The Syntax of Discourse Functions in Greek: a Non-Configurational Approach

Theodora Alexopoulou

A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy
to the
University of Edinburgh
1999

Abstract

This thesis offers an account of the syntactic properties of Focus-movement, Topicalisation and Clitic Left Dislocation (CLLD) in Greek. As these phenomena are central to discussions of the syntax-discourse interface, a significant part of this study pursues the question of the representation of the discourse functions of *topic* and *focus* and their relation to syntax.

For the most part, the literature on the syntax of Focus-movement, Topicalisation and CLLD advocates that *focus* and *topic* are encoded in the Phrase Structure by distinct Functional Projections: Focus Phrase (FP) and Topic Phrase (TP). *Foci* and *topics* move to the Specifier of the relevant Projection to check their discourse features. The term *Discourse Configurational Languages* has been recently coined for languages that encode *focus* and *topic* through Phrase Structure configurations.

With respect to the syntactic properties of the relevant structures, the *Discourse Configurational* approach assumes that Focus-movement, Topicalisation and CLLD instantiate three distinct syntactic operations; A-bar-movement, A-movement and base-generation respectively. This complex syntax enables a simple view of the syntax-discourse interface; there is an isomorphic relation between syntax and discourse, as each discourse function is associated with a distinct syntactic operation. Further, *focus* and *topic* are treated as syntactic features, specifying heads of Functional Projections.

This thesis, in contrast, argues for a *non-configurational* approach. It shows that the claim that Focus-movement and Topicalisation instantiate A-bar-movement and A-movement respectively is based on insufficient evidence. This claim is motivated by the absence of weak crossover effects in Topicalisation and their presence in Focus-movement. However, this study argues that the weak crossover effect is not a valid diagnostic of the A/A-bar distinction, since some cases of Wh-questions, the prototypical instance of A-bar movement, do not give rise to weak crossover effect. Further, in the *Discourse Configurational* approach, CLLD is

treated as an instance of base-generation rather than movement, because it does not license parasitic gaps. In this thesis, CLLD is analysed as adjunct extraction and it is shown that the unavailability of parasitic gaps is a general property of adjunct extraction. Further, this study demonstrates that Focus-movement, Topicalisation and CLLD exhibit the same syntactic properties and instantiate the same extraction mechanism. Thus, they are given a unified syntactic treatment.

The argument that Focus-movement, Topicalisation and CLLD share the same syntax has implications for the architecture of the discourse-syntax interface. Unlike the *Discourse Configurational* approach, this syntactic analysis implies a non-isomorphic relation between syntax and discourse, as a single syntactic structure corresponds to more than one discourse function. Thus, the syntax of discourse constructions is independent of the discourse functions encoded. It is argued that the discourse evidence does not justify the incorporation of discourse functions in Phrase Structure or their treatment as syntactic features. Rather, *focus* and *topic* should be represented at a distinct level, independent of syntax, *Information Structure*.

The analysis is couched in the framework of Head-driven Phrase Structure Grammar. The syntactic properties of extractions in Greek are readily captured by the HPSG mechanism of Unbounded Dependencies. The multidimensional nature of HPSG signs allows for the representation of discourse functions and a flexible mapping between syntax and discourse.

Declaration

This thesis has been composed by myself and it has not been submitted in any previous application for a degree. The work reported within was executed by myself, unless otherwise stated.

March 1999

Acknowledgements

Just a few nights —no, I would not leave this until the last minute— before this study takes its final shape, it still feels unreal that it was indeed possible to 'finish' a thesis. What now feels impossible is to thank all the people who made this moment real and this thesis possible. But I will do my best.

I first wish to thank my principal supervisor, Caroline Heycock. I am deeply indebted for her genuine interest, the energy she put into teaching me how to convert vague intuitions about the data into linguistic arguments, the constant criticism and encouragement, the inspiration she provided at all stages and the meticulous comments on close-to-final drafts of this study. I learnt a lot.

I am also grateful to my second supervisor, Ronnie Cann. His expertise and insightful criticism has influenced this work while his thorough comments on earlier drafts have contributed to a, hopefully, clearer discussion.

Of course, none of the two is responsible for any mistakes or infelicities.

I gratefully acknowledge the financial support I received from the State Scholarship Foundation of Greece (I.K.Y.).

This study has benefited immensely from the Focus seminar at the University of Edinburgh (1995-96) organised by D.R.Ladd and B.Hoffman. I would like to thank all the participants of this seminar. In particular, I am grateful to D.R.Ladd and I.Mennen who, long after the end of the seminar, had the patience to explain to me basic facts about the phonology of Greek.

I would also like to thank the participants of the syntax group and the HPSG group at the University of Edinburgh. They helped me understand crucial papers and provided a friendly environment for informal presentations of my work. I would like to mention Colin Mathieson and, in particular, Claire Grover for her patience in explaining aspects of the HPSG formalism and useful feedback on parts of my analysis.

Many thanks to our Departmental secretary, Ethel Jack, for finding the easiest way out bureaucratic labyrinths, but most important, for taking care and being an endless source of fun.

I would also like to thank friends and colleagues in the Department of Linguistics for sharing anxiety, frustration, insecurity, but also enthusiasm, hope and the bright side of life. Special thanks to Miriam Eckert for *deixis* to positive thinking, Kook-hee Gill for *binding* me in space and time the last month of this project, Ineke Mennen for always being there to listen and Marta Robinson for being a caring officemate.

I am grateful to Iraide Ibarretxe, Kook-hee Gill, Max Louwerse, Natasa Mangana and Satu Manninen for the care with which they proof-read parts of this thesis. LaTeX is to be blamed for any remaining typos.

Many thanks and love to Dimitra Kolliakou, who, despite the distance, has offered insightful feedback in the course of developing the ideas that appear in this thesis, meticulous comments on earlier drafts, advice, support and has been an invaluable friend throughout.

Many thanks and love to my friends in Edinburgh, Sofia Antoniadou, Stavroula Fylahtou, Yanis Karayorgos, Nikos Lambaras, Eleni Manolakaki, Demetra Papakostantinou and Takis Voilas. The last five years we shared a lot: dissapointments, anxieties, visiting parents from Greece, flues, hangovers, Christmas, love disasters and love triumphs, World Cup 98...we had a lot of FUN! Special thanks to Nikos for his care, to Sofia for endless sharing and to Takis for his wit, good taste and dance company.

Many thanks and love also to my friends outside Edinburgh who kept in touch, Yorgos Antonis, Eleni Argyriou, Mariana Kitsona, Camilla Knox-Peebles, Kostas Stamatis and Loida Vlahava. In particular to Eleni, for proving that distance and circumstances cannot affect a real friendship.

My deepest thanks and love to Marcelo Fiore for reminding me that research is fun, his constant emotional support, practical tips for LaTeX and, above all, his invaluable love.

Last, but definitely MOST, I would like to thank my parents, Μαρίνα and Παναγιώτης, and my sister, Γιώτα, for their unconditional love, the financial support towards the end of this project and the reassuring belief that they are always there for me, no matter the distance. I dedicate this thesis to them with my love.

Δώρα Αλεξοπούλου

Abbreviations and Symbols

Text

CLLD	:	Clitic Left Dislocation
COMP	:	the head position of CP
compl	:	complementiser
DL-position	:	Discourse-linked position
HFP	:	Head Feature Principle
LPR	:	Linear Precedence Rule
p-gaps	:	parasitic gaps
QR	:	Quantifier Raising
SLAC	:	Slash Amalgamation Constraint
SLIP	:	Slash Inheritance Principle
UDCs	:	Unbounded Dependencies
VALP	:	Valence Principle
wco effects	:	weak crossover effects

Glosses

ACC	:	accusative
CL	:	clitic
FUT	:	future
GEN	:	genitive
NOM	:	nominative
PL	:	plural
SG	:	singular
SMALL CAPITALS	:	nuclear accent placement

The symbol $*$ indicates ungrammatical sentences whereas @ indicates infelicitous ones.

Feature Structures

AVM	:	attribute-value matrix
[ ]	:	feature structure
< >	:	list
{ }	:	set
$\oplus$	:	append
$\bigcirc$	:	sequence union (*shuffle*)
$\cup$	:	(set) union
$x < y$	:	x precedes y
numbered boxes $\boxed{n}$	:	token-identity, structure-sharing
italics	:	sorted feature structures
SMALL CAPITALS	:	attributes in a feature structure
elist	:	empty list
nelist	:	non-empty list

Contents

Chapter 1

Introduction

1.1 Unbounded Dependencies and discourse functions

Unbounded Dependencies are available crosslinguistically and play a significant role in the grammar. For example, they are central in the formation of Wh-questions and Relative constructions. In recent years it has been acknowledged that, in a number of languages, Unbounded Dependencies are employed for the realisation of the discourse functions of topic and focus (Kiss 1995a). In this study I focus on such uses of these constructions and, in particular, on Topicalisation, Clitic Left Dislocation (CLLD) and Focus-movement illustrated in the following examples from Greek:

(1.1) a. **Topicalisation**

tin parastasi skinothetise o Dimitris POTAMITIS

the performance-ACC directed-3SG the Dimitris Potamitis-NOM
'Dimitris Potamitis directed the performance.'

b. **Clitic Left Dislocation (CLLD)**

tin parastasi *ti* skinothetise o Dimitris POTAMITIS

the performance-ACC it-CL directed-3SG the Dimitris Potamitis-NOM
'Dimitris Potamitis directed the performance.'

c. **Focus-movement**

tin PARASTASI skinothetise o Dimitris Potamitis

the performance-ACC directed-3SG the Dimitris Potamitis-NOM
'Dimitris Potamitis directed the performance.'

All three constructions in (1.1) exhibit typical properties of Unbounded Dependencies. However, they differ in interpretation, as suggested by their names. The extracted object in Focus-movement conveys the new/updating information of the sentence. On the other hand, the extracted object in Topicalisation and CLLD is the topic/link, the locus-of-update of the sentence (I will return to the definitions of topic/link and focus in Section 2.2). The focused object in (1.1c) bears the main or nuclear accent of the sentence, as indicated by the small capitals. By contrast, the extracted topics in (1.1a-b) bear no accent. Further, unlike Topicalisation and Focus-movement, the extracted object in CLLD is coindexed with a pronominal clitic.

The representation of these constructions has direct implications for the way in which Unbounded Dependencies are treated in different syntactic frameworks. In Minimalism, Unbounded Dependencies involve movement of the preposed constituent to a specific position at the syntactic tree (Chomsky 1995; Chomsky 1996; Marantz 1995; Radford 1997). Movement is driven by feature checking. Very roughly, specific positions in the syntactic tree are associated with a feature (e.g. C is specified for the [wh] feature (Radford 1997)). Constituents marked with a feature move to the corresponding position in the tree to check this feature. In Wh-questions, wh-phrases are marked with the [wh] feature and move to [Spec,CP] to check their feature. Under this view, Unbounded Dependencies can only be motivated by feature checking. It follows, that Focus-movement, Topicalisation and CLLD should also be motivated by feature-checking. The obvious way is to assume that the features involved in these constructions are discourse ones, the focus and topic feature. Indeed this is the dominant view in the literature. Various authors extend Phrase Structure with distinct functional projections, the Focus Phrase (FP) and the Topic Phrase (TP) (Brody 1990; Agouraki 1993; Kallulli 1997; King 1995; Kiss 1995a; Rizzi 1995; Tsimpli 1995). The Focus and Topic head are specified for the focus and topic feature respectively. Topics and foci move to the Specifier position of the corresponding functional projection to check their discourse features. *Discourse Configurational Languages* is a term recently coined for languages that encode focus and topic through Phrase Structure configurations. Kiss (1995a) is a collection of representative papers of the *Discourse Configurational* approach to Topicalisation, CLLD and Focus-movement.

Further, *Discourse Configurational* approaches propose that Topicalisation, CLLD and Focus movement involve three distinct syntactic operations. While all three constructions

exhibit some properties typical of Unbounded Dependencies—crosslinguistically, they all involve long-distance extraction and they all obey subjacency restrictions—they differ in various ways. Focus-movement and Topicalisation license parasitic gaps (p-gaps) but CLLD does not. Focus-movement gives rise to weak crossover effects (wco) whereas Topicalisation and CLLD do not. CLLD involves coindexing with a pronominal clitic whereas Topicalisation and CLLD involve a trace/gap. In order to account for their differences various authors propose that these constructions involve three distinct syntactic operations. Focus-movement involves A-bar-movement which is quantificational in nature and gives rise to wco effects. Topicalisation instantiates A-movement which is anaphoric in nature and does not give rise to wco effects. Finally, CLLD does not involve movement. Rather, the dislocated element is base-generated at its surface position. This analysis is supported by the unavailability of p-gaps in CLLD constructions.

Thus, the *Discourse Configurational* approach offers a rather complex syntax, since the three constructions are associated with three distinct variants of Unbounded Dependencies. However, the complex syntax is compensated for by a simple, isomorphic view of the discourse-syntax interface. There is a one-to-one mapping between discourse functions and syntactic constructions (modulo the difference between Topicalisation and CLLD).

In this thesis I will argue against the *Discourse Configurational* approaches to Topicalisation, CLLD and Focus-movement. My objections are based on two main arguments.

First, I will present evidence indicating that the discourse functions of focus and topic should be represented independently of syntax. By encoding focus and topic in Phrase Structure configurations, *Discourse Configurational* approaches fail to capture the independence of discourse functions from syntax and make the wrong predictions about the empirical domain.

Second, I will show that Topicalisation, CLLD and Focus-movement do not involve three distinct syntactic constructions. Rather they all instantiate the same syntactic structure. I will therefore propose a uniform syntactic treatment which results in a reduced and more elegant syntax.

The analysis is couched in the framework of Head-driven Phrase Structure Grammar (HPSG) (Pollard & Sag 1987; Pollard & Sag 1994). Unlike the Minimalist approach, the HPSG mechanism of Unbounded Dependencies does not involve any movement (a view originally proposed in Gazdar (1981) and inherited in HPSG by GPSG (Pollard & Sag 1994)).

Crucially, it does not involve any feature checking. As a result, there is no need to assume that focus and topic instantiate some abstract syntactic head or feature. That is, there is no need to incorporate focus and topic in the syntax and extend Phrase Structure. Rather, the multidimensional nature of the HPSG *sign* allows a flexible accommodation of the discourse functions of focus and topic independently of syntax. Further, the HPSG mechanism of Unbounded Dependencies can capture readily the syntactic properties of the relevant constructions in Greek.

The organisation of this thesis is as follows. In the remainder of this chapter I present some basic assumptions about the structure of Greek clauses (Section 1.2) and a brief introduction to HPSG (Section 1.3). In this study, I adopt the view of Vallduví (1992) that there is a distinct level of grammar, *Information Structure*, encoding the organisation of new/focus vs. given/ground information in a sentence (Vallduví 1992). In Chapter 2 I introduce the primitives of the focus-ground articulation and show how Focus-movement, Topicalisation and CLLD are employed for the realisation of *Information Packaging* in Greek. In Chapter 3, I discuss in detail *Discourse Configurational* approaches to Topicalisation, CLLD and Focus-movement. Next, in Chapter 4, I present a unified syntactic analysis of these constructions, employing the HPSG mechanism of Unbounded Dependencies. In Chapter 5, I offer an alternative view of the discourse-syntax interface that captures the independence of Information Structure from syntax (and phonology). Finally, I conclude in Chapter 6.

1.2 Basic assumptions about Greek clause structure

VSO as the basic order of Greek

For the most part, the Greek literature assumes VSO as the basic order of Greek (Catsimali 1990; Philippaki-Warburton 1982; Philippaki-Warburton 1985; Tsimpli 1995; Tsimpli 1996)[1]. Though VSO is statistically rare—in Lascaratou (1989:p.42) VSO represents only 1.1% of the corpus—there are theoretical arguments for taking VSO to be the basic order of Greek.

Philippaki-Warburton (1985) observes that there is a group of subordinate adjunct clauses in which SVO is impossible:

[1] By contrast, Horrocks (1983) and Horrocks (1994) proposes that Greek has two basic ordering patterns, VSO and SVO.

(1.2) a. svisame ta fota [ja na filisi o janis ti maria]
switched-off-we the lights [for that kiss-he [the] John [the] Mary]
'We switched off the lights so that John would kiss Mary.'

b. *[ja o janis na filisi ti maria]

(Philippaki-Warburton 1985:ex.39)

Further, Philippaki-Warburton (1985) notes that VSO sentences are the most natural answer to a question like *What happened?*:

(1.3) a. ti ejine
'What happened?'

b. filise o janis ti maria
kissed-he the-NOM John-NOM the-ACC Mary-ACC
'John kissed Mary.'

(Philippaki-Warburton 1985:ex.39)

Philippaki-Warburton (1985) assumes that answers to a question like (1.3a) are pragmatically neutral because all the information conveyed by them is new. No theme/topic is present in (1.3b). Thus, this sentence should instantiate the basic word order of Greek. I will return to this issue in Section 2.4.1, after presenting the definitions of focus and topic adopted in this study.

Following the literature, I will also assume that VSO is the basic word order of Greek. Additional evidence comes from the fact that VSO appears as the only unambiguous order in the absence of morphological cues. In the following examples the two NPs are morphologically ambiguous; they can be either nominative or accusative. Example (1.4a), in which the nuclear accent falls on the rightmost element can only have a VSO reading. By contrast, example (1.4b) is ambiguous between an SVO and OVS reading. Example (1.4c) has an unambiguous SVO reading. However, as will be shown in Section 2.4.3, if this example had an OVS structure, the object should be coindexed with a clitic. Thus, unlike the V-NP-NP order, the NP-V-NP order cannot be considered unambiguous.

(1.4) a. htipise to koritsi to AGORI
hit-3SG the girl the boy
'The girl hit the boy.'

b. to KORITSI htipise to agori
the girl hit-3SG the boy
'The girl hit the boy/The boy hit the girl.'

c. to koritsi htipise to AGORI
the girl hit-3SG the boy
'The girl hit the boy.'

On the configurationality of Greek clauses

The question of the configurationality of Greek clauses has been a matter of controversy. Below I briefly present the picture Greek presents with respect to various diagnostics. [I omit the theoretical assumptions on which the diagnostics used are based. For a detailed discussion see Catsimali (1990), Horrocks (1994) and Tsiplakou (1998) among others.]

i) **Nominative reflexives-binding asymmetries**

Typically, non-configurational languages allow both nominative and accusative reflexives (Kroeger 1994; Tsiplakou 1998). Greek exhibits both as shown in (1.5):

(1.5) a. o eaftos tu_j katestrepse ton $Petro_j$
the self-NOM his-CL.GEN destroyed-3SG the Petros-ACC

b. o $Petros_j$ katestrepse ton eafto tu_j
the Petros-NOM destroyed-3SG the self-ACC his-CL.GEN

(Tsiplakou 1998:ex.19,20)

In relation to this, Horrocks (1994) presents evidence indicating that NPs like *o eaftos tu*, *ton eafto tu* etc. do not exhibit typical properties of reflexives. For example, they can appear without an overt antecedent:

(1.6) o eaftos tu ftei
the self of-him is-to-blame
'He has only himself to blame.'

(Horrocks 1994:ex.14)

On the basis of (1.6), Horrocks (1994) concludes that NPs like *o eaftos tu* are not typical reflexives. As a result, any evidence involving these NPs is irrelevant to the question of the

existence of a VP constituent in Greek.

ii) **Dummy subjects**

The presence of dummy subjects is a characteristic property of configurational languages. Their absence in Greek (1.7) points to a flat clause structure:

(1.7) vrehi

rain-3SG
'It rains.'

iii) **Subject-object asymmetries with respect to extraction**

In configurational languages there are asymmetries between subject and object extraction from a *that-clause*. In English, (1.8b) is grammatical only if *that* is deleted:

(1.8) a. Who$_j$ do you think (that) John met t$_j$?

b. Who$_j$ do you think (*that) t$_j$ met John?

(Tsiplakou 1998:ex.17)

Non-configurational languages do not exhibit such asymmetries (Kroeger 1994). In Greek *oti* ('that') deletion is unavailable. However, subjects are extracted from *oti-clauses* in the same manner objects are:

(1.9) a. Pion nomizis oti sinandise o Yanis?

who-ACC think-2SG that met-3SG the Yanis-NOM
'Who do you think that John met?'

b. Pios nomizis oti sinandise to Yani?

who-NOM think-2SG that met-3SG the Yanis-ACC
'Who do you think met John?'

(Tsiplakou 1998:ex.18)

iv) **Subject-verb idioms**

Another property distinguishing configurational from non-configurational languages is the availability of subject-verb idioms, indicative of lack of a VP constituent. Greek has a variety of such idioms, a list of which is offered in Tsiplakou (1998).

v) **VP ellipsis**

The availability of VP ellipsis indicates the existence of a VP constituent. In English, the complex *so+do* may replace a VP as shown in (1.10a). By contrast, example (1.10b), in which *so did* is supposed to replace only the verb, is ungrammatical:

(1.10) a. John gave the book to Mary and so did Peter

b. *John gave the book to Mary and so did Peter the letter to Tom

(Tsiplakou 1998:ex.58)

In the absence of such a diagnostic in Greek, Tsiplakou (1998) proposes that examples similar to (1.10) can be constructed in Greek with the word *episis*, which could be roughly translated as *as well,too*[2]. She offers the following examples which, as she admits, are slightly awkward in Greek:

(1.11) a. o Yanis edose tis Marias to molivi ke o Vasilis
the Yanis-NOM gave-3SG the Maria-GEN the pencil-ACC and the Vasilis-NOM
episis
too
'John gave Mary the book and so did Bill.'

b. *o Yanis edose tis Marias to molivi ke o Vasilis
the Yanis-NOM gave-3SG the Maria-GEN the pencil-ACC and the Vasilis-NOM
episis tis Elenis to vivlio
too the Eleni-GEN the book-ACC
'*John gave Mary the book and so did Bill Helen the book.'

(Tsiplakou 1998:ex.59)

Though both of the above examples are awkward, the second is worse. She takes the unavailability of (1.11b) as evidence that VP ellipsis is available in Greek and that VP is a distinct constituent.

It is not clear whether *episis* is an equivalent to the *so+do* complex in English. Even so, in pragmatically more plausible contexts, *episis* is acceptable in examples corresponding to (1.11b):

[2]Very often it may mean *in addition,also.*

(1.12) a. o Yanis estile to diaziyio sti Maria ke episis o Petros
the Yanis-NOM sent-3SG the divorce-ACC to-the Maria-ACC and too the Petros-NOM
stin Eleni
to-the Eleni-ACC
'Yanis sent the divorce to Maria and 'so-did' Petros to Eleni.'

b. o Yanis estile to diaziyio sti Maria ke episis tin adia
the Yanis-NOM sent-3SG the divorce-ACC to-the Maria-ACC and too the permit
tu gamu stin Eleni
the wedding-GEN to-the Eleni-ACC
'Yanis sent the divorce to Maria and 'so-did' the wedding permit to Eleni.'

In (1.12a) *episis* replaces the verb with its direct object, whereas in (1.12b) it replaces the verb with the subject.

Finally, Greek allows the following examples in which subparts of the 'VP' are elliptical [the examples are from Catsimali (1990), cited in Tsiplakou (1998)]:

(1.13) a. o Yanis edose tis Marias to molivi ke o Vasilis tis
the Yanis-NOM gave-3SG the Maria-GEN the book-ACC and the Vasilis-NOM the
Elenis
Eleni-GEN
'*John gave Mary the pencil and Bill Helen.'

b. o Yanis evale to vivlio sto trapezi ke o Vasilis sto
the Yanis-NOM put-3SG the book-ACC to-the table and the Vasilis-NOM to-the
rafi
shelf
'*John put the book on the table and Bill on the shelf.'

Unlike the examples constructed with *episis*, which are relatively awkward, the examples in (1.13) are quite natural. Note that the translations in English are ungrammatical. In this respect, Greek allows 'ellipsis' in a context where it is not available in a prototypically configurational language like English.

In conclusion, the evidence from VP ellipsis does not indicate the existence of a VP constituent in Greek.

vi) **Weak crossover effect**

In configurational languages like English, the wh-phrase may not be coindexed with the possessive pronoun in examples like (1.14):

(1.14) *Who$_j$ does his$_j$ mother love?'

The phenomenon is known as weak crossover effect (Haegeman 1991). Non-configurational languages are supposed to exhibit no wco effects. Thus, examples like (1.14) are expected to be grammatical (Kroeger 1994; Tsiplakou 1998). This is indeed the case in Greek, as shown in (1.15)—see Tsiplakou (1998) for similar examples:

(1.15) pion$_j$ ton$_j$ agapai i mana tu$_j$?
who-ACC him-CL.ACC love-3SG the mother-NOM his-CL.GEN
'Who does his (own) mother love?'

Note, though, that wco is eliminated in (1.15) due to the presence of the clitic. In the absence of the clitic, the example is ungrammatical:

(1.16) *pion$_j$ agapai i mana tu$_j$?
who-ACC love-3SG the mother-NOM his-CL.GEN
'Who does his mother love?'

In view of the contrast between (1.15) and (1.16), no conclusion may be drawn for the configurationality of Greek clauses from data involving wco. Wco effects will be discussed in more detail in Chapter 3. However, as suggested in Alexopoulou (1997), wco is subject to discourse rather than syntactic constraints, and it is unlikely that it may serve as a diagnostic for constituency.

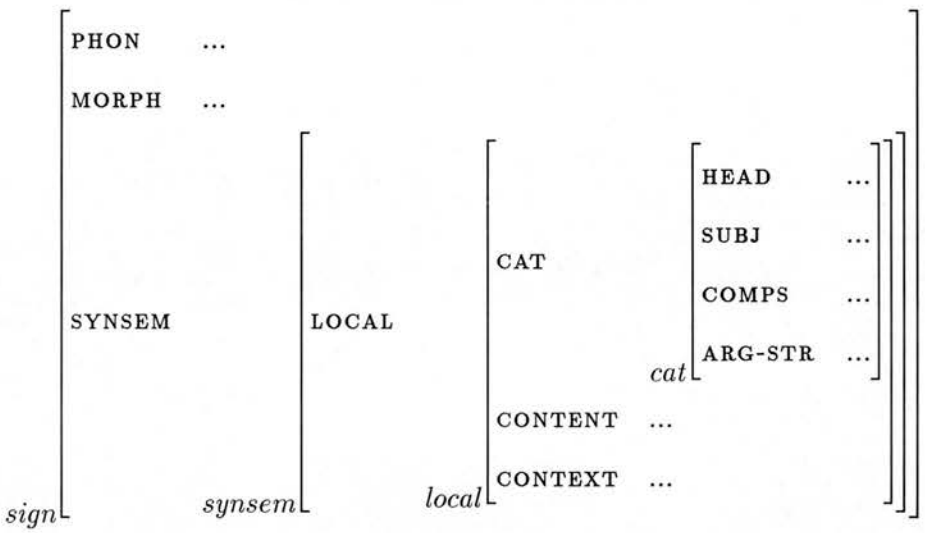
To conclude, it seems that, in every respect, Greek exhibits properties of a non-configurational language. Thus, I will assume a flat structure for Greek with VSO as basic order.

1.3 Head-Driven Phrase Structure Grammar: introduction

1.3.1 Basics

In HPSG, linguistic expressions are *signs*, structured objects encoding information about familiar levels of linguistic organisation: *phonology, morphology, syntax, semantics* (1.17). Signs are modelled by *feature structures* which, by convention, are displayed in the form of ATTRIBUTE-VALUE MATRICES (AVMs) as in (1.17) (Pollard & Sag 1987; Pollard & Sag 1994):

(1.17)



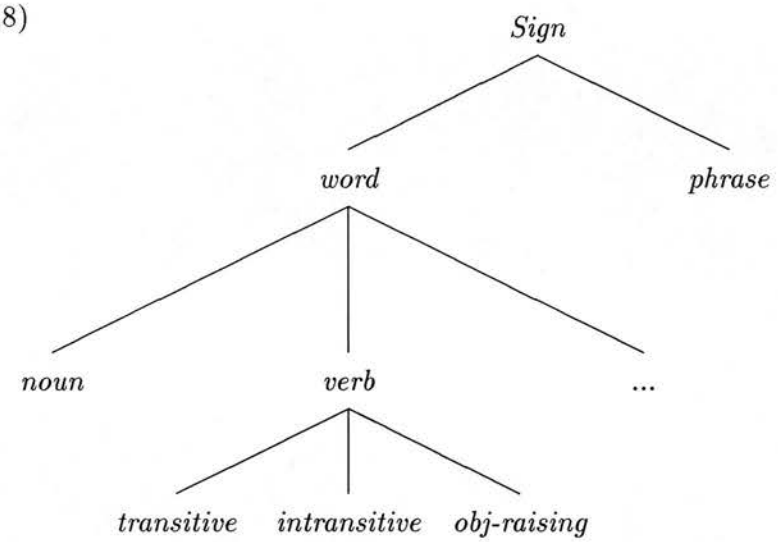
In the above AVM, the attributes PHONOLOGY (PHON), MORPHOLOGY (MORPH), SYNTAX-SEMANTICS (SYNSEM) take as values feature structures that encode information about the phonological, morphological and syntactic-semantic features of an expression.

All feature structures are *sorted.* They are labelled with a *sort* symbol indicating which type of object this feature structure is modelling. By convention, *sortal* labels are indicated by *italics* outside the feature-structure (here, bottom-left). The feature structure in (1.17) illustrates the structure of a *sign.* The attribute SYNSEM takes as its value a feature-structure of sort *synsem*, encoding general information about the syntax and semantics. I will present the various subsorts of feature structures (*synsem, local, category etc.*) in Section 1.3.2. Throughout, small capitals will be used for ATTRIBUTES in a feature structure and italics will be used for *sorted feature structures.*

Signs have various subtypes (*word, phrase, etc.*), organised in a *hierarchical* way. By way

of illustration consider the following example of a *type-hierarchy*:

(1.18)



A grammar is a set of constraints on the organisation of features in a feature structure (which feature/attribute is relevant in a feature-structure in a language) and on the appropriate values for these features. The constraints of the grammar are imposed on types. For example, English *verbs* are subject to the following constraint (Sag 1997):

(1.19) $$\textit{Verb} \rightarrow \left[\textsc{subj} \left\langle \textbf{NP} \right\rangle\right]$$

The above states that a *word* of sort *verb* must have a subject. Similarly, the following constraint states that a *transitive* verb has, at least, one object appearing in its list of complements (COMPS):

(1.20) $$\textit{transitive-verb} \rightarrow \left[\textsc{comps} \left\langle \textbf{NP,...} \right\rangle\right]$$

On the other hand, an *intransitive verb* takes no object (its COMPS is empty) as stated in (1.21):

(1.21) $$\textit{intransitive-verb} \rightarrow \left[\textsc{comps} \langle\rangle\right]$$

Type-hierarchies are organised in such a way that subtypes inherit the properties or constraints of their supertype. Thus, both *transitive* and *intransitive verbs* inherit the constraint in (1.19) from their supertype *verb*. In this way, redundancy is avoided in the lexical entries and intermediate-level generalisations/constraints are expressed where relevant.

Finally, linguistic expressions are grammatical if they are modelled by feature structures which are *totally-well-typed* and *sort-resolved*. A feature structure is *well-typed* if it has the feature attributes appropriate for its type. A feature structure is *totally-well-typed* if the appropriate feature for each node is present. A feature-structure is *sort-resolved* if it is *totally-well-typed* and, in addition, every attribute takes as its value the *maximum sort* that is appropriate for that attribute. For example, a feature structure modelling a *noun*, is *sort-resolved* if the value of its CASE attribute is *nominative*, *accusative* etc. not just *case*. Thus, partial/underspecified feature-structures do not correspond to grammatical linguistic expressions. They are used though, frequently, in the formulation of constraints.

1.3.2 Words

Feature Structures in Lexical entries

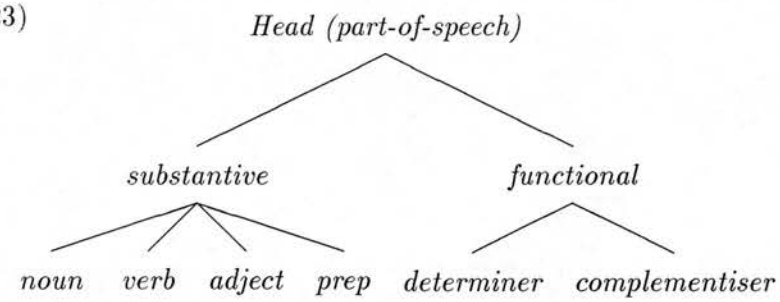
Consider the lexical entry of the *verb vlepi* (*'sees'*-3sg.indic.present):

(1.22)
$$
\textit{word}\left[\begin{array}{ll}
\textsc{phon} & \textit{vlepi} \\
 & \left[\begin{array}{ll}\textsc{i-form} & \textit{vlepi}\\ \textsc{root} & \textit{vlep}\end{array}\right] \\
\textsc{synsem}\,|\,\textsc{local} & \left[\begin{array}{ll}
\textsc{cat} & \left[\begin{array}{ll}\textsc{head} & \textit{verb}\\ \textsc{comps} & \langle \boxed{3} \rangle \\ \textsc{arg-str} & \boxed{3}\langle \text{NP}\boxed{1}[\textsc{3sg-nom}]_j,\ \text{NP}\boxed{2}[\textsc{acc}]_i \rangle\end{array}\right] \\
\textsc{content} & \textit{psoa}\left[\begin{array}{ll}\textsc{relation} & \textsc{see}\\ \textsc{seer} & j\\ \textsc{seen} & i\end{array}\right] \\
\textsc{context} & \ldots
\end{array}\right]
\end{array}\right]
$$

MORPH encodes information about the morphology of the word. The verb *vlepi* has a morphological root (ROOT), *vlep*, and an INFLECTED-FORM (I-FORM), *vlepi*.

