic must be repeated with each verb. The examples below are from SMG:

(8.4) *To* *ekopsa* *ke* *to* *maγirepsa*
it.CL-ACC cut and it.CL-ACC cooked
'I cut it and cooked it.'

(8.5) **Ekopsa* *ke* *to* *maγirepsa*
cut and it.CL-ACC cooked
'I cut it and cooked it.'

Similar behaviour of clitics is exhibited in CG and GSG and is quite common among Romance languages (see Monachesi , 1993; Miller & Sag, 1997). However, PG clitics are an exception to this, since clitics can in fact take scope over coordination, showing that what has been taken to be criterial of clitic pronouns is in fact subject to variation:[2]

[2]Similar behaviour is found in European Portuguese, see Miller & Monachesi, 2003.

(8.6) *Ekops* *ke* *emaiγreps* *ato*
cut.1SG CONJ cooked.1SG it.CL
'I cut it and cooked it.'

(8.7) *Ekops* *ato* *ke* *emaiγreps* *ato*
cut.1SG it.CL CONJ cooked.1SG it.CL
'I cut it and cooked it.'

In order to provide an account of the above facts, one must first look at the behaviour of the languages in ellipsis constructions, since structures like the above are elliptical structures (8.6). Such a detailed study cannot however be done in this thesis, so this particular distribution is left for further research.

8.2.2 Sequences of Two Dative Clitics

The last outstanding issue that is left as a subject of further research, concerns sequences of two dative clitics in a limited number of constructions in SMG. In these constructions, one of the clitics is obligatorily interpreted as an ethical dative. These structures have been already been reported by Michelioudakis (2007) and were discussed in chapter 6. Two examples of such constructions are given below:

(8.8)

a. *%Na* *mu* *tis* *δosis* *pola* *filia*
SUBJ me.CL-GEN her.CL-GEN give.2SG many kisses
'Give her many kisses, on behalf of me.'

b. *%Na* *min* *mu* *tis* *aγorasis* *tipota*
SUBJ NEG me.CL-GEN her.CL-GEN buy.2SG nothing
'Do not buy her anything (the ethical dative is reinforcing the command in that example).'

As already noted, the account proposed will rule out examples like the ones above, since assuming that dative clitics will project locally unfixed nodes, a sequence of dative clitics will always be ruled out given the "one unfixed node at a time" constraint. However, as already said, structures like the ones shown above are surprisingly sensitive to the form

of the 3rd person dative clitic used. In that sense, substituting the dative feminine clitic *tis* 'her.CL-GEN' with the masculine equivalent *tu* 'him.CL-GEN', considerably reduces acceptability:

(8.9)

a. **Na mu tu δosis pola filia*
SUBJ me.CL-GEN him.CL-GEN give.2SG many kisses
'Give him many kisses, on behalf of me.'

b. *???Na min mu tu aγorasis tipota*
SUBJ NEG me.CL-GEN him.CL-GEN buy.2SG nothing
'Do not buy him anything (I strictly order you).'

These data are indeed intriguing and seem to resist any principled analysis, since it is very hard to imagine a principled explanation that will allow feminine dative clitics but not masculine ones in these limited set of contexts only. More data are needed in order to elucidate what exactly is going on in these constructions. But whatever the outcome, it is clear that this will have to be analysed as some form of routinised lexically triggered macro of actions. This issue though is left open for further research.

Bibliography

Adger, D., Harbour, D., 2007. Syntax and Syncretisms of the Person Case Constraint. *Syntax 10*, 2-37.

Agouraki, G., 1993. *Spec-Head Licensing: The Scope of the Theory*. Phd Thesis, University College London.

Agouraki, Y., 1997. On the Enclisis-Proclisis Alternation. In: Drachman, G., Malikouti-Drachman, A., Fykias, J., & Klidi, C., (eds), *Proceedings of the 2nd International Conference on Greek Linguistics*. Graz, W. Neugebauer Verlag, 393-404.

Agouraki, Y., 2001. The position of clitics in Cypriot Greek. In: Ralli, A., Joseph, B., & Janse, M., (eds), *Proceedings of the First International Conference on Greek Linguistics*, 1-18.

Agouraki, Y., 1997. On the Enclisis-Proclisis Alternation. In: Drachman, G., Malikouti-Drachman, A., Fykias J., & Klidi, C., *Proceedings of the 2nd International Conference on Greek Linguistics*. Graz, W. Neugebauer Verlag, 393-404.

Albizu, P. 1997. Generalized Person-Case Constraint: A Case for a Syntax-Driven Inflectional Morphology. In: Mendikoetxea, A. & Uribe-Etxebarria, M., (eds), *Theoretical Issues at the Morphology-Syntax Interface*. Donostia, ASJU, 1-33.

Alexiadou, A., & Anagnostopoulou, E., 1998. Parametrizing AGR: Word order, V Movement and EPP-Checking. *Natural Language and Linguistic Theory 16*, 491-539.

Alexopoulou, T., 1999. *The Syntax of Discourse Functions in Greek: A Non-Configurational Approach*. PhD thesis, University of Edinburgh.

Alexopoulou, D., & Kolliakou, D., 2002. On Linkhood, Topicalization and Clitic Left Dislocation. *Journal of Linguistics*38, 193-245.

Anagnostopoulou, E., 1994. *Clitic Dependencies in Greek*. Doctoral dissertation, University of Salzburg.

Anagnostopoulou, E., 1997. Clitic Left Dislocation and Contrastive Left Dislocation. In: Anagnostopoulou, E., Van Riemsdjik, H., & Zwarts, F. (eds), *Materials on Left Dislocation* (Linguistik Aktuell 14). Amsterdam, John Benjamins, 151-192.