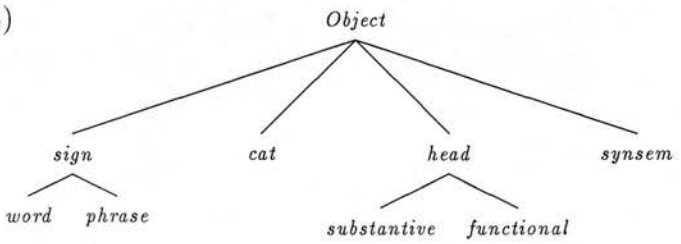
The attribute SYNSEM stands for SYNTAX-SEMANTICS. Its value, a feature structure of sort *synsem*, contains basic information about the category, the subcategorisation requirements and the semantics of the verb which are encoded locally (LOCAL) or non-locally (NONLOCAL; see Section 4.3 for the attribute NONLOCAL). LOCAL, in turn, is divided into CATEGORY (CAT), CONTENT and CONTEXT. The value of CAT is a feature-structure of sort *category*, with appropriate attributes HEAD, COMPS and ARG-STR. The value of HEAD provides the part of speech which in (1.22) is *verb*. Appropriate values for the attribute HEAD are objects of sort *head*, sorted according to the following hierarchy:

(1.23)



Head has two major subtypes: *substantive* and *functional*. Examples of *substantive* heads are *noun, verb, adjective and preposition*, while *determiner* and *complementiser* are instances of *functional heads*[3]. Note that the *head/part-of-speech* hierarchy is distinct from the *word* hierarchy (1.18). *Word* and *part-of speech* are subtypes of *Object* as illustrated in (1.24) (Pollard & Sag 1994):

(1.24)



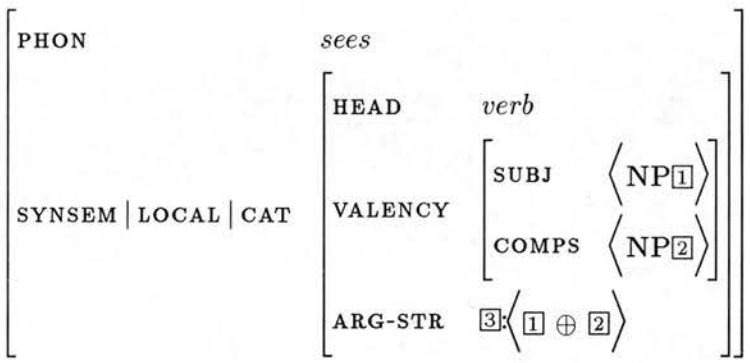
The ARG-STR is a *list* in which the arguments of the verb appear in order of obliqueness.

[3]In Section 4.4.2, I will introduce a finer organisation of functional heads, according to which, complementisers are a subtype of *marker*.

The members of the ARG-STR are (feature-structures) of sort *synsem*. This restriction captures the fact that heads select only the syntax-semantic features of their arguments. Their phonological and morphological features are not relevant for argument selection. *Synsem* feature structures correspond, roughly, to NPs, VPs, PPs etc. In (1.22), the verb has two arguments. The first argument corresponds to the subject and it is a nominative NP, whereas the second argument is the object, an accusative NP.

ARG-STR encodes a 'deeper' level of the organisation of arguments. Crosslinguistic differences in the actual (syntactic and/or morphological) realisation of the arguments are subject to *linking* constraints between ARG-STR and Phrase-Structure or Morphology (Manning & Sag 1995). In Greek, verbal arguments are realised as complements of the verbal head. This information is encoded in the COMPLEMENTS-LIST (COMPS). As Greek exhibits properties of a non-configurational language, the subject and the object are treated uniformily and both appear in COMPS. Thus, in Greek, the list of complements (COMPS) and the list of arguments (ARG-STR) are identical (indicated by the tag 3). However, in configurational languages like English, subject and objects are represented separately, as illustrated in (1.25):

(1.25)



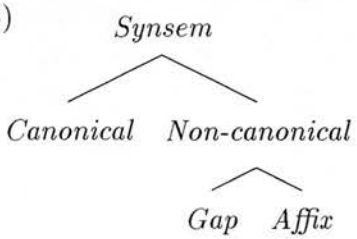
In (1.25), a new attribute, VALENCY, is introduced. VALENCY is divided into SUBJECT (SUBJ) and COMPS. Note that Greek and English verbs have the same members in their ARG-STR. In English the ARG-STR is obtained by appending (⊕) the list of COMPS to SUBJ, whereas in Greek the ARG-STR is identical with the COMPS[4].

Within the same language, the arguments of a verb may be realised in different ways. For example, in Greek, arguments may appear as 'in-situ' complements or may be extracted to

[4]The treatment of adjuncts will be discussed in detail in Section 4.5.1, where it will be argued that adjuncts should be analysed as complements and appear in COMPS.

the left periphery of the clause. The various possibilities of argument realisation are captured by subtypes of *synsem* (and constraints mapping ARG-STR to Phrase Structure—see Section 4.3). *Synsems* are subtyped with respect to their canonicity, as illustrated in the following hierarchy (Sag & Miller 1997):

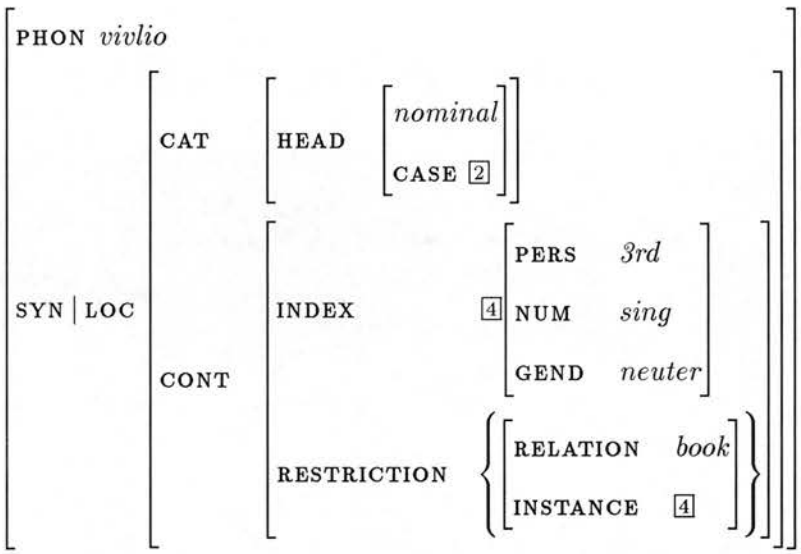
(1.26)



NPs of synsem sort *canonical* in the ARG-STR of a verb (1.22) correspond to in-situ complement NPs, present in COMPS. *Gaps* are relevant in the account of extractions, where it is assumed that the extracted argument is realised in the ARG-STR as a *gap* (see Section 4.3). *Affixes* will be employed in the analysis of pronominal clitics (Section 4.6.2).

Along with information for the category of the head and its argument structure, LOCAL encodes information about the semantics of the verb and some aspects of its contextual function. This information is represented in the values of the attributes CONTENT and CONTEXT respectively. CONTEXT will be employed for the representation of discourse functions and will be discussed at length in Chapter 5. The CONTENT attribute of *vlepi* ('sees') takes as its value a *parametrised-state-of-affairs* (*psoa*), which, in this case, describes a relation of *seeing* with a *seer* and a *seen*. Note that in (1.22) the two semantic roles have the same *index* (*j,i*) with the two arguments in the ARG-STR of the verb. The *index* of a NP contains information about PERSON, NUMBER and GENDER. In order to understand its function, a small digression to the internal structure of *nouns* and NPs is necessary. I will briefly describe the lexical entry of the Greek noun *vivlio* ('book'):

(1.27)



CAT|CASE is an appropriate attribute for a *sign* with a *nominal* head. *Nominal* is a subtype of *head/part-of-speech* and a supertype of *noun* and *adjective*—see Kolliakou (1995) for the details of the hierarchy of *nominal* in Greek. The value of the attribute CASE is of sort *case* with subtypes, *nominative, accusative, genitive* and *vocative* (for Greek). The attribute CONTENT of a *nominal* takes as a value a feature structure with the attributes INDEX and RESTRICTION. The value of INDEX, a structure of sort *index*, is subdivided in PERSON, GENDER and NUMBER. The *index* of a *nominal* represents the entity the *nominal* refers to. The attribute RESTRICTION takes as its value a *psoa* that poses semantic conditions on the *index*. In (1.27), the restriction states that, when the noun is used referentially, its *index* should be anchored to an entity which is a *book*. The details of the structure of nouns and NPs will not concern us here more. For the remainder, the abbreviation NP$\boxed{1}nom_i$ will refer to a NP with LOCAL value $\boxed{1}$, INDEX value i and CASE value *nominative*.

Let us return to (1.22). The semantic roles of SEER and SEEN are anchored to the indices of the corresponding arguments in the ARG-STR. This is a case of *structure-sharing*. *Structure-sharing* arises when two paths in a feature structure lead to one and the same node, i.e. they share their value. In (1.22) the SYNSEM|LOC|CONTENT|INDEX path in the argument NP$\boxed{1}$ and the SYNSEM|LOC|CONTENT|SEER have the same structure (j) as their common value. It is important to distinguish *structure sharing* from *type* or *structural identity*. The latter involves values that are identical feature structures. By contrast, *structure-sharing* involves *token-identity*. The lexical entry in (1.22) contains one more instance of *structure-sharing*. The LOCAL value of the members of the ARG-STR is *token-identical* with the LOCAL value of

the members of COMPS. Throughout, *token-identity* of LOCAL values will be indicated by tags with the same number (e.g. [1], [2]).

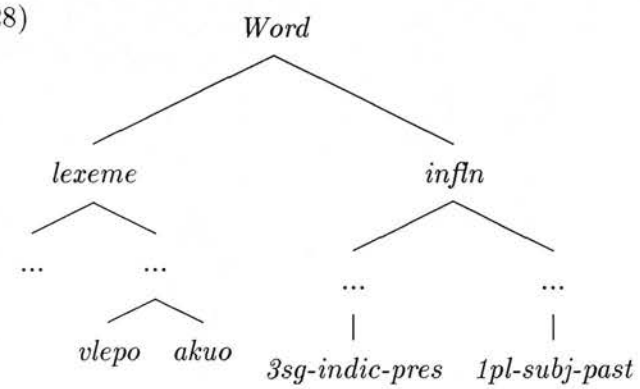
As will become evident, *structure sharing* is the main explanatory mechanism of HPSG. Various phenomena, such as the relationship between an extracted XP and its 'trace'/gap or between a word and the category of its phrasal projection will be analysed as instances of *structure-sharing.*

Constraints on Multiple Inheritance Hierarchies

The information appearing in a lexical entry is not listed separately for each entry. Rather, it arises from combined constraints on the various supertypes a *word* may belong to simultaneously. An example of the interaction between constraints on type-hierarchies and lexical entries was discussed in Section 1.3.1. The lexical entries of *verbs* inherit the restriction that they should take a subject from their supertype *verb* (1.19). The lexical entries of *transitive* and *intransitive verbs* inherit restrictions on the structure of their COMPS from their supertypes *transitive* and *intransitive-verb* respectively. In this section, I discuss some more aspects of the interaction between type-hierarchies and lexical entries.

Following Sag & Miller (1997), I assume that each inflected *word* belongs simultaneously to two compatible subtypes of *word*; a *lexeme* type and an *inflectional* (*infln*) type (these partitions are in addition to the *word* subtypes in 1.18):

(1.28)



It should be noted that, while an inflected *word* may belong simultaneously to more than one compatible subtype of *word*, there is nothing in the sortal hierachies specifying which types are compatible. For example, *noun* and *verb* are subtypes of *word* (1.18) along with *lexeme*

and *infln* (1.28). Nothing in the hierarchy states that *3sg-indic-pres* is compatible with a *verb* but not with a *noun*. These kinds of restrictions are ensured by the specifications of particular lexical entries. Roughly, no lexical entry will be specified as *3sg-indic-pres & noun* whereas many verb entries will be specified as *3sg-indic-pres & verb*.

The *lexeme* type of a *verb* provides basic information about its category, its arguments and the semantic roles of its arguments. Recent work (Davis 1997; Wechsler 1995) suggests that, to a large extent, the linking between ARG-STR members and semantic roles can be predicted on the basis of the semantics of the verb. Thus, *lexemes* are hierarchically organised on semantic grounds so that the mapping between ARG-STR members and semantic roles in CONTENT follows from the lexeme type of the verb.

In addition to syntactic-semantic information, the *lexeme* type of a verb provides its morphological root as a value to the attribute MORPH|ROOT. Consider (1.29), which illustrates the lexemic description of *vlep* (the root of *'see'* in Greek):

(1.29)
$$\left[\begin{array}{ll} \textsc{phon} & \textit{vlep} \\ \textsc{morph} & \left[\textsc{root} \quad \textit{vlep-}\right] \\ \textsc{synsem}\,|\,\textsc{local} & \left[\begin{array}{ll} \textsc{cat} & \left[\begin{array}{ll} \textsc{head} & \textit{verb} \\ \textsc{comps} & \langle \boxed{3} \rangle \\ \textsc{arg-str} & \boxed{3}\langle \text{NP}\boxed{1}[\textsc{nom}]_j, \text{NP}\boxed{2}[\textsc{acc}]_i \rangle \end{array}\right] \\ \textsc{content} & \left[\begin{array}{ll} \textsc{relation} & \textsc{see} \\ \textsc{seer} & j \\ \textsc{seen} & i \end{array}\right] \\ \textsc{context} & \ldots \end{array}\right] \end{array}\right]$$

The *inflectional* type specifies an inflectional form (I-FORM) for a given lexeme. Compare the lexemic description of *vlep* in (1.29) with (1.22) which describes an *inflected* form of the

lexeme *vlep*, *vlepi* (3rd sing). For convenience the lexical entry of *vlepi* is repeated below as (1.30):

(1.30)
$$\left[\begin{array}{ll}
\textsc{phon} & \textit{vlepi} \\
\textsc{morph} & \left[\begin{array}{ll} \textsc{i-form} & \textit{vlepi} \\ \textsc{root} & \textit{vlep} \end{array}\right] \\
\textsc{synsem}\,|\,\textsc{local} & \left[\begin{array}{ll}
\textsc{cat} & \left[\begin{array}{ll} \textsc{head} & \textit{verb} \\ \textsc{comps} & \langle \boxed{3} \rangle \\ \textsc{arg-str} & \boxed{3}\langle \text{NP}\boxed{1}[\textsc{3sg-nom}]_j,\ \text{NP}\boxed{2}[\textsc{acc}]_i \rangle \end{array}\right] \\
\textsc{content} & \left[\begin{array}{ll} \textsc{relation} & \textsc{see} \\ \textsc{seer} & j \\ \textsc{seen} & i \end{array}\right] \\
\textsc{context} & \ldots
\end{array}\right]
\end{array}\right]$$

The two AVMs differ in two ways. First, while the *lexeme* has only a MORPH|ROOT value, the *inflected word*, takes an I-FORM value (in addition to MORPH|ROOT). The second difference involves the first argument in the ARG-STR. In (1.29) the first argument is just a nominative NP. In (1.30), this NP is additionally marked *3rd person singular*. This is the consequence of the following constraint on verbs of sort *3rd-person-sing-verb* (Sag 1997)[5]:

(1.31)
$$\textit{3rd-person-sing \& verb} \rightarrow \left[\textsc{comps}\ \left\langle \text{NP}\boxed{1}[\textsc{3sg}], \ldots \right\rangle\right]$$

Thus, subject-verb agreement is captured by a constraint on the relevant inflected verb form.

As has become evident in this section, the lexical entries of *words* in HPSG are feature structures of considerable complexity since they encode information about different levels of grammar. However, this complexity is compensated for by the hierarchical organisation of the lexicon, which allows the expression of cross-cutting generalisations on appropriate types and their inheritance by their subordinate types. Thus, most of the information present in a

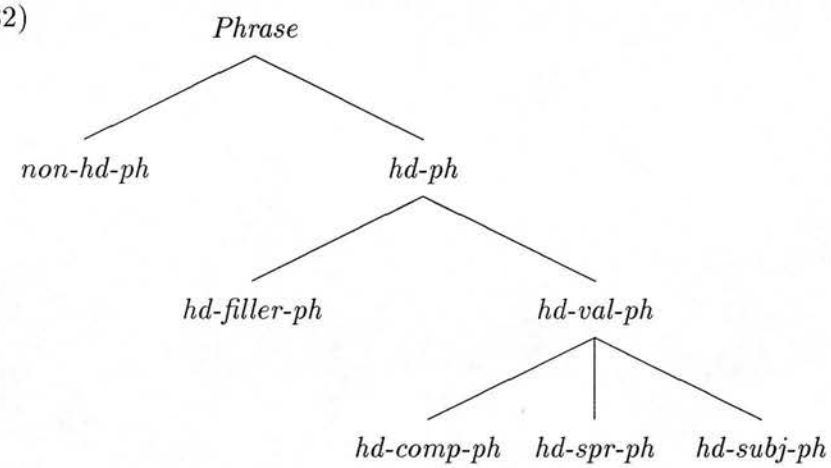
[5] In English this constraint involves the SUBJ whereas in Greek it involves the first member of COMPS.

lexical entry of a *word* is predicted by constraints on its supertypes.

1.3.3 Phrases

Syntactic constituents, *phrases*, are a subtype of *sign*. *Phrase*, in turn, has its own subtypes. The hierarchy in (1.32) illustrates some subtypes of *phrase* for English (Sag 1997):

(1.32)



Phrase is subdivided into *non-headed* (*non-hd-ph*) and *headed-phrase* (*hd-ph*). The latter, is broken into *head-filler-phrase* and *head-valence-phrase*. *Head-filler-phrases* are constituents where an extracted XP, the *filler*, combines with a clause containing a *gap* corresponding to the *filler*. *Head-valence-phrases* are constituents arising from a head combining with a complement (*head-complement-phrase*), a subject (*head-subject-phrase*) or a specifier (*head-specifier-phrase*). In addition to PHON and SYNSEM attributes, feature structures modelling *phrases* have attributes whose values describe the DAUGHTERS (DTRS) of the *phrase*. All *headed-phrases* have a HEAD-DTR attribute and one or more NON-HEAD-DTRS. Depending on the type of *phrase*, NON-HEAD-DTRS may be COMPLEMENT-DTRS, FILLER-DTRS etc. DTRS in turn, take *phrases* as values. Note that as a flat clause structure is assumed for Greek, *head-subj-phrase* is not a subtype of Greek *phrase*.

There are two general constraints on *headed-phrases*, the HEAD-FEATURE-PRINCIPLE (HFP) and the VALENCE PRINCIPLE (VALP) (Pollard & Sag 1994):

- HEAD-FEATURE-PRINCIPLE

  A HEAD-DTR's HEAD value is token-identical to that of its MOTHER.

- VALENCE-PRINCIPLE

  For each VALENCE feature F (COMPS, SUBJ..), the value of F in a *headed-phrase*, is the HEAD-DTR's F value minus the realised NON-HEAD-DTRS (COMPS-DTR, SUBJ-DTR).

HFP allows a MOTHER to inherit the HEAD value of the HEAD-DTR. In this way, *phrases* are 'projections' of their HEAD-DTRS. VALP ensures that a complement, subject or specifier, once realised as a COMPS-DTR, SUBJ-DTR or SPR-DTR is not inherited by the MOTHER.

Phrases are licensed through IMMEDIATE DOMINANCE SCHEMATA (ID-SCHEMATA) which specify the appropriate DAUGHTERs for each constituent and their properties. For example, the HEAD-COMPS-ID-SCHEMA licenses a *head-comp-phrase* (Pollard & Sag 1994). According to this schema, the HEAD-DTR is constrained to be of sort *word*. In addition, the LOC value of the COMPS-DTRs should be *token-identical* with the LOC value of the corresponding members in the COMPS of the HEAD-DTR. This is illustrated in AVM (1.33):

(1.33)

$$\begin{bmatrix} \textsc{synsem} & \begin{bmatrix} \textsc{head} & \boxed{1} \\ \textsc{comps} & \langle\rangle \end{bmatrix} \\ \textsc{head-dtr} & \begin{bmatrix} \textsc{head} & \boxed{1}(\textit{word}) \\ \textsc{comps} & \left\langle \boxed{2} \right\rangle \end{bmatrix} \\ \textsc{comps-dtr} & \begin{bmatrix} \textsc{loc} & \boxed{2} \end{bmatrix} \\ \ldots\ldots & \end{bmatrix}$$

In the above, the LOC value of the COMPS-DTR 2 is *token-identical* with the LOC value of the complement 2 in the COMPS of the HEAD. HFP guarantees that the MOTHER inherits the HEAD value 1 of the HEAD-DTR. VALP ensures that 2 does not appear in the MOTHER's COMPS.

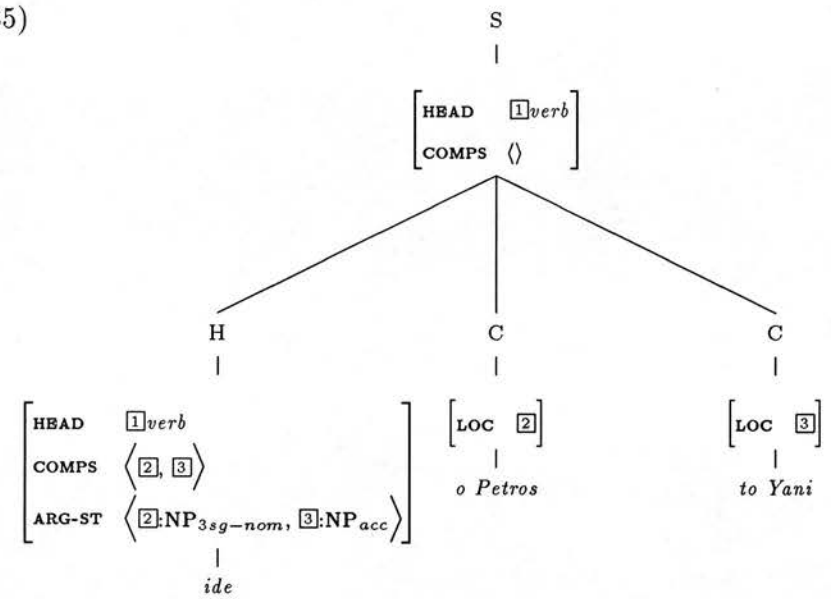
The HEAD-COMP-SCHEMA licenses an example of a VSO sentence in Greek (1.34):

(1.34) ide o Petros to Yani

saw-3SG the Petros-NOM the Yani-ACC
'Petros saw Yanis.'

The structure of (1.34) is shown in (1.35)

(1.35)



The structure in (1.35) is depicted as a tree for expository purposes. In 'reality', the structure of a *phrase* is an AVM, similar to AVMs describing *words*, (with some extra DTRS attributes). The tree in (1.35) corresponds to the AVM representation in (1.36):

(1.36)

$$\begin{bmatrix} \textbf{PHON} & \textit{ide o Petros to Yani} \\ \textbf{SYNSEM} \mid \textbf{LOC} \mid \textbf{CAT} & \begin{bmatrix} \textbf{HEAD} & \boxed{1}\textit{verb} \\ \textbf{COMPS} & \langle\rangle \end{bmatrix} \\ \textbf{HEAD-DTR} & \begin{bmatrix} \textbf{HEAD} & \boxed{1}\textit{verb} \\ \textbf{COMPS} & \langle \boxed{2}, \boxed{3} \rangle \\ \textbf{ARG-ST} & \langle \boxed{2}:\mathrm{NP}_{3sg-nom}, \boxed{3}:\mathrm{NP}_{acc} \rangle \end{bmatrix} \\ \textbf{COMPS-DTR} & \begin{bmatrix} \textbf{LOC} & \boxed{2} \end{bmatrix} \\ \textbf{COMPS-DTR} & \begin{bmatrix} \textbf{LOC} & \boxed{3} \end{bmatrix} \end{bmatrix}$$

However, for expository purporses, *phrases* will be displayed in the form of trees.

In (1.35), the H node stands for HEAD-DTR and the C nodes for COMPS-DTRS. The top node is an S. Technically speaking, there is no S in an HPSG grammar. For convenience, S

will be used throughout to refer to a *phrase* (of Greek) with a verbal head and empty COMPS list (S→HEAD:*verb*, COMPS: <>). This is an underspecified description, since it only refers to the value of the HEAD and COMPS. As a result, this description may satisfy various subtypes of *phrase*. For example, S may be a *head-comp-phrase* or, as it will be shown in Chapter 4, a *head-filler-phrase* or *head-marker-phrase*.

In configurational languages like English, S corresponds to a phrase with a verbal head and empty VALENCY (S→HEAD:*verb*, SUBJ: <>, COMPS: <>). On the other hand, VP corresponds to a phrase with empty COMPS and a non-empty SUBJ.

It is worth pointing out that, in HPSG, *phrases* are in essence no different from *words*. Both are subsorts of *sign* and their well-formedness depends on whether they satisfy constraints imposed on their various supertypes. In this respect, HPSG belongs to the family of *lexicalist* frameworks (Pollard & Sag 1994).

Chapter 2

Information Packaging

2.1 Introduction

The constructions this study is concerned with play an important role in the realisation of Information Packaging in Greek. To a large extent, their analysis depends on the understanding of the discourse functions they encode and the representation of these discourse functions in grammar. In this chapter I discuss various aspects of the discourse phenomena encoded through Focus-movement, Topicalisation and CLLD. In Section 2.2 I present the primitives of Information Packaging and adopt the four instruction types proposed in Vallduví (1992). In Section 2.3 I illustrate the linguistic realisation of the four instruction types in English and Catalan. English relies on accent shift for the realisation of Information Packaging while Catalan employs word order for this purpose. The realisation of Information Packaging in Greek, presented in Section 2.4, exploits both accent shift and word order. In this respect, Greek appears a mixed-type language, in common with various West Germanic languages and Turkish.

In addition to Information Packaging, accent placement and word order are often employed to encode types of pragmatic phenomena independent of Information Packaging. I discuss some examples of such uses of accent and word order in Section 2.5, with particular reference to NP de-accenting and marking of unexpected/surprising information. In the same section, I also present some interactions between the *cognitive status* of NPs and their potential to function as links or foci. In particular, I discuss the case of weak pronouns in English and object clitics in Greek and the potential of indefinite NPs to function as links.

In Sections 2.6 & 2.7 I argue for the independence of Information Structure from syntax and semantics and its representation at a distinct level of grammar. Drawing evidence from the Information Structure of sentences with embedded clauses, I argue that the ground-focus partition is not organised in a recursive way. The non-recursive nature of the ground-focus partition lends support to the view that Information Structure should be represented independently of syntax. Further, I discuss the relation of Information Structure with semantics with particular reference to the phenomenon known as *association with focus*. I adopt the argumentation of Vallduví (1992) and Vallduví & Zacharski (1994) who show convincingly that the ground-focus partition does not affect the truth conditions of a sentence. Information Structure should therefore be represented independently of semantics.

Finally, in Section 2.8, I present some properties wh-phrases share with focus, that will be important for the discussion that will follow in Chapter 3.

2.2 Information Packaging

Information Packaging refers to the structuring of sentences according to what speakers assume hearers would like/need to know at the time of the utterance. In particular, speakers use various linguistic cues to distinguish what they believe is new/informative material for the hearer from what they assume the hearer already knows. Thus, a sentence is partitioned into a *focus* segment, containing new-updating information and a *ground* segment anchoring the new information to already given/known/old/background information (Vallduví 1992). All sentences in (2.1) convey the same information. However, this information is *packaged* in different ways, according to which part of the sentence is focused each time (the focus part is indicated by square brackets):

(2.1) a. The pipes are [$_F$ RUSTY].

b. The pipes [$_F$ are RUSTY].

c. [$_F$ The PIPES are rusty].

d. [$_F$ The PIPES] are rusty.

e. The pipes [$_F$ ARE] rusty.

(Vallduví & Engdahl 1996:ex.2)

All examples in (2.1) have the same truth conditions. However, they do not have the same *felicity conditions.* That is, they are not interchangeable in a given context. A standard way of illustrating the differences in felicity conditions is by using question-answer pairs. Each of the examples in (2.1) is a felicitous answer to a different question:

(2.2) a. What about the pipes? In what condition are they?
The pipes are [$_F$ RUSTY].

b. What about the pipes? What's wrong with them?
The pipes [$_F$ are RUSTY].

c. Why does the water from the tap come out brown?
[$_F$ The PIPES are rusty].

d. I have some rust remover. You have any rusty things?
[$_F$ The PIPES] are rusty.

e. I wonder whether the pipes are rusty.
The pipes [$_F$ ARE] rusty.

(Vallduví & Engdahl 1996:ex.3)

This view of Information Packaging is the one proposed in Chafe (1976), Prince (1986) and Vallduví (1992) among others. Vallduví (1992), building on Chafe (1976) and Prince (1986), views each sentence as an instruction to the hearer on how to update his/her knowledge-store or *information state.* Before explaining the nature of these instructions, let us first consider the organisation of information states.

Information states are viewed as databases organised in *files*, a metaphor used in Heim (1982) and Heim (1983). Each file is a collection of *filecards.* Each filecard denotes an entity and contains a number of *records/conditions* listing attributes and relations holding for that entity. Communication involves updating of filecards.

According to Vallduví (1992), Information Packaging aims at signalling the new information conveyed by the sentence and, at the same time, ensuring that this information is anchored to the appropriate locus in the hearer's information state—in other words, each sentence is an instruction ensuring that the right information is added to the right filecard.

Vallduví (1992) proposes that each sentence has a level of organisation, *Information Structure*, encoding an instruction for the updating of the current information state. The Information Structure of a sentence arises from different combinations of the Information Packaging primitives in (2.3):

(2.3) S = { focus, ground }
Ground = { link, tail }

The focus part of a sentence contains the updating information, the information that is to be added on a specific filecard. Since all sentences have an update potential, all sentences have a focus segment. The ground contains already known information that acts as an anchor for focus, indicating where and how the new information should be added. In particular, the *link* points to a specific filecard in a file and so designates WHERE the new information should be added. While links specify WHERE, tails specify HOW updates should take place. The presence of a tail indicates that the new information cannot just be added to the filecard as a new condition/record. Rather, it should either complete or alter an already existing condition/record in the current filecard.

Different realisations of *the ground-focus* partition yield four different *instruction-types*:

1. focus
2. link-focus
3. focus-tail
4. link-focus-tail

The following examples illustrate the four instruction-types:

(2.4) a. **focus**
The president has a weakness.
[$_F$ He hates CHOCOLATE].

b. **focus**
So, did anything happen while I was away?
[$_F$ The PRESIDENT called].

c. **link-focus**

Tell me about the people in the White House. Anything I should know?

The president [$_F$ hates CHOCOLATE].

d. **focus-tail**

You shouldn't have brought chocolates for the president.

[$_F$ He HATES] chocolate.

e. **link-focus-tail**

And what about the president? How does *he* feel about chocolate?

The president [$_F$ HATES] chocolate.

(Vallduví & Engdahl 1996:ex.15-19)

Examples (2.4a&b) illustrate all-focus sentences. In (2.4a) there is no need for a link, a locus of update, because the locus of update is inherited from the previous discourse. The new information conveyed by (2.4a) is added to the filecard denoting the *president* which is already 'opened' in the previous sentence. In a similar manner, the locus of update is inherited from the previous discourse in (2.4d), which is also a linkless instruction. On the other hand, in examples like (2.4b), it is assumed that the sentence is anchored to the time and space information that is either mentioned in the previous discourse or is generally implied from the context of the conversation. Example (2.4c) instructs the hearer to go to the filecard denoting the *president* and add a new condition, *hates chocolate*. Finally, example (2.4e) instructs the hearer to look in the filecard *president* for a condition of the form *feels-like-about-chocolate* and replace the predicate *feels-like* with *hates*.

Locus of update and aboutness

An important characteristic of the model proposed in Vallduví (1992) is the trinominal nature of the ground-focus articulation. This model conflates the *focus-ground* partition (2.5c) with that of *topic-comment* (2.5b). Consider the following example:

(2.5) a. What about John? What does he drink?

b. [$_T$ John] [$_C$ drinks BEER]

c. [$_G$ John drinks] [$_F$ BEER]

d. [$_G$ [$_L$ John] drinks] [$_F$ BEER]

(Vallduví & Engdahl 1996:ex.11-14)

According to some proposals—Halliday (1967); Gundel (1988); Reinhart (1982) cited in Vallduví (1992)— example (2.5b) consists of a *topic* and a *comment*. One of the distinctive properties of topics is the *aboutness* feeling (Reinhart 1982). Topics are what the sentence is about. Typically, topics belong to the ground (2.5c) part of the utterance [but see Reinhart (1982)]. However, in the topic-comment articulation in (2.5b), the comment contains both ground and focus information (*drinks* and *beer* respectively). On the other hand, the ground-focus articulation in (2.5c) does not capture the intuition that *John* has a different status from the rest of the ground material. Vallduví (1992) circumvents this, by dividing ground into link and tail and proposing a tripartite hierarchical structure (2.5d). In (2.5d) the link *John* is distinguished structurally from the rest of the material in ground. However, a link is defined as the locus of update, not as what the sentence is about. Vallduví (1992) views the *aboutness* feeling of links as a result of their informational role. He notes: '...aboutness is treated as an epiphenomenon resulting from the very relation of links as address pointers with the informative part of the sentence: if the information is retrieved and entered under a given address, that information will be felt as being about the denotation of that address.' (Vallduví 1992:p.48).

Broad vs. narrow focus

A distinction often drawn in the discussion of focus, is that between *narrow* and *wide* or *broad* focus. Example (2.6c) illustrates a case of narrow focus while (2.6a-b) illustrate wide focus. Examples like (2.6a) are also described as *all-focus* sentences:

(2.6) a. Any news?
[$_F$ Mary bought a CAR].

b. What did Mary do?
Mary [$_F$ bought a CAR].

c. What did Mary buy?
Mary bought [$_F$ a CAR].

Similar distinctions can be drawn within NPs, as shown in (2.7):

(2.7) a. I didn't give him your camera, I gave him [$_F$ five FRANCS].

b. I didn't give him your three francs, I gave him [$_F$ FIVE] francs.

c. I didn't give him your five marks, I gave him five [$_F$ FRANCS].

(Ladd 1996:ex.5.1-4)

Contrast

A notion directly linked with discussions of focus is that of *contrast.* Some authors incorporate the notion of contrast in their definition of focus (Kiss 1998; Rooth 1996). However, Vallduví & Vilkuna (1997) and Vilkuna (1995) show convincingly that contrast is orthogonal to the ground-focus partition. Both foci and links may allow a contrastive reading once the context establishes a contrast between two or more discourse entities (Vallduví & Engdahl 1996; Vallduví & Vilkuna 1997). In this section I will limit the discussion to descriptive facts and I will not present the various accounts of the phenomenon, as it is not directly relevant to the arguments developed in this thesis.

Contrastive links

A typical case of contrastive links arises when two or more entities belong to a class already mentioned in the discourse and an explicit contrast is built between them. Example (2.8) illustrates such a case.

(2.8) Where can I find the cutlery?

a. *The forks* are in the CUPBOARD...

b. but *the knives* I left in the DRAWER.

(Vallduví & Engdahl 1996:ex.22)

The knives and *the forks* are members of *cutlery* and are explicitly contrasted in (2.8). The examples in (2.9) illustrate some more cases of contrastive links in English:

(2.9) a. **The first 100 meters** she ran [$_F$ in a record TIME].

b. **Beer** I [$_F$ LIKE].

(Vallduví & Vilkuna 1997:ex.7)

Vallduví & Vilkuna (1997) explain: 'The kontrastive nature of the *first 100 meters* and *beer* implies that membership sets play a role in their intrepretation. Indeed, in (7a) [2.9a here] the kontrastive element denotes a member of the set of parts of a track race. A natural follow-up would be, for example, a statement about the racer slowing down later in the race. In (7b) [2.9b here] the kontrastive element denotes a member of, for example, the set of alcoholic drinks. The sentence could be followed by something like *but whisky I hate*' (Vallduví & Vilkuna 1997).

Note that membership does not obligatorily lead to a contrastive interpretation. For example, consider (2.4c) repeated below as (2.10):

(2.10) **link-focus**

Tell me about the people in the White House. Anything I should know?