Anagnostopoulou, E., 1999. Conditions on Clitic Doubling in Greek. In: Riemsdijk, H. (ed.), *Clitics in the languages of Europe*. Berlin, New York, Mouton De Gruyter, 761-798 .

Anagnostopoulou, E., 2001. Two Classes of Double Object Verbs: The Role of Zero Morphology. In: van Oostendorp, M. & Anagnostopoulou, E. (eds.), *Progress in Grammar: 20 Years of Grammaticamodellen.* Meertens Institute, electronic book.

Anagnostopoulou, E., 2003. *The Syntax of Ditransitives: Evidence from Clitics*. Mouton de Gruyter.

Anagnostopoulou, E., 2005. Strong and Weak Person Restrictions: A Feature Checking Analysis. In: L., Heggie & F., Ordóñez (eds), *Clitics and Affixation*. Amsterdam, John Benjamins, 199-235.

Anagnostopoulou, E. 2006. Clitic doubling. In: Everaert, M., & van Riemsdijk, H., (eds.), *The Blackwell Companion to Syntax, vol. 1*. Oxford, Blackwell, 519-581.

Anagnostopoulou, E., & Giannakidou, A., 1995. Clitics and prominence, or why specificity is not enough. *Chicago Linguistic Society 31: Parasession on Clitics*, 1-14.

Androulakis, A., 2001. Clitics and doubling in Greek. In: Georgiafentis, M., Kerswill, M.P., & Varlokosta S., (eds.), *Reading Working Papers in Linguistics* 5, 85-111.

Bejar, S. & Rezac, M. 2003. Person Licensing and the Derivation of PCC Effects. In: Perez-Leroux, A., & Roberge, Y. ed., *Romance Linguistics: Theory and Acquisition*. Amsterdam, John Benjamins.

Blackburn, P. & Meyer-Viol, W., 1994. Linguistics, Logic and Finite Trees. *Bulletin of Interest Group of Pure and Applied Logics* 2, 2-39.

Bianchi, V., 2006. *On the Syntax of Personal Arguments*. Lingua 116.

Bonet, E., 1991. *Morphology after Syntax: Pronominal Clitics in Romance*. Doctoral dissertation, MIT.

Bonet, E., 1994. The Person Case Constraint: A Morphological Approach. In: Harley, H., & Philips, C., (eds), *The morphology-Syntax Connection*. Cambridge, MA, MIT press, 33-52.

Bonet, E., 2007. The Person Case Constraint and Repair Strategies. To appear in: Alessasndro, R., Fischer, S., & Hrafnbjanson, G. H. (eds), *Person Restrictions*. Mouton de Gruyter.

Bošković, Ž., 1995. Participle Movement and Second Position Cliticization in Serbo-Croatian. *Lingua 96*, 245-266.

Bošković, Ž., 2001. *On the Nature of the Syntax-Phonology Interface: Cliticization and Related Phenomena*. Elsevier.

Bouzouita, M., 2002. *Clitic Placement in Old and Modern Spanish: A Dynamic Account.* Msc thesis, Kings college, London.

Bouzouita, M., 2008a. *The Diachronic Development of Spanish Clitic Placement*. Phd Thesis, King's College, London.

Bouzouita, M., 2008b. At the Syntax-Pragmatics Interface: Clitics in the History of Spanish. In: Cooper, R., Kempson, R. (eds), *Language Evolution and Change*. Cambridge, Cambridge University Press.

Bouzouita, M., & Chatzikyriakidis, S., 2009. Clitics as Calcified Processing Strategies. *Proceedings of the 14th International LFG conference*, Cambridge, July 2009. CSLI online publications, 188-207.

Branigan, H.P., Pickering, M.J., & McLean, J.F., 2005. Priming Prepositional-Phrase Attachment During Comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 31, 468-481.

Bresnan, B., 2001. *Lexical-Functional Syntax.* Oxford, Blackwell.

Cann, R., 1994. *Formal Semantics: An Introduction*. Cambridge, Cambridge university press.

Cann, R., forthcoming. Towards an Account of the English Auxiliary System: Building Interpretations Incrementally. In: Kempson R., Gregoromichelaki, E., & Howes C., (eds), *The Dynamics of Lexical Interfaces*. Stanford, CSLI publications.

Cann, R., & Kempson, R., 2008. Production Pressures, Syntactic Change and the Emergence of Clitic Pronouns. In: Cooper, R. & Kempson, R., (eds), *Language in Flux: Dialogue Coordination, Language Variation, Change and Evolution*, London, College Publications, 221-263.

Cann, R., Kempson, R., & Marten, L., 2005. *The Dynamics of Language*. Oxford, Elsevier.

Cann, R., Kempson, R. & Purver, M., 2007. Context and Well-formedness: the Dynamics of Ellipsis. *Research in Language and Computation* 5.

Cardinaletti, A., & Starke M., 1999. The typology of structural deficiency: A case study of the three classes of pronouns. In: Riemsdijk, H., (ed.), *Clitics in the Languages of Europe. Empirical Approaches to Language Typology*, 145-233.

Cardinaletti, A., & Shlonsky, U., 2004. Clitic Positions and Restructuring in Italian. *Linguistic Inquiry* 35, 519-557.

Carpenter, R., 1991. The generative power of CG and HPSG with lexical rules. *Computational linguistics*, 17: 3, 301-314.

Cecchetto, Carlo. 1999. A Comparative Analysis of Left and Right Dislocation in Romance. *Studia Linguistica 1* 53, 40-67.