The president [$_F$ hates CHOCOLATE].

In (2.10) *the president* is a member of *the people in the White House.* However, *the president* is not explicitly contrasted to any other member of the group of *people in the White House.* Thus, it is not a case of contrastive link.

Contrastive vs. presentational focus

As in the case of contrastive links, a focused element is associated with a contrastive reading when it is contrasted with one or more entities in the discourse. Consider the following examples:

(2.11) a. Who did you see?

I saw [$_F$ John].

b. Who did you see, Peter or John?

I saw [$_F$ John].

c. I saw [$_F$ John], not Peter.

All of the above examples instantiate cases of narrow focus for *John.* In (2.11a) the focused element is not contrasted with any other entity in the discourse. This is a case of non-contrastive or presentational focus (King 1995). By contrast, in (2.11b-c) the focused NP,

John, is opposed to another NP, *Peter*. Example (2.11c) is referred to as *metalinguistic correction* (Ladd 1996; Vallduví & Engdahl 1996). In English there is no structural difference between contrastive and presentational focus. The examples in (2.11) are ambiguous between a contrastive and a non-contrastive reading. However, other languages distinguish the two cases structurally. As will be shown in Section 2.4, Greek is such a language [see also Kiss (1995b) for Hungarian, King (1995) for Russian, Choi (1996) for Korean, Vilkuna (1995) for Finnish and Rizzi (1995) for Italian].

2.3 The linguistic realisation of Information Packaging: plastic and non-plastic languages

Crosslinguistically, languages exploit various aspects of grammar to encode Information Packaging: accent placement, word order, morphology. In this section, I will briefly discuss two examples, English and Catalan. English relies on accent placement for the realisation of Information Packaging while Catalan exploits word order. The former is described as a *plastic* language whereas the latter as a *non-plastic* one. As will be shown in Section 2.4, Greek appears a hybrid case, as it uses both intonation and word order for the realisation of Information Packaging. Thus, the discussion of English and Catalan will prove relevant for the presentation of the Greek data.

English: a plastic language

Vallduví (1992), Vallduví (1995) and Vallduví & Engdahl (1996) show that in English Information Packaging is encoded through intonation. All examples in (2.12) have the same syntax. Shifting nuclear accent on different positions on the same string is enough to signal different ground-focus articulations (2.12):

(2.12) a. The pipes [$_F$ are [$_F$ RUSTY]].

b. [$_F$ [$_F$ The PIPES] are rusty].

c. The pipes [$_F$ ARE] rusty.

The Information Structure of examples (2.12a&b) is ambiguous between narrow and broad focus readings. The ambiguity of these examples is resolved by the context in which they appear.

In order to understand the role of intonation in the realisation of Information Packaging in English, let us briefly consider the notion of nuclear accent. Nuclear accent or *primary stress* is 'used to describe the pitch accent which stands out as the most prominent in an intonation-group' (Cruttenden 1986:p.49) or refers to 'the syllable in a tone unit which carries maximal prominence, usually due to a major pitch change' (Crystal 1991:p.238). These definitions imply that a sentence may contain more than one pitch accent. However, only the one which is 'most prominent' is associated with the nuclear accent. In English, nuclear accent is associated with a high tone (H*), referred to in the literature as **A-accent** (Steedman 1991; Vallduví 1992; Vallduví & Engdahl 1996). The words in SMALL CAPS in the above examples are associated with an A-accent and they are all focused.

In addition to A-accent, English allows a complex of a high tone preceded by a distinctive low level accent (L+H*), which is referred to as **B-accent**. Unlike the A-accent, the B-accent cannot be the nuclear accent of a sentence. It is a secondary accent that serves for the identification of links (Steedman 1991; Vallduví 1992). In the following example, the B-accented elements appear in italics:

(2.13) Where can I find the cutlery?

a. *The forks* are in the CUPBOARD...

b. but I left *the knives* in the DRAWER.

(Vallduví & Engdahl 1996:ex.23)

The answer to the question in (2.13) has a link-focus structure (the link is contrastive). Example (2.13a) is an instruction to add a condition in the filecard denoting *the forks* while in example (2.13b) the locus of update is the filecard denoting *the knives*. Thus, *the forks* and the *knives* are the links of each instruction. Both appear in situ, but they are marked intonationally with a B-accent.

In addition to being intonationally marked, links may be also marked syntactically. The link, *knives*, in (2.14b) bears B-accent and has undergone Topicalisation:

(2.14) Where can I find the cutlery?

a. *The forks* are in the CUPBOARD...

b. but *the knives* I left in the DRAWER.

(Vallduví & Engdahl 1996:ex.22)

Vallduví & Vilkuna (1997) argue that it is usually contrastive links that undergo Topicalisation in English.

In general, non-subject links, in situ or preposed, contrastive or not, are obligatorily associated with a B-accent. However, non-contrastive subject links may not be associated with a B-accent. Compare (2.14a) with (2.4c) repeated as (2.15):

(2.15) Tell me about the people in the White House. Anything I should know?
The president [$_F$ hates CHOCOLATE].

In both examples the link is the grammatical subject of the sentence. However, *forks* is a contrastive link wheras *the president* is a non-contrastive one. Since *the president* is a non-contrastive subject link it does not bear B-accent.

Tails in English are not marked either by means of syntax or phonology. They appear in situ and do not bear any kind of accent. For example, consider the tail, *chocolate*, in (2.4d&e). The examples in (2.16) illustrate some more cases of tails:

(2.16) a. [$_L$ *John*] loves [$_F$ BEER] (and *Mary* loves CIDER).

b. John [$_F$ LEFT].

(Vallduví & Engdahl 1996:ex.25-26)

In (2.16a) the tail is the verb, *loves*. In (2.16b) the subject, *John* may be interpreted as a tail; for example, (2.16b) may answer the question 'Why didn't she come with John?'. Thus, (2.16b) is ambiguous between a tail-focus and a link-focus interpretation.

Constituents bearing A-accent in English are focused (2.1a), (2.1d) & (2.1d). They are either the only focused constituent in cases of narrow focus, or part of the focus in cases of wide focus. On the other hand, unaccented constituents may also be focused in cases of broad focus (2.1c), (2.4a-c). Narrow and broad focus differ in a crucial way. Narrow focus can appear in any position/constituent of a sentence, as long as the relevant constituent has nuclear accent. Unlike narrow focus, wide focus may arise from accent on specific positions in a string (2.1c), (2.4a-c). Engdahl & Vallduví (1996) propose that in English wide focus arises from accent on the most oblique complement. For example, *the pipes* in (2.1c), *chocolate* in (2.4a&c) and

the president in (2.4b) instantiate the most oblique complement of the verb. By contrast, an accented verb (2.4e) or subject of a transitive verb (2.17) can only be interpreted as narrowly focused:

(2.17) Anything I should know?
[$_F$ The PRESIDENT hates chocolate].

This generalisation captures a large amount of cases. However, there are some cases that do not seem to be predicted. For example, consider (2.18):

(2.18) S_1: [narrating] Then, after lunch I laid out all the gifts on the table.
S_2: [interrupting] Oh, by the way, John had left a NOTE on the table. Did you see it?

(Vallduví & Engdahl 1996:ex.34)

The information structure of S_1 & S_2 in (2.18) is illustrated in (2.19):

(2.19) a. S_1: [$_F$ ...I laid out all the gifts on the TABLE].
b. S_2: [$_F$ John had left a NOTE] on the table.

Though the status of the locative is controversial, let us tentatively assume that it is the most oblique complement of the verb. In (2.19a) the accent falls on the most oblique complement and allows an all-focus interpretation of the sentence. In (2.19b) the accent does not fall on the most oblique complement. However, accent on the object gives rise to wide focus interpretation for the material preceding nuclear accent.

In sum, accent placement is the main strategy for the realisation of the ground-focus partition in English. With the exception of Topicalisation, ground and focus elements remain in situ. Shifting the two accents available in the language at different positions of invariable strings is enough to express different ground-focus articulations. Because of the flexibility of the accent placement, Vallduví (1992) calls English a *plastic* language.

Catalan: a non-plastic language

In contrast to English, Catalan is a *non-plastic* language. The nuclear accent must always fall on the rightmost clause boundary. Nuclear accent shift to any other position results in ungrammaticality:

(2.20) a. Ficarem el ganivet al CALAIX.
we-will-put the knife in-the drawer
'We'll put the knife in the drawer.'

b. *Ficarem el GANIVET al calaix.

c. *FICAREM el ganivet al calaix.

(Vallduví 1995:ex.8)

Instead, Catalan resorts to word order to signal discourse information. Ground information, the locative in (2.20b), the object and the locative in (2.20c) cannot remain within the core clause. The relevant constituents are obligatorily detached to either the left or the right periphery of the clause:

(2.21) a. Hi_1 ficarem el GANIVET t_1 al $calaix_1$.
locative we-will-put the knife t in-the drawer
'We'll put the KNIFE in the drawer.'

b. Al $calaix_1$ hi_1 ficarem el GANIVET t_1.

(Vallduví 1995:ex.9)

The detached ground constituents are coindexed with a clitic pronoun that remains within the core clause. In (2.21) the locative, *al calaix*, is coindexed with a locative clitic *hi*.

Links are detached to the left of the clause, tails to the right. Example (2.22) illustrates a case of contrastive links. The object link, *els ganivets*, is detached to the left of the core clause (2.22b). Example (2.22c), in which *els ganivets* is detached to the right, is not a felicitous answer to the question in (2.22), as indicated by the symbol @:

(2.22) On són, esl coberts?
'Where's the cutlery?'

a. Les forquilles són a l'armari, però...
'The forks are in the cupboard, but...'

b. ...els ganivets$_1$ els$_1$ vaig-ficar t$_1$ al **CALAIX**.
...the knives ogj 1s-pst-put t in-the drawer

c. @...els$_1$ vaig ficar al **CALAIX** els ganivets$_1$.

(Vallduví & Engdahl 1996:ex.39)

Non-contrastive links are also detached to the left:

(2.23) How come she's all alone?

a. El Joan [$_F$ se'n va **ANAR**].
the John [$_F$ refl.loc 3s-pst leave].
'John left.'

b. @[$_F$ Se'n va **ANAR**], el Joan.

(Vallduví & Engdahl 1996:ex.41)

According to Vallduví & Engdahl (1996), the question in (2.23) gives rise to a link-focus answer. The link *el Joan* has to appear to the left (2.23a), as indicated by the infelicity of (2.23b).

By contrast, tails should be detached to the right:

(2.24) Why didn't she come with John?

a. [$_F$ Se'n va **ANAR**], el Joan.

b. @El Joan [$_F$ se'n va **ANAR**].

The question in (2.24) requires a focus-tail answer. The tail has to appear to the right (2.24a)[1].

In sum, in Catalan accent always appears at the right boundary of the core clause, while links and tails are detached to the left and right respectively. All-focus sentences arise when all constituents remain in situ and accent falls on the rightmost boundary of the clause.

[1] It is not clear why *el Juan* in (2.23a) is a link and not a tail, while in (2.24a) it is a tail. In both cases the locus of update is the filecard referring to *she*. The link *el Juan* in (2.23a) seems to point to the filecard designating *Juan*, not a condition in the filecard of *she*.

2.4 Information packaging in Greek

Greek employs both accent placement and word order for encoding Information Packaging. Focus is realised either in situ or preverbally, through Focus-movement, and is associated with the nuclear accent of the sentence. Ground elements are not marked intonationally. Their realisation relies on word order. Links are encoded through Topicalisation and CLLD while tails are typically dislocated to the right edge of the clause. In addition, contrastive foci are distinguished structurally from non-contrastive ones. I will therefore consider the two cases of focus seperately.

2.4.1 Non-contrastive focus

In Greek non-contrastive focus appears postverbally, within the core clause. I will illustrate the facts using question-answer pairs. I should point out, though, that this paradigm is quite artificial. Answers repeating part of the question are quite unnatural. Even so, there are acceptability differences between the candidate answers that will serve for demonstrating the points made in this section.

Let us first consider cases of narrow focus. Presentational narrow focus appears postverbally, irrespective of grammatical function. In addition, the focused constituent bears the nuclear accent of the sentence. The following examples show cases of narrow focus for the subject (2.25), the object (2.26) and the adverb (2.27):

(2.25) **subject focus:**

pios apelise ti Maria;

'Who fired Maria?'

a. ti Maria tin apelise [$_F$ o YANIS]

the Maria-ACC her-CL fired-3SG [$_F$ the Yanis-NOM]

'Yanis fired Maria.'

b. @[$_F$ o YANIS] tin apelise ti Maria

@[$_F$ the Yanis-NOM] her-CL fired-3SG the Maria-ACC

(2.26) **object focus:**

pion apelise i Elena?

'Who did Elena fire?

a. i Elena apelise [$_F$ to YANI]
the Elena-NOM fired-3SG [$_F$ the Yanis-ACC]
'Elena fired Yanis.'

b. @[$_F$ to YANI] apelise i Elena
@[$_F$ the Yanis-NOM] fired-3SG the Elena-NOM

(2.27) **adverb focus:**

pote efiye o Yanis?
'When did Yanis leave?

a. o Yanis efiye [$_F$ HTES]
the Yanis-NOM left-3SG [$_F$ yesterday]
'Yanis left yesterday.'

b. @[$_F$ HTES] efiye o Yanis
@[$_F$ yesterday] left-3SG the Yanis-NOM

In the above, the (a) examples are felicitous because the focused XP appears postverbally. However, this does not mean that the preverbal appearance of the focused XP is absolutely infelicitous. But, in such an answer, the preposed XP is associated with a surprise or contrastive reading. I will return to this kind of reading in Section 2.4.2.

Narrow focus on the verb is obtained by accent on the verb:

(2.28) **verb-focus:**

'ti tha to kani o Yanis to aftokinito?
'What will Yanis do with the car?'
o Yanis [$_F$ tha to PULISI] to aftokinito
the Yanis-NOM [$_F$ will it-CL sell] the car
'Yanis will sell the car.'

Let us turn to cases of broad focus. As I have assumed that Greek has a non-configurational clause structure, I do not distinguish cases of VP focus from S focus. Instead, I consider cases of verb-subject and verb-complement focus, which, as will become evident, are realised in a similar manner.

First, consider cases of verb-subject focus:

(2.29) **verb-subject focus:**

a. ti eyine i efimerida?

'What happened to the newspaper?/where is the newspaper?'

[$_F$ tin pire o YANIS]

[$_F$ her-CL took-3SG the Yanis-NOM]
'Yanis took it.'

b. ti ehi to nero ki ine etsi?

'Why is the water like this?

[$_F$ tha 'hi skuriasi o SOLINAS]

[$_F$ will have got-rusty the pipe-NOM]
'The pipe must be rusty.'

In the above examples the subject appears postverbally and bears nuclear accent. In these examples the subject is the only complement of the verb. The following examples illustrate cases involving more than one complement:

(2.30) a. ti kani o Yanis?

'What is Yanis doing?

[$_F$ milai me ton Petro sto TILEFONO]

[$_F$ talk-3SG with the Petros-ACC at-the telephone]
'He is talking on the phone with Petros.'

b. ihes kana neo apo to Niko?

Have you had any news from Nikos?

[$_F$ ton idame me ti Maria sto SINEMA]

[$_F$ him-CL saw-1PL with the Maria at-the cinema]
'We saw him with Maria at the cinema.'

c. ti eyine i efimerida?

'What happened to the newspaper?/where is the newspaper?'

[$_F$ tin evale i Maria sto SIRTARI]

[$_F$ her-CL put-3SG the Maria-NOM at-the drawer]
'Maria put it in the drawer.'

As in (2.29), the verbal complements appear postverbally in (2.30). The accent falls on the rightmost boundary of the clause which is, usually, the position where the most oblique complement appears[2]. It is, typically, accent on this position that can give rise to a broad focus reading. Accent on the verb or on a constituent other than the rightmost/most-oblique does not give rise to wide focus readings:

(2.31) ti kani o Yanis?
'What is Yanis doing?

a. @[$_F$ MILAI me ton Petro sto tilefono]
@[$_F$ talk-3SG with the Petros-ACC at-the telephone]
'He is talking on the phone with Petros.'

b. @[$_F$ milai me ton PETRO sto tilefono]
@[$_F$ talk-3SG with the Petros-ACC at-the telephone]

The above examples are infelicitous because the accent falls on the verb (2.31a) and the least oblique complement (2.31b). So, they involve narrow focus readings for the accented constituents, *milai* and *Petro* respectively.

In sum, in cases of wide focus, verbal complements appear in situ and accent falls on the most oblique complement. Let us, finally, consider examples of all-focus sentences (examples (2.32c-d) are adapted from Philippaki-Warburton (1982)):

(2.32) **all-focus:**
kana neo?
Any news?

a. [$_F$ tilefonise o YANIS] ke ipe...
[$_F$ phoned-3SG the Yanis-NOM] and said-3SG...
'Yanis phoned and said...'

b. @[$_F$ o Yanis TILEFONISE] ke ipe...
@[$_F$ the Yanis-NOM phoned-3SG] and said-3SG...

[2]This is not always true. Quite often, the rightmost constituent bearing accent is not the most oblique verbal complement. I will return to this issue in Chapter 5.

c. [$_F$ pire tilefono o Yanis ti MARIA] ke tis ipe...
[$_F$ took-3SG phone the Yanis-NOM the Maria-ACC] and her-CL told-3SG
'Yanis phoned Maria and she told her...'

d. [$_F$ o Yanis pire tilefono ti MARIA] ke tis ipe...
[$_F$ the Yanis-NOM took-3SG phone the Maria-ACC and her-CL told-3SG

The question in (2.32) requires an all-focus answer. Examples (2.32a&c) in which the subject appears postverbally are the most natural answers to the question. Example (2.32b) with an SV order could be felicitous only if a phone call from Yanis was expected. However, example (2.32d) with an SVO order is a felicitous answer without implying that Yanis was expected to phone Maria. Vallduví & Engdahl (1996) note that questions like the one in (2.32) may give rise to VP rather than S focus. The subject may be interpreted as a link. It remains an open question why such a possibility is not available in (2.32b). [See also Ladd (1996) for similar contrasts in the felicity conditions of SV and SVO sentences in Italian.]

2.4.2 Contrastive focus

Preverbal focus in Greek is typically associated with a contrastive reading (Agouraki 1993; Tsimpli 1995; Tsiplakou 1998). As shown in the previous section, focused XPs corresponding to the wh-phrase in a Wh-question appear postverbally. However, if the Wh-question introduces an overt contrast between two entities, the focused XP may appear preverbally in the answer (and bear the nuclear accent). Compare (2.25,2.26&2.27) with (2.33,2.34&2.35):

(2.33) **subject contrastive focus**:
pios apelise ti Maria; o Petros i o Yanis?
'Who fired Maria; Petros or Yanis?'

a. [$_F$ o YANIS] tin apelise ti Maria
[$_F$ the Yanis-NOM] her-CL fired-3SG the Maria-ACC
'Yanis fired Maria.'

b. ti Maria tin apelise [$_F$ o YANIS]
the Maria-ACC her-CL fired-3SG [$_F$ the Yanis-NOM]

(2.34) **object contrastive focus:**

pion apelise i Elena; to Yani i to Niko?
'Who did Elena fire; Yanis or Nikos?

a. [$_F$ to YANI] apelise i Elena
[$_F$ the Yanis-NOM] fired-3SG the Elena-NOM
'Elena fired Yanis.'

b. i Elena apelise [$_F$ to YANI]
the Elena-NOM fired-3SG [$_F$ the Yanis-ACC]
'Elena fired Yanis.'

(2.35) **adverb contrastive focus:**

pote efiye o Yanis; htes i prohtes?
'When did Yanis leave; yesterday or the day before yesterday?

a. [$_F$ HTES] efiye o Yanis
[$_F$ yesterday] left-3SG the Yanis-NOM
'Yanis left yesterday.'

b. o Yanis efiye [$_F$ HTES]
the Yanis-NOM left-3SG [$_F$ yesterday]
'Yanis left yesterday.'

In the above, the (a) examples are not only acceptable, but preferred over the (b) ones with the postverbal XP.

In addition, the focused XP tends to appear preverbally in *metalinguistic corrections*:

(2.36) [$_F$ to YANI] apelise i Elena (ohi ton Petro)
[$_F$ the Yanis-NOM fired-3SG the Elena-NOM (not the Petros-ACC)
'Elena fired Yanis (not Petros).'

Note that preverbal focus cannot give rise to wide focus readings:

(2.37) *[$_F$ o YANIS tin apelise ti Maria]
*[$_F$ the Yanis-NOM] her-CL fired-3SG the Maria-ACC]
'Yanis fired Maria.'

In sum, Focus-movement is employed in Greek to encode narrow contrastive focus.

Though preverbal foci are typically associated with contrastive readings, they can also serve other pragmatic functions. In some contexts, they may imply surprise and various kinds of emphasis as shown in Tzanidaki (1994). The details of these phenomena will not concern us here. What is of interest with respect to this thesis, is that in all these cases, preverbal focus conveys the updating information of the sentence. In this respect, focus in Greek is realised both by means of accent placement and word order.

2.4.3 Links and tails

The question of the realisation of links and tails, as defined in Vallduví (1992), has not been addressed in the Greek literature. However, a significant part of the literature has been concerned with the pragmatic aspects of constructions involving what would be links and tails in Vallduví's terms. In this section I will briefly present the observations made in the Greek literature and I will then try to identify the realisation of links and tails in Greek.

Various authors argue that XPs undergoing Topicalisation (2.38a&c-d) or CLLD (2.38b) function as the topic/theme of the sentence (Agouraki 1993; Anagnostopoulou 1994; Philippaki-Warburton 1982; Philippaki-Warburton 1985; Schneider-Zioga 1994; Tzanidaki 1994; Tsimpli 1995; Tsiplakou 1998):

(2.38) a. tin parastasi skinothetise o Karolos KOUN
the perfomance-ACC directed-3SG the Karolos-NOM Koun
'Karolos Koun directed the perfomance.'

b. tin parastasi *tin* skinothetise o Karolos KOUN
the perfomance-ACC her-CL directed-3SG the Karolos-NOM Koun
'Karolos Koun directed the perfomance.'

c. me to treno taksideuoun oi fitites ke i TOURISTES
with the train travel-3PL the students-NOM and the tourists-NOM
'Tourists and students travel by train/Trains are used by students and tourists.'

d. sta nisia stelnun osous den ehun MESO
to-the islands send-3PL those-who-ACC not have-3PL 'connections'
'They send to the islands those who do not have any connections (to be sent to a better place).'

The preposed XPs are thought of as conveying old/given/background/discourse-linked information. In this respect, they can be viewed as links pointing to the locus of update. In addition, the literature shares the intuition that the sentence is *about* the preposed XP. Thus, the literature agrees that links in Greek tend to appear preverbally, involved in Topicalisation or CLLD structures.

On the other hand it is assumed that subjects and doubled objects dislocated to the right (2.39a) convey given information (Agouraki 1993; Anagnostopoulou 1994; Schneider-Zioga 1994; Tsimpli 1995; Valiouli 1994):

(2.39) a. ta **ESTILE** ta luludia o Yanis
them-CL send-3SG the flowers the Yanis-NOM
'Yanis sent the flowers.'

b. efiye **NORIS** o Manolis
left-3SG early the Manolis-NOM
'Manolis left early.'

Examples like (2.39) involve *Right Dislocation* of the object and subject NPs. In particular (2.39a) involves Clitic Right Dislocation since the dislocated object is coindexed with a clitic. Right Dislocated XPs appear after the nuclear accent. (This is so if the accent falls on a postverbal complement or on the verb. In Focus-movement, where the accent falls preverbally, postverbal constituents are not considered dislocated to the right.)

Let us consider the realisation of links and tails in more detail. Consider first the examples of contrastive links in (2.14):

(2.40) pu ine ta maheropiruna;
Where are the cutlery?

a. ta maheria ine sto **PROTO** sirtari ke ta pirounia sto **DEFTERO**
the knives are in-the first drawer and the forks in-the second
'The knives are in the first drawer and the forks in the second one.'

b. ta maheria ta vazume sto **PROTO** sirtari ke ta pirounia sto
the knives them-CL put-1PL in-the first drawer and the forks in-the
DEFTERO
second
'We put the knives in the first drawer and the forks in the second one.'

c. @ta vazume sto **PROTO** sirtari ta maheria ke sto **DEFTERO** ta
@them-CL put-1PL in-the first drawer the knives and in-the second the
pirounia
forks

The answer to the question in (2.40) builds a contrast between *knives* and *forks*. The two contrastive links appear preverbally and bear no accent. In addition, (2.40c), in which the links are dislocated to the right, is infelicitous. The following examples illustrate a similar case with indirect objects:

(2.41) pote tous ta stiles;
When did you send them to them?

a. ston Petro ta stila htes to **PROI** ke sto Dimitri
to-the Petros them-CL send-1SG yesterday the morning and to-the Dimitris
htes to **MESIMERI**
yesterday the midday
'I send them to Petros yesterday morning and to Dimitris yesterday afternoon.'

b. @ta stila htes to **PROI** ston Petro ke htes to
@them-CL send-1SG yesterday to-the Petros the morning and yesterday the
MESIMERI sto Dimitri
midday to-the Dimitris

Felicity judgements vary from context to context. For example, (2.41b) is somewhat better than (2.40c). However, there is a strong tendency for contrastive links to appear preverbally.

Let us turn to non-contrastive links and tails. The questions used to identify links and tails in English are those in (2.5a) and (2.4d-e) repeated in (2.42):

(2.42) a. **link-focus**
What about John? What does he drink?
[$_G$ [$_L$ John] drinks] [$_F$ **BEER**].

b. **focus-tail**
You shouldn't have brought chocolates for the president.
[$_F$ He **HATES**] chocolate.

c. **link-focus-tail**

And what about the president? How does *he* feel about chocolate?

The president [$_F$ HATES] chocolate.

The first problem with these diagnostics is their artificial nature. Links point to a filecard. However, the filecard in question is already open by the time the question is uttered. There is no obvious need for it to be repeated in the answer. The same holds for tails. Tails indicate a specific condition in a filecard (e.g. *feels about chocolate*). This condition is already present in the question.

The second problem involves the assumption that (2.42b) gives rise to a focus-tail sentence whereas (2.42c) to a link-focus-tail answer. There is no obvious reason why *the president* is more felicitous in (2.42c) than in (2.42b). In Greek, the corresponding examples take the same set of answers/follow-up sentences. I will therefore use only one of them.

Let us begin by considering link-focus utterances:

(2.43) **link-focus**

ya pes mu ya to Yani; pos ta pai me tus sinadelfus tu?

Tell me about Yanis; how is he getting along with his colleagues?

a. to Yani [$_F$ ton agapane OLI]; ine hriso PEDI

the Yani-ACC [$_F$ him-CL love-3PL all-NOM]; is gold guy

'Everybody loves Yanis; he is a nice guy.'

b. [$_F$ ton agapane OLI] to Yani; ine hriso PEDI

[$_F$ him-CL love-3PL all-NOM] the Yani-ACC; is gold guy

c. o Yanis [$_F$ aresi s' OLUS]; ine hriso PEDI

the Yanis-NOM [$_F$ like-3SG to all]; is gold guy

'Everybody likes Yanis; he is a nice guy.'

d. [$_F$ aresi s' OLUS] o Yanis ; ine hriso PEDI

[$_F$ like-3SG to all] the Yanis-NOM; is gold guy

Though there is a preference for the (a) examples, in which the link is preposed (through Topicalisation or CLLD), the (b) examples (which involve Right Dislocation and Clitic Right Dislocation) are also felicitous.

Let us consider now focus-tail examples:

(2.44) den eprepe na tis feris SKILADIKA tis Elenis
'You shouldn't have brought 'dog-songs' to Eleni.'

a. **focus-tail**
[$_F$ ta SIHENETE] ta skiladika
[$_F$ them-CL detest-3SG] the dog-songs
'She detests 'dog-songs'.'

b. @ta skiladika [$_F$ ta SIHENETE]
@the dog-songs [$_F$ them-CL detest-3SG] the dog-songs

The infelicity of (2.44b) indicates that tails in Greek should appear after focus, dislocated to the right.

The sentence in (2.44) may also have the sentences in (2.45) as a follow-up:

(2.45) a. **link-focus-tail**
(i kopela) [$_F$ ta SIHENETE] (i kopela) ta skiladika (i kopela)
(the girl-NOM) [$_F$ them-CL detest-3SG] (the girl-NOM) the dog-songs (the girl-NOM)
'The girl detests 'dog-songs'.'

b. i kopela ta skiladika [$_F$ ta SIHENETE]
the girl-NOM the dog-songs [$_F$ them-CL detest-3SG] the dog-songs

c. ?@ta skiladika [$_F$ ta SIHENETE] i kopela
?@the dog-songs [$_F$ them-CL detest-3SG] the girl-NOM

The supposed link, *i kopela*, may appear either preverbally or postverbally. However, the preferred order is the one in which *i kopela* is dislocated to the right. Note as well that, unlike (2.44b), example (2.45b) is felicitous and (2.44c) is better.

It is unclear what is the right interpretation of the data in (2.45). It is possible that the appearance of *i kopela* and *ta skiladika* encodes pragmatic functions independent of the ground focus partition. As already mentioned, in these examples, there is no need for either a link or a tail. However, the repetition of the relevant NPs is not without reason. By repeating those two NPs, the speaker adds emphasis to the content of the sentence and emphasizes the mistake of the hearer. Of course, this might explain their repetition, but not their order. Valiouli (1994) discusses the pragmatic content of Right Dislocation of full NPs in Greek. She

suggests that Right Dislocated phrases, in addition to introducing a *latent discourse topic*, often serve various other pragmatic functions, such as the speaker's empathy or contempt. Since there is no need for a link and a tail, it is possible that, in these examples, word order is used to encode pragmatic effects of the type described in Valiouli (1994). As will be shown in Section 2.5.5, accent in English and both accent and word order in Greek are often used to encode pragmatic distinctions orthogonal to the ground-focus partition.

The study of the factors affecting word order in examples like (2.43), (2.44) & (2.45) requires systematic research. However, in this thesis I assume that links in Greek are realised preverbally and precede focused elements, while tails appear after focus, dislocated to the right. I base this assumption on the evidence from contrastive links which are preverbal, and the data in (2.43) and (2.44) which suggest that links are preferred preverbally and tails dislocated to the right.

Note finally, that ground elements are preferably dislocated outside of the focus part of the utterance. This is indicated by the infelicity of (2.46) as the answer to (2.43):

(2.46) @[$_F$ ton agapane to Yani OLI]; ine hriso PEDI
@[$_F$ him-CL love-3PL the Yani-ACC all-NOM]; is gold guy

Some crosslinguistic comparisons

As mentioned in Section 2.3, Vallduví (1992) draws a distinction between plastic and non-plastic languages, the prototypical examples of which are English and Catalan respectively. Greek appears as a hybrid case since it uses both accent shift and word order to encode Information Packaging. Links as well as narrow foci may appear preverbally (2.33,2.34,2.35) while tails tend to be dislocated to the right. Like English, accent may fall on any position within the core clause (2.28) or even on a preverbal focused XP (2.33,2.34,2.35). Greek is not unique in its hybrid properties with respect to the realisation of Information Packaging. Various other languages exploit both accent shift and word order for this purpose; Dutch and Turkish are two such examples discussed in Vallduví & Engdahl (1996). See also Eckert (1998a) and Eckert (1998b) for German.

In cases of wide focus, accent falls on the rightmost boundary of the clause in Catalan while in English it falls on the most oblique complement. Though these two generalisations

appear to be the most economical for the description of facts in English and Catalan respectively, they seem to obscure the fact that the two languages share more than these two generalisations suggest. Crosslinguistically, wide focus appears to be associated with accent on the most oblique complement of the clause. In English and Greek, in the absence of any adjuncts, the object of a transitive verb is the most oblique complement that coincides with the rightmost boundary of the clause. In this respect, English and Greek are not radically different from Catalan (modulo the fact that Catalan is VOS; thus, the rightmost complement is the subject which is not the most oblique one). In sentences involving intransitive verbs, broad focus is associated with accent on the subject which is the most oblique complement of an intransitive verb. In Greek, Catalan and Italian these accented subjects appear postverbally as the rightmost element of the clause. In this case, associating accent with the most oblique complement is equivalent to associating accent with the rightmost boundary of the clause. By contrast, in English, the subject of an intransitive verb appears preverbally and accent is associated with the leftmost rather than rightmost element.

Note, finally, that wide focus is also associated with accent on the object in SOV sentences for example in Hungarian, Turkish, German and Korean (Choi 1996; Kiss 1995b; Ladd 1996; Horvath 1995; Hoffman 1995)[3].

2.5 Pragmatic effects independent of the ground-focus partition

2.5.1 The Cognitive Status of NPs

Independently of their function as ground or focus elements, the NPs participating in a discourse encode an inherent degree of novelty/familiarity (Heim 1983; Gundel *et al.* 1993; Vallduví & Engdahl 1996). In (2.47), all object NPs are focused; they contribute the update information of (2.47):

(2.47) Who did you see?
I saw [$_F$ a boy/the boy/John].

[3] I also owe the crosslinguistic data and some of the observations made here to the Focus seminar at Edinburgh University (1995-96) organised by D.R.Ladd and B.Hoffman.

Though all focused, the object NPs in (2.47) differ in their *cognitive status.* The indefinite NP is thought of as more novel/less familiar than the definite NP and the proper name. Vallduví & Engdahl (1996) note: 'simultaneously, yet independently, the marking of cognitive status (familiarity/novelty) of discourse referents is responsible for providing the hearer with instructions for file-card management (very roughly, an indefinite NP instructs the hearer to create a new file card [denotes a novel referent], while a definite NP instructs the hearer to activate a dormant, already existing file card [denotes a familiar referent])' [ibid, p.469].

Though Information Packaging and cognitive status are distinct, they often interact in interesting ways. Weak pronouns in English provide an example of such an interaction. Gundel *et al.* (1993) propose a *givenness hierarchy* which correlates with particular forms of referring expressions. Weak pronouns occupy the highest position in this hierarchy, as the most *given* entities in a discourse. They are followed by demonstratives, definite NPs, referential indefinite NPs and indefinite NPs. Consider now examples (2.48) in which the weak pronouns are de-accented (2.48):

(2.48) a. I SAW him.

b. I'll go WITH him.

Vallduví (1992) argues that weak pronouns in English do not contribute to the Information Structure of a sentence. Instead, they function as syntactic placeholders. Both examples in (2.49) have the same syntax and intonation contour. However, they have different Information Structures:

(2.49) a. John [$_F$ LOVES] beer.

b. [$_F$ He LOVES it].

(Vallduví & Engdahl 1996:ex.30)

While (2.49a) has a non-canonical prosody, (2.49b) has a canonical one, as suggested by the awkwardness (2.50b):

(2.50) a. John loves BEER.

b. ?He loves IT.