Chatzikyriakidis, S., 2006. *Clitics in Modern Greek: A Dynamic Account*. MSc thesis, King's college London.

Chatzikyriakidis, S., 2009a. Clitics in Grecia Salentina Greek: A Dynamic Account. *Lingua* 119. Elsevier, 1939-1968.

Chatzikyriakidis, S., 2009b. The Person Case Constraint in Modern Greek: A Unified Dynamic Syntax Account. *Newcastle Working Papers in Linguistics* 15, 40-62.

Chatzikyriakidis, S., 2009c. *A Dynamic Account of Clitic Climbing*. Ms, King's College, London.

Chatzikyriakidis, S., forthcoming. A Dynamic Account of Clitic Climbing: A First Sketch. In: Kempson R., Gregoromichelaki, E. (eds), *The Dynamics of Lexical Interfaces*. Stanford, CSLI publications.

Chatzisavvidis, S., 1995. Ta pontiaka ston elladiko horo. *Arhion pontou, tomos 46*: I pontiaki dialektos allote kai tora.

Chierchia, G., 1995. *Dynamics of Meaning: Anaphora, Presupposition, and the Theory of Grammar.* The University of Chicago Press.

Chomsky, N., 1986. Barriers. *Linguistic Inquiry Monograph* 13, Cambridge, MA, MIT Press.

Chomsky, N., 1995. *The minimalist Program.* Cambridge, MA, MIT Press.

Cinque, G., 1977. The Movement Nature of Left Dislocation. *Linguistic Inquiry 8*: 397-412.

Cinque, G., 1990. *Types of A' Dependencies*. Cambridge, MA, MIT press.

Cinque, G., 1999. *Adverbs and Functional heads: A Cross- Linguistic Perspective*. New York, Oxford University press.

Cinque, G., 2001. Restructuring and the Order of Aspectual and Root Modal Heads. In: Cinque, G. and Salvi, G. (eds.), *Current studies in Italian Syntax.* Elsevier, 137-155.

Cinque, G., 2006. *Restructuring and Functional Heads*. *Oxford Studies in Comparative Syntax*, Oxford university press.

Condoravdi, C., & Kiparksy, P., 2002. Clitics and Clause Structure. *Journal of Greek Linguistics* 2, 1-39.

Contossopoulos, N.G., 1994. *Dialects and Idioms of Modern Greek.* Athens: Academy of Athens.

Contossopoulos, N., 1997. *Themata Kritikis Dialektologias* [Themes of Cretan Dialectology]. Athens: Anastasakis.

Cuervo, M.C, 2002. *Spanish Clitics: Three of a Perfect Pair.* Ms, MIT.

Dawkins, R.M., 1916. *Modern Greek in Asia Minor. A study of the dialects of Silly, Cappadocia and Pharasa.* Cambridge, Cambridge University Press.

Dowty, D.R., 1981. Introduction to Montague Semantics. *Studies in Linguistics and Philosophy*, Springer publications.

Fodor, J.A., 1975. *Language of Thought*. Cambridge: Harvard University Press.

Fodor, J.A, 1981. *Representations: Philosophical Essays on the Foundations of Cognitive Science.* Cambridge, MA, MIT Press.

Fodor, J. (1983). *The Modularity of Mind*. Cambridge, MA, MIT Press.

Fostiropoulou, D., 1938. Paramithia Imeras [Stories from Imera]. *Pontus Archive 8*, 181-202.

Franco, J., & Huidobro, S., 2008. Ethical Datives, Clitic Doubling and the Theory of pro. In: Bruhn de Garavito, J., & Valenzuela, E., (eds). *Selected Proceedings of the 10th Hispanic Linguistics Symposium*, Somerville, MA, Cascadilla Proceedings Project, 215-224.

Franks, S., & King., T., H., 2000. *A Handbook of Slavic Clitics*. Oxford University Press.

Garrod, S. & Doherty, G., 1994. Conversation, Co-ordination and Convention: an Empirical Investigation of How Groups Establish Linguistic Conventions. *Cognition* 53, 181-215.

Garrett, A., 1996. Wackernagel's Law and Unaccusativity in Hittite. In: Halpern, A.L, & Zwicky, A.M., *Approaching Second: Second Position Clitics and Related Phenomena*. Stanford, CSLI, 85-133.

Garrod, S. & Pickering, M., 2004. Why is conversation so easy? *Trends in Cognitive Sciences* 8, 811.

Giannakidou, A., 2007. *A temporal semantics for the subjunctive.* Ms., University of Chicago.

Ginzburg, J. & Sag, I.A., 2000. *Interrogative investigations: The form, meaning and use of English interrogatives.* Stanford, CA: CSLI.

Gregoromichelaki, E., 2006. *Conditionals in Dynamic Syntax*. Phd thesis, King's College, London.

Gregoromichelaki, E., 2010. A Dynamic Perspective on Left-Right Asymmetries: CLLD and Clitic Doubling in Greek. In: *Rightward Movement in a Comparative Perspective: Selected Papers from the DGfS-AG Rightward Movement Workshop*. Amsterdam, John Benjamins.

Giannakidou, A., 1995. Subjunctive, Habituality and Negative Polarity items. *Proceedings of SALT IV*: 132-150.

Grimshaw, J., 2001. Optimal Clitic Positions and the Lexicon in Romance Clitic Systems. In: Legendre, G., Grimshaw, J., & Vikner, S., (eds.), *Optimality-Theoretic Syntax.* Cambridge, MA, MIT Press, 205-240.

Haspelmath, M., 2004. *Explaining the Ditransitive Person-Role Constraint*. Electronic form article. Available at: http://email.eva.mpg.de/ haspelmt/2004dpr.pdf.