(Vallduví & Engdahl 1996:ex.31)

Vallduví (1992) proposes that unlike (2.49a) which consists of a ground and a focus part, example (2.49b) instantiates an all focus sentence in which the pronouns are just syntactic (argument) placeholders. This is an idiosyncrasy of English. According to Vallduví (1992), languages that allow subject and object drop (2.49b) would just contain the verb. Catalan is an example of such a language.

Examples like the following are interpreted along similar lines:

(2.51) Have you seen the newspaper?
[$_F$ JOHN took it].

At first sight, example (2.51) is a counterexample to the generalisation that, in English, wide focus is associated with accent on the most oblique complement. Technically speaking, the most oblique complement in (2.51) is the object which appears in the form of a weak pronoun. Because weak pronouns do not participate in the Information Structure of a sentence and resist accent, the accent shifts to the subject and gives rise to an all-focus interpretation.

Greek is a pro-drop language. Thus, as Vallduví (1992) predicts for pro-drop languages, example (2.52) may have an all-focus reading:

(2.52) EFIYE
left-3SG
'S/he left.'

There is also a class of object clitic pronominals which can neither be accented (2.53b) nor appear in focus contexts (2.53a):

(2.53) Pjon idhe?
whom saw-he
'Who did he see?'

a. @Me idhe
@me-CL saw
'He saw me.'

(Anagnostopoulou 1994:ex.39)

b. *ME idhe
*me-CL saw

Like its English counterpart (2.49b), the clitic in (2.54) refers to a given entity in the previous discourse:

(2.54) Ti kanate me tis valitses?
What did you do with/arrange for the luggage?
[$_F$ tha erthun na *tis* parun avrio to **PROI**]
[$_F$ will come-**3PL** to them-**CL** take-**3PL** tomorrow the **MORNING**]
'They'll come to pick them up tomorrow morning.'

The clitic *tis* in (2.54) refers to *valitses* (luggage). However, it does not contribute to the Information Structure of (2.54), as it neither acts as a link, pointing to a filecard, nor as a tail, pointing to a condition in a filecard. At best, it refers to the current filecard, *valitses*. Thus, following Vallduví (1992), I will also assume that clitics in Greek do not contribute to the Information Structure of a sentence. However, their presence cannot be reduced to syntactic requirements only. Object drop is also possible in Greek. Keller & Lapata (1998) show that the presence/drop of object clitics is subject to discourse restrictions [see also Dimitriadis (1994a)]. However, the nature of these discourse restrictions is independent of the ground-focus partition and, therefore, will not concern us here.

In conclusion, weak pronouns crosslinguistically, do not contribute to the Information Structure of the sentences they appear in.

2.5.2 The independence of Clitic Doubling from Information Packaging

The Greek literature treats doubled objects as ground elements (Agouraki 1993; Anagnostopoulou 1994; Iatridou 1995; Philippaki-Warburton 1982; Schneider-Zioga 1994; Tsimpli 1995; Tsiplakou 1998; Tzanidaki 1994; Valiouli 1993). This view is based on the fact that doubled objects resist accent and cannot function as foci[4]:

(2.55) *(to **YANI**) *ton* idame (to **YANI**)
*(the Yani-**ACC**) *him*-**CL** saw-**1PL** (the Yani-**ACC**)
'We saw Yanis.'

As was shown in Section 2.4.3, doubled objects may function as links in CLLD or as tails in Clitic Right Dislocation. (In Chapter 5, it will be shown that also objects in Clitic Doubling

[4] In some contexts they can marginally be accepted as part of wide focus.

structures are typically interpreted as tails. The differences between Clitic Doubling and Clitic Right Dislocation will be discussed in Sections 3.3.5 & 5.2.4). However, object links are not necessarily doubled. In (2.56a) the object link is not coindexed with a clitic:

(2.56) a. ton prothupurgo upodehtike o ipurgos **EKSOTERIKON**
the prime-minister-**ACC** accepted-**3SG** the minister-**NOM** foreign-affairs-**GEN**
'The minister of foreign affairs welcomed the prime minister.'

b. ton prothupurgo *ton* upodehtike o ipurgos **EKSOTERIKON**
the prime-minister-**ACC** *him*-**CL** accepted-**3SG** the minister-**NOM** foreign-affairs-**GEN**

Both examples in (2.56) have the same Information Structure, instantiating a link-focus instruction. The preposed object functions as a link in both cases, irrespective of whether it is doubled or not. Examples (2.56) indicate that doubling is distinct from Information Packaging.

It is not clear what are the interpretational differences between (2.56a&b) and what is the exact effect of doubling. A plausible hypothesis is that doubling affects the cognitive status of the doubled NP. Let us consider the case of doubled definite NPs. According to the givenness hierarchy of Gundel *et al.* (1993), definite NPs are less given/familiar than pronouns but more given/familiar than indefinite NPs. A doubled definite is an ordinary definite NP construed with a pronoun. It is possible that doubling increases the givenness/familiarity of the relevant NP. However, this hypothesis will not be pursued here. For the purposes of this study, the following two generalisations suffice as a conclusion:

- Doubled NPs are ground elements and resist accent.
- Doubling is independent of the ground-focus partition.

2.5.3 Doubled objects, links and referentiality

In this section I will consider another example of the interaction between Information Packaging and the cognitive status of the NPs involved. The literature suggests that there are semantic restrictions on the kinds of NPs that can function as topics. These restrictions involve the cognitive status of link/topic NPs. Interestingly, the same restrictions have been claimed to apply to Clitic Doubling Constructions. Given that doubled elements are ground elements, and, in particular, that CLLDed objects are links, it is not surprising that the same

restrictions have been claimed for both cases. However, I will consider the two generalisations independently:

a Topics (links,themes) should be definite/referential/specific or generic (Kiss 1995b; Philippaki-Warburton 1985).

b Doubled objects are definite/referential/specific or generic (Anagnostopoulou 1994; Cinque 1990; Dobrovie-Sorin 1990; Iatridou 1995; Niño 1994; Philippaki-Warburton 1985).

Thus, non-referential indefinites are excluded from Clitic Doubling constructions and cannot function as links.

Let us first consider Clitic Doubling constructions:

(2.57) a. *ena doro theli na to vri

a gift wants-he to it find-he
'He wants to find a gift.'

b. ena filo tu psaxni na ton vri apo htes

a friend his searches-he that him fiends-he since yesterday
'A certain friend, he is looking for him since yesterday.'

(Philippaki-Warburton 1985:ex.26,28)

Philippaki-Warburton (1985) attributes the acceptability of (2.57b) to the fact that the indefinite has a specific/referential reading. By contrast, in the unacceptable (2.57a), the doubled object cannot be interpreted as specific. On the basis of (2.57), Philippaki-Warburton (1985) concludes that indefinites cannot be doubled. However, in the same paper, she offers the following examples which involve grammatical cases of doubled indefinites:

(2.58) a. ena ouzaki tha to pina efharistos

an ouzo would it drink-I gladly
'As for an ouzo, I would have one gladly.'

b. ena mpanaki tha to kana efharistos

a bath would it do-I gladly
'As for an bath, I would love to take one.'

(Philippaki-Warburton 1985:ex.22-23)

As she admits in a footnote, the doubled objects in (2.58) are not referential/specific. She notes, though, that they are quite rare and function as the second clauses of hypothetical clauses of the type *if you were to offer an ouzo, I wouldn't say no to it.* They are, therefore, associated with modal readings which, according to Philippaki-Warburton (1985), render the examples irrelevant for the discussion of the availability of indefinites in Clitic Doubling constructions. In a similar spirit, Anagnostopoulou (1994) discusses the Clitic Doubling counterpart of (2.58a):

(2.59) tha to pina efharistos ena ouzaki
would it drink-I gladly an ouzo
'I would gladly have an ouzo.'

(Anagnostopoulou 1994:ex.1,fn 4)

She suggests that (2.59) has a modal reading and, as a result, it is not relevant to the discussion of the semantic restrictions on Clitic Doubling. However, neither Philippaki-Warburton (1985) nor Anagnostopoulou (1994) explain why modal readings render these examples irrelevant.

The following two examples illustrate some more cases of doubled indefinites:

(2.60) a. (mia ekdromi) tin eho poli anagki (mia ekdromi)
(a trip-ACC) her-CL have-1SG very necessity (a trip-ACC)
'I need a lot a trip.'

b. KANENA$_i$ den ton$_i$ agapai i pethera tu$_i$
nobody-acc not him-cl love-3sg the-nom mother-in-law-nom his-gen
'His (own) mother-in-law loves nobody/ Nobody is loved by his own mother-in-law.'

Undoubtedly the modality of the verb affects the interpretation of their indefinite objects (in general, not just doubled ones). However, examples (2.58,2.59&2.60) are instances of doubled indefinites which, unless an explanation is provided, cast serious doubts on the validity of the generalisation in (b).

Let us turn to the claim that indefinite links/topics are impossible. A natural question to ask is whether the preposed indefinites in (2.58&2.60a) can satisfy the definition of link,

as described in Vallduví (1992)[5]. In Vallduví's view, the distinguishing property of links is that they designate the locus of update. They instruct the hearer to open the appropriate filecard to add the information conveyed by the sentence. According to Vallduví & Engdahl (1996), a definite link should point to an already existing filecard, whereas an indefinite link instructs the hearer to create a new filecard. For example, in (2.61a) the hearer should open the already existing filecard of the speaker's husband whereas in (2.61a) the hearer should create a new filecard denoting 'a friend of the speaker' and add the relevant information.

(2.61) a. ton antra mu ton epiasan na pernai me kokino ke...
the man-ACC my-GEN him-CL caught-3PL to cross-3SG with red and...
'They caught my husband crossing the street while the red light was on and... My husband was caught crossing the street while the red light was on and...'

b. ena filo mu ton epiasan na pernai me kokino ke...
a friend-ACC my-GEN him-CL caught-3PL to cross-3SG with red and...
'They caught a friend of mine crossing the street while the red light was on and... A friend of mine was caught crossing the street while the red light was on and...'

Despite their differences, both NPs in (2.61) act as pointers to the locus of update[6]. As mentioned in Section 2.2, Vallduví's links also incorporate the traditional notion of aboutness associated with topics. The links in (2.61) satisfy this aspect of linkhood as well, as they are what the sentence is about.

Let us turn to the doubled indefinites in (2.58&2.60a) and compare them with the links in (2.61). The indefinites in (2.58&2.60a) have the syntactic trappings of links since they participate in CLLD. However, it is unlikely that these NPs designate the locus of update and instruct the hearer to create a new filecard denoting *ouzo* or *excursion*. Intuitively, the locus of update is the speaker who informs the hearer about his/her wishes and needs. At

[5]In (2.60b) the doubled quantifier bears the nuclear accent. I will leave open the question whether *kanenas* in this example should be analysed as a link or as a focused element. Some Greek quantifiers (e.g. *oli=all, everybody*) attract accent even though they undergo doubling, and can function as foci.

[6]Of course, it is possible to imagine a context in which the hearer does not know that the speaker is married and has a husband. In such a case, the definite link would involve the creation of a new filecard, just in the same way an indefinite/referential link does. In relation to this, Hendriks & Dekker (1995) note that the filecard representations proposed in Heim (1982) and Heim (1983) cannot accommodate the differences between definite, indefinite and various quantified NPs. Instead, they propose DRT representations for the organisation of information states.

best, the preposed indefinites designate part of a specific condition/record in the speaker's filecard (e.g. *feel-like-about-ouzo*). If this is so, then, in Information Packaging terms, they are tails rather than links. However, these indefinites satisfy the more traditional notion of topichood, that is, they give rise to the aboutness relation. According to Vallduví's definition, this is only a secondary and epiphenomenal property of links. However, the examples in (2.58&2.60a) suggest that this might not be so. While it is reasonable to assume that the locus-of-update is what the sentence is about, that is the topic, it is not obvious why the topic of a sentence should always be the locus-of-update (given that very often sentences are linkless). The examples in (2.58&2.60a) illustrate cases in which the sentence topic is not the locus of update and does not point to any filecard. Rather, what seems to be the case is that non-referential indefinites cannot function as the locus-of-update but they can function as topics, that is the entity the sentence is about.

The notion of topic is a very controversial one and its proper definition stays beyond the aims of this thesis. In this section I have tried to present some interactions between the cognitive properties of NPs and their potential to function as links. In the remainder of this study, I will use the terms topic/link invariably in either sense, that is as the entity being the topic of the sentence or designating the locus-of-update and, most often, both. I will also tentatively assume that the indefinite NPs in (2.58&2.60a) are links, since they satisfy the aboutness relation which, though secondary, is a property of Vallduví's definition of link. In any case, this assumption will not affect the syntactic arguments presented in Chapters 3 & 4. Note, finally, that the assumption that the indefinite NPs in (2.58&2.60a) are links/topics implies that the generalisation in (a) is an invalid one.

2.5.4 NP de-accenting

In English NPs, the accent may shift from the noun to the adjective in examples like Clara's answer in (2.62b):

(2.62) a. Ann: What did you get Ben for Christmas?
Clara: I got her [$_F$ a blue SHIRT.]

b. Ann: What did you get Diane?
Clara: I got him [$_F$ a RED shirt]

(Vallduví & Zacharski 1994:ex.22)

In both answers in (2.62) the focus segment of the sentence is the object NP. In (2.62a) the accent falls on the noun, whereas in (2.62b), it falls on the adjective. The phenomenon is known as *de-accenting* (Ladd 1980; Ladd 1996; Steedman 1991; Vallduví & Zacharski 1994). The de-accenting in Clara's answer in (2.62b) is triggered by the presence of the NP *blue shirt* in (2.62a).

In Greek, NP de-accenting is realised by accent shift, as in English, or by the use of a so-called *polydefinite* (Kolliakou 1998b):

(2.63) Zoe: Ti pires tu Yanni gia ta Xristugena?
'What did you get Yanis for Christmas?'
Daphne: Tu pira [$_F$ tin asimenia PENA].
Daphne: him-CL got-1SG [$_F$ the silver pen]
'I got him the silver PEN.'

(Kolliakou 1998b:ex.4a-b)

(2.64) Zoe: Ti pires tis Marias?
'What did you get Maria?'

a. Daphne: Tis pira [$_F$ ti HRISI pena].
Daphne: her-CL got-1SG [$_F$ the gold pen]
'I got her the GOLD pen.'

b. Daphne: Tis pira [$_F$ tin pena ti HRISI].
Daphne: her-CL got-1SG [$_F$ the pen the gold]
'I got her the GOLD pen.'

(Kolliakou 1998b:ex.4c-d)

In (2.64a) the accent has shifted to the adjective, as in the English example in (2.62b). By contrast, in (2.64b) the accent falls on the rightmost boundary of the NP, as in the canonical

case (2.63). However, the NP in (2.64b) has the structure of a polydefinite. The adjective appears postnominally and is marked with the definite article.

While NP de-accenting involves distinguishing new from given information, the aspect of new/given it encodes is distinct from Information Packaging. All examples, the canonical and the de-accented or polydefinite ones, have the same Information Structure. It is worth noting that the encoding of this aspect of the new/given distinction relies on the linguistic cues used for Information Packaging, that is accent placement and word order (in Greek).

De-accenting is distinguished from narrow NP focus:

(2.65) Zoe: Pia pena tis pirate?
'Which pen did you buy for her?'

a. Daphne: Tis pirame [$_F$ ti HRISI] pena
Daphne: her-CL got-1PL [$_F$ the gold] pen
'We got her the GOLD pen.'

b. Daphne: ?@Tis pirame tin pena [$_F$ ti HRISI].
Daphne: ?@her-CL got-1PL the pen [$_F$ the gold]

(Kolliakou 1998b:ex.4c-d)

In (2.65a) the focus of the sentence is the adjective, *hrisi.* This example is ambiguous between a narrow focus reading and a NP de-accenting reading (2.64a). On the other hand, example (2.65b) is infelicitous in a context requiring a narrow focus reading.

2.5.5 Encoding unexpected/surprising information

Accent placement does not always indicate update focus. Quite often, elements conveying unexpected or surprising information may attract accent. Consider the following question-answer pair:

(2.66) a. PIOS tus ipe oti tha stilume to YANI stis Vrikseles?
who-NOM them-CL said-3SG that will send-1PL the Yani-ACC to-the Brussels?
'Who told them we are going to send Yanis to Brussels?'

b. [$_F$ o **PETROS**] (tus ipe oti tha stilume to **YANI** stis
[$_F$ the Petros-**NOM**] (them-**CL** said-**3SG** that will send-**1PL** the Yani-**ACC** to-the
Vrikseles)
Brussels)
'Petros (told them we are going to send Yanis to Brussels). '

The focus segment of (2.66b) is the matrix subject, *o Petros*, which is stressed[7]. However, *Petros* is not the only stressed phrase. The embedded object, *to Yani* is also perceived as stressed. *To Yani* is also perceived as stressed in the question (2.66a). However, *to Yani* cannot be identified as (part of) the focus of either sentence in (2.66). There is no sense in which it updates the hearer's information state. The question in (2.66) could express the speaker's surprise that somebody said they are going to send *Yani* to Brussels, while, for example, it has been decided that they will send somebody else.

Similar phenomena are attested in English. Vallduví & Zacharski (1994) draw attention to the example in (2.67):

(2.67) That Ann—she's such an **IN**teresting **PER**son.
She dances taran**TEL**a with a **PAS**sion:
she grew up in **SOUTH DA**kota,
and she studied classical Chi**NESE** at **HAR**vard.

(Vallduví & Zacharski 1994:ex.13)

The focus of update in the last sentence of (2.67) is the VP. The accent on *Harvard* is enough to give rise to the VP focus. However, the direct object, *Chinese*, is also accented. Vallduví & Zacharski (1994) argue that the accent on *Chinese* bears no relevance to the ground-focus articulation of this sentence. They follow Bolinger (1989) and Zacharski (1993) who suggest that elements that are interesting, informative or depart from some semantic or cultural stereotype attract accent. They point out that no accent is needed in (2.68) where *Chinese* is replaced with relatively uninteresting words:

(2.68) a. She took courses at **HAR**vard.

[7]Even though the focused constituent is not associated with a contrastive reading, that is, it encodes presentational focus, it appears preverbally. Example (2.83) in Section 2.6.2 illustrates a similar case.

b. She studied English at HARvard.

c. She got a degree at HARvard.

(Vallduví & Zacharski 1994:ex.14)

To sum up, accent placement is employed in English and Greek to mark unexpected, informative, surprising elements. This use of accent placement is independent of the ground-focus partition. In (2.67) the accented element appears within focus while in (2.66b) it appears within the ground part of the sentence. [See also Ladd (1996) who notes that in Romanian pitch accents at the end of questions are often associated with surprise readings.]

In a similar manner, accent placement is sometimes used to express the speaker's emotions towards the conveyed message. Example (2.69) may be an all-focus sentence, despite the fact that the accent falls on the verb which, as argued in Section 2.4.1, is typically interpreted as narrow focus:

(2.69) [$_F$ tin PIRE tin ipotrofia i Sofia]
[$_F$ her-CL got-GOT the scholarship the Sofia-NOM]
'Sofia got the scholarship.'

Example (2.69), in addition to conveying the event, expresses the speaker's enthusiasm about the event. Of course, example (2.69) presupposes background knowledge that *Sofia* was expecting to hear about the scholarship. Such background knowledge may be present for (2.70). However, if the speaker knows that Sofia expects to hear about the scholarship and would like her to get it, then the speaker has no reason to be enthusiastic about the news that 'Yanis got the scholarship'. As a result, example (2.70b) with the accent on the verb is an infelicitous utterance in such a context:

(2.70) a. [$_F$ tin pire o YANIS tin ipotrofia]
[$_F$ her-CL got-3SG the Yanis-NOM the scholarship]
'Somebody else got the scholarship.'

b. @[$_F$ tin PIRE o Yanis tin ipotrofia]
@[$_F$ her-CL got-3SG the Yanis-NOM the scholarship]

2.6 The independence of Information Structure from Syntax

The representation of Information Structure in the grammar has been a matter of controversy. In this section, I will present arguments supporting the view that Information Structure should be represented independently from the other levels of the grammar. The discussion will focus on the relation of Information Structure with syntax (see Steedman (1991) and Steedman (1998) for the independence of Information Structure from intonational structure). This issue is of particular importance for the study of a language like Greek, that relies on syntactic operations for the encoding of the ground-focus partition.

The diversity of the crosslinguistic realisation of Information Packaging

Crosslinguistically, Information Packaging is realised in diverse ways. As already discussed, English employs intonation while Catalan relies on syntax. Mixed-type languages like Greek and West Germanic exploit both accent placement and word order. Languages like Vatu rely on morphology (Vallduví 1995)—see also Bresnan & Mchombo (1987). Vallduví (1992) and Vallduví (1995) argues that the crosslinguistic variation can be captured in an elegant way if it is assumed that Information Structure is represented in an independent level of grammar. In the analysis he offers in Vallduví (1992), Information Structure is a distinct level, similar to LF. It is crosslinguistically uniform, in the same manner LF is. S-structure representations are mapped onto Information Structure representations, parallel to LF ones.

Subjacency

As has become evident from the data presented in the previous sections, there is no close correspondence between the ground-focus partition and syntactic constituents. Any kind of constituent may function either as a focus or ground element. In addition, focus is insensitive to syntactic constraints like subjacency (Giannakidou 1997; Rooth 1996). In (2.71) the focused element appears within a strong island:

(2.71) svisane ta fota ya na filisi o Yanis ti **MARIA**
switched-off-3PL the lights for PART kiss-3SG the Yanis-NOM the Maria-ACC
'They switched off the lights so that Yanis would kiss Maria.'

As will be shown in Chapter 3, Focus-movement out of a strong island is ungrammatical in Greek (as all extractions are in general). Any representation of the Information Structure of the sentence in (2.71) would involve the inheritance of some focus value (or non-overt movement of some contituent marked with some focus feature) from the lower clause to the top node. This inheritance/movement is not constrained by subjacency. On the other hand, subjacency does constrain the extraction of a focused element. Thus, the two cases are distinct. This mismatch between the syntax of Focus-movement and focus with respect to subjacency suggests that the representation of focus should be independent of syntax.

Finally, a question central to the relation of Information Structure with syntax, is whether Information Structure is organised in a recursive way, as syntax is. Next, I will consider this question with respect to matrix clauses (Section 2.6.1) and embedded ones (Section 2.6.2).

2.6.1 The non-recursive nature of the ground-focus partition

Based on examples like (2.72) various authors have proposed that focus is recursive (Krifka 1991; Partee 1991; Rooth 1996). Krifka (1991) interprets the pitch accent on *youngest* as an instance of focus within the link/topic:

(2.72) What did Bill's sisters do?
[$_L$ Bill's [$_F$ YOUNGEST] sister] [$_F$ kissed JOHN.]

Vallduví & Zacharski (1994) argue that such an interpretation is problematic. First, *youngest* does not convey the new/update information of the sentence. In this respect, it deviates from standard definitions of focus. Second, *youngest* carries a B-accent which is indicative of linkhood rather than focushood. Vallduví & Zacharski (1994) propose that (2.72) has the Information Structure shown in (2.73), identical to (2.74):

(2.73) What did Bill's sisters do?
[$_L$ Bill's YOUNGEST sister] [$_F$ kissed JOHN.]

(Vallduví & Zacharski 1994:ex.33)

(2.74) What did Bill's siblings do?
[$_L$ Bill's SISTER] [$_F$ kissed JOHN.]

(Vallduví & Zacharski 1994:ex.33)

The difference between the two examples above is that in (2.73) the B-accent has shifted to the adjective. Following Steedman (1991), they interpret this shift as a case of de-accenting, triggered by the notions of informativeness/giveness discussed in Section 2.5.5. They also note that in the corresponding examples in Catalan, a language that does not display any de-accenting strategy, there is no accent in any part of the link. They argue that, if the accent on *youngest* indicated focus, this focus should be present crosslinguistically. Thus, Vallduví & Zacharski (1994) conclude that focus is not recursive in nature.

Recursive occurrences of focus are impossible in Greek (Tsimpli 1995; Tsiplakou 1998), as the ungrammaticality of (2.75) shows:

(2.75) a. *to YANI ida sto SINEMA
*the Yani-ACC saw-1SG at-the cinema
'I saw Yanis at the cinema.'

b. *o PETROS petakse ta VIVLIA
*the Petros-NOM threw-3SG the books
'Petros threw the books.'

The sentences in (2.75) are ungrammatical because they contain two accented elements. Rizzi (1995) offers similar examples from Italian.

The examples from English and Greek considered in this section involve simple matrix sentences, that do not contain any embedded clauses. In the next section I will discuss the Information Structure of sentences with embedded clauses.

2.6.2 The Information Structure of subordinate clauses

The Information Structure of embedded clauses has not received much attention in the literature. Vallduví & Zacharski (1994) leave open the possibility that embedded clauses may have their own focus-ground structure. They acknowledge that, in this respect, the focus-ground partition is recursive. By contrast, Heycock (1993) observes that a recursive syntactic structure may correspond to a non-recursive Information Structure unit. For example, the link in (2.76) contains a whole embedded clause:

(2.76) A: What do you think of the allegations that John is a liar?
B: [$_L$ The allegations that John is dishonest] [$_F$ are FALSE].

(Heycock 1993:ex.31)

With respect to Greek, Tsimpli (1995) and Tsiplakou (1998) note that recursive foci are ungrammatical in sentences containing a subordinate clause:

(2.77) a. *to YANI ipe oti ide sto SINEMA
*the Yani-ACC said-3SG that saw-3SG at-the cinema
'S/he said that s/he saw Yanis at the cinema.'

b. *o PETROS ipe oti petakse ta VIVLIA
*the Petros-NOM said-3SG that threw-3SG the books
'Petros said that he threw the books—S/he said that Petros threw the books.'

In the ungrammatical (2.77a) both focused constituents belong to the embedded clause and one of them has undergone long distance extraction. The same ungrammaticality arises in (2.77b) where, under one reading, there is one focused element per clause: the matrix subject and the embedded object. The examples in (2.76-2.77) show that embedded clauses do not have their own Information Structure. Rather, the whole utterances have a single Information Structure. In this section I will discuss in more detail the Information Structure of sentences containing embedded clauses and show that the ground-focus partition is not sensitive to the recursive syntax of these constructions.

Let us consider some of the possibilities of the ground-focus articulation for a main clause containing an embedded one. In some cases, the focus of the utterance is the embedded clause:

(2.78) A: What did he say?
B: He said [$_F$ they are going to fire JOHN].

(2.79) A:ti ipe?
'What did s/he say?'
B: ipe [$_F$ oti tha dioksun to YANI].
B: said-3SG [$_F$ that will fire-3PL the Yani-ACC]
'S/he said they are going to fire Yani.'

In these examples, it is possible to claim that the embedded clause has its own Information Structure and that it instantiates an all-focus instruction. However, consider an example in which the focus segment of the utterance is a constituent within the embedded clause:

(2.80) A: Who did he say they are going to fire?
B: He said that they are going to fire [$_F$ JOHN].

(2.81) A: pion ipe oti tha dioksun?
'Who did s/he say they are going to fire?'
B: ipe oti tha dioksun [$_F$ to YANI].
B: said-3SG that will fire-3PL [$_F$ the Yani-ACC]
'S/he said they are going to fire Yani.'

In (2.80&2.81) the Information Structure of the subordinate clause is indistinguishable from the Information Structure of the matrix clause. The ground part is all the material up to the embedded object containing elements from both the matrix and the embedded clause.

Further, the embedded clause may function as ground:

(2.82) A: Who said they are going to fire John?
B: [$_F$ PAUL] said that they are going to fire John.

(2.83) A: se PION ipe oti tha dioksun to Yani?
'To whom did s/he say they are going to fire John?'
B: [$_F$ ston ALEKSI] ipe oti tha dioksun to Yani.
B: [$_F$ to-the Alexis] said-3SG that will fire-3PL the Yani-ACC
'S/he said to Alexis that they are going to fire Yani.'

Note again that the ground part contains the whole embedded clause as well as material from the matrix clause.

Thus, in a significant number of cases, the Information Structure of the subordinate clause is indistinguishable from that of the matrix clause. The next question is, whether there are any cases where the two can be distinguished. In order to test this, I will consider a question forcing recursive foci in the answer. Consider a situation where there is a discussion about two members of a committee that decided to send some employees to Brussels. Various names have arisen and there is confusion as to which member of the committee wanted to send which employee to Brussels. In (2.84a) the first wh-phrase belongs to the matrix clause and the second to the embedded:

(2.84) a. A: mporis na mu pis pios ithele na stilume pion
A: can-2SG to me-CL tell-2SG who-NOM wanted-3SG to send-1PL who-ACC
stis Vrikseles?
to-the Brussels?
'Can you tell me who wanted to send who to Brussels?'

b. B: o Yanis ithele na stilume ti MARIA (stis Vrikseles)
B: the Yanis-NOM wanted-3SG to send-1PL the Maria-ACC (to-the Brussels)
ke o Petros tin ELENI
and the Petros-NOM the Eleni-ACC
'Yanis wanted to send Maria (to Brussels) and Petros Eleni.'

c. B:*o YANIS ithele na stilume ti MARIA stis Vrikseles

Despite the fact that the question (2.84a) contains two wh-phrases, only one phrase is focused in the answer. Example (2.84c) which has two accented phrases corresponding to the wh-phrases in the question is ungrammatical. Note that example (2.84c) is parallel to (2.66b) repeated below as (2.85):

(2.85) [$_F$ o PETROS] tus ipe oti tha stilume to YANI stis
[$_F$ the Petros-NOM] them-CL said-3SG that will send-1PL the Yani-ACC to-the
Vrikseles
Brussels
'Petros told them we are going to send Yanis to Brussels. '

Example (2.85) is grammatical even though it contains two accented elements. By contrast, examples like (2.84c) have been claimed to be ungrammatical. The difference in grammaticality judgments seems to depend on the strength of the accent on *Yanis*. According to my intuitions, the accent on *Yanis* in (2.85) is perceived as slightly weaker than the one on *Petros*. By contrast, both accents in (2.84c) are meant to be equal in intensity. The phonological details of these examples are not of crucial importance here. What is of interest is that, even though Greek does allow recursive accents, this possibility is not exploited for contexts that require recursive foci. Thus, even in context that requires recursive foci, the ground-focus partition is not organised in a recursive way.

To conclude, the examples in (2.76-2.83) show that, in cases where recursive foci are not present, the Information Structure of subordinate clauses cannot be distinguished from

the Information Structure of the matrix clause. The ungrammaticality/infelicity of (2.84c) indicates that recursive foci are not possible at all. Thus, recursive syntactic structures are always associated with a non-recursive Information Structure. The non-recursive nature of the ground-focus partition strongly supports the view that Information Structure is represented independently of syntax.

Let us return to the Information Structure of (2.84b). This example seems to instantiate a link-focus instruction. This hypothesis is supported by the fact that it is a felicitous answer to (2.86) which can take a (contrastive) link-focus answer:

(2.86) Pion ithelan na stilume stis Vrikseles ta meli tis epitropis?
who-ACC wanted-3PL to send-1PL to-the Brussels the members the committee-GEN?
'Who did the committee members want to send to Brussels?'

In English as well, questions (2.84a) and (2.86) take the same answer:

(2.87) a. Who did the committee members want to send to Brussels?
b. *John* wanted to send MARY and *Peter* wanted to send JAMES.

(2.88) a. Who wanted to send who to Brussels?
b. *John* wanted to send MARY and *Peter* wanted to send JAMES.

Examples (2.87b&2.88b) have the same prosody[8]. In both examples *John* is associated with a B-accent. Thus, like in Greek, recursive foci are not available in English. [For some authors, in multiple Wh-questions like (2.87a&2.88a), the first wh-phrase is a contrastive topic—see Bolinger (1978) cited in Erteschik-Shir (1986).]

2.6.3 Embedded links and tails

As argued in the previous section, embedded clauses do not have an autonomous Information Structure. However, there are some cases which, at first sight, might suggest that subordinate clauses have their own, independent Information Structure. Consider the following examples:

(2.89) Tu tilefonises; Ti ipe?
'Did you phone him? What did he say?'

[8] I owe this information to D.R.Ladd.

a. ipe [$_F$ oti [$_L$ tis afises] tha tis ehi etimes stin
said-3SG [$_F$ that [$_L$ the posters-ACC] will them-CL have-3SG ready-ACC.PL on
ORA-tus]
time-their]
'He said that he'll have the posters ready on time/that the posters will be ready on time'

b. ipe [$_F$ oti tha ton DIOKSUN [$_T$ to Yani]].
said-3SG [$_F$ that will him-CL fire-3PL [$_T$ the Yani-acc]]
'He said that they will fire Yanis.'

In the above examples, the focused part is the embedded clause which seems to have a ground-focus rather than an all-focus structure. The embedded object in (2.89a) looks like a link. It has the syntactic trappings of a link, as it has undergone CLLD. It also functions as a link. *Tis afises* can be thought of as designating the locus of update. The aboutness feeling is also present. Similarly, *to Yani* in (2.89a) appears as a tail. It is doubled and dislocated to the right. Valiouli (1994) also offers examples involving Right Dislocated NPs in all-focus sentences. The embedded clauses in (2.89) appear to have their own Information Structure which is independent of the matrix one. This interpretation though, is incompatible with the facts presented in the previous section, where it was shown that embedded clauses do not have their own Information Structure.

Alternatively, the word order variation of the embedded clauses in (2.89) could be viewed as a phenomenon similar in nature to the ones discussed in Sections 2.5.4 & 2.5.5. That is, word order is used in (2.89) to express a relation of given-familiar/new-novel independent of the ground-focus partition of the sentence. As has already been discussed in Sections 2.5.4 & 2.5.5 English uses intonation to encode both the ground-focus partition of a sentence as well as other pragmatic aspects of 'new', independent of Information Packaging. Since Greek uses word order for Information Packaging, it is not surprising that word order is also used to encode given/new distinctions orthogonal to Information Packaging.

Let us now turn to the use of word order for encoding Information Packaging in sentences with a subordinate clause. In a context where the embedded link functions as the topic of the whole utterance, the best answer is the one in which the embedded topic is extracted at the beginning of the matrix clause:

(2.90) Ti ipe ya to Yani?
what said-3SG for the Yani-acc
'What did s/he say about Yani?'

a. ?@ipe oti to Yani tha ton **DIOKSUN**
said-3SG that the Yani-acc will CL-ACC fire
'S/he said they are going to fire Yani'

b. to Yani ipe oti tha ton **DIOKSUN**
the Yani-acc said-3SG that will CL-ACC fire
'S/he said they are going to fire Yani'

Example (2.90b) is the preferred answer to the question in (2.90). Again, the Information Structure of the subordinate clause is indistinguishable from that of the matrix clause. In addition, unlike (2.90a) or (2.89a), example (2.90b) is an infelicitous answer to the question in (2.89). The contrasts in the felicity conditions of (2.90a&2.89a) and (2.90b), indicate that, while word order is used both for Information Packaging and other pragmatic purposes, the actual permutations used for each case are not identical.