Heap, D., 2005. Constraining optimality: Clitic sequences and feature geometry. In: Heggie, L., & Ordonez, F., (eds.), *Clitic and Affix Combinations*. Amsterdam, John Benjamins, 81-102.

Heim, I., 1990. E-Type Pronouns and Donkey-Anaphora. *Linguistics and Philosophy* 13, 137-178.

Heim, I., & Kratzer, A., 1997. *Semantics in Generative Grammar*. Oxford, Blackwell.

Hilbert, D. & Bernays, P., 1939. *Die Grundlagen der Mathematik II*. Second edition. Reprint 1970. Berlin, Heidelberg, New York, Springer.

van Hoecke, W., 1996. The Latin dative. In: van Belle, W., & Langendonck, van, W. (eds), *The Dative: Volume I Descriptive Studies*. Amsterdam, John Benjamins, 3-38.

Horrocks, G., 1997. *Greek: A History of the Language and its Speakers*. London: Longman.

Iatridou, S., 1995. Clitics and Island Effects. In: Bhatt, R., Garrett, S., Han, C.H., & Izvorsky. R. (eds), *Penn working papers in Linguistics* 2, 11-31.

Iatridou, S., 1996. To Have and Have Not. *Proceedings of WCCFL XIV.*

Janse., M., 1993. La Position des Pronoms Personnels Enclitiques en Grec Néo-Testamentaire à la Lumière des Dialectes NéoHellénique. In: Claude, Brixhe (ed.), *La Koiné Grecque Antique. I. Une Langue Introuvable.* Nancy, Presses Universitaires de Nancy, 83-121 (Etudes anciennes, 10).

Janse, M., 1994. Son of Wackernagel. The Distribution of Object Clitic Pronouns in Cappadocian. In: Philippaki-Warburton, I., Nicolaidis, K., & Sifianou, M., (eds.), *Themes in Greek Linguistics. Papers from the First International Conference on Greek Linguistics*. Reading, September 1993. Amsterdam, John Benjamins, 435-442.

Janse, M., 1998. Grammaticalization and Typological Change. The Clitic Cline in Inner Asia Minor Greek. In: Janse, M., (ed.), *Productivity and Creativity. Studies in General and Descriptive Linguistics in Honor of E.M. Uhlenbeck.* Berlin: Mouton de Gruyter, 1998, 521-547 (Trends in Linguistics. Studies and Monographs, 116).

Janse, M., 2000. Convergence and Divergence in the Development of the Greek and Latin Clitic Pronouns. In: Sornicola, R., Poppe, E., & Shisha H.A, (eds.), *Stability, Variation and Change of Word-Order Patterns over Time*. Amsterdam, John Benjamins, 2000, 231-258 (Current Issues in Linguistic Theory, 213).

Janse, M., 2006. I kappadokiki dialektos. In: Tzitzilis, Ch. (eds), *Neoellinikes dialektoi*. Instituto Neoellinikon spoudon (Idrima Manoli Triantafillidi). In press.

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*. Orlando, Florida Academic Press, 15-42.

Jannaris, A.N., 1968. *Historical Greek Grammar*. Georg Olms Verlagsbuchhandlung: Hildesheim .

Joseph, B., 1983. The Synchrony and Diachrony of the Balkan Infinitive. A Study in Areal, General, and Historical Linguistics. *Cambridge Studies in Linguistics, Supplementary Volume.* Cambridge, Cambridge University Press.

Joseph, B., 1988. Pronominal affixes in Modern Greek: The case against clisis. In: Brentari D., et al., (eds) *Papers from the 24th regional meeting, Chicago Linguistic Society.* Chicago, CLS, 203-215.

Joseph, B., 1989. The Benefits of Morphological Classification: On Some Apparently Problematic Clitics in Modern Greek. In: Dressler, W., Luschutzky, H., Pfeiffer, O., & Rennison, J., (eds.), *Contemporary Morphology*. Berlin, Mouton de Gruyter, 171-181.

Joseph, B., 2002. Defining word in Modern Greek: A response to Philippaki-Warburton & Spyropoulos 1999. *Yearbook of morphology 2002*.

Jouitteau, M., & Rezac, M., 2008. The French Ethical Dative: 13 Syntactic Tests. *Bucharest Working Papers in Linguistics*, IX (1): 97-108.

Kallulli, D., 2000. Direct object clitic doubling in Albanian and Greek. In: Beukema, F., & Den Dikken. M., (eds.), *Clitic Phenomena in European Languages*. Amsterdam, John Benjamins, 209-248.

Karanastasis, A., 1997. *Gramatiki ton Elinikon Idiomaton tis Kato Italias (Grammar of the Greek Dialects of Southern Italy)*. Academy of Athens.

Katsoyannou, M., 1995. *Le Parler Greco de Galliciano (Italie): Description d'une Langue en Voie de Disparition*. PhD thesis, University of Paris VII.

Kayne, R.S, 1975. *French syntax : The Transformational Cycle* . Cambridge, MA, MIT Press.

Kayne, R.S., 1994. *The Antisymmetry of Syntax*. Linguistic Inquiry Monograph 25. Cambridge, MA, The MIT Press.

Kayne, R.S., 1989. Null Subjects and Clitic Climbing. In: Jeggli, O., & Safir, K.J, (eds.), *The Null Subject Parameter*. Dordrecht, 239-261.

Kempson, R., Meyer-viol, & W., Gabbay, D., 2001. *Dynamic Syntax: The Flow of Language Understanding*. Oxford, Blackwell.