Finally, it should be noted that, unlike foci, recursive topics are possible:

(2.91) a. ta vivlia tu Yani tu ta stilame [$_F$ htes
the books-ACC the Yani-GEN him-CL.GEN them-CL.ACC sent-3PL [$_F$ yesterday
to **PROI**]
the morning]
'We sent the books to Yanis yesterday morning.'

b. tu Yani tu ipe oti ta vivlia tu ta
the Yani-GEN him-CL.GEN said-3SG that the books-ACC him-CL.GEN them-CL.ACC
stilame [$_F$ htes to **PROI**]
sent-3PL [$_F$ yesterday the morning]
'S/he told Yanis that we sent him the books yesterday morning.'

Example (2.91a) has two preposed topics whereas in (2.91b) there is one topic in the matrix and one in the embedded clause. Compare (2.91) with (2.75&2.77) which illustrate the ungrammaticality of recursive foci. However, even though recursive topics are possible, the Information Structure of both examples (2.91) remains non-recursive. In (2.91b) the two

topics may appear in two different clauses but they both belong to the ground part of the sentence, which contains material from both the matrix and the subordinate clause.

2.7 Association with focus: the independence of Information Structure from semantics

It has been argued that the focus part of a sentence provides the nucleus of various semantic operators (Krifka 1991; Partee 1991; Rooth 1996). This argument is based on examples like the following:

(2.92) a. Mary always took JOHN to the movies.

b. Mary always took John to the MOVIES.

c. MARY always took John to the movies.

(Partee 1991:ex.7)

(2.93) a. John only introduced Bill to SUE.

b. John only introduced BILL to Sue.

(Partee 1991:ex.8)

In (2.92&2.93), the operators *always* and *only* are *associated* with the accented/focused element that functions as the nucleus of the operator [similar examples can be constructed with *even*, *also*, counterfactuals, frequency adverbs and modals (Partee 1991; Rooth 1996)]. One of the main arguments for the independence of Information Structure from semantics is that Information Packaging does not affect the propositional content of a sentence. However, in the examples in (2.92&2.93) different associations with focus yield different propositional content. These examples suggest that focus can affect truth conditions.

The belief that pragmatic focus interacts directly with the propositional content of sentences has led to the proposal that focus is a semantic operator that interacts in various ways with other semantic operators (Krifka 1991; Rooth 1996). These analyses differ from the one proposed in Vallduví (1992) in an important way. Focus is treated as a semantic rather than pragmatic phenomenon. Roughly, focus is represented at the level of LF together with other logico-semantic phenomena. Vallduví (1992) and Vallduví & Zacharski (1994) acknowledge

that very often, focus provides the nucleus of various semantic operators. However, they show that this is not always the case and the two phenomena are distinct—see also Kuhn (1997) who reaches the same conclusion based on quantificational adverbials in German. They first note that *only* may be associated with a subpart of focus, as in the following example:

(2.94) [$_F$ There's only a month till CHRISTMAS now].

(Vallduví & Zacharski 1994:ex.11)

In (2.94) *only* is associated with *month. Month* is within the focus, but it is not the focus of the sentence. Moreover, it does not carry any pitch accent.

Further, semantic operators may be associated with ground elements. In (2.95) *only* is associated with the link, *John*:

(2.95) John and Mary know the Amazon quite well,
but only John's [$_F$ been to the CITIES in Brazil.]

(Vallduví & Zacharski 1994:ex.28)

Similarly, in the context of (2.96), the operator *only* is associated with a ground element, *Bill.* Compare (2.96) with (2.93a) where *only* is associated with *Sue*:

(2.96) Who did John only introduce *Bill* to?
John only introduced *Bill* to SUE.

(Vallduví 1992:ex.286)

Finally, *always* in (2.97) is associated with *John*, rather than *Mary.* Compare (2.97) with (2.92c):

(2.97) Who always took JOHN to the movies?
MARY always took John to the movies.

(Vallduví & Zacharski 1994:ex.9)

In conclusion, examples (2.94-2.97) show that the generalisation that pragmatic focus provides the nucleus for semantic operators is not a valid one. In effect, the analyses based on this generalisation, that treat focus as a semantic operator are not adopted here. By contrast, examples (2.94-2.97) indicate the independence of Information Structure from semantics.

2.8 Wh-questions and focus

In this section I will discuss the relation of wh-phrases and focused elements and present some facts that will be important for the analyses presented in Chapter 3. For a discussion of similar phenomena in English see, among others, Chafe (1983); Culicover & Rochemont (1983); Erteschik-Shir (1986); Rooth (1996).

2.8.1 Similarities and differences

Phonological facts

Like focused elements, wh-phrases in Greek Wh-questions are associated with the nuclear accent. In this respect, the wh-phrase in (2.98b) is on a par with the focused phrase in (2.98a):

(2.98) a. to YANI ida sto sinema
the Yani-ACC saw-1SG at-the cinema
'I saw Yanis at the cinema.'

b. PION ides sto sinema?
who-ACC saw-2SG at-the cinema
'Who did you see at the cinema?

Echo Wh-questions have a different prosodic pattern, the main characteristic of which is that accent does not fall on the wh-phrase. Unlike canonical Wh-questions (2.99a), the accent in the echo question in (2.99b) falls at the rightmost element, *Yanis*:

(2.99) a. Canonical Wh-question:
PION ide o Yanis?
who-ACC saw-3SG the Yani-NOM
'Who did Yanis see?'

b. Echo Wh-question:
pion ide o YANIS?
who-ACC saw-3SG the Yani-NOM
'Yanis saw who?'

Recursiveness in canonical Wh-questions

Recursive wh-phrases are, in general, ungrammatical:

(2.100) a. *PIOS ide pion/PION?

who-NOM saw-3SG who-ACC
'Who saw who?'

b. *PIOS ipe oti edosan pia/PIA vivlia?

who-NOM said-3SG that gave-3PL which books
'Who said that they gave which books?'

The unavailability of recursive wh-phrases is on a par with the unavailability of recursive foci. Compare (2.100) with (2.101) displaying recursive foci:

(2.101) a. *o YANIS ide ton PETRO

the Yanis-NOM saw-3SG the Petros-ACC
'Yanis saw Petros.'

b. *o YANIS ipe oti edosan ta MITHISTORIMATA

the Yanis-NOM said-3SG that gave-3PL the novels
'Yanis said that they gave the novels.'

However, recursive appearance of wh-phrases is possible in examples like the following:

(2.102) PIOS kseri pia/*PIA vivlia tus edosan?

who-NOM know-3SG which books them-CL gave-3PL
'Who knows which books they gave them?'

Note though, that the embedded wh-phrase in (2.102) does not take matrix scope. The answer to (2.102) can only be *Yanis* not *Yanis the novels.* The presence of the embedded wh-phrase seems to satisfy subcategorisation requirements. This holds even if both wh-phrases involved belong to the embedded clause[9]:

(2.103) PIO VIVLIO kseris pios sto stile?

which book know-2SG who-NOM you-it-CL send-3SG
'Which book do you know who sent?'

Thus, even when recursive wh-phrases are present, only one of them can acquire wide scope.

[9]Example (2.103) is rather awkward. Despite its oddness, it is only the extracted wh-phrase that takes matrix scope.

Recursive readings in echo questions

Echo questions present an exception to the generalisation that only one wh-phrase may take wide scope. Multiple wh-readings are possible in the echo-questions in (2.104). Note as well that in the echo questions the accent falls on the last wh-phrase (2.104):

(2.104) a. pios ide **PION**?

who-**NOM** saw-3SG who-**ACC**
'Who saw who?'

b. pios ipe oti edosan **PIA** vivlia?

who-**NOM** said-3SG that gave-3PL which books
'Who said that they gave which books?'

c. pios ithele na stilume **PION** stis Vrikseles?

who-**NOM** wanted-3SG to send-1PL who-**ACC** to-the Brussels?
'Who wanted to send who to Brussels?'

Compare the above with (2.100a&b). A felicitous answer to (2.104a) is *John saw Maria and Peter Suzan* not just *John*. Similarly for (2.104b&c).

However, not all instances of echo-questions allow multiple readings:

(2.105) pios kseri pia vivlia tus **EDOSAN**?

who-**NOM** know-3SG which books them-**CL** gave-3PL
'Who knows which books they gave them? (echo)'

The answer to (2.105) is just *John*. As in (2.102), the embedded wh-phrase satisfies subcategorisation requirements and cannot take matrix scope.

Embedded Wh-questions

Let us now turn to embedded Wh-questions. Consider the following examples:

(2.106) rotise pios ide **PION**

asked-3SG who-**NOM** saw-3SG who-**ACC**
'S/he asked who saw who.'

(2.107) Mporis na mu pis pios ithele na stilume pion stis Vrikseles?

can-2SG to me-**CL** tell-2SG who-**NOM** wanted-3SG to send-1PL who-**ACC** to-the Brussels?

'Can you tell me who wanted (us) to send who to Brussels?'

A felicitous continuation of (2.106) is *they told him that John saw Mary and Peter Suzan*, but not just *John*. Similarly for (2.107).

However, the embedded questions below do not allow multiple readings:

(2.108) thelo na mathis pios kseri pia omada tha KERDISI
want-1SG to learn-2SG who-NOM know-3SG which team will win-3SG
'I want you to find out who knows which team will win.'

(2.109) rotise pios kseri pia vivlia tus STILANE
asked-3SG who-NOM know-3SG which books them-CL send-3PL
'S/he asked who knows which books they sent.'

Example (2.108) could be paraphrased as *I want you to find out who knows the winning team*, not as *I want you to find out who knows the winning team and which is the winning team.* Similarly, (2.109) could be followed by *and they told him John does.*

The contrast between (2.106&2.107) and (2.108&2.109) could be explained on the tentative assumption that the embedded questions in (2.106&2.107) are echo ones. This assumption is supported by the fact that the matrix echo questions in (2.104), which correspond to the embedded questions in (2.106&2.107) allow multiple readings. By contrast, the matrix echo question in (2.105), which corresponds to (2.109) does not allow a multiple reading.

In conclusion, wh-phrases in canonical Wh-questions share various properties with focus. In particular, like foci, they attract the nuclear accent and are not recursive. By contrast, wh-phrases in echo Wh-questions do not behave as foci. They are not accented and may appear recursively.

2.8.2 Incompatibility of focus and Wh-questions

Rizzi (1995) and Tsimpli (1995) observe, for Italian and Greek respectively, that focus is impossible in Wh-questions:

(2.110) a. *A GIANNI che cosa hai detto (, non a Piero)?
'To Gianni what did you tell (, not Piero)?'

(Rizzi 1995:ex.25a)

b. *pios nomizi to YANI oti ide?
who-NOM think-3SG the-ACC Yani-ACC that saw-3SG
'Who thinks that s/he saw Yani?'

c. *se pion nomizis to VIVLIO oti edose?
to who-ACC think-2SG the-ACC book-ACC that gave-3SG
'To whom do you think that s/he gave the BOOK?'

(Tsimpli 1996:ex.13a-b, p.7))

d. *pios milise sto YANI
who-NOM talked-2SG to-the Yani-ACC
'Who spoke to Yanis?'

However, compare (2.66a) repeated here as (2.111) with (2.110):

(2.111) PIOS tus ipe oti tha stilume to YANI stis Vrikseles?
who-NOM them-CL said-3SG that will send-1PL the Yani-ACC to-the Brussels?
'Who told them we are going to send Yanis to Brussels?'

According to native speakers' intuitions *Yani* is stressed in (2.111). Note though, that, as pointed out in Section 2.5.5, the accent on *Yani* does not convey new/updating information. Rather, it expresses a surprise or possibly contrastive reading which, as already argued, is independent of the ground focus partition. In fact, it is not clear what it would mean for a Wh-question to have update-focus, since a question seems to require rather than contribute information in the discourse. Thus, what Rizzi (1995) and Tsimpli (1995) seem to interpret as potential focus in a Wh-question is surprise/contrastive readings like the one present in (2.111). The question that arises then, is why this kind of reading is unavailable for (2.110). The badness of the Greek examples at least, cannot be attributed to the incompatibility of Wh-questions with surprise/contrastive readings ('focus' for Rizzi and Tsimpli) since such a reading is present in the grammatical (2.111). The explanation seems to lie in the phonology[10]. Ladd *et al.* (1998) argue that questions in Greek may involve a main, nuclear accent and a secondary one. Note that the strings in (2.110) are much shorter than those in (2.111). It seems that, at least the Greek examples in (2.110) are bad because there is not enough

[10]I owe most of the material presented in this section to a very helpful discussion with D.R.Ladd and Ineke Mennen. Any misinterpretations of their suggestions remain mine.

'phonological space' to allow the development of a secondary accent. By contrast, this space is available in (2.111).

We are left with the question of what is the nuclear accent of (2.111). This example contains a main, nuclear accent and a secondary one. Recall that, according to the definitions mentioned in Section 2.3, nuclear accent refers to 'the syllable in a tone unit which carries maximal prominence, usually due to a major pitch change' (Crystal 1991:p.238). Definitions of this kind do not clarify whether prominence is based on phonetic facts or perception. If prominence is decided on the basis of phonetic facts, then deciding what the nuclear accent is in (2.111) amounts to deciding which of the two accents, is 'phonetically' more prominent. From this point of view, it is perhaps best to assume that the nuclear accent in (2.111) is associated with the wh-phrase.

However, in this example, speakers perceive *Yani* as the most stressed/prominent phrase, not the wh-phrase. If the notion of *prominence* relies on speakers' perception, then the nuclear accent in (2.111) should be associated with the secondary accent that falls on *Yani.* This assumption would imply that prominence, arises from the interaction of both phonetic factors and interpretational ones—which entity is viewed as the most informative/expected etc. Though the phonological details are of no immediate interest here, a clear notion of nuclear accent is crucial to any theory that associates update focus with nuclear accent placement.

Wh-phrases as the focus of Wh-questions

The close correspondence between the wh-phrase and the focused elements in question-answer pairs has given rise to the hypothesis that wh-phrases instantiate the focus of Wh-questions (Culicover & Rochemont 1983; Brody 1990). This hypothesis is particularly appealing in the case of Greek, given that, like foci, wh-phrases in Greek bear the nuclear accent. However, this hypothesis is faced with problems, some of which I will discuss below with particular reference to Greek.

The assumption that wh-phrases in Greek are focused lies on their association with nuclear accent. However, this hypothesis could not be extended to English Wh-questions. In principle, Wh-questions in both languages should have the same Information Structure. If wh-phrases are focused in Greek, they should be focused in English. An explanation should be provided for why wh-phrases in English are not accented. Thus, the association of wh-phrases with nuclear

accent does not support their interpretation as focused constituents. Rather, it suggests that accent should not be always associated with informational focus. Wh-questions are not the only case in which accent is not associated with update focus. As shown in Section 2.5.5, accent in both English and Greek is often employed to mark informative/unexpected elements. It is possible that accent in Greek Wh-questions is used for similar purposes. However, this use of accent is orthogonal to its use for encoding the ground-focus partition.

The parallel between wh-phrases and foci with respect to recursiveness could also be explained on the assumption that wh-phrases are foci. However, this assumption would not explain why recursive wh-phrases are possible in echo questions (and in English canonical Wh-questions).

The final objection to the hypothesis that wh-phrases realise the focused constituent of a Wh-question is of a theoretical nature. According to the definition of focus adopted in this study, questions are not expected to have a focus-segment. Focus contributes updating information to the hearer's information state. Questions, at least in most cases, do not aim at updating the hearer's information state. In this respect, they are not expected to contain a focus segment.

2.9 Conclusions

In this chapter I have shown how word order and accent placement are employed in Greek to encode Information Packaging. The three constructions that will concern us in the following chapters, Focus-movement, Topicalisation and CLLD, are central to the realisation of specific instruction types. In particular, Focus-movement instantiates a (link)-focus-ground instruction. It involves the extraction of the focused constituent to a preverbal position and accent shift to this position. Topicalisation and CLLD are both employed for the realisation of links and involve a link-focus-(ground) instruction.

A conclusion that will be of crucial importance in the discussion that follows in Chapter 3 is that of the independence of Information Structure from the other levels of grammar. The evidence drawn from embedded clauses shows that Information Structure is not sensitive to the recursive organisation of syntactic structures. In addition, a closer investigation of the role of information focus in phenomena of *association with focus* leads to the conclusion that information focus does not affect the truth conditions of sentences. Information Structure is

therefore independent of the semantics of sentences. Thus, the view adopted in this study is that Information Structure involves a distinct level of grammar. The analysis of the discourse-syntax interface I will present in Chapter 5 builds on this assumption.

Chapter 3

Discourse Configurational Approaches

3.1 Introduction

Work on the syntax of Topicalisation, CLLD and Focus-movement has widely acknowledged the discourse import of these structures and is characterised by an attempt to integrate the discourse functions of topic and focus into the grammar as well as to account for the syntactic properties of these structures. *Discourse Configurational* approaches represent the dominant tendency in the literature.

With respect to the first objective, the integration of topic and focus in the grammar, Discourse Configurational approaches hold that topic and focus are encoded in Phrase Structure through distinct functional projections. Topics and foci move to preverbal positions of the syntactic tree to check their discourse features.

With respect to the second objective, the account of the syntactic properties of the relevant structures, Discourse Configurational approaches associate Topicalisation, CLLD and Focus-movement with three distinct syntactic operations. Both Focus-movement and Topicalisation involve movement but they differ in that Focus-movement is quantificational in nature, whereas Topicalisation is anaphoric. No movement is involved in CLLD; rather, the dislocated phrase is base-generated at its surface position.

In this chapter I present some representative examples of the Discourse Configurational approaches. In order to facilitate the presentation, I first summarise some syntactic properties

of Topicalisation, CLLD and Focus-movement that have been central to the discussion of these structures[1].

3.1.1 Similarities

Topicalisation, Focus-movement and CLLD have in common the following syntactic properties (the examples in this section are from Greek, unless stated otherwise):

i) Crosslinguistically, all three constructions involve **long distance extraction**:

(3.1) a. **Focus-movement**

tin PARASTASI ipan oti skinothetise o Dimitris Potamitis
the-ACC performance said-3PL that directed-3SG the-NOM Dimitris-NOM Potamitis-NOM

'They said that Dimitris Potamitis directed the performance.'

b. **Topicalisation**

tin parastasi anakinosan oti tha skinothetisi o Dimitris
the-ACC performance announced-3PL that will direct-3SG the-NOM Dimitris-NOM
POTAMITIS
Potamitis-NOM
'They announced that Dimitris Potamitis will direct the performance.'

c. **Clitic Left Dislocation**

tin parastasi ipan oti tha *ti* skinothetisi o Dimitris
the-ACC performance said-3PL that will her-CL direct-3SG the-NOM Dimitris-NOM
POTAMITIS
Potamitis-NOM
'They said that Dimitris Potamitis will direct the performance.'

For similar examples in other languages see Kiss (1995a), Cinque (1990), Dobrovie-Sorin (1990), Sanfilippo (1990), Hoffman (1995) and Balari (in press) among others.

ii) Crosslinguistically, all three constructions obey **strong islands**:

[1] Part of the material presented in this chapter appears in Alexopoulou (1998a).

(3.2) a. **Focus-movement**

*to YANI efigan noris gia na apofigun

the Yani-ACC left-3PL early so-as to avoid-3PL
'They left early so as to avoid Yanis.'

b. **Topicalisation**

*tin parastasi agnoun ton kalitehni pu SKINOTHETISE

the-ACC performance not-know-3PL the-ACC artist-ACC that directed-3SG
'They do not know the artist that directed the performance.'

c. **Clitic Left Dislocation**

*to Yani sinadisa tin kopela pu *ton* IDE

the Yani-ACC met-1SG the-ACC girl-ACC that him-CL saw-3SG
'I met the girl that saw Yanis.'

In (3.2a) the object is extracted out of a sentential adjunct clause, while (3.2b-c) illustrate Complex Noun Phrase (CNP) violations. See also Kiss (1995a), Cinque (1990), Dobrovie-Sorin (1990), Sanfilippo (1990), Hoffman (1995) and Balari (in press).

iii) In Greek, in all three constructions the extracted XP may appear on either side of the complementiser *oti*:

(3.3) a. **Focus-movement**

ipe (to YANI) oti (TO YANI) ide sto sinema

said-3SG (the Yani-ACC) that (the Yani-ACC) saw-3SG at-the cinema
'S/he said that s/he saw Yanis at the cinema.'

b. **Topicalisation**

mas ipe (me kokini mpoya) oti (me kokini mpoya) tha vapsoume

us-GEN said-3SG (with red paint) that (with red paint) will paint-1PL

ta PARATHIRA

the windows
'S/he said that we will paint the windows with red paint.'

c. **Clitic Left Dislocation**

mu ipe (ta klidia) oti (ta klidia) *ta* edose sti
me-GEN said-3SG (the keys-ACC) that (the keys-ACC) them-CL gave-3SG to-the
MARIA
Maria
'S/he told me that she gave the keys to Maria.'

Though topics/links in Topicalisation or CLLD constructions can appear before *oti*, they are preferred after the complementiser.

iv) None of the three constructions blocks selection by a higher verb as illustrated in (3.3) and (3.4):

(3.4) a. **Focus-movement**

rotise to YANI pios ide sto sinema
asked-3SG the Yani-ACC who-NOM saw-3SG at-the cinema
'S/he asked who saw Yanis.'

b. **Topicalisation**

rotise me auta ta lefta ti mporis na AGORASIS
asked-3SG with these the money what can-2SG to buy-2SG
'S/he asked what you can buy with this money.'

c. **Clitic Left Dislocation**

rotise ta isitiria pios ta PETAKSE
asked-3SG the tickets-ACC who-NOM them-CL threw-3SG
'S/he asked who threw away the tickets.'

The matrix verb *rotise* can select the indirect question despite the intervening XP between *rotise* and *pios*. The same is true of the following examples from Italian:

(3.5) a. Mi domando, il premio Nobel, a chi lo potrebbero dare
'I wonder, the Nobel Prize, to whom they could give it.'

b. ?Mi domando a GIANNI che cosa abbiano detto (, non a Piero)
'I wonder to GIANNI what they said (, not to Piero).'

(Rizzi 1995:fn 18)

v) None of the three constructions creates islands for extraction:

(3.6) a. **Focus-movement**

rotise pios nomizis ti MARIA oti tha psifize

asked-3SG who-NOM think–2SG the Maria-ACC that would vote-3SG
'S/he asked who you think would vote for Maria.'

b. **Topicalisation**

pios ipane afti tin askisi oti de tha tin kataferi me tipota

who-NOM said-3PL this the exercise-ACC that not will it-CL make with nothing

'Who did they say will not be able to solve this exercise by any means?'

c. **Clitic Left Dislocation**

pios nomizis ti Maria oti tha *tin* psifize

who-NOM think-2SG the Maria-ACC that her-CL vote–3SG
'Who do you think would vote for Maria?'

(Iatridou 1995)

In examples (3.6) the extraction of the embedded XP does not affect Wh-movement. This is also true of CLLD in Italian (Section 3.3.3).

Thus, with respect to i-v, Focus-movement, Topicalisation and CLLD exhibit the same properties.

3.1.2 Some differences

Weak Crossover Effects

Crosslinguistically, Focus-movement gives rise to weak crossover effects, while Topicalisation and CLLD do not (Cinque 1990; Dobrovie-Sorin 1990; Iatridou 1995; Kiss 1995a; Lasnik & Stowell 1991; Tsimpli 1995; Tsiplakou 1998; Vallduví 1995). Examples (3.7a) and (3.7c) show Focus-movement and CLLD in Greek while (3.7b) shows Topicalisation in English:

(3.7) a. **Focus-movement**

to YANI$_{*i/j}$ agapa i mana tu$_i$

the-ACC Yani-ACC love-3SG the-NOM mother-NOM his
'His mother loves Yani.'

b. **Topicalisation**

John$_i$, his$_i$ mother really likes t$_i$.

(Rizzi 1995:ex.30)

c. **Clitic Left Dislocation**

to Yani$_i$ ton$_i$ AGAPAI i mana tu$_i$

the-ACC Yani-ACC him-ACC love-3SG the-NOM mother-NOM his
'His mother LOVES Yani.'

In (3.7a) the possessive pronoun *tu* cannot be coindexed with the focused NP *to Yani*. By contrast, the possessive *tu* in (3.7c) and *his* in (3.7b) are coindexed with the preposed object NPs.

It is worth pointing out that the order of constituents does not affect wco in Greek. In situ focus equally gives rise to wco (3.8a), whereas coindexing is allowed in Clitic Doubling constructions (3.8b):

(3.8) a. **Focus**

I mana tu$_i$ agapai to YANI$_{*i/j}$

the-NOM mother-NOM his love-3SG the-ACC Yani-ACC
'His mother loves Yani.'

b. **Clitic Doubling**

I mana tu$_i$ ton$_i$ AGAPAI to Yani$_i$

the-NOM mother-NOM his him-ACC love-3SG the-ACC Yani-ACC
'His mother loves Yani.'

The examples in (3.8) are on a par with (3.7) with respect to wco.

Parasitic gaps

Focus-movement and Topicalisation license parasitic gaps (p-gaps). The following show p-gaps in Focus-movement in Greek (3.9), Topicalisation in English (3.10a) and Topicalisation in Hungarian (3.10b):

(3.9) **Focus-movement**

to YANI apelisan horis na (ton) proidopiisun

the Yani-ACC fired-3PL without PART him-CL warning-3PL
'They fired Yani without warning (him).'

(3.10) **Topicalisation**

a. The paper$_i$ we filed t_i before we could read t_i.

b. Egy iratot$_i$ elvesztettunk t_i, mielott elolvastunk-volna t_i

a paper-ACC we-lost t_i before we-had-read t_i
'A paper, we lost before we had read.'

(Kiss 1995b:ex.81)

Unlike Focus-movement and Topicalisation, CLLD does not license p-gaps (Iatridou 1995; Kolliakou 1991; Schneider-Zioga 1994; Tsimpli 1995). In (3.11) below the omission of the clitic is ungrammatical:

(3.11) **Clitic Left Dislocation**

ton Petro ton apelisan horis na *(ton) PROIDOPIISUN

the Petros-ACC him-CL fired-3PL without PART him-CL warn-3PL
'They fired Petros without warning.'

As in Greek CLLD does not license p-gaps in Italian (Section 3.3.3) and Romanian (Section 3.4.1).

Recursiveness

As already discussed in Sections 2.6.1—2.6.3, recursive topics are available , but recursive foci are not.

Compatibility with Wh-questions

Various authors claim that focused constituents are illicit in matrix Wh-questions. This claim is based on examples like (3.12a) from Italian and (3.12b) from Greek:

(3.12) a. **Focus-movement:**

A GIANNI che cosa hai detto (, non a Piero)?

'To Gianni what did you tell (, not Piero)?'

(Rizzi 1995:ex.25a)

b. *ti ipes sto YANI

what said-2SG to-the Yani-ACC
'What did you say to Yanis?

As discussed in Section 2.8.2, what Rizzi (1995) and Tsimpli (1995) describe as potential focus in (3.12) is the availability of a contrastive/surprise reading for *Gianni* or *Yanis*. As argued, such readings are available in Greek Wh-questions involving longer strings (2.111). In Section 2.8.2, I attributed the ungrammaticality of (3.12b) to phonological reasons rather than to the incompatibility of Wh-questions with focus/contrastive readings. However, for the sake of the presentation of the analysis of Rizzi (1995) and Tsimpli (1995), I will assume in the subsequent sections that examples (3.12) suggest that focus is incompatible with Wh-questions.

Rizzi (1995) and Tsimpli (1995) note that, unlike matrix Wh-questions, indirect ones allow focused constituents as shown in (3.4a) and in (3.5a).

On the other hand, topics are compatible with both matrix and indirect Wh-questions. The following illustrate Topicalisation (3.13a) and CLLD (3.13b) in Greek matrix Wh-questions:

(3.13) a. **Topicalisation**

me kokini mpoya pia parathira tha vapsume

with red paint which windows will paint-1PL
'Which windows will we paint with red paint?'

b. **CLLD**

ton Aleksandro pios tha ton pari tilefono

the-acc Aleksandro-acc who-nom will him-CL take-3sg telephone
'Who is going to phone Aleksandros?'

CLLD is compatible with Wh-questions in Italian as well:

(3.14) A Gianni, che cosa gli hai detto?

'To Gianni, what did you tell him?'

(Rizzi 1995:ex.24a)

3.1.3 Left Dislocation vs. Topicalisation

A distinction important for the discussion that follows is that between Left Dislocation and Topicalisation. In English, the former involves a resumptive pronoun (3.15a) but the latter a

gap (3.15b):

(3.15) a. (As for) those books, I gave them to my brother long ago.

b. Those books I gave to my brother long ago.

In both constructions the preposed XP functions as a link/topic (Cinque 1990; King 1995). However, they differ in their syntactic properties. Crucially, unlike Topicalisation, Left Dislocation is insensitive to strong islands:

(3.16) a. (As for) books like that, I would be surprised to meet anyone who liked them.

b. *Books like that I would be surprised to meet anyone who liked.

In languages with case marking, there are often case mismatches in Left Dislocation but not in Topicalisation. For example, in Russian, left dislocated objects may appear in nominative and be coindexed with resumptive pronouns in accusative (King 1995). In languages where resumptive pronouns may appear either as clitics/weak or strong/tonic pronouns, it is harder to distinguish Left Dislocation from CLLD/Topicalisation. In Italian Left Dislocation, the left dislocated element may be coindexed with a strong/tonic pronoun whereas CLLD obligatorily involves a clitic (Cinque 1990). In Greek, it is hard to substantiate the distinction between CLLD and Left Dislocation, as objects dislocated to the left cannot be coindexed with a strong/tonic pronoun alone. CLLD can be distinguished from Left Dislocation only on the basis of intonation. Example (3.2c), indicating that CLLD obeys strong islands, can improve to full acceptability if uttered with a comma intonation break after the dislocated object.

However, the contrast between Topicalisation and Left Dislocation becomes clearer once adjuncts are considered:

(3.17) a. *sto Londino ksero kapion pu meni
to-the London know-1SG somebody-ACC that live-3SG
'I know somebody who lives in London.'

b. sto Londino ksero kapion pu meni eki
to-the London know-1SG somebody-ACC that live-3SG there
'I know somebody who lives in London.'

The adjunct, *sto Londino*, cannot be extracted out of the relative clause in (3.17a). In (3.17b), where the adverbial is coindexed with *eki* the example is grammatical. Again, there

is an intonational break after *Londino* in (3.17b). However, a similar intonation break in (3.17a) cannot improve the acceptability of this example.

To summarise Sections 3.1.1–3.1.3, topics (in Topicalisation and CLLD constructions) and foci exhibit the same properties with respect to extraction possibilities. However, they differ with respect to wco effects, the licensing of p-gaps, recursive occurence and, according to some authors, compatibility with matrix Wh-questions. On the other hand, all three constructions differ from Left Dislocation with respect to subjacency violations. Topicalisation, CLLD and Focus movement are sensitive to strong islands, but Left Dislocation is not. Finally, as in Topicalisation and CLLD, the preposed XP in Left Dislocation is a link/topic.

My presentation of the literature on this subject consists of two parts. In Section 3.2 I discuss work on Focus-movement and Topicalisation. In Section 3.3 I present various accounts for CLLD and its relation to Topicalisation.

3.2 Focus-Topicalisation

By and large, the literature shares the intuition that Focus-movement instantiates A-bar movement, which is quantificational in nature while Topicalisation involves A-movement which is anaphoric in nature[2]The distinction between quantificational and anaphoric movement is drawn in Lasnik & Stowell (1991). As this distinction underlies the analysis of Focus-movement and Topicalisation, I start with the proposal of Lasnik & Stowell (1991) in Section 3.2.1. Next, in Sections 3.2.2–3.2.5, I present Kiss (1995b), Rizzi (1995), Tsimpli (1995) who discuss Focus-movement and Topicalisation in Hungarian, Italian and Greek respectively.

3.2.1 Quantificational vs. Referential Operators: Lasnik & Stowell 1991

The Weak Crossover Effect (wco) has been considered a syntactic reflex of quantification (Larson & Segal 1995; Lasnik & Stowell 1991). This generalisation belongs to Lasnik & Stowell (1991) who drew attention to the fact that weak crossover effects are present only in a subset of instances of A-bar movement (i.e. movement to non-argument positions (Haegeman 1991)). In particular, Wh-movement and Quantifier Raising (QR) are sensititive to wco, wheras Topicalisation, Tough Movement and Parasitic Gap constructions are not. Examples (3.18&3.19)

[2]As will be shown in the following sections, A-movement corresponds to A-bar-anaphoric movement in Lasnik & Stowell's terms and to NP-movement in Kiss's terms. A-bar-movement is A-bar-quantificational movement for Lasnik & Stowell; Kiss and Tsimpli refer to A-bar-movement as Operator-movement. Crucially, A-bar/Operator/A-bar-quantificational movement involves quantification whereas A/NP/A-bar-anaphoric movement involves an anaphoric operator.

show some of the cases they discuss; examples (3.18) show the LF representations of a Wh-question and a QR construction:

(3.18) a. *Who$_i$ [$_{IP}$ does [$_{NP}$ his$_i$ boss] [$_{VP}$ dislike t$_i$]]?

b. *No man$_i$ [$_{IP}$ [$_{NP}$ his$_i$ friends] should [$_{VP}$ mistreat t$_i$]]

(Lasnik & Stowell 1991:ex.13b-d)

(3.19) a. [John$_i$ [NO$_i$ [I believe his$_i$ mother loves e$_i$]]]

b. [John$_i$ should be easy for [his$_i$ wife] [NO$_i$ [PRO to love e$_i$]]]

(Lasnik & Stowell 1991:ex.33a,28a)

In (3.18) the pronoun *his* cannot be bound by the wh-phrase *who* or the phrase *no man*. By contrast, *John* can bind *his* in (3.19). While all structures in (3.18&3.19) instantiate A-bar movement, only Wh-movement and QR induce wco effects. Topicalisation (3.19a) and Tough Movement (3.19b) allow binding.

Lasnik & Stowell (1991) propose that wco arises only in the presence of a *True Quantifier*, which is defined as follows (QP stands for Quantifier Phrase):

- '...a true QP is composed of a quantifier Q and a nominal term T defining a range R that Q quantifies over, such that R is a **possibly nonsingleton set**. For instance, in the true QP *which man*, Q is *which*, T is *man*, and R is a set of two or more men...'

The wh-phrase in (3.18a) and the quantifier in (3.18b) are True Quantifiers and, therefore, give rise to wco. The ungrammaticality of coindexing in (3.18) is a consequence of the following principles:

I. In a configuration where a pronoun P and a trace T are both bound by a quantifier Q, T must c-command P.

II. There is a bijective correspondence between variables and A-bar positions
(i.e. each operator must A-bar bind exactly one variable, and each variable must be A-bar bound by exactly one operator).

They also assume Chomsky's definition of a variable:

III. α is a variable iff α is locally A-bar bound and in an A-position.

Note that (III) makes no distinction between empty categories and pronouns. So, in (3.18) both the pronoun and the trace qualify as variables, as they appear in A-positions. As the trace does not c-command the pronoun, coindexing cannot be licensed by I. Further, the Q cannot bind both the trace and the pronoun because this would violate II. Thus, the only grammatical structure for (3.18) is one in which the Q binds the empty category and the pronoun is locally free.