Kempson, R., Cann, R., & Kiaer, J., 2004. *Topic, Focus and the Structural Dynamics of Language*. Ms.

Kempson, R., Kiaer, J., & Cann, R., forthcoming. Periphery effects and the dynamics of tree growth. In: Shaer, B., Cook, P., Frey, W., & Maienborn, C., (eds.), *Dislocation: syntactic, semantic and discourse perspectives.* Routledge,

Kempson, R., & Kurosawa, A., forthcoming. At the Syntax Pragmatics Interface: Japanese Relative Clause Construal. In: Hoshi., H. (ed.), *Language Mind and Brain*. Kuroshio publishers.

Kenesei, I., 1993. *A Minimalist Program for the Syntax of Focus*. Ms., Department of English, University of Szeged.

Kordoni, V., 2004. Between shifts and alternations: Ditransitive constructions. In: Müller, S., *Proceedings of the HPSG04 Conference*. Stanford, CSLI publications,

Kratzer, A., 1981. The notional category of modality. In: Eikmeyer, H., & Rieser, H., (eds.), *Words, Worlds, and Contexts. New Approaches in Word Semantics.* Berlin, de Gruyter, 3874.

Kratzer, A., 1991. Modality. In: von Stechow, A. & Wunderlich, D. (eds), *Semantics: An International Handbook of Contemporary Research.* Berlin, de Gruyter, 639-50.

Kroch, A., 1989. Reflexes of Grammar in Patterns of Language Change. *Language Variation and Change* 1981, 199-244.

Lapesa, R., 1978. Sobre dos tipos de subordinación causal. In: *Estudios ofrecidos a Emilio Alarcos Llorach*. Oviedo, Universidad de Oviedo, 173-205.

Lemmon, E.J., 1965. *Beginning Logic*. London, Nelson.

Lewis, D., 1986. *Philosophical Papers, volume II*. New York, Oxford University Press.

Lifschitz, V, 1997. On the Logic of Causal Explanation. *Artificial Intelligence*96, 451-465.

Marten, L., 2002. *At the Syntax-Pragmatics Interface. Verbal Underspecification and Concept Formation in Dynamic Syntax.* Oxford University Press.

Mackridge, P., 1985. *The Modern Greek Language: a Descriptive Analysis of Standard Modern Greek*. Oxford University Press, Oxford.

Manolessou, I., 2005. The Greek dialects of Southern Italy: an Overview. *Kambos, Cambridge papers in Modern Greek.*, 103-125.

Manolessou, I., & Beis, S., 2006. Syntactic Isoglosses in Modern Greek Dialects: The Case of the Indirect Object. In: Janse, M., Brian, J.D., & Angela Ralli (eds), *Proceedings of the Second International Conference of Modern Greek Dialects and Linguistic Theory*. Mytilene, University of Patras Press, 220-235

Manzini, R., 1983. *Restructuring and Reanalysis*. PhD thesis. Cambridge, MA, MIT press.

Manzini, R., & Savoia, L., 2004. Clitics: Cooccurence and Mutual Exclusion Patterns. In: Rizzi, L. (ed.), *The Structure of CP and IP. The Cartography of Syntactic Structures Vol. 2*. Oxford University Press,

Martins, A.M., 2002. The Loss of IP-Scrambling in Portuguese: Clause Structure, Word Order Variation and Change. In: Lightfoot, D., (ed.), *Syntactic Effects of Morphological Change.* Oxford University Press.

McCormack, A., 2008. *Subject and Object Pronominal Agreement in Southern Bantu Languages: A Dynamic Syntax Perspective.* PhD Thesis, School of Oriental and Africal Studies, University of London.

Menéndez Pidal, R., 1980. *Cantar de Mio Cid: Texto, gramática y vocabulario.* Madrid: Espasa Calpe.

Michelioudakis, D., 2007. *Ethical Datives and Argument Structure in Modern Greek*. Mphil thesis, University of Cambridge.

Michelioudakis, D., 2009. What does the Person-Case Constraint tell us about the licensing of dative arguments? In: Karasimos, A., (ed.), *Proceedings of the First Patras International Conference of Graduate Students in Linguistics*, University of Patras Press.

Michelioudakis, D., & Sitaridou, I., 2010. Syntactic micro-variation in Pontic: Dative constructions. In Karasimos, A., & Ralli, A., (eds.), *Proceedings of the 4th International conference on Modern Greek dialects and Linguistic Theory*, University of Patras, 131-147.

Miller, P., & Monachesi, P., 2003. Les Pronoms Clitiques dans les Langues Romanes. In: Godard, D., (ed.), *Les Langues Romanes: Problemes de la Phrase Simple*. Paris: Editions du CNRS, 67123.

Miller, P., & Sag, I., 1997. French Clitic Movement without Clitics or Movement. *Natural Language and Linguistic Theory* 15, 573–639.

Mavrogiorgos, M., To appear. C Agreement and the EPP: Deriving Enclisis in Modern Greek. In: Tsoulas, G. (ed), *Proceedings of the 7th International Conference on Greek Linguistics.*

Meyer-Viol, W., 1995. *Instantial Logic*. PhD Dissertation, University of Utrecht.

Monachesi, P., 1993. Object Clitics and Clitic Climbing in Italian HPSG Grammar. In: *Proceedings of the Sixth Conference of the European Chapter of the Association for Computational Linguistics*. Utrecht, 437 - 442.

Monachesi, P., 1998a. The Morphosyntax of Romanian Cliticization. In: Coppen P., Van Halteren, H., Teunissen L. (eds.), *Proceedings of Computational Linguistics in The Netherlands* 1997. Rodopi, Amsterdam-Atlanta, 99-118.