By contrast, no Q is present in (3.19). For these cases Lasnik & Stowell (1991) assume a Null Operator which is of *referential* nature. Unlike the wh-phrase or the QP in (3.18), the NP *John* in (3.19) has specific reference, the individual 'John'. In the absence of a Q, none of the principles of the grammar blocks coindexing between *John*, *his* and *e*.

Lasnik & Stowell (1991) discuss the presence of wco effects in focus constructions from English:

(3.20) a. His$_i$ mother SHOT John$_i$.

b. *His$_i$ mother shot JOHN$_i$.

c. His$_i$ mother bought a PICTURE of John$_i$.

(Lasnik & Stowell 1991:ex.82a-c)

They note: 'From the perspective of our theory 82b [3.20b here] is surprising...the focused NP does not seem to be a true Qunantifier'[p.716]. To circumvent this, they propose that the focused NP contains a covert operator *only*, 'which carries the semantic import of focusing' [p.716].

In sum, Lasnik & Stowell (1991) propose that wco is expected only in the presence of a quantifier. Wco in Focus-structures indicates that focus is a quantifier on a par with wh-phrases and quantifiers. The absence of wco effects in Topicalisation structures indicates absence of a quantifier and is attributed to a referential operator.

3.2.2 Kiss (1995): Focus movement and Topicalisation in Hungarian

Topicalisation and Focus-movement is also attested in Hungarian:

(3.21) a. Janos EVAT varta a mozi elott

John Eve-acc waited the cinema in-front-of

'John waited for EVE in front of the cinema.'

b. Evat **JANOS** varta a mozi elott

Eve-acc John waited the cinema in-front-of
'Eve was waited for by **JOHN** in front of the cinema.'

c. a mozi elott **JANOS** varta Evat

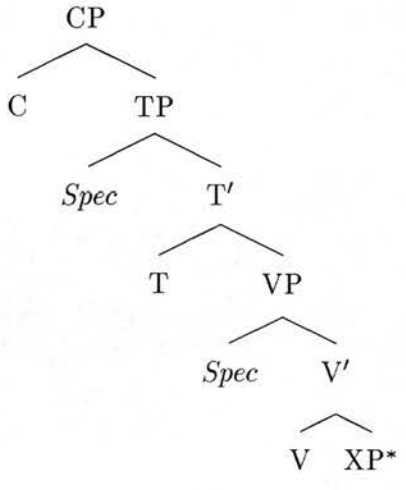
the cinema in-front-of John waited Eve-acc
'In front of the cinema, **JOHN** waited for Eve.'

(Kiss 1995b:ex.2)

As shown in (3.21), XPs can be topicalised or undergo Focus-movement irrespective of grammatical function.

Kiss associates focus and topic with preverbal positions in the syntactic tree. She assumes the following clause structure for Hungarian:

(3.22)



(Kiss 1995b:ex.9)

Topics appear at the Specifier position of the Tense Phrase (TP) and focus at the Specifier of VP. As both Topicalisation and Focus-movement obey strong islands and license p-gaps in Hungarian, Kiss assumes that they both instantiate movement. Topicalisation involves a relation of *predication* between the moved XP and the clause; no quantifier/operator is present in this structure. Focus-movement is Operator-movement[3] on a par with Wh-movement and Quantifier Raising. The quantificational nature of focus is supported by the presence of

[3] A-bar-quantificational movement in the terms of Lasnik & Stowell (1991).

wco effects in Focus-movement (Kiss 1995b). Further, the quantificational nature of focus is evident in examples like (3.23) in which focus creates scope:

(3.23) a. Janos **MARIVAL** valoszinu hogy talalkozott
John Mary-with likely that met
'It is Mary that John is likely to have met.'

b. Janos valoszinu hogy **MARIVAL** talalkozott
John likely that Mary-with met
'John is likely to have met Mary.'

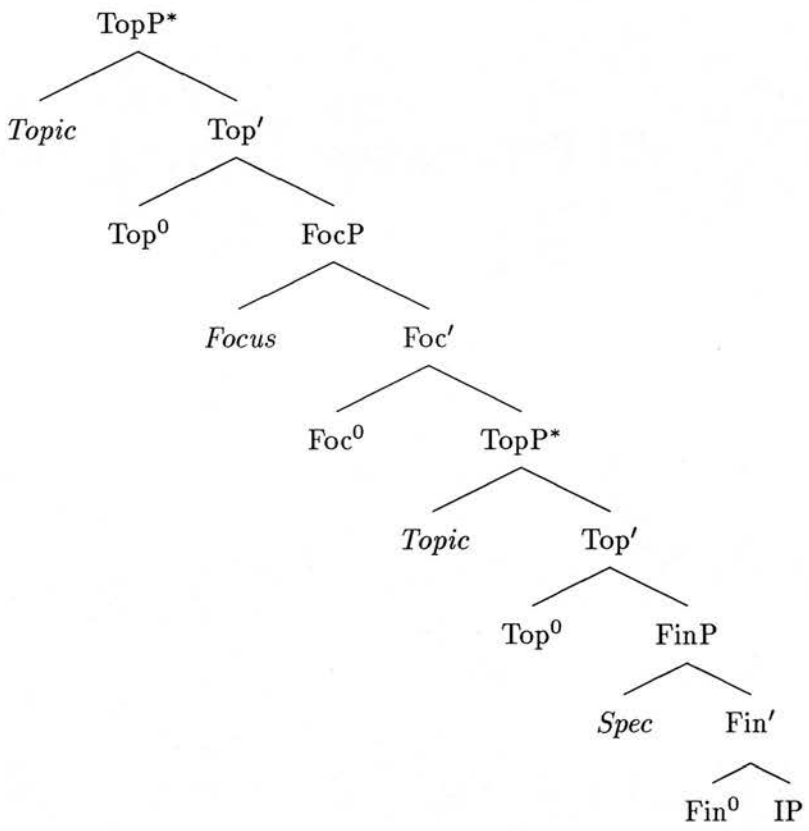
(Kiss 1995b:ex.48)

In (3.23a) the focused NP, *Marival*, takes scope over *valoszinu* whereas in (3.23b) it appears within the scope of *valoszinu*. She proposes that 'Focusing.....moves a constituent from an argument position into an operator position from which the moved constituent will c-command its scope' (Kiss 1995b:p.223).

3.2.3 Rizzi (1995): Discourse Functional Projections

Rizzi (1995) extends Phrase Structure with two preverbal Functional Projections hosting topics and foci [see also Brody (1990) for similar views]. The following tree illustrates his analysis:

(3.24)



(Rizzi 1995:ex.41)

Topics and foci move to the Specifier of the Topic Phrase (TP) and the Focus Phrase (FP) forced by the Topic and the Focus Criterion respectively [see Rizzi (1995), Agouraki (1993) and Brody (1990) for variants of the Focus Criterion]. The Topic Phrase can either precede or follow the FP:

(3.25) A Gianni, QUESTO, domani, gli dovrete dire
'To Gianni, THIS, tomorrow, you should tell him.'

(Rizzi 1995:ex.23)

Focus-movement instantiates quantificational A-bar movement in the sense of Lasnik & Stowell (1991). The sensitivity of Focus-movement to wco is the crucial evidence in support of the quantificational nature of focus:

(3.26) a. Gianni$_i$, sua$_i$ madre lo$_i$ ha sempre apprezato
'Gianni, his mother always appreciated him.'

b. GIANNI$_i$, sua$_{*i/j}$ madre ha sempre apprezato t (non Piero)
'GIANNI his mother always appreciated, not Piero.'

(Rizzi 1995:ex.17-18)

He also observes that focus resists resumptive pronouns in Italian:

(3.27) *IL TUO LIBRO lo ho comprato (non il suo)
the your book him-cl have-1sg bought (not the his)
'YOUR BOOK I bought it.'

(Rizzi 1995:ex.16a)

The unavailability of resumptives in (3.27) follows from the quantificational nature of focus. In (3.27) both the trace and the clitic are potential bindees. The trace does not c-command the clitic pronoun. Thus, principle I of Lasnik & Stowell (1991) cannot license coindexing. The violation of the Bijection Principle renders (3.27) ungrammatical. Note that the ungrammaticality of (3.18) and (3.27) follow from exactly the same principles. [In essence, Rizzi rules out (3.27) as a case of crossover. As it will be shown in Section 3.3.1, the clitic is an operator in his analysis and appears at Top^0. Rizzi does not clarify why the clitic qualifies as a variable in (3.27) since, according to his assumption, it does not appear in an A-position.]

Topicalisation is A-bar movement **anaphoric** (referential in Lasnik and Stowell's terms) in nature, as the insensitivity to wco suggests. The Top^0 instantiates an **Anaphoric Operator**, which rather than assigning a range to its bindee 'seeks for an antecedent to which it connects its bindee' [p.11]. In English Topicalisation the Anaphoric Operator is null:

(3.28) Your book, [OP[I bought t]]

(Rizzi 1995:ex.29)

A *predication* relationship connects the topic with the clause: 'Top^0 defines a kind of higher predication...its function is analogous to the function of AgrS ....which also configurationally connects a subject and a predicate' [p.5].

FP and TP have the same structure but differ in interpretation. FP instantiates a *focus-presupposition* sequence whereas TP a *topic-comment* one. The [Spec,FP] hosts the focused XP, the new information. The complement of Foc^0 is the *presupposition*, the given information; Rizzi's *presupposition* corresponds to Vallduví's ground (Vallduví & Engdahl 1996). Since the complement of focus contains the presupposition/ground information, it follows that it cannot contain new information at the same time, that is, it cannot contain another focused element. In this way, Rizzi rules out recursive foci. For example, in the ungrammatical (3.29), the complement of the FP hosting *Gianni* contains a focused phrase, *il libro*; *il libro* cannot convey given information (as part of the presupposition) and new information (as a focused XP) at the same time:

(3.29) a. * A GIANNI IL LIBRO daro (non a Piero, l'articolo)
'TO JOHN THE BOOK I'll give, not to Piero, the article.'

(Rizzi 1995:ex.22)

On the other hand, the complement of Top^0 is the *comment*. As discussed in Section 2.2, the topic-comment partition may very often overlap with but does not correspond to the ground-focus partion. It is possible that the *comment* contains both given/ground and new/focus information. Thus, the comment may contain another topic. Rizzi notes: 'nothing excludes that a comment... may be articulated in turn as a topic-comment structure, so topic phrases can undergo free recursion' [p.15]. In (3.30) the comment following the topic *il libro*, instantiates another topic-comment articulation with *a Gianni* as topic. The comment of *a Gianni* is another topic-comment articulation where *domani* functions as a topic:

(3.30) il libro, a Gianni, domani glielo doro senz'altro
'The book, to John, tomorrow, I'll give it to him for sure'

(Rizzi 1995:ex.21)

Finally, the incompatibility of focus with Wh-movement is accounted for by the assumption that in matrix questions the wh-phrase moves to [Spec,FP]. Wh-phrases and foci compete for the same position [in a similar spirit Brody (1990) proposes replacing the Wh-criterion with the Focus-criterion whereas Kiss (1995b) assumes that wh-phrases are subtypes of Focus]. However, in indirect clauses, foci are acceptable (3.5b). Rizzi speculates that foci and

wh-phrases appear in distinct positions in indirect clauses, but does not elaborate on this.

3.2.4 Discussion

Discourse Configurational approaches view the realisation of discourse functions as a purely structural issue and encode discourse functions directly in the syntax, through Phrase Structure Configurations. Implicit in this approach are several assumptions about the nature of discourse related phenomena as well as the Syntax-Discourse Interface:

1 Inasmuch as discourse functions are encoded in the Phrase Structure, they are expected to display syntactic properties: to be recursive in the way syntactic structures are and to be subject to syntactic constraints (e.g. subjacency).

2 The relative order between topics, foci and wh-phrases is captured through the order of the relevant projections in the syntactic tree, predicting a rigid ordering of these elements.

3 Topicalisation and Focus-movement involve movement to distinct syntactic positions and instantiate two distinct syntactic operations: anaphoric and quantificational movement respectively.

4 Focus is treated as a semantic quantifier, on a par with wh-phrases and quantifiers. Thus, discourse functions are dealt with at LF along with semantic ones.

5 There is a one-to-one relation between syntactic positions and discourse functions: each discourse function maps to a distinct Phrase Structure Configuration.

Points 1 & 2 make direct predictions about the empirical domain whereas points 3-5 are more theory internal. In this section, I will briefly comment on points 1 & 2 and then discuss in detail 3 & 4. The relation between syntax and discourse (5) will be discussed in Chapter 5.

The independence of Information Structure from syntax has been argued for in detail in Chapter 2. It was shown there that focus is neither recursive nor does it obey subjacency. Thus, by encoding focus in Phrase Structure, Discourse Configurational approaches make the wrong predictions with respect to the recursiveness of focus and its sensitivity to subjacency (1). Rizzi (1995) rules out recursive foci on the basis of inconsistent interpretation.

However, extending Phrase Structure with a functional projection and then alluding to interpretational devices to rule out its recursive appearance yields a rather inelegant account. This analysis introduces unnecessary complexity to the grammar and does not capture the independent nature of discourse functions. Even if this problem were solved, the absence of subjacency constraints on focus remains a problem for this analysis.

Kiss (1995b) gives no explanation for why or how foci appear at [Spec,VP] and topics at [Spec, TenseP]. It is not clarified whether there is something intrinsic in TenseP that matches with topic features or whether the choice of TenseP as the landing site of Topicalisation is a stipulation necessary to predict the surface word order.

Crosslinguistically, preverbal foci tend to appear adjacent to the verb: this has been noted for Hungarian (Brody 1990), Greek (Tsimpli 1995), Spanish (Vallduví 1995), Turkish (Hoffman 1995), Russian (King 1995). Italian and Catalan do not require adjacency (Rizzi 1995; Vallduví 1995). However, as will be discussed in Section 5.3.2, at least in Greek, adjacency is not as rigid a requirement as configurational accounts predict. It will also be argued there that discourse constraints on the relative order between topics and focus are weaker than syntactic/Phrase Structure ones. Thus, the Discourse Configurational approach is too rigid for the data.

Let us consider the view that Topicalisation and Focus-movement involve distinct syntactic operations (3). This analysis emphasizes the differences between the two structures but does not reveal their similarities. Note that, to a large extent, this analysis is based on the assumption that Focus-movement involves a semantic quantifier (4) whereas Topicalisation does not. Below, I will review this assumption and argue that it is not a valid one.

As discussed in Cann (1993) and Larson & Segal (1995), quantifiers, rather than having specific reference, that is, referring to an identifiable individual, range over a set of individuals that satisfy the *restriction* of the quantifier (Cann 1993). In the terms of Lasnik & Stowell (1991), a quantifier ranges over a possibly *non singleton set.* It is hard to see how focus satisfies these requirements. Consider the following:

(3.31) I saw JOHN/MY MOTHER/THE STUDENT.

In (3.31) the NPs *John, my mother, the student*, which are focused, have specific reference. Consider also examples (3.32). In (3.32a) *il tuo libro* is focused whereas in (3.32b) it functions as a topic:

(3.32) a. IL TUO LIBRO ho letto (non il suo)
the your book have-1sg read (not the yours)
'Your book I read (, not his).'

b. Il tuo libro, lo ho letto
the your book cl have-1sg read
'Your book I read it.'

(Rizzi 1995:ex.4,3 p.5)

In both examples *il tuo libro* is a definite NP with specific reference. This NP can only be used in (3.32a) if, in the speaker's understanding, there is only one book that belongs to the hearer. That is, it cannot range over a non singleton set of books that belong to the hearer. In this respect, there is no obvious sense in which focus can be analysed as a quantifier. In addition, the NP *il tuo libro* has the same non-quantificational interpretation in both examples in (3.32). Thus, not only are focused XPs not quantificational, but they do not differ from topicalised ones in this respect.

As already mentioned, Lasnik & Stowell (1991) do notice this problem and attempt to circumvent it by proposing that focus involves *only* as a covert operator. However, this assumption is not well motivated. Consider for example (3.33b), in which *Melina* is narrowly focused:

(3.33) a. pius ides sto party htes?
who-ACC.PL saw-2SG at-the party yesterday
'Which people did you see at the party yesterday?'

b. ida ti Melina ke m' epiasan ta nevra mu pali
saw-1SG the-acc Melina-acc and me-cl cought the nerves my again
'I met Melina and I got angry again.'

In (3.33a) the wh-phrase *pius* is plural, introducing the presupposition that the speaker of (3.33b) has seen more than one person. However, (3.33b) is a felicitous answer to (3.33a). The example (3.33b) cannot be paraphrased as *Melina was the only person I saw or I only saw Melina.* The speaker in (3.33b) is concerned with informing his/her interlocutor of the fact that s/he saw Melina. S/he is not concerned with whether or not Melina was the only person s/he met [see Heycock (1993) for similar observations].

In view of this, the solution proposed by Lasnik & Stowell (1991) cannot be sustained and, in effect, the quantifier analysis of focus cannot be sustained either.

Let us now turn to Kiss's claim that focus creates scope. Consider her examples, repeated below:

(3.34) a. Janos MARIVAL valoszinu hogy talalkozott
John Mary-with likely that met
'It is Mary that John is likely to have met.'

b. Janos valoszinu hogy MARIVAL talalkozott
John likely that Mary-with met
'John is likely to have met Mary.'

Whatever the exact nature of the interpretational differences between (3.34a) and (3.34b), both examples have the same truth conditions. In classical cases of scope interaction between more than one quantifiers/semantic operators (e.g. existential-universal quantifier) scope affects the propositional content. This is not so in (3.34). As a result, these examples do not provide evidence that focus is a quantifier.

In Greek, Topicalisation, by subtracting material from the VP, gives rise to interpretational differences similar to those in (3.34); consider examples (3.35) from Greek:

(3.35) a. Pistevo oti ide TO YANI
believe-1sg that saw-3sg the-acc Yani-acc
'I believe that s/h saw Yani.'

b. to Yani pistevo oti ton IDE, (ala ton Petro ohi)
the-acc Yani-acc believe-1sg that him-cl saw-3sg (but the-acc Petros-acc not)
'Yani, I believe s/he saw him, but Petros s/he didn't.'

In both examples in (3.35) *pistevo* is 'associated' with the embedded clause. In (3.35b) the object, *Yani*, has been extracted outside the embedded clause. However, the interpretational difference between these two examples is not one of propositional content.

Finally, let us consider wco facts. An unstated motivation for the analysis of focus as quantification is to preserve the generalisation that wco is a syntactic reflex of quantification. But, leaving aside focus, there is other evidence that this is not the correct generalisation. Consider the following examples displaying Wh-questions and QR in Greek. No wco is attested

in the (b) examples which differ minimally from the (a) ones by the presence of a clitic; the examples in (3.37) show the surface structure:

(3.36) a. pion$_i$ agapai i mana tu$_{j/*i}$?
who-ACC love-3SG the-NOM mother-NOM his-GEN
'*Who does his mother love?'

b. pion$_i$ ton$_i$ agapai i mana tu$_i$?
who-ACC him-CL love-3SG the-NOM mother-NOM his-GEN
'Who does his (own) mother love?'

(3.37) a. Kanena$_i$ den agapai i pethera tu$_{*i/j}$
nobody-ACC not love-3sg the-NOM mother-in-law-NOM his-GEN
'His mother-in-law loves nobody.'

b. Kanena$_i$ den ton$_i$ agapai i pethera tu$_i$
nobody-ACC not him-CL love-3SG the-NOM mother-in-law-NOM his-GEN
'*His (own) mother-in-law loves nobody/ Nobody is loved by his own mother-in-law.'

Similar data are attested in Romanian (Dobrovie-Sorin 1990) and Italian (Cinque 1990). The abscence of wco in the (b) examples above casts serious doubts on the association of wco with quantification and, in effect, on the structural distinction (quantificational/anaphoric A-bar movement) based on this association. This has two implications for the analysis of Focus. a) The presence of wco in Focus-structures does not constitute an argument that focus is a quantifier; b) The structural distinction between quantificational and anaphoric movement cannot be sustained. Consequently, the differences between Focus-movement and Topicalisation cannot be derived from the purported difference between quantificational and anaphoric movement.[In Alexopoulou (1997) I argue that wco is subject to discourse constraints. Thus, I attribute the difference between Topicalisation and Focus-movement with respect to wco effects to the differences in their Information Structure.]

The discussion in this section leads to the conclusion that a structural distinction between Focus-movement and Topicalisation cannot be sustained. In particular:

— The analysis of focus as a quantifier lacks empirical evidence and is inconsistent with current assumptions about the properties of quantifiers.

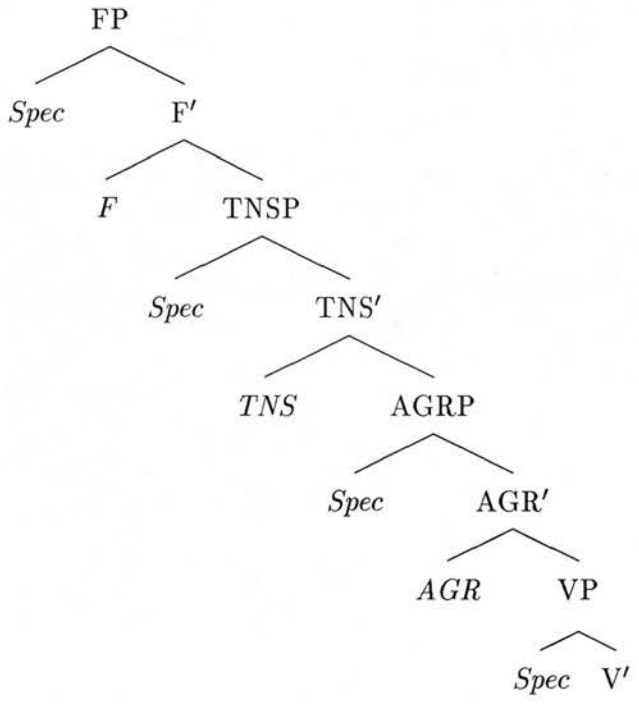
— Wco does not unambiguously indicate quantification; thus, the presence of wco in Focus-structures does not constitute evidence that focus is a quantifier.
— To the extent that wco is the only diagnostic distinguishing quantificational from anaphoric movement, this distinction is unmotivated.

Finally, Discourse Configurational approaches, as they stand, provide no account for cases of broad focus. In addition, though the importance of nuclear accent placement is acknowledged, no attempt is made for the integration of phonological facts in the grammar.

3.2.5 Tsimpli 1995-96: Topicalisation and Focus in Greek

Tsimpli (1995) and Tsimpli (1996) provide an account of Focus-movement for Greek in the same spirit as Rizzi (1995), though developed independently. She proposes the following clause structure for Greek:

(3.38)

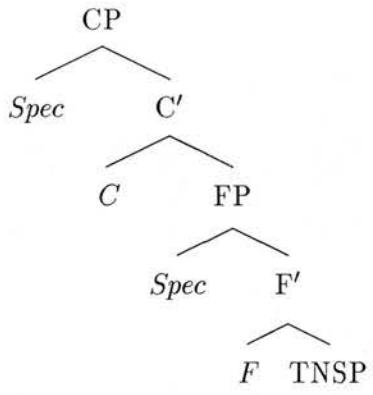


The tree in (3.38) illustrates the structure of matrix clauses. Tsimpli takes VSO to be the basic order for Greek. In the VSO order, the subject appears in its canonical position, [Spec,AGRP] and the verb in TNS. Tsimpli assumes that in SVO and VOS orders the subject is a topic. In both orders it occupies a Topic Position adjoined to TNSP. The difference between the two

structures is the directionality of the adjunction. Focused elements appear at the Specifier of the Focus Phrase (FP). Movement of foci to [Spec,FP] is forced by the Focus Criterion. Focus-movement can be delayed until LF; hence, in situ focus is grammatical. By LF however, Focus-movement is obligatory. In addition to the [f] feature, the head of FP is specified for the [wh] feature. Thus, in matrix Wh-questions, the wh-phrase competes with the focused phrase for the the [Spec,FP] position. This explains the ungrammaticality of focus in matrix Wh-questions. The [Spec,FP] is occupied by the wh-phrase leaving no room for Focus-movement.

The structure of embedded clauses, on the other hand, includes a CP projection:

(3.39)



Like FP, CP is specified for both the [f] and the [wh] feature. Example (3.40) provides evidence that C is specified for the [f] feature (in addition to [wh]). In (3.40) the focused phrase appears before the complementiser *oti* which occupies C. Thus, *to Yani* appears in [Spec,CP], which suggests that C is specified for the [f] feature:

(3.40) ipe to YANI oti ide i Maria
said-3SG the-ACC Yani-ACC that saw-3SG the-NOM Maria-NOM
'S/he said that Maria saw Yanis.'

As both CP and FP are specified for [f], foci may appear before or after the complementiser *oti*, at [Spec,CP] or [Spec,FP] respectively.

The semantic treatment of Focus is based on the following assumptions (Tsimpli 1996):

- Focus is a semantic operator on a par with wh-phrases and sentential operators like negation.
- It forces an individual reading.
- It always takes the widest scope.

Evidence for the quantificational nature of focus comes from its interaction with other operators. For example, whenever focus is within the scope of a wh-word the result is ungrammatical:

(3.41) a. *pios nomizi to YANI oti ide?
who-NOM think-3SG the-ACC Yani-ACC that saw-3SG
'Who thinks that s/he saw Yani?'

b. *se pion nomizis to VIVLIO oti edose?
to who-ACC think-2SG the-ACC book-ACC that gave-3SG
'To whom do you think that s/he gave the book?'

(Tsimpli 1996:ex.13a-b)

In matrix Wh-questions, where the domains of focus and wh-phrases are identical (since there is only one landing site for both foci and wh-phrases, [Spec,FP]), the result is ungrammatical:

(3.42) *pion ide o YANIS?
who-acc saw-3sg the-nom Yanis-nom
'Who did Yanis see?'

The scope interactions between wh-phrases and focus follow from interpretational differences between the two. Wh-phrases allow a functional (non-individual) reading whereas focus forces an individual (non-functional) reading. Tsimpli follows Chierchia (1992)—cited in (Tsimpli 1996:p.10)—in considering wh-phrases as 'quantifiers whose variable is bound by a function, the [wh] feature' (Tsimpli 1996:p.10). The availability of the functional reading is tested by the availability of (3.44b) below, as an answer to (3.43):

(3.43) who does everyone love x?

(3.44) a. Mary

b. John,Mary, Peter,Susan, Bill,Patricia.

As long as (3.44b) is a felicitous answer, (3.44a) can be interpreted as a special case of the paired-list reading involving a one-membered set. Consider now the echo question in (3.45):

(3.45) everyone loves who?

Question (3.45) cannot take (3.44b) as an answer, that is, the functional reading is blocked and only the individual is available. Tsimpli follows the suggestion of May (1985)—cited in Tsimpli (1996)—that wh-phrases in situ are focused. She then assumes that in (3.45) the functional reading is blocked because the wh-phrase is focused. In sum, the semantics of focus is essentially that of wh-phrases in echo questions.

Tsimpli argues further that wh-phrases and focus also differ in the way they acquire scope. The relative scope of a wh-phrase is determined by the position in which it is spelled-out in syntax. Overt Wh-movement is obligatory. Focus, on the other hand, may acquire scope irrespective of its position in syntax. As a result, Focus-movement is optional at syntax. However, in the presence of sentential operators, a focused phrase can only acquire wide scope through overt movement. The following data illustrate the interaction between focus and negation:

(3.46) a. den ida mono to YANI
not saw-1sg only the-acc Yani-acc
'I didn't see only Yani.'

b. mono to YANI den ida
only the-acc Yani-acc not saw-1sg
'It is only Yanis I did not see.'

(Tsimpli 1996:ex. 27b-c)

A natural continuation of (3.46a) is *I also saw Eleni* whereas (3.46b) can be followed by *I saw everybody else.* Tsimpli claims that in (3.46b) *mono to Yani* is extracted out of the scope domain of negation in order to acquire wide scope. The operator *mono*, which displays association with focus, is used to illustrate the effect.

Yes/no questions provide another case in which a sentential operator, the Question operator, blocks the wide scope reading of focus:

(3.47) a. ides to YANI?
saw-2SG the-ACC Yani-ACC
'Did you see Yani?'

b. to YANI ides?

the-ACC Yani-ACC saw-2SG
'Was it Yani that you saw?'

(Tsimpli 1996:ex. 26a-b)

In (3.47a) the nuclear accent is on *Yani.* Unlike declaratives, in which narrow focus can appear in situ, (3.47a) can only have a broad focus interpretation. In (3.47b), *Yani* is out of the scope of the question operator so that the narrow focus reading for the object is available.

Finally, the ungrammaticality of recursive foci follows from the fact that focus takes the widest scope. Consider the following:

(3.48) a. *TA VIVLIA agorase ya TI MARIA

the books bought-3sg for the-acc Maria-acc
'S/he bought the books for Maria.'

b. *STI MARIA ipe TON PETRO oti sinantise

to-the Maria said-3sg the-acc Petro-acc that met-3sg
'S/he said to Maria that s/he met Petro.'

(Tsimpli 1996:ex. 15)

The focused phrases *ti Maria* and *ton Petro* cannot take wide scope because they are within the scope of *ta vivlia* and *sti Maria* respectively. In this respect, (3.48) are on a par with (3.46a) and (3.47a).

Unlike foci, wh-phrases, need not take wide scope. Thus, recursive wh-phrases are grammatical (3.49):

(3.49) a. pios agorase ti ya pion?

who-nom bought-3sg what for who-acc
'Who bought what for whom?'

b. pios se rotise ya pion agorases to roloi?

who-nom you-acc asked-3sg for who-acc bought-2sg the watch
'Who asked you for whom you bought the watch?'

(Tsimpli 1996:ex. 14)

In sum, the syntactic distribution of foci is accounted for by allowing two preverbal projections, CP and FP, to be specified for the [f] feature. Semantically, focus is analysed as a

quantifier forcing an individual reading. In the absence of sentential quantifiers, focus takes the widest scope irrespective of its spell-out at syntax. In the presence of sentential operators like the Question Operator or Negation, overt movement at syntax is obligatory so that focus can take wide scope.

3.2.6 Discussion

To the extent that Tsimpli's analysis encodes discourse functions in Phrase Structure, it is subject to the overall criticism of Discourse Configurational accounts. It fails to capture the independence of Information Structure and cannot account for discourse-syntax mismatches. Furthermore, no analysis is offered for broad focus and there is no integration of phonology. In this section I will leave these issues aside and discuss the core of her syntax-semantic analysis.

Tsimpli postulates two distinct projections, FP and CP, both specified for the [wh] and [f] features. This assumption seems to undermine the notion of a functional projection. CP and FP look more like two positions in a tree stipulated in order to account for surface order. In addition, [f] and [wh] are conflicting features, as they imply an individual and a functional reading respectively. Tsimpli implicitly suggests that only one feature is checked each time. In particular she notes: 'the [wh] feature is suppressed when a [+wh] phrase also bears the [+f] feature' (Tsimpli 1996:p.13). Even so, the nature of the FP looks like a stipulation. It is meant to instantiate the [f] feature and then, it is specified for the [wh] feature which conflicts with the [f] feature.

Let us consider the idea that focus necessarily involves the individual reading. This generalisation is based solely on the semantics of wh-phrases in echo Wh-questions, which are assumed to be focused. Unlike ordinary wh-phrases (which are not focused), echo-focused wh-phrases are associated with an individual reading and do not allow the functional one (3.44a-b). However, the evidence from echo Wh-questions is very weak. First, recall that wh-phrases in echo questions in Greek are exactly the ones that do not pattern with focus (Section 2.8.1). Second, even if wh-phrases in echo questions are focused, Tsimpli does not explain why the individual reading in echo questions is a consequence of focus. A very plausible alternative explanation can be given in terms of the pragmatics of echo-questions. In an echo question the wh-phrase refers to an entity misconveyed in the previous discourse, which the utterer of the question would like to restore. Thus, the individual reading is expected. If the speaker feels

that the misconveyed entity is a paired-list, the question may take a different form, e.g. *who loves who?*.

Third, in some cases, wh-phrases in canonical Wh-questions may be associated with the individual reading. For example, the functional reading is available in (3.50a) but not in (3.50b); that is, a list answer is felicitous for (3.50a) but not for (3.50b):

(3.50) a. pion_i agapai i mana $\text{tu}_{j/*i}$?
who-ACC loves-3SG the-NOM mother-NOM his-GEN
'Who does his mother love?'

b. pion_i ton_i agapai i mana tu_i?
who-ACC him-CL loves-3SG the-NOM mother-NOM his-GEN
'Who does his (own) mother love?'

According to Tsimpli's assumptions, *pion* in (3.50b) should be focused. However, unlike (3.50a), (3.50b) does not display wco. In addition, *pion* is doubled in this example. Lack of wco and doubling make (3.50b) pattern with Topicalisation and CLLD rather than Focus-structures.

Finally, the generalisation that focus always forces an individual reading does not hold in environments other than echo Wh-questions:

(3.51) a. ena MATHITI agapai kathe mathitria
a-ACC.MASC student-ACC.MSC loves-3SG each student-NOM.FEM
'Each female student loves a male student.'

b. ENA mathiti agapai kathe mathitria
one-ACC.MASC student-ACC.MSC loves-3SG each student-NOM.FEM
'Each female student loves ONE male student.'

In (3.51a-b) *ena mathiti* can take either wide or narrow scope. Tsimpli predicts that only the wide scope/individual reading is available, that is, the one in which every female student loves the same male student. Contra to her predictions, the narrow scope reading is not only available but the preferred one (maybe due to the fact that the second quantifier is *each* rather than *every*). A possible continuation of (3.51a) would be *not a teacher*. Under such a contrast it is hard to get a reading in which every female student loves the same male student as opposed to the same male teacher. The same applies to (3.51b) in which the wide scope

reading is even harder than in (3.51a). We can thus conclude that the association of focus with the individual reading is not only based on weak evidence, but it is wrong.

Let us turn to the interaction of focus with sentential quantifiers. According to Tsimpli, in the presence of the Question operator, the focused phrase can only acquire wide scope through overt movement to a scope position (3.47b). However, it is not always the case that the sentential Question operator blocks the narrow focus reading:

(3.52) a. telika ides to MARCELO?
in-the-end saw-2SG the-ACC Marcelo
'In the end, was it Marcelo that you saw?'

b. tis agorase i DIMITRA tis sokolates?
them-ACC bought-3SG the-NOM Dimitra-NOM the-ACC chocolates-ACC
'Was it Dimitra who bought the chocolate?'

In (3.52a & b) narrow reading for *Marcelo* or *Dimitra* is possible, despite the focused phrase being within the scope of the sentential operator.[Note that the availability of a *narrow focus* reading means that the focus quantifier acquires *wide scope* over the Question operator.]

The examples illustrating interaction between focus and negation involve the operator *mono* ('only') (3.46). First, as discussed in Section 2.7, operators like *mono* are not always associated with the focus part of a sentence. Second, the presence of *mono* is a confounding factor, as it is not clear if (3.46) show the interaction between focus and negation or between *mono* and negation. In relation to this, note that no scope effects are induced when *mono* is not present:

(3.53) a. den ida to YANI
not saw-1SG the-ACC Yani-ACC
'I did not see Yanis.'

b. to YANI den ida
the-ACC Yani-ACC not saw-1SG
'I did not see Yanis.'

Both examples in (3.53) convey the same propositional content.

Tsimpli claims that when focus is within the scope of a wh-phrase (3.41) the result is ungrammatical. This claim is not a valid one as it is based on the assumptions that focus is

a quantifier forcing an individual reading and acquiring wide scope, both of which have been shown not to hold. In addition, compare (3.41a), repeated below as (3.54), with (3.55):

(3.54) *pios nomizi to YANI oti ide?
who-NOM think-3SG the-ACC Yani-ACC that saw-3SG
'Who thinks that s/he saw Yani?'

(3.55) rotise pios nomizi to YANI oti ide
asked-3sg who-nom think-3sg the-acc Yani-acc that saw-3sg
'S/he asked who thinks that s/he saw Yani.'

In (3.55) focus appears within the scope of the wh-phrase but ungrammaticality is not affected. An important difference beteween (3.55) and (3.54) is that the latter is an indirect question whereas the former is a direct one.

The availability of multiple wh-phrases is more restricted than Tsimpli claims. First, the question (3.49a) is an echo question. There is also a strong tendency for a comma intonation break after the verb and *what*, indicating separate intonational phrases. Example (3.49b) does not involve an echo question. However, as discussed in Section 2.8.1, in examples like this, the embedded wh-phrase does not take matrix scope. Rather it satisfies subcategorisation requirements.

Finally, recursive foci are excluded because focus has to take widest scope. As shown previously, this assumption is untenable and, consequently, so are explanations based on this assumption.

In conclusion, Greek provides no evidence that focus is a quantifier assigning an individual reading. The interpretation of NPs as quantificational is independent of focus and depends on the inherent properties of the NP attributed to their cognitive status. Furthermore, the association of focus with two preverbal projections (CP,FP) leads to a complicated account of the syntactic properties of Focus-movement while it fails to capture the structural similarities between Focus-movement and Topicalisation.

3.3 Clitic Left Dislocation-Topicalisation

CLLD was first noticed and studied by Cinque (1990). This structure has been of particular interest due to its paradoxical behaviour with respect to two diagnostics of movement: it obeys

subjacency but does not license p-gaps. This paradox, known as *Cinque's paradox*, has given rise to a controversy as to whether CLLD involves movement or not. Furthermore, in recent years, it has been realised that Topicalisation and CLLD are employed for the realisation of topics/links and, in addition, they behave similarly with respect to wco effects.

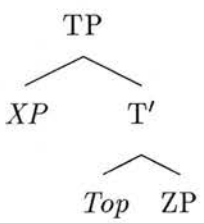
In the following sections I present the dominant views on the syntax and interpretation of CLLD. First, in Section 3.3.1, I review Rizzi (1995), who offers a parametric account of English Topicalisation and Italian CLLD. In Sections 3.3.3 & 3.3.4 I present Cinque (1990) and Iatridou (1995) who discuss CLLD in Italian and Greek and argue for a base-generation analysis. These papers are chosen as the most relevant to the issues discussed here. For elaborated discussions on Clitic Constructions in Greek and Romance see Agouraki (1993); Agouraki (1992); Anagnostopoulou (1994); Kallulli (1997); Niño (1994); Schneider-Zioga (1994) among others.

3.3.1 Rizzi (1995): a parametric account of *topic-comment*

Rizzi (1995) draws attention to the fact that the *topic-comment* structure is instantiated by Topicalisation in English and CLLD in Italian. He shows that, despite their superficial dissimilarity, Topicalisation and CLLD share interpretational as well as structural properties. As discussed above, both structures are insensitive to wco effects. Neither topicalised nor CLLDed constituents can bear the main sentential accent; topicalised and CLLDed constituents are interpreted as topics/links.

Building on these facts, Rizzi (1995) proposes a parametric analysis of *topic-comment* structures. Recall the structure of the Topic Phrase repeated below:

(3.56)



The English *topic-comment* structure has a null anaphoric operator:

(3.57) Your book, [OP[I bought t]].

(Rizzi 1995:ex.29)

In Romance, this operator is realised by the clitic. For example, direct-object-topics are coindexed with a clitic:

(3.58) il tuo libro, lo ho comprato
the your book, it have bought
'Your book, I bought it.'

(Rizzi 1995:ex.15a)

The parametric variation beteween English and Romance is explained as follows: 'The parameter differentiating English and Romance topic-comment structures resides in the non-availability of the null anaphoric operator in Romance topic-comment. Null operators and clitics are functionally equivalent here in that they establish the connection between the topic and the open position in the comment; Romance has the second device freely available while English, which lacks clitics in general, reverts to the first device' (Rizzi 1995:p.11-12).

While quantifiers license variables that they bind at LF, the anaphoric operator licences a null constant (Rizzi 1995:ex.28):

- A null constant is licensed by an anaphoric operator.

The distinction between *variables* and *null constants* explains the obligatoriness of clitics in the Italian *topic-comment* structure:

(3.59) *Il tuo libro ho comprato t
'Your book, I bought.'

(Rizzi 1995:ex.15b)

Compare (3.59) with (3.58). In (3.58) the empty category is a null constant licensed by the anaphoric operator (the clitic). In (3.59) there is neither a quantificational operator (e.g. focus) nor an anaphoric one. The empty category can be neither a variable nor a null constant; it is illicit and the example is ungrammatical.

3.3.2 Discussion

In his attempt to relate English Topicalisation and Italian CLLD, Rizzi brings to light several structural and interpretational similarities between the two phenomena. He seems to have

been the first to attempt a refined syntactic account of *topic-comment* articulations, at a *crosslinguistic* level. However, the parametric account he puts forward has various problems.

As mentioned above, clitics are treated as topic operators. However, Rizzi does not present any independent motivation for this analysis. Admittedly object NPs coindexed with a clitic (in CLLD, Clitic Doubling or Clitic Right Dislocation) receive a topic/ground interpretation. This, however, is not sufficient evidence for claiming that clitics are topic operators. First, clitics are pronouns that can appear on their own:

(3.60) a. ton ida
him-cl saw-1sg
'I saw him.'

b. lo visto
him-cl saw-1sg
'I saw him.'

No doubling takes place in (3.60) and no topic is involved. In addition, the clitic has the same interpretation the pronoun *him* has in the English translations. In every respect it behaves as an ordinary personal pronoun. It is not obvious how Rizzi's *topic operator* can account for these data.

Second, clitics do not resemble well known *topic markers* like the Japanese *wa*. *Wa* is attached to the noun (not the verb); it cannot function as a pronoun referring to a NP. In addition, it is not sensitive to the grammatical category of the NP it attaches to. *Wa* does not show case-marking and can attach to either subjects or objects (Hoji 1985:p.130-133). On the other hand, clitics are marked for case and can only corefer with grammatical categories marked for the same case (accusative, genitive and nominative in some dialects of Italian (Sanfilippo 1990)).

Third, in Rizzi's analysis, the clitic occupies a head position (Top^0). However, the clitic is a pronominal argument. Assuming that an argument is the head of a functional projection is a rather awkward hypothesis.

Thus, the properties of the pronominal clitics do not support the view that the clitic is a topic operator. Rizzi's assumption that the clitic is a topic operator explains the fact that 'the clitic seeks for an antecedent'. However, it is a property of pronouns in general to seek for an antecedent in the previous discourse. In the case of CLLD, the antecedent appears

within the clause.

As shown in Section 3.1, Topicalisation and CLLD share many structural properties but differ in that Topicalisation but not CLLD can license p-gaps. Rizzi's account does not predict this asymmetry.

Finally, in comparing Topicalisation with CLLD, Rizzi occasionally brings in English Left Dislocation. Examples (3.57) and (3.58) instantiate the same structure, *topic-comment.* The gloss/translation of (3.58) is expressed through Left Dislocation rather than Topicalisation. In addition, he uses English Topicalisation to gloss/translate the ungrammatical Italian example in (3.59). As shown in Section 3.1.3, Left Dislocation and Topicalisation are distinct constructions.

In sum, Rizzi's observation that Topicalisation and CLLD instantiate a *topic-comment* articulation is important and crucial for the understanding of these constructions. However, the analysis of the clitics as topic operators lacks empirical evidence.

3.3.3 Cinque (1990): a Base-Generation approach to CLLD

Cinque (1990) was the first to systematically study the Italian structure he called CLLD. He argues that CLLD involves no movement; rather, the dislocated elements are base-generated at their surface position. He considers two possibilities for a movement analysis for CLLD: a) CLLD is derived through movement from an underlying Clitic Doubling structure or b) CLLD instantiates Wh-movement and the clitic functions as a trace.

According to the first possibility, the dislocated phrase originates lower in the clause, it is doubled and then moves to some higher position. He rejects this option by observing that there are languages like Italian, which display CLLD but no Clitic Doubling, indicating that they are two distinct phenomena.

According to the second possibility, CLLD is the result of Wh/A-bar movement and the clitic is an overt wh-trace. Indeed, in some languages, resumptive pronouns behave as variables/wh-traces; however, such resumptives licence p-gaps and give rise to wco effects (Engdahl 1985). Clitics in Italian CLLD neither give rise to wco effects (Section 3.3.1) nor do they license p-gaps (3.61):

(3.61) *Gianni, l'ho cercato per mesi, senza trovare *e*
Gianni I have looked for for months without finding.

(Cinque 1990:p.62 ex.6)

The ungrammaticality of the p-gap in (3.61) and the absence of wco effects indicates that the clitic in CLLD is not a variable/wh-trace and that CLLD is not an instance of Wh/A-bar movement.

Another piece of evidence against movement is provided by the grammaticality of (3.62):

(3.62) Loro, il libro, credo che a Carlo sia sicuro che non glielo daranno
them, the book I-think that to Carlo it-is certain that they-will-never give it
mai
to-him.

(Cinque 1990:p.63,ex.10)

In (3.62) three constituents are left-dislocated from the lower clause. Cinque assumes some Topic Position adjoined to the left of CP [p.65]. *Loro* & *il libro* appear adjoined to the left of the matrix CP, whereas *a Carlo* is adjoined to the embedded CP. If *il libro* moves to its surface position, it should cross two CP barriers, violating subjacency. The grammaticality of (3.62) indicates that no movement is involved in this example.

So, Cinque concludes that no movement takes place in CLLD. Clitic Left Dislocated phrases are base-generated at their surface position. However, if CLLD does not involve movement, it should not obey island constraints. Surprisingly, CLLD does obey strong islands. Violation of the Complex NP constraint results in ungrammaticality in (3.63):

(3.63) *[PP A Carlo], ti parlero solo del [NP le persone [CP che gli
[PP to Carlo], I-will-talk to-you only about [NP the people [CP that to-him
piacciono]]
appeal]]

(Cinque 1990:p.59,ex.1f)

Cinque's answer to this paradox is that islands are a *condition on representation* rather than on movement. Islands constrain 'the possibility of entering binding chains...., which must be able to arise in either of two ways: via movement or base generation' [p.56-57]. See Cinque (1990:Chapter 1) for a detailed discussion of this view.

3.3.4 Iatridou (1995): CLLD in Greek

Iatridou (1995) follows Cinque's basic assumptions and argues that, in Greek, the dislocated object is base-generated to the left of CP. She rejects movement on similar grounds to Cinque (1990). As in Italian, Greek CLLD does not license p-gaps (3.64) and does not give rise to wco effects (3.65):

(3.64) *Afto to arthro i Maria to arhiothetise horis na-diavasi
this the article the Mary it filed without reading
'Mary filed this article without reading.'

(Iatridou 1995:ex.17)

(3.65) a. ton Kosta$_i$ i mitera tu$_i$ ton$_i$ agapa
the Kosta-acc the mother his him loves
'His mother loves Kosta.'

b. kathe pedi$_i$ i mitera tu$_i$ to$_i$ agapa
each child the mother his it loves
'His mother loves each child.'

(Iatridou 1995:ex.14)

She concludes that '... from the absence of wco violations and the unacceptability of parasitic gaps, we can conclude that there is no A-bar trace after the verb in a CLLD construction' (Iatridou 1995:p.14).

Following Cinque (1990) she argues that CLLD and Clitic Doubling are two distinct phenomena and provides additional evidence from Greek. Though Greek has both CLLD and Clitic Doubling, there are semantic classes of NPs that can appear in CLLD but not in Clitic Doubling:

(3.66) a. tria provlimata mono o-Kostas ta elise
three problems only Kosta them solved
'Three problems are such that only Kostas solved them.'

b. mono o-Kostas (*ta) elise tria provlimata
only Kosta them solved three problems.

(Iatridou 1995:ex.4)

Anagnostopoulou (1994) has taken this argument further by showing that CLLD and Clitic Doubling, even when grammatical, differ in felicity conditions:

(3.67) a. O kathigitis glosologias edose stous fitites bibliografia ke ton Chomsky
The professor linguistics-gen gave the students bibliography and the Chomsky
ton vrikan poli diskolo
him they-found very difficult
'The professor of linguistics gave a bibliography to the students and Chomsky they found very difficult.'

b. @O kathigitis glosologias edose stous fitites bibliografia ke ton vrikan
The professor linguistics-gen gave the students bibliography and him they-found
ton Chomsky poli diskolo
the Chomsky very difficult
'The professor of linguistics gave a bibliography to the students and Chomsky they found very difficult.'

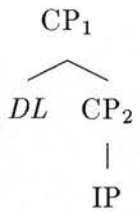
(Anagnostopoulou 1994:ex.13)

In the above, the definite NP *ton Chomsky* can appear both in CLLD and Clitic Doubling. In the context of *the professor gave a bibliography to the students*, CLLD is felicitous (3.67a) but Clitic Doubling is not (3.67b).

Finally Iatridou notes that there are languages like Rio Plates Spanish that impose an animateness restriction on Clitic Doubling but not on CLLD.

Having established that no movement is involved in CLLD, she proposes the following analysis. The dislocated phrase is base-generated at its surface position and is linked with the clause through *predication* (Iatridou 1995:p.20). The dislocated phrase is the subject of the predication and the clause is the predicate. The clitic appears because the verb must project its argument. It functions as the predicate variable and is adjoined to I^0. The dislocated DP is adjoined to the CP. The following tree illustrates her proposal:

(3.68)



(Iatridou 1995:p.20)

DL stands for discourse-linked position. The name reflects the fact that the dislocated XPs are discourse-linked.

Further, Iatridou argues that the dislocated phrase does not appear at the Specifier of a distinct maximal projection, but is adjoined to CP. She draws evidence from the fact that CLLD neither creates islands for extraction (3.69) nor does it block access of a higher verb to COMP (3.70):

(3.69) pios nomizis ti Maria oti tha tin psifize
who you-think the Maria-acc that FUT her vote
'Who do you think would vote for Mary?'

(Iatridou 1995:ex.24)

(3.70) anarotieme ton Kosta pios ton ide
wonder the Kostas who him saw
'I wonder who saw Kostas.'

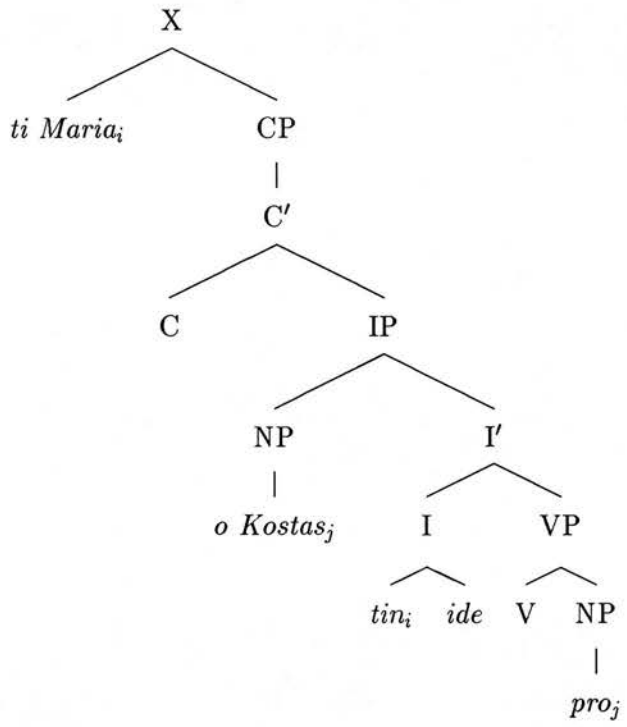
(Iatridou 1995:ex.25)

The grammaticality of (3.69) indicates that *ti Maria* cannot occupy [Spec, CP] (or any maximal projection); otherwise, extraction of *pios* would be impossible. In (3.70) the matrix verb *anarotieme* selects the maximal projection containing the wh-phrase, despite the intervention of *ton Kosta.* If the dislocated phrase appears at a maximal projection, the higher verb would not be able to govern the CP containing the wh-phrase and its subcategorisation requirements could not be satisfied. Thus, the dislocated NP must be adjoined to CP. The tree in (3.72) shows the structure of (3.71):

(3.71) ti Maria o Kostas tin ide

the Maria-ACC the Kostas-NOM her-CL saw-3SG
'Kostas saw Maria.'

(3.72)



(Iatridou 1995:ex.21)

As in Italian, CLLD in Greek obeys strong islands:

(3.73) a. *ton Kosta sinandisa tin kopela pu ton ide
the K. (I)-met the girl who him saw

b. *tin efimerida apokimithike diabazondas tin
the newspaper (he)-fell-asleep reading it

c. *ton Kosta diabasa tin idisi oti ton apelisan
the K. (I)-read the news that him (they)-fired

(Iatridou 1995:ex.28-31)

In (3.73a) *ton Kosta* belongs to the relative clause (Complex NP constraint), in (3.73b) *tin*

efimerida is the object of the adjunct participle and in (3.73c) *ton Kosta* is extracted out of a NP-island.

However, despite obeying strong islands, CLLD does not license p-gaps, as already mentioned. Iatridou's solution to this paradox is slightly different from Cinque's. She argues that islands do not constrain CLLD, that is, adjunction to the lower CP in (3.73). Rather, they constrain the movement from the adjoined position of the lower CP to the higher CP. She draws a parallel between (3.73) and examples of long distance CLLD:

(3.74) ton Kosta nomiza oti i Maria ton ide
the K. (I)-thought that the M. him saw
'I thought that Maria saw Kostas.'

(Iatridou 1995:ex.27)

The source of (3.74) is (3.75) below:

(3.75) nomiza ton Kosta oti i Maria ton ide
(I)-thought the Kosta that the M. him saw

(Iatridou 1995:ex.34)

Ton Kosta is base-generated at the lower CP and then it is adjoined to the matrix CP in (3.74). She concludes: 'Islands constrain the relationship between the position in which *ton Kosta* is generated (as in 34) [here 3.75] and the position in which it appears in (27) [here 3.74]. This is a movement relationship. This is movement out of an adjoined position and extraction out of such a position over an island is predicted to have the "heavy" feeling of an ECP violation, as in the case of adjunct extraction out of an island, and not a subjacency violation as when an object is extracted out of an island".

Finally she explains the grammaticality of (3.69) which indicates that CLLD does not create islands for extraction in the following way: 'This is because the DL-position and all the traces that the CLLDed element might leave on its way up are adjunction sites, and adjunction does not create islands, unlike A-bar movement through [Spec,CP], which does create islands by blocking up "escape hatches". This explains the superficially odd combination of properties that movement involved in long distance CLLD has: it obeys, but does not create, islands' (Iatridou 1995:p.24).

3.3.5 Discussion

Cinque (1990) and Iatridou (1995) argue that CLLD and Clitic Doubling are two distinct phenomena. The first argument they put forward is that there are languages like Italian which display CLLD but not Clitic Doubling. Note, though, that similar facts are true of Wh-movement across languages. There are languages with obligatory Wh-movement and others with Wh-in-situ available. Despite such cross-linguistic asymmetries, Wh-movement and Wh-in-situ have been thought as instantiations of the same phenomenon (Haegeman 1991:p.450). It is not obvious why the relation between CLLD and Clitic Doubling is different from the relation between Wh-movement and Wh-in-situ.

The second argument brought in support of the claim that CLLD and Clitic Doubling do not share the same underlying structure is semantic/interpretational in nature. Some semantic classes of NPs can undergo CLLD but not Clitic Doubling (Iatridou 1995). Moreover, CLLD and Clitic Doubling do not have the same felicity conditions (Anagnostopoulou 1994). These facts are no doubt true. However, they would only be relevant to the syntactic analysis, if these semantic/interpretational differences correlated with structural ones. But there is no evidence that they do. Leaving aside ordering facts, CLLD, Clitic Doubling and Clitic Right Dislocation behave uniformly. First, in all three structures, the doubled NP is optional. Second, none of these constructions gives rise to wco effects (3.76):

(3.76) ton$_i$ AGAPAI i mana tu$_i$ to Yani$_i$
him-CL love-3SG the mother-NOM his-GEN the Yani-ACC
'His (own) mother loves John.'

Third, more than one doubled XP is available:

(3.77) a. tis to DOSE htes, to vivlio tis Marias
her-CL-GEN it-CL-ACC gave-3SG yesterday, the book-ACC tis Marias-GEN
'S/he gave the book to Maria yesterday.'

b. tis to dose to vivlio tis Marias HTES
her-CL-GEN it-CL-ACC gave-3SG the book-ACC tis Marias-GEN
'S/he gave the book to Maria yesterday.'

Fourth, all three constructions are compatible with Wh-questions:

(3.78) a. pios tu tilefonise tu Aleksandrou htes to vradi
who him-CL-GEN phoned-3SG the Aleksandros-GEN yesterday the evening
'Who phoned Aleksandros yesterday evening?'

b. pios tha tou tilefonisi aurio to proi tu Aleksandrou
who-NOM will him-CL-GEN call tomorrow the morning the Aleksandros-GEN
'Who is going to call Aleksandros tomorrow morning?'

Finally, with respect to discourse import, all doubled NPs are ground elements.

Returning to *Cinque's paradox*, Iatridou claims that islands constrain movement from the adjoined position of the lower clause to the adjoined position of the higher clause (3.73). Note though, that examples (3.79), in which the dislocated phrase is adjoined to the left of the adjunct CP, are ungrammatical as well:

(3.79) a. *sinandisa tin kopela *ton Kosta* pu ton ide
met-1sg the girl-ACC the Kosta-ACC that him-CL saw-3sg
'I met the girl that saw Kostas.'

b. *apokimithike *tin efimerida* diabazondas tin
fell-asleep-3SG the newspaper-ACC reading her-CL
'S/he fell asleep reading the newspaper.'

c. *diabasa tin idisi *ton Kosta* oti ton apelisan
read-1SG the news-ACC the Kosta-ACC that him-CL fired-3PL
'I read the news that they fired Kostas.'

The ungrammaticality of (3.79) indicates that islands also constrain adjunction to the lower CP, that is islands constrain CLLD.

Finally, Iatridou argues that the dislocated phrase is adjoined to CP because CLLD does not create islands and does not block access of a higher verb to COMP. Further, she also distinguishes CLLD from Focus-movement which is A-bar movement and is predicted to create islands for extraction. However, as shown in Section 3.1.1, Focus-movement and Topicalisation are no different from CLLD: they do not create islands (3.6) and do not block access to COMP (3.4). Compare (3.69&3.70) with the following examples of Focus-movement:

(3.80) rotisa pios nomizis ti MARIA oti tha psifize
asked-1SG who-NOM think-2SG the Maria-ACC that would vote-3SG
'I asked who you think would vote for Maria.'

(3.81) anarotieme ton KOSTA pios ide
wonder-1SG the Kosta-ACC who-NOM saw-3SG
'I wonder who saw Kostas.'

Contrary to her assumptions about Focus-movement, Iatridou's argumentation entails that foci, rather than moving to [Spec,CP], must also be adjoined to CP.

To summarise, the claim that CLLD and Clitic Doubling are two distinct phenomena lacks syntactic motivation and does not capture the similarities between CLLD, Clitic Doubling and Clitic Right Dislocation. The base-generation analysis of CLLD can account for some structural properties of CLLD (e.g. unavailability of p-gaps). However, it cannot be extended to Focus-movement and Topicalisation which show the same properties as CLLD. Finally, the idea that islands constrain movement from an adjoined position on the lower CP to the higher one does not provide a satisfactory answer to *Cinque's paradox*, since adjunction to the lower CP is also ungrammatical.

3.4 Clitics and Quantifiers

3.4.1 Clitics and Quantifiers in Romance

A question that has received attention in the literature of Topicalisation/CLLD is the distribution of clitics in structures containing wh-phrases and quantifiers. Wh-phrases and quantifiers undergo movement at LF to a scope position from which they bind the trace/variable left in base position. Since clitics are not variables (Cinque 1990; Rizzi 1995), they should be excluded from Wh-questions and QR structures. However, crosslinguistic evidence does not verify this prediction.

The problem has been addressed in Cinque (1990), Dobrovie-Sorin (1990), Lasnik & Stowell (1991), Postal (1993) and Rizzi (1995) among others. Dobrovie-Sorin (1990) provides a detailed discussion. She concludes that wh-phrases and quantifiers, when coindexed with a resumptive pronoun, rather than undergoing QR, participate in a CLLD structure. When dislocated, wh-phrases receive a *specific/referential* or *d-linked* reading.

Romanian has two types of wh-phrases. *Cine=who* & *ce* (N')=*what* behave like wh-phrases in English; they cannot be coindexed with a resumptive pronoun, they licence p-gaps and display wco effect (Dobrovie-Sorin 1990:ex.2a-b,12,14).

On the other hand, the wh-phrase *care=which*, not only allows clitics (3.82a), but requires them (3.82b):

(3.82) a. Pe care (baiat) l-ai vazut?
pe which (boy) him-have (you)-seen
'Which one (which boy) did you see?'

b. *Pe care (baiat) ai vazut?
pe which (boy) have (you)-seen
'Which one (which boy) did you see?'

(Dobrovie-Sorin 1990:ex.3)

In Romanian, the doubled NP is accompanied by a dummy preposition *pe*. Data like (3.82a) exhibit the structural properties of CLLD: they do not licence p-gaps, do not give rise to wco but are sensitive to islands (Dobrovie-Sorin 1990:p.354-8).

The ungrammaticality of clitics in *ce*-structures is explained on similar grounds to Rizzi (1995). In both examples below the *ce*-NP undergoes Wh-movement/QR:

(3.83) a. Ce elev ai putea tu suporta?
what student could you stand

b. *Ce elev l-ai putea tu suporta?

(Dobrovie-Sorin 1990:ex.21)

The underlying structure of (3.83b) is shown in (3.84):

(3.84) wh_i cl_i e_i.

(Dobrovie-Sorin 1990:ex.9a)

Dobrovie-Sorin (1990) adopts the definition of variables of Chomsky (1981) according to which, variables are empty categories in an A-position, bound by an A-bar position and Case-marked. She also assumes, following Jaeggli (1982) [cited in Dobrovie-Sorin (1990)], that the clitic absorbs the Case of the object NP. As a result, the empty category in (3.84) is not case-marked and cannot be a variable. There is no variable to be bound by the wh-phrase. The ungrammatical (3.84) thus violates the principle of Full Interpretation.

For *care*-structures, in which the wh-phrase is coindexed with a clitic, she proposes that the wh-phrase does not undergo Wh-movement/QR. This is because, as Dobrovie-Sorin (1990) assumes, the [wh] features of *care* do not percolate to the dominating NP. Consider the following:

(3.85) a. Pe care baiat l-ai vazut?
pe which boy him-have (you)-seen
'Which boy did you see?'

b. Pe care l-ai vazut?
pe which-(one) him-have (you)-seen
'Which (one) did you see?'

(Dobrovie-Sorin 1990:ex.19)

In (3.85) *care* appears in [Spec,NP]. As *care*'s wh-features do not percolate to the NP, the dominating NP does not undergo Wh-movement, does not scope over the clause and does not bind a variable. The examples in (3.85) do not instantiate Wh-movement, but CLLD. The clitic is, therefore, grammatical. Dobrovie-Sorin claims that *care* takes scope only over the NP; it is a *restricted quantifier*, its restriction being the NP it appears in.

In sum, the difference between *ce* and *care* is one of scope. *Ce* takes scope over the clause, but *care* only over the NP. This is captured by their LF representations; (3.86) and (3.87) show the LF representations of (3.83) and (3.85) respectively:

(3.86) for what x, x is a student, you saw x.

(Dobrovie-Sorin 1990:ex.22)

(3.87) a. [NP_i for which x, x is a boy] you saw him_i.

b. [NP_i for which x, x is e] you saw him_i.

(Dobrovie-Sorin 1990:ex.20)

(In 3.87b **e** is identified by its antecedent).

According to Dobrovie-Sorin, the *ce*-NP ranges over the set of individuals that can satisfy the restrictions posed by the predicate (somebody you have seen). *Care* ranges over the set of individuals that can satisfy the referential properties defined by the lexical properties of N′,

the set of boys in (3.87a) or of the antecedent in (3.87b). This correlates with a systematic interpretational difference. *Care* is associated with a specific/d-linked reading. As Dobrovie-Sorin explains '*care* structures can be used only if a certain set of students has already been mentioned or is implicit in a given dialogue; *ce* structures suppose no such shared knowledge between the two speakers...in other words, the quantifying domain [of *care*] is independent of the sentence in which the quantified NP is used....on the other hand, the quantifying domain of *ce* (N′) is defined both by the lexical properties of N′ and by the properties of the rest of the sentence' Dobrovie-Sorin (1990:p.362).

This analysis is extended to indefinite NPs. Indefinite NPs can optionally undergo Clitic Doubling as shown in (3.88):

(3.88) a. Caut o secretara
(I)-look-for a secretary.

b. O caut o secretara
her (I)-look-for a secretary.

(Dobrovie-Sorin 1990:ex.52)

Example (3.88a) is ambiguous between a non-specific and a specific reading. Under the non-specific reading, the speaker is looking for any person who is qualified as a secretary, whereas under the specific reading s/he is looking for a certain secretary. When the indefinite NP is doubled (3.88b) the non-specific reading is not available.

The indefinite determiner is specified for *qu(antifier)*-features which may optionally percolate to the dominating NP. Percolation takes place when the non-specific reading is selected. The NP undergoes QR (at LF) and takes scope over the clause. When the specific reading is selected, no percolation takes place. The Q (the indefinite determiner) does not take clausal scope. It is restricted by the lexical properties of the nominal head (the class of secretaries, pupils etc.). The contrast between the specific and non-specific reading is captured by the LF representations in (3.89a-b):

(3.89) a. **non-specific**: there is an x such that x is a secretary and I look for x.

b. **specific**: there is an x such that x is a secretary and I look for *her*.

(Dobrovie-Sorin 1990:p.379)

In sum, when wh-phrases/quantified NPs undergo CLLD and appear coindexed with a clitic, their scope is restricted within the NP they appear in. In this respect, they behave like names rather than clausal quantifiers.

There is a group of quantifiers that behaves similarly to *care* in not transferring their *qu*-features to their dominating NP. *Toti=all, fiecare=each, oricare=whichever,any* obligatorily undergo Clitic Doubling, not QR.

Similar data are attested in Italian. The picture from Italian presents some interesting differences. Unlike Romanian, the compatibility of clitics with the quantifiers *nessuno* ('no-one') and *tutto* ('everything') does not depend on the lexical properties of the quantifier, but on whether or not the quantifier has a lexical restriction. In general, *nessuno* and *tutto*, when bare, resist CLLD (Rizzi 1995):

(3.90) a. *Nessuno, lo ho visto
'No-one, I saw him.'

b. *Tutto, lo ho fatto
'Everything, I did it.'

(Rizzi 1995:ex.19a-b)

However, when the same quantifiers appear with a lexical restriction, the examples improve to full acceptability:

(3.91) a. ?Nessuno dei tuoi libri, lo conosco veramente bene
'None of your books, I know it really well.'

b. Tutti i tuoi libri, li ho rimessi a posto
'All your books, I put them back.'

c. Molti libri, li ho buttati via
'Many books, I threw them away.'

(Rizzi 1995:ex.34)

Recall that in Rizzi's analysis, the clitic is the anaphoric operator involved in Topicalisation. Thus, the ungrammaticality of (3.90) is due to the clash between QR, which is quantificational, and Topicalisation, which is anaphoric. In (3.90) the quantifiers *tutto & nessuno* appear in [Spec, TP]. They need to bind a variable at LF. As already discussed (Sections 3.3.1&3.3.3) clitics can not be variables. The clitic trace[4] cannot qualify as a variable because it is a null constant licenced by the anaphoric operator (the clitic). If the quantified expressions are further moved by QR, they will leave a trace in the Topic position [Spec, TP], which is an A-bar position. So, no licit variable is available and the examples are ungrammatical.

To account for the acceptability of (3.91), Rizzi assumes that QR extracts the quantifier from the DP. The LF representation of (3.91c) is shown in (3.92):

(3.92) Molti [ec libri] TOP^0, [li ho buttati via].

(Rizzi 1995:ex.35)

In (3.92) the quantifier binds the variable within the [Spec,TP][5].

Cinque (1990) discusses the quantifiers *qualcuno* ('someone') and *qualche* ('some'). These quantifiers, even without a lexical restriction, can be coindexed with a clitic:

(3.93) a. Qualcuno, (lo) troveremo
someone we-(him) will-find.

b. Qualche errore, Carlo *(lo) ha fatto
some error Carlo (it) has made.

(Cinque 1990:ex.43-44)

Cinque points out that the presence of the clitic is not optional. The absence of the clitic implies an 'non-referential' interpretation of the quantifier (someone or other). The presence of the clitic implies that the quantifier is used 'referentially'. The speaker has someone/something specific in mind. Like Dobrovie-Sorin, he reaches the conclusion that, when coindexed with a clitic, these quantifiers participate in CLLD (Cinque 1990:p. 14-17).

[4] Rizzi does not clarify why the empty category is a clitic trace. The clitic appears at Top^0. It cannot have moved from an argument position to a head position.

[5] It is not clear why a variable can appear in [Spec,TP] since this is not an A-position.

3.4.2 Discussion

Dobrovie-Sorin (1990) brings to light very interesting data which are crucial to our understanding of the structural properties of quantification and CLLD. Further, she points out the importance of discourse effects in these structures. However, her analysis suffers from the fact that discourse effects (d-linking) and syntactic-semantic ones (Wh-movement, QR) are dealt with at the same level, that is LF. Such a treatment fails to capture the independence of Information Structure. In addition, it leads to the somewhat awkward conclusion that discourse linked wh-phrases/quantifiers do not take scope over the clause, but only over the NP they appear in.

First, it is unclear what it means for a wh-phrase/quantifier (of the kind discussed in this paper) to take scope over the clause. It is hard to establish any parallel with other semantic operators displaying scope effects. For example, the scope of negation (NP vs. clause) is typically associated with differences in propositional content which can straightforwardly be attributed to scope. It is unlikely that the data discussed in Dobrovie-Sorin (1990) can be explained along similar lines.