Monachesi, P., 1998b. Italian Restructuring Verbs: a Lexical Analysis. In: Hinrichs, E., Kathol, A. & Nakazawa, T. (eds.), *Complex Predicates in Nonderivational Syntax*. Syntax and Semantics 30. San Diego Academic Press, 313-368.

Monachesi, P., 1999. The Syntactic Structure of Romanian Auxiliary (and Modal) Verbs. In: G., Bouma et al. (eds), *Constraints and resources in natural language syntax and semantics*. CSLI publications, Stanford.

Montague, R., 1974. *The Proper Treatment of Quantification in Ordinary English*. Dordrecht, D. Reidel.

Morrill, G.V., 1994. *Type logical grammar: Categorial logic of signs.* Dordrecht, Kluwer.

Nevins, A., 2007. The Representation of Third Person and its Consequences for Person-Case Effects. *Natural Language and Linguistic Theory* 25 (2), 273-313.

Nevins, P. & Savescu O. (2008). An Apparent Number-Case Constraint in Romanian. In: Silvina Montrul et al. (eds.), *Romance Linguistics 2008*. Amsterdam, John Benjamins, 185-200.

Newton, B., 1972. *Cypriot Greek: Its Phonology and Inflections.* The Hague: Mouton.

Nicholas, N., 1999. *The story of pu: The Grammaticalisation in Space and Time of a Modern Greek Complementizer.* PhD thesis, University of Melbourne, Australia.

Nicol, F., 2005. Romance clitic clusters: On diachronic changes and cross-linguistic contrasts. In: Heggie, L, & Ordonez, F., *Clitic and affix combinations*. Amsterdam, John Benjamins, 141-197.

Ormazabal, J., & Romero, R., 2007. The Object Agreement Constraint. *Natural Language and Linguistic Theory* 25. Springer, 315-347.

Ouhalla, J., 1997. Remarks on focus in Standard Arabic. In: Mushira, E., & Robert R.R, (eds.), *Perspectives on Arabic Linguistics X*. Amsterdam, John Benjamins, 9-45.

Özkan, H., 2010. Inflection and Tenses in the Sürmene Varieties of Pontus. *Workshop on the Greek language in Pontus: Romeyka in contemporary Trebizond*, Cambridge, 13 March 2010.

Papadopoulos, A., 1955. *Istoriki gramatiki tis pontikis dialektou* (Historical Grammar of the Pontic Dialect). The Committee for Pontic Studies, Athens, supplement 1.

Papangeli, D. 2000. Clitic Doubling: A Head-Complement Relation. *UCL Working Papers in Linguistics*, London.

Pappas, P., 2001. Weak Object Pronoun Placement in Later Medieval Greek: Intralinguistic Parameters Affecting Variation. In: *Ohio State University Working Papers in Linguistics* 56, 79-106.

Pappas, P., 2004. Variation and Morphosyntactic Change in Greek: From Clitics to Affixes. *Palgrave Studies in Language Change and Language History*, Palgrave.

Pappas, P., 2006. Object clitic pronouns in the dialects of Medieval Greek. In: Janse, M., Joseph B., & Ralli, A. (eds.), *Proceedings of the 2nd International Conference on Modern Greek Dialects and Linguistic Theory*.

Paraskeua, M., 2007. *A Dynamic Syntax Account of Object Clitics in Cypriot Greek*. Msc Thesis, King's College, London.

Philippaki-Warburton, I., 1977. Modern Greek Clitic Pronouns and the 'Surface Structure Constraints' Hypothesis. *Journal of Linguistics*, 13, 259-281

Philippaki-Warburton, I., 1987. The Theory of Empty Categories and the pro-drop Parameter in Modern Greek. *Journal of linguistics* 23, 289-318.

Philippaki - Warburton, I., 1994a. V Movement and Clitics in Modern Greek. In: Philippaki-Warburton, I., Nikolaidis, K. & Sifianou, M. (eds), *Themes in Greek Linguistics*. Amsterdam, John Benjamins, 53-60.

Philippaki-Warburton, I., 1994b. The Subjunctive Mood and the Syntactic Status of the Particle na in Modern Greek. *Folia Linguistica* XXVIII/3-4: 297-328

Philippaki-Warburton, I., 1998. Functional Categories and Modern Greek Syntax. *The Linguistic Review* 15, 159-186.

Philippaki-Warburton, I., & Veloudis, J., 1984. The Subjunctive in Complement Clauses. *Studies in Greek Linguistics* 5: 87-104. Aristotle University, Thessaloniki.

Philippaki-Warburton, I., & Spyropoulos, V., 1999. On the Boundaries of Inflection and Syntax. *Yearbook of morphology 1998*, 45-72.

Philippaki-Warburton, I., & Spyropoulos, V., 2004. *A Change of Mood: the Development of the Greek Mood System.* Linguistics 42, 791-817.

Philippaki-Warburton, I., Varlokosta, S., Georgiafentis, M., & Kotzoglou, G., 2002. On the status of clitics and their doubles in Greek. *Reading Working Papers in Linguistics 6*, 57-84.

Pickering, M., & Garrod, S., 2004. Towards a Mechanistic Psychology of Dialogue. *Brain and behavioral science* 27.

Pickering, M. J. & Garrod., S., 2005. Establishing and Using Routines during Dialogue: Implications for Psychology and Linguistics. In: Cutler, A. (ed.), *Twenty-first Century Psycholinguistics: Four Cornerstones.* London, Erlbaum, 85-101.

Progovac, L., 1996. Clitics in Serbian/Croatian: Comp as the Second Position. In: Halpern, A.L, & Zwicky, A.,M., *Approaching Second: Second Position Clitics and Related Phenomena.* CSLI publications: 411-428.

Profili, O., 1983. *Le parler Grico de Corigliano d'Otrante (Province de Lecce, Italie). Phenomene d'interference entre ce parler grec et les parlers romans environnants, ainsi qu'avec l'italien*. Phd thesis, Universite des Langues et Lettres. Grenoble, France.

Purver, M., & Otsuka, M., 2003. Incremental generation by incremental parsing: Tactical generation in Dynamic Syntax. In: *Proceedings of the 9th European Workshop in Natural Language Generation (ENLG-2003)*.

Purver, M., Kempson, R., 2004. Incremental Parsing, or Incremental Grammar? In: *Proceedings of the ACL Workshop on Incremental Parsing*. Barcelona, Spain, July 2004, 74-81.

Purver, M.,Cann, R., Kempson, R., 2006. Grammars as Parsers: Meeting the Dialogue Challenge. Research on Language and Computation 4(2-3), October 2006, 289-326.

Ralli, A., 2006. Syntactic and Morphosyntactic Phenomena in Modern Greek Dialects: The State of the Art. *Journal of Greek Linguistics 2006*, 121-159.

Revithiadou, A., 1999. *Headmost Accent Wins: Head Dominance and Ideal Prosodic Form in Lexical Accent Systems*. Doctoral dissertation, LOT Dissertation Series 15 (HIL/Leiden University). The Hague, Holland Academic Graphics.

Revithiadou, A., 2006. Prosodic Filters on Syntax: An Interface Account of Second Position Clitics. *Lingua 116*, 79-111.

Rezac, M., 2008a. *Phi across modules*. Unpublished monograph, Laboratoire de Linguistique de Nantes.

Rezac, M., 2008b. The Syntax of Eccentric Agreement: The Person Case Constraint and Absolutive Displacement in Basque. *Natural Language and Linguistic Theory* 26, 61-106.

Rivero, L., 1994. Clause Structure and V-movement in the Languages of the Balkans. *Natural Language and Linguistic Theory* 12, 63-120.

Rivero, L., & Terzi, A., 1995. Imperatives, V-movement and Logical Mood. *Journal of Linguistics* 31, 2, 301-32.

Rizzi, L., 1982. Issues in Italian Syntax. *Studies in Generative grammar* 11. Dordrecht, Foris.

Rizzi, L., 1997. The fine structure of the left periphery. In: Haegemann, L., (ed.), *Elements of Grammar: Handbook of Generative Syntax*. Dordrecht, Kluwer, 281-337.

Rohlfs, G., 1950. *Historische Grammatik der unteritalienischen Gräzität*, Halle.

Rohlfs, G., 1977. *Grammatica Storica dei Dialetti Italogreci*. C. H. Becksche, Munchen.

Rooryck, J., 1988. Formal aspects of French nonlexical datives. *Folia Linguistica 22*: 373-386.

Roussou, A., 2000. On the Left Periphery: Modal Particles and Complementizers. *Journal of Greek Linguistics* 1, 65-94.

Roussou, A., 2008. *Selecting complementizers*. Ms., University of Patras.

Ruh, N., Cooper, R.P., & Mareschal, D., 2005. The Time Course of Routine Action. In: *27th Annual Meeting of the Cognitive Science Society*, Stresa, Italy.

Savescu, O., 2007. Challenging the Person Case Constraint: Evidence from Romanian. In: Camacho, J., Flores-Ferran, N., Sanchez, L., Deprez, V., & Jose Cabrera, M. (eds), *Romance Linguistics 2006. Selected Papers from the 36th Linguistic Symposium on Romance Languages* (LSRL), New Brunswick, March-April 2006. Amsterdam, John Benjamins, 255-269.

Săvescu, O., 2009. *A Syntactic Analysis of Pronominal Clitic Clusters in Romance: The View from Romanian.* Phd Dissertation, New York University.

Shafer, G., 1998. Causal Logic. *Proceedings of the 13th European Conference on Artificial Intelligence* (ECAI 98), John Wiley & Sons, 1998, 711-720.

Schütze, C., 1994. Serbo-Croation Second Position Clitic Placement and the Phonology-Syntax Interface. In: Carnie, A., & Harley, H., (eds), *Papers in Phonology and Morphology,* MITWPL 21. Cambridge, MA, 373-465.

Sitaridou, I., 1998. *Ethical Datives in Romance and in Greek.* Unpublished MA dissertation. UCL

Sitaridou, I., 2010a. Fieldworking on Romeyka in Pontus. Invited talk at the workshop: *The Greek language in Pontus: Romeyka in contemporary Trabzon*. Queens College, Cambridge, 13/3/2010.

Sitaridou, I., 2010b. Survival against all odds: The infinitive in the Hellenic varieties of Pontus. Invited talk at the workshop: *The Greek language in Pontus: Romeyka in contemporary Trabzon*. Queens College, Cambridge, 13/3/2010.

Sperber, D., Wilson, D., 1986. *Relevance: Communication and Cognition*. Oxford, Blackwell

Sportiche, D., 1993. *Clitic Constructions*. Ms, UCLA.

Taylor, A., 1994. The change from SOV to SVO in Ancient Greek. *Language Variation and Change* 6, 1-37.

Taylor, A. 1995. The distribution of object clitics in Koine Greek. *Penn Working Papers in Linguistics* 2, 121-139.

Taylor, A., 1996. A Prosodic Account of Clitic Position in Ancient Greek. In: Halpern, A.L., & Zwicky, A.M., *Approaching Second: Second Position Clitics and Related Phenomena*. Stanford, CSLI publications, 411-428.

Taylor, A., 2002. The distribution of object clitics in Koine Greek. In: Southern R.V, S., (ed.), *Indo-European Perspectives*. Washington DC, Institute for the Study of Man, 285-316.

Tenny, C., & Pustejovsky., J., 2000. A History of Events in Linguistic Theory. In: Tenny, C., & Pustejovsky., J. (eds.), *Events as Grammatical Objects*. Stanford, CSLI publications.

Terzi, A., 1999a. Cypriot Greek clitics and their Positioning Restrictions. In: Alexiadou A., Horrocks, G. & Stavrou, M. (eds.), *Studies in Greek Syntax*. Dordrecht, Kluwer, 227-240.

Terzi, A., 1999b. Clitic combinations, their Hosts and their Ordering. *Natural Language and Linguistic Theory* 17, 85-121.

Thavoris, A., 1977. Morfologika merikon idiomaton tis Ditikis Makedonias. *First Symposium for Northern Greek Linguistics.* Thessaloniki: Institute for Balkan Studies, 75-95.

Tombaidis, D.E, 1996. *Meletimata pontiakis dialektou*. [Pontic Dialect Studies]. Thessaloniki, Kodikas.

Triandafyllidis, M., 1981. *Neoeliniki gramatiki: Istoriki isagogi* [Modern Greek Grammar: Historical Introduction]. Thessaloniki, Aristotle University.

Tsakali, V., 2008. Double Floating Quantifiers in Modern Greek and Pontic. *Syntax and Semantics 36.* Emerald Group, 189-204.

Tsimpli, I.M., 1990. The Clause Structure and Word Order in Modern Greek. *UCL Working Papers in Linguistics 2*. 226-255.

Tsiplakou, S., 2009. Code-switching and Code-mixing between Related Varieties: Establishing the Blueprint. *The International Journal of the Humanities* 6, Issue 12, 49-66.

Tsiplakou, S., in press. Competing grammars in Cypriot Greek? Evidence from language alternation. *Proceedings of the 18th International Conference on Historical Linguistics*, Montreal, 6-11 August 2007. Amsterdam, John Benjamins.

Tsiplakou, S., Panagiotidis, P. & Grohmann, K.K., 2007. Properties of Cypriot Greek Wh-question formation. In: *Proceedings of the 7th International Conference on Greek Linguistics*, 83-98.

Tzartzanos, A., 1940. *Sintaktiko tis Arxeas Elinikis Glosis.* [Syntax of the Ancient Greek Language]. Kiriakidis, Thessaloniki, 1994.

Zvolensky, Z., 2002. Is a possible-worlds semantics of modality possible? A problem for Kratzer's semantics. In: Jackson, B. (ed.), *Proceedings of SALT XII*. Ithaca, Cornell University Press, 339-358

Appendix A

Sources and Texts Consulted

Makhairas = Eksigisis tis glikias xoras Kipru. In: Dawkins, R.M, (ed.), Leontios Makhairas: Recital concerning the Sweet Land of Cyprus entitled ‘Chronicle’, Oxford 1932 .

Appendix B

Additional SMG Doubling Examples

B.1 Indefinite Clitic Doubling in SMG

(B.1) *Ego eftixos θa ime ektos γia to δiimero. Eftixos γiati*
I fortunately FUT be out for the weekend fortunately because
tin xreiazome mia ekδromula
her.CL-ACC need one excursion
'Fortunately, I will be away for the weekend. That is fortunate, since I need an excursion. [From the net, as retrieved on 24 Apr 2010 23:25:38 GMT http://webcache.googleusercontent.com/search?q=cache:QipnExbIlh0J:www.mybike.gr/topic/14253

(B.2) *Leo na tin proslavo mia γramatea. Ti les?*
say.1SG SUBJ her.CL-ACC hire a.ACC secretary.ACC what say.2SG
' I'm thinking of hiring a secretary. What do you think?'

(B.3) *Θa to aγoraso ena autokinitaki kapia stiγmi*
FUT it.CL-ACC buy a.ACC car.ACC some time
'I will buy a car one day.'

(B.4) *Δen ksero esis ti kanate xtes, pados eγo to*
NEG know.1SG you what did.2PL yesterday anyway I it.CL-ACC
δiavasa ena vivlio
read a.ACC book.BOOK
'I do not know what you did yesterday, but I read a book.'

B.2 Bare Quantifier CD in SMG

(B.5) *Ine oli tus kala peδia. Tus ksero merikus*
are.3PL all them good boys them.CL-ACC know some.ACC
'All of them are goo guys. I know some of them'

(B.6) *Oli itan pitsirikia. To ksero γiati tus iδa kapius*
all were young it.CL-ACC know because them.CL-ACC saw.1SG some.ACC
'They were all young. I know it because I saw some of them'

(B.7) *Tus iδa kana δio na fevγun ala δen katalava oti itan afti*
them.CL-ACC saw.1SG one two SUBJ leave.3PL but NEG understood.1SG that were they
'I saw a couple of them leaving but I did not realize they were these people.'

B.3 Partitive CD in SMG

(B.8) *Tus vlepo kapius apo aftus na benun sto maγazi*
them.CL-ACC see.1SG some.ACC of them SUBJ enter to-the store
'I saw some of them entering the store'

(B.9) *Mas enδiaferi polus apo mas*
us.CL interests many.ACC from us
'Many of us are interested.'

(B.10) *Sas ksero merikus apo sas. Θa sas katagilo*
you.CL know.1SG some.ACC from you FUT you.CL sue
'I know some of you, I will sue you.'