Second, the interpretational differences between doubled and non-doubled wh-phrases/quantifiers are differences of d-linking. Dobrovie-Sorin does not explain how or why d-linking affects the 'scope' of wh-phrases/quantifiers or why it should block Wh-movement/QR. All d-linking does is to restrict the set over which a wh-phrase/quantifier ranges. In Greek, doubling may reinforce the existential presupposition for the person/entity the wh-phrase refers to. For example, *nobody* is an infelicitous answer to (3.50b) repeated as (3.94):

(3.94) pion$_i$ ton$_i$ agapai i mana tu$_i$?
who-ACC him-CL love-3SG the-NOM mother-NOM his-GEN
'Who does his (own) mother love?'

However, doubling or d-linking does not affect the quantificational nature of the wh-phrase. Doubled wh-phrases still lack specific reference, which is the main motivation for treating them as quantificational (Cann 1993).

As mentioned in the previous section, the non-specific reading is blocked for Romanian doubled indefinite NPs (3.88b). Though this is very often the case in Greek, there are examples of Clitic Doubling constructions involving indefinite NPs (Section 2.5.3). Thus, this analysis could not be extended to Greek.

Finally, note that the LF representation for (3.89a), rather than showing the non-specific reading, as Dobrovie-Sorin claims, shows the specific one. Example (3.89a) presupposes the existence of a specific secretary.

3.4.3 Clitics and Wh-questions in Greek

The distribution of clitics in Wh-questions has received attention in the literature on Greek as well. In general, wh-phrases resist clitics in Greek:

(3.95) Pion (*ton) idhes?
who (him-CL) saw
'Who did you see?'

(Iatridou 1995:ex.46a)

However, wh-phrases coindexed with clitics are tolerated in examples like the following:

(3.96) pia pedhia (ta) maloses
which children them scolded
'Which children did you scold?'

(Iatridou 1995:ex.48)

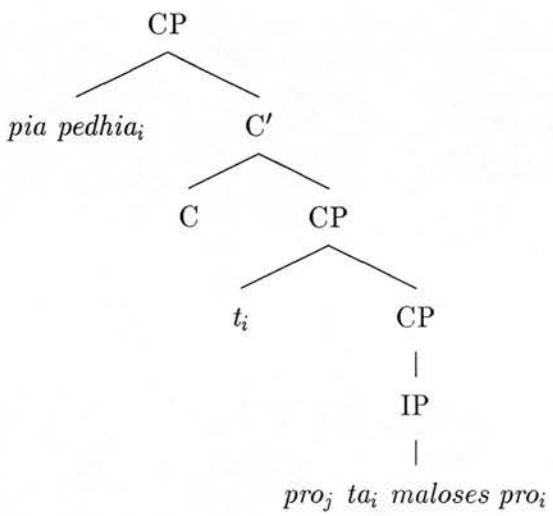
(3.97) [Pion apo tus dhio tus]$_i$ ipes oti dhen *ton*$_i$ simpathi i mitera
whom from the two of-them said-2,sg that not him(CLITIC) like-3,sg the mother-nom
su
of-you?
'Which of them did you say that your mother does not like?'

(Kolliakou 1991:ex.2-5a)

Iatridou (1995) observes that in (3.96), when *pia pedhia* is coindexed with the clitic it receives a d-linked interpretation. She notes: 'The two expansions of (48a) [here 3.96] (with and without a clitic) are not synonymous. Without the clitic, the sentence means something like "In the group of scolded people, which children fit?", while with the clitic it means "of the mentioned children, which ones did you scold?". In other words, the expansion with the clitic

has a different domain of discourse" (Iatridou 1995:p.27). She proposes the structure in (3.98) for (3.96)—the version with the clitic:

(3.98)



(Iatridou 1995:p.26)

In the above, the wh-phrase has been extracted from the DL-position.

The ungrammaticality of (3.95) is a consequence of independent (semantic) constraints on the d-linking of wh-words. She assumes that it is very hard to d-link a word like *who* (*pios*).

Her account explains the following contrast in the licensing of parasitic gaps:

(3.99) a. pion andhra pandreftike xoris na agapa
which man married without loves
'Which man did she marry without loving?'

b. pion andhra ton pandreftike (*xoris na agapa)
which man him(clitic) married without loves.

(Iatridou 1995:ex.56)

The grammaticality of the p-gap in (3.99a) indicates A-bar movement. The unavailability of p-gaps in (3.99b) suggests that the wh-phrase has been extracted from the DL-position. Additional evidence for the analysis in (3.98) comes from examples like (3.100). In this example the doubled wh-phrase is extracted from the DL-position of the embedded CP. The extraction from this position licenses a p-gap in the adjunct clause of the matrix:

(3.100) pion andhra ipe xoris na agapa PG [DL [oti tha ton pandrefti]]
which man say without loves that FUT him marry
'Which man did she say that she would marry without loving?'

(Iatridou 1995:ex.57)

Anagnostopoulou (1994) makes the interesting observation that wh-phrases which tolerate clitics do not obligatorily appear adjacent to the verb:

(3.101) posa apo ta abstracts i epitropi to Glow (ta) aperipse omofona?
how-many of the abstracts the committee the Glow rejected unanimously?
'How many of the Glow abstracts did the Glow committee reject unanimously?'

(Anagnostopoulou 1994:ex.47)

By contrast, wh-phrases which resist clitics are obligatorily adjacent to the verb:

(3.102) a. *pion o Jannis idhe?
*whom the John saw?
'Who did John see?'

b. *pion ton idhe o Jannis?
*whom him saw the John?
'Who did John see-him?'

(Anagnostopoulou 1994:ex.43)

In this respect, doubled wh-phrases pattern with adjuncts. In Greek, there is an asymmetry between wh-arguments and wh-adjuncts; while wh-arguments should appear adjacent to the verb (3.102a), wh-adjuncts can tolerate intervening subjects:

(3.103) a. jati o Peros ine lipimenos?
why the Peter is sad?
'Why is Peter sad?'

b. Pote o Petros protognorise tin Ilektra?
When the Peter-NOM first-met the Ilektra-ACC
'When did Peter first meet Ilektra?'

(Anagnostopoulou 1994:ex.12,14)

Further, Anagnostopoulou (1994) argues that the distribution of clitics in wh-phrases depends on the following *referentiality hierarchy*:

- *Referentiality Hierarchy*:
  overt partitive wh-phrases (*which of your books*); which-phrases (*which book*); what-phrases (*what book*); bare wh-phrases (*who, what*)

(Anagnostopoulou 1994:p.173-4)

The acceptibility of doubling reduces from the left to the right: overt partitives are fully grammatical; which-phrases are marginal; what-phrases and bare wh-phrases are ungrammatical (Anagnostopoulou 1994:p.173-174,ex.37-40). The *referentiality hierarchy* reflects a hierarchy of d-linking as well; overt partitives are typically the 'most' and bare wh-phrases the 'least' d-linked.

3.4.4 Discussion

Iatridou (1995) offers an analysis of doubled wh-phrases which differs from Cinque (1990) and Dobrovie-Sorin (1990) in an important way. She does not consider Wh-movement and CLLD incompatible. In other words, she allows a constituent (wh-phrase) to undergo CLLD and A-bar movement. The result is quite similar to Dobrovie-Sorin (1990). The variable bound by the wh-quantifier in (3.96) is not the actual argument *in* the clause, but it is adjoined to the CP and coindexed with a clitic in the clause. In this respect, the wh-phrase does not 'take scope' over the clause.

Iatridou claims that (3.95) is ungrammatical because *pios* is a wh-word hard to d-link. This should be true of (3.104) below:

(3.104) pion$_i$ ton$_i$ agapai i mana tu$_i$
who-acc him-CL love-3sg the mother-nom his
'Who does his (own) mother love?'

Pion in (3.104) is d-linked. Iatridou's analysis cannot explain the contrast between (3.95) and (3.104). Example (3.104) is a counterexample also for Anagnostopoulou's *referential hierarchy*. Contrary to her predictions (3.104) with a bare wh-phrase is much better than (3.96) which is equivalent to a which-phrase.

There is agreement in the literature that d-linked wh-phrases can be doubled. However, with the exception of (3.104), the acceptability varies from speaker to speaker. Moreover, the question in (3.105) is much better than (3.96):

(3.105) pia pedhia ipes oti ta maloses
which children said/2/SG that them scolded
'Which children did you say you scolded?'

(Iatridou 1995:ex.54)

It is unlikely that this difference in acceptability can be attributed to d-linking alone.

3.5 Conclusions

In this chapter, I have reviewed Discourse Configurational approaches to Focus-movement, Topicalisation and CLLD and tried to evaluate the proposals made within this paradigm. These proposals have revealed various syntactic and interpretational aspects of these structures: wco facts, p-gaps, extraction possibilities, the interpretational similarities between Topicalisation and CLLD, the question of the compatibility of quantifiers with clitics. However, various assumptions underlying these proposals have been shown to be problematic. In particular:

1 By encoding the discourse functions of topic and focus in Phrase Structure, Discourse Configurational approaches fail to capture the independence of Information Structure from Syntax.

2 They wrongly predict a rigid order between foci, topics and wh-phrases (see Chapter 5).

3 The association of Focus-movement, Topicalisation and CLLD with three distinct syntactic operations does not capture the structural similarities between them. Furthermore, this distinction has been shown to be not well motivated. In particular, the distinction between Focus-movement and Topicalisation is based on the assumption that focus is a quantifier and that wco indicates quantificational A-bar movement which is to be distinguished from anaphoric A-bar movement. Both assumptions have been shown to be invalid. Further, the base-generation hypothesis for CLLD, though it

may account for some properties of CLLD (e.g.unavailability of p-gaps) does not explain the sensitivity of CLLD to strong islands. In addition, it cannot be extended to Focus-movement and Topicalisation which show the same properties as CLLD. Finally, the base-generation analysis relies on the assumption that CLLD, Clitic Doubling and Clitic Right Dislocation are distinct syntactic phenomena. This claim was shown to lack syntactic evidence and to introduce unecessary complexity to the grammar of Clitic Constructions in Greek. In sum, the structural distinctions between Topicalisation, CLLD and Focus-movement cannot be sustained.

4 These proposals do not seem to account for cases of broad focus (e.g. VP) focus.

5 It is not obvious how these analyses can be extended to account for the interaction between Information Packaging and phonology.

In the next chapter I will argue for a unified syntactic treatment of Focus-movement, Topicalisation and CLLD and present an analysis of these constructions in HPSG.

Chapter 4

A Unified Grammar of Unbounded Dependencies

4.1 Introduction

In this chapter, I provide an alternative analysis to the one suggested by Discourse Configurational accounts. In particular, I propose a unified syntactic treatment of Focus-movement, Topicalisation and CLLD, which results in a reduced and more elegant syntax. This analysis is strongly supported by the syntactic properties shared by Topicalisation, CLLD and Focus-movement (Section 3.1), summarised below:

i) Crosslinguistically, all three constructions involve **long distance extraction**.

ii) All three constructions obey **strong islands**.

iii) In Greek, in all three constructions the extracted XP may appear on either side of the complementiser *oti*.

iv) None of the three constructions blocks selection from a higher verb.

v) None of the three constructions creates islands for extraction.

vi) Focus-movement and Topicalisation allow adjunct extraction. (In Section 4.6.5 I will argue that CLLD is an instance of adjunct extraction.)

These facts show that the syntactic properties of extracted XPs are not affected by their discourse import. Rather, the same extraction mechanism is employed to encode diverse discourse functions. The discourse function is disambiguated by intonation. For example, the following, which display Focus-movement and Topicalisation respectively, differ minimally in accent placement:

(4.1) a. tin **PARASTASI** skinothetise o Dimitris Potamitis
the performance-ACC directed-3SG the Dimitris Potamitis-NOM
'Dimitris Potamitis directed the performance.'

b. tin parastasi skinothetise o Dimitris **POTAMITIS**
the performance-ACC directed-3SG the Dimitris Potamitis-NOM
'Dimitris Potamitis directed the performance-The performance was directed by Dimitris Potamitis.'

Extractions may apply recursively:

(4.2) to party i **ELENI** to 'thele, (o Yanis den ihe oreksi)
the party the Eleni-NOM it-CL wanted-3sg (the Yanis not had appetite)
'Eleni wanted the party (Yanis did not feel like it).'

This analysis is reminiscent of CP-recursion (Iatridou & Kroch 1992) in that it allows recursive occurrence of extracted XPs at the left periphery of the clause. However, CP-recursion exhibits two properties that distinguish it from Greek extractions: it creates islands for extraction (Frisian) and blocks selection of an embedded clause by a matrix verb (English). Since neither of these properties holds for Greek, the analysis of Iatridou & Kroch (1992) cannot be adopted for the Greek data.

Extractions in Greek exemplify some universal properties: long-distance extraction, sensitivity to island constraints and licensing of p-gaps. Crosslinguistically, adjunct extraction also is quite a common phenomenon (Bouma *et al.* 1997). As will be shown in the following sections, these properties are readily captured by the mechanism of Unbounded Dependencies in HPSG. The organisation of this chapter is as follows. I first discuss some facts about the similarities and differences between argument and adjunct extraction in Section 4.2. In Section 4.3 I present the basics of the HPSG analysis of UDCs. Next, I present an account of long distance extractions and adjunct extraction in Sections 4.4&4.5. I then turn to the

syntax of Clitic Doubling and CLLD in Greek, in Section 4.6. Finally I discuss Wh-questions in Section 4.7 and conclude in Section 4.8.

4.2 Argument vs. adjunct extraction: some facts

In Greek, adjunct extraction shares various properties with argument extraction. In particular, adjuncts can undergo long-distance extraction and appear on either side of *oti* or before an indirect question (Section 3.1.1, examples 3.3b&3.4b). In addition, adjunct extraction is sensitive to strong islands; in (4.3) the adjunct, *me treno*, is extracted out of an adjunct-clause:

(4.3) *me treno eftasan pio grigora epidi taksidepsan
with train arrived-3PL more fast because travelled-3PL
'They arrived sooner because they travelled by train.'

The close resemblance of argument and adjunct extraction is not an idiosyncracy of Greek. Bouma *et al.* (1997) present evidence indicating that, crosslinguistically, adjunct and argument extraction are subject to the same constraints [see Hukari & Levine (1995), McCloskey (1979) and McCloskey (1989) cited in Bouma *et al.* (1997)]. For example, Irish, which distinguishes morphologically complementisers introducing a clause containing a gap from those introducing gapless clauses, shows no asymmetry between adjunct and complement extraction. The crosslinguistic similarities between adjunct and argument extraction point to a uniform treatment of the two.

However, adjunct extraction differs from argument extraction in two important ways. Crosslinguistically, sentential adjuncts are islands for extraction (4.3&3.2), whereas sentential complements are not.

Further, adjunct extraction does not license p-gaps. As already shown in Section 3.1.2 (examples 3.9), argument extraction licenses p-gaps:

(4.4) to YANI apelisan horis na proidopiisun
the Yani-ACC fired-3PL without PART warning-3PL
'They fired Yanis without warning (him).'

Adjunct extraction, however, does not:

(4.5) *me to TRENO taksidepse horis na theli na taksidepsi
with the train traveled-3sg without PART want-3sg PART travel-3sg
'S/he travelled by train without wanting to travel (by train).'

Example (4.5) is grammatical, but the phrase *by train* is not 'recovered' in the adjunct clause. P-gaps are licensed only in the presence of a proper gap. An in situ complement cannot license a p-gap. Unlike (4.4), the p-gap in (4.6) is ungrammatical:

(4.6) *apelisan to YANI horis na proidopiisun
fired-3PL the Yani-ACC without PART warning-3PL
'They fired Yanis without warning (him).'

Thus, there is a grammaticality contrast between (4.4) and (4.6). No such contrast is present between (4.5) and (4.7) in which the adjunct phrase in the matrix clause is in situ; both examples are equally bad:

(4.7) *taksidepse me to TRENO horis na theli na taksidepsi
traveled-3sg with the train without PART want-3sg PART travel-3sg
'S/he travelled by train without wanting to travel (by train).'

The data suggest that a uniform analysis of argument and adjunct extraction is needed, so that redundancy in the grammar of extractions is avoided. However, this analysis should be fine-grained enough to distinguish between the two at some level of representation, so that their differences are also accounted for.

4.3 Unbounded Dependencies: basics

The basic idea underlying the analysis of UDCs in HPSG is that the extracted XP, the *filler*, combines with a clause containing a *gap* corresponding to the *filler*. In the following example, *to Yani* is the *filler* combining with the S, *ide o Petros*.

(4.8) to Yani ide o Petros
the Yani-ACC saw-3SG the Petros-NOM
'Petros saw Yanis.'

Each *synsem* has a NONLOCAL feature which functions as a 'placeholder' for *gaps* (4.9):

(4.9)

$$\textit{synsem}\left[\begin{array}{ll} \textsc{local} & \left[\textsc{cat}\ \left[\begin{array}{ll}\textsc{head} & \ldots\\ \textsc{comps} & \ldots\\ \textsc{arg-str} & \ldots\end{array}\right]\right]\\ \textsc{nonlocal} & \left[\begin{array}{ll}\textsc{que} & \{\ldots\}\\ \textsc{rel} & \{\ldots\}\\ \textsc{slash} & \{\ldots\}\end{array}\right]\end{array}\right]$$

NONLOCAL takes three features as values; QUE and REL are used for the analysis of Wh-questions and Relative clauses respectively (Pollard & Sag 1994:p.159). The third feature, SLASH, is relevant for the analysis of Topicalisation and Focus-movement. In these structures, the information that a phrase contains a *gap* is encoded in its NONLOCAL|SLASH value. SLASH takes a *set* as value, so that multiple extractions are accounted for. In (4.8) the S *ide o Petros* is a saturated phrase (COMPS <>) with a *gap* in its NONLOCAL|SLASH value (4.10):

(4.10)

$$\left[\begin{array}{ll}\textsc{local}\,|\,\textsc{cat}\,|\,\textsc{comps} & \langle\rangle\\ \textsc{nonlocal}\,|\,\textsc{slash} & \{\boxed{1}\}\end{array}\right]$$

|

ide o Petros

As mentioned in Section 1.3.2 (ex. 1.26), *synsems* may be of sort *canonical* or *gap*. In a UDC, the ARG-STR of the verb contains a *gap* corresponding to the extracted XP. *Gaps* have the following structure (Sag & Miller 1997; Bouma *et al.* 1997):

(4.11)

$$\textit{gap-ss} \rightarrow \left[\begin{array}{ll}\textsc{local} & \boxed{1}\\ \textsc{nonlocal} & \left[\textsc{slash}\ \{\boxed{1}\}\right]\end{array}\right]$$

Gaps are synsems whose SLASH value is non-empty and, in addition, it is structure-shared with the LOC value of the *gap*. The ARG-STR of *ide* in (4.8) contains a *gap* corresponding to

the object NP (4.12):

(4.12)
$$\left[\begin{array}{ll} \textsc{syn}\,|\,\textsc{loc}\,|\,\textsc{cat} & \left[\begin{array}{ll} \textsc{head} & \textit{verb} \\ \textsc{comps} & \langle \boxed{1} \rangle \\ \textsc{arg-st} & \langle \boxed{1}\!:\!\mathrm{NP}\,\textit{canon}_{nom},\ \boxed{2}\!:\!\mathrm{NP}\,\textit{gap}_{acc} \rangle \end{array}\right] \end{array}\right]$$

In the above, 1 corresponds to the nominative NP, *Petros*, which is *canonical*. Arguments of sort *canonical* are realised locally as COMPS-DTRs and appear in COMPS. The accusative NP is a *gap*, to be bound by the *filler*, *to Yani*. Earlier accounts of UDCs in HPSG assume the existence of a *trace* that appears as a COMPS-DTR (Pollard & Sag 1994). However, recent work in the HPSG literature proposes a traceless account of extractions (Sag & Fodor 1994; Sag 1997). Thus, arguments of sort *gap* are not realised locally as COMPS-DTRs and do not appear in COMPS. The various possibilities for the realisation of the ARG-STR and COMPS of a *word* is captured by the ARGUMENTS REALISATION CONSTRAINT:

(4.13) ARGUMENTS REALISATION
$$\textit{Word} \rightarrow \left[\begin{array}{ll} \textsc{comps} & \boxed{1}\textit{list}(\textit{canon-ss}) \\ \textsc{arg-str} & \boxed{1} \bigcirc \textit{list}(\textit{gap-ss}) \end{array}\right]$$

The ARG-STR is obtained by *shuffling* COMPS 1 with the list of *gaps*. The symbol '○' stands for *shuffle* or *sequence union*. The *sequence union* of two lists A and B, yields a list C which is the set union of A and B, and preserves the order of the members in the original lists (i.e. if X precedes Y in A or in B, then X precedes Y in C) (Reape 1994; Kathol 1995; Sag & Miller 1997).

Since ARG-STR encodes information about the arguments of a verb irrespective of their realisation—e.g. locally as complements, nonlocally as fillers—, the number of its members is constant. For example, the ARG-STR of a verb like *vlepo* ('see') has always two members. When no extraction is involved, the ARG-STR is identical to COMPS (Section 1.3.2). The arguments are realised locally, as COMPS-DTRs of a *head-comp-phrase*. Extracted arguments on the other hand, do not appear in COMPS, since they do not realise any COMPS-DTR. This is ensured by the restriction that the members of COMPS are of sort *canonical*. In such cases,

the ARG-STR has the members of COMPS, which are of synsem sort *canonical*, plus the *gap* arguments.

In essence, ARGUMENTS REALISATION constrains the linking between ARG-STR and Phrase Structure (Manning & Sag 1995). Roughly, this constraint is analogous to the way movement relates Deep and Surface Structure in Principles&Parameters (Haegeman 1991). In HPSG, the linking between ARG-STR and Phrase Structure is handled by a declarative constraint rather than a tranformational mechanism[1].

The propagation of the *gap* from ARG-STR to higher positions of the tree is ensured by two constraints, the SLASH AMALGAMATION CONSTRAINT (SLAC) and the SLASH INHERITANCE PRINCIPLE (SLIP). SLAC ensures that a *gap* appearing in the ARG-STR of a verb will also appear in its NONLOCAL|SLASH value:

(4.14) SLASH AMALGAMATION CONSTRAINT (*preliminary version*)

$$\left[\begin{array}{ll} \text{LOC} & \left[\text{CAT} \left[\text{ARG-STR} \left\langle \left[\text{SLASH } \boxed{1}\right], \ldots, \left[\text{SLASH } \boxed{n}\right] \right\rangle\right]\right] \\ \text{NONLOCAL} & \left[\text{SLASH } \boxed{1} \cup \ldots \cup \boxed{n}\right] \end{array}\right]$$

According to the above, the SLASH value of a verbal head is the union ($\cup$) of the SLASH values of its arguments.

On the other hand, the SLASH INHERITANCE PRINCIPLE (SLIP) handles the configurational propagation of SLASH values from DTRs to MOTHERs:

(4.15) SLASH INHERITANCE PRINCIPLE

$$\textit{Head-val-phrase} \rightarrow \left[\begin{array}{ll} \text{NONLOC} & \left[\text{SLASH } \boxed{1}\right] \\ \text{HD-DTR} & \left[\text{SLASH } \boxed{1}\right] \end{array}\right]$$

SLIP states that the MOTHER of the phrase inherits the NONLOCAL|SLASH value of the HEAD-DTR. This constraint applies to phrases of sort *head-valence-phrase* (1.32), but not on *head-filler-phrases* where the *gap*[2] is bound by the *filler* (see below). By virtue of SLAC and SLIP,

[1]The HPSG treatment is rather inelegant in that two intuitively distinct levels—deep/argument structure and phrase-structure/constituency— are represented in one level. HPSG is forced into this solution due to its commitment to monostratal representations.

[2]Technically, the *filler* binds the SLASH value of the *gap* appearing in the ARG-STR. In the remainder, I will informally use the term *gap* to refer both to a synsem of sort *gap* and to the SLASH value of a *gap*.

heads inherit *gaps* from their ARG-STR and phrasal MOTHERs from their HEAD-DTRs. Note that a head can pick the *gaps* in its ARG-STR in examples like (4.8) as well as in examples of long-distance extraction (4.16):

(4.16) to Yani ipe oti ide o Petros
the Yani-ACC said-3SG that saw-3SG the Petros-NOM
'Petros said that he saw Yanis.'

The ARG-STR of *ipe* in (4.16) is shown in (4.17):

(4.17)
$$\left[\text{ARG-ST} \left\langle \boxed{3}\text{:NP}pro_{nom}, \boxed{4}\text{:S}\left[\begin{array}{ll}\text{HEAD} & verb \\ \text{SLASH} & \left\{\boxed{2}_{acc}\right\}\end{array}\right]\right\rangle\right]$$

By SLAC (4.14), the SLASH value of $\boxed{4}$ (*gap*$\boxed{2}$) appears in the SLASH value of the matrix head, *ipe*:

(4.18)
$$\left[\begin{array}{ll}\text{ARG-ST} & \left\langle \boxed{3}\text{:NP}pro_{nom}, \boxed{4}\text{:S}\left[\begin{array}{ll}\text{HEAD} & verb \\ \text{SLASH} & \left\{\boxed{2}_{acc}\right\}\end{array}\right]\right\rangle \\ \text{NONLOC}\,|\,\text{SLASH} & \left\{\boxed{2}\right\}\end{array}\right]$$

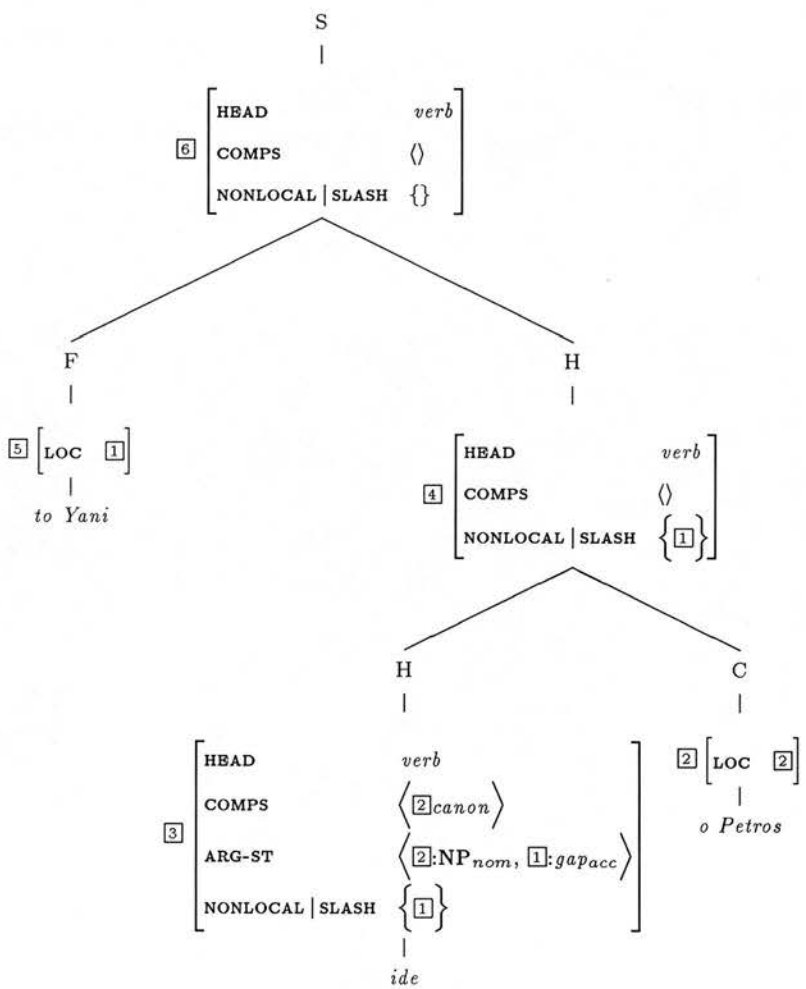
Finally, the *head-filler-phrase*, where the *gap* is bound by the corresponding *filler*, is licensed by the following ID-SCHEMA (adapted from (Bouma *et al.* 1997)):

(4.19) HEAD-FILLER-PHRASE ID-SCHEMA
$$\left[\begin{array}{ll}\text{NONLOC} & \left[\text{SLASH} \left\{\boxed{2},\ldots\boxed{n}\right\}\right] \\ \text{HD-DTR} & \left[\begin{array}{ll}\text{HEAD} & verb \\ \text{COMPS} & \langle\rangle \\ \text{SLASH} & \left\{\boxed{1},\boxed{2},\ldots,\boxed{n}\right\}\end{array}\right] \\ \text{NON-HD-DTR} & \left\langle\left[\text{LOC}\ \boxed{1}\right]\right\rangle\end{array}\right]$$

This schema licenses a *head-filler-phrase*. The **HEAD-DTR** of such a phrase is a saturated S (**HEAD**:*verb*, **COMPS** <>). The **LOC** value of the *filler* is token identical with that of the **SLASH** value of the **HEAD-DTR**, which is the **LOC** value of the *gap* in the **ARG-STR**. **LOC** contains information about the syntax and semantics of the relevant NP (case, index, restriction etc.). Since the *filler*'s **LOC** value is token-identical to the **HEAD-DTR**'s, the filler satisfies the subcategorisation requirements and the semantic restrictions imposed by the verb. Finally, the **SLASH** value of the bound *gap*, 1, is not inherited by the **MOTHER**. However, unbound *gaps* (2,...,n) 'pass-up' to their **MOTHER**, to be bound at some higher level of the tree.

The following tree illustrates example (4.8):

(4.20)



The verbal head, *ide*, has two arguments, shown in its **ARG-STR**. The subject NP2 is *canonical* and appears in the **COMPS** list. The object NP1 is realised *non-locally*, as a *filler*. Locally,

the ARG-STR contains a *gap*, the LOC value of which is token identical with that of the *filler*. The SLASH value of *gap* [1] appears in the NONLOCAL|SLASH value of the verbal head, *ide*, by SLAC (4.8). Then, it is inherited by the phrasal MOTHER [4] (SLIP). The *head-filler-phrase* is licensed by the HEAD-FILLER-ID-SCHEMA (4.19).

In sum, the HPSG analysis of UDCs is based on the following four constraints:

a) ARGUMENTS REALISATION: allows the members of ARG-STR to be realised as *gaps* and constrains the members of COMPS to be of sort *canonical*.

b) SLASH AMALGAMATION CONSTRAINT: ensures that *gaps* appear in the NONLOCAL|SLASH value of their head.

c) SLASH INHERITANCE PRINCIPLE: ensures that a MOTHER inherits the *gaps* appearing in the NONLOCAL|SLASH value of the HEAD-DTR.

d) HEAD-FILLER-PHRASE-ID-SCHEMA: licenses the *head-filler-phrase*.

4.4 Long distance extractions

The account of long-distance UDCs depends on the treatment of the complementiser *oti*. Two possibilities have been proposed in the HPSG literature. According to the first, the complementiser is a phrasal head and the clause is its complement (Sag 1997). Alternatively, the clause is the phrasal head and the complementiser is a syntactic *marker* (Pollard & Sag 1994). In the next sections I will consider briefly the two proposals with respect to the analysis of long-distance UDCs.

4.4.1 Compl as head

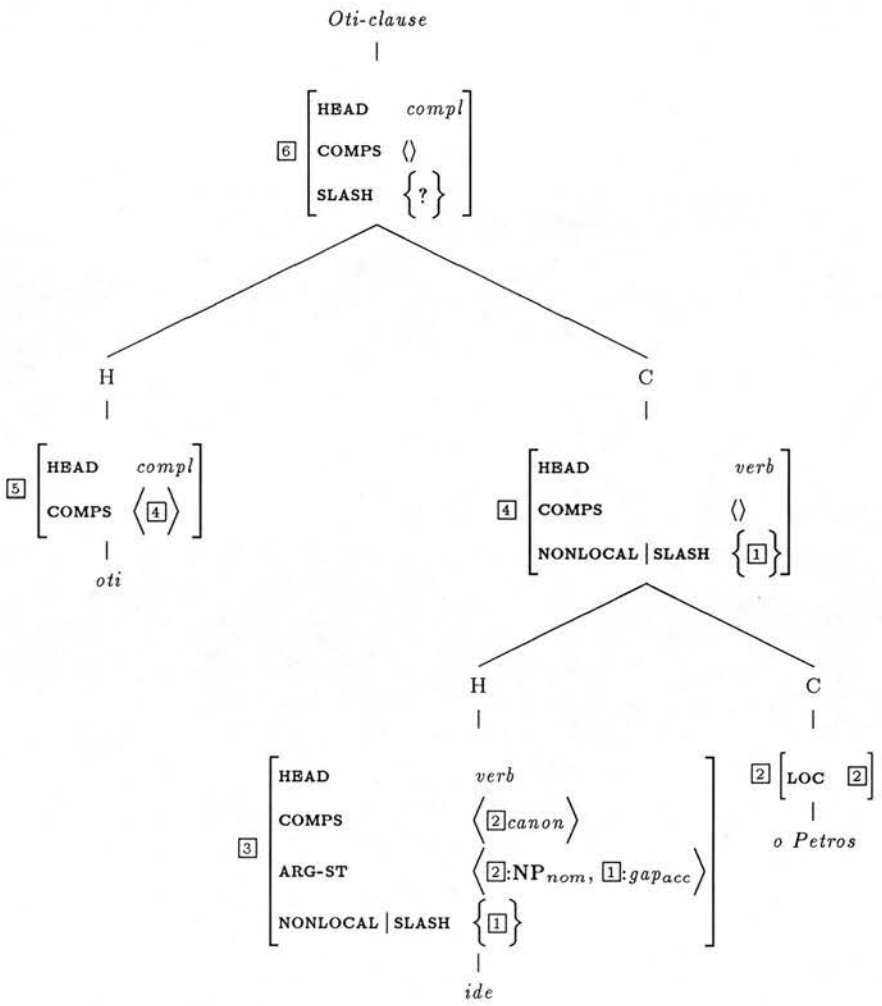
Consider first, the lexical entry for the complementiser *oti*, shown in (4.21):

(4.21)
$$\left[\begin{array}{ll} \textsc{phon} & \textit{oti} \\ \textsc{synsem}\,|\,\textsc{loc}\,|\,\textsc{cat} & \left[\begin{array}{ll} \textsc{head} & \textit{compl} \\ \textsc{comps} & \left\langle \left[\begin{array}{ll} \textsc{head} & \textit{verb} \\ \textsc{comps} & \langle\rangle \end{array}\right] \right\rangle \end{array}\right] \end{array}\right]$$

As shown in (4.21), *oti* takes a saturated S as a complement. The *oti*-clause is licensed by the HEAD-COMP-ID-SCHEMA (Section 1.3.3). *Oti* is the HEAD-DTR and the clause is the COMPS-DTR (4.22).

The head analysis of the complementiser faces two problems. The first problem involves the propagation of *gaps* in long-distance extractions. Consider the structure of the *oti-clause* in (4.16) shown in (4.22):

(4.22)



The SLASH value of the object *gap* 1 should be inherited by 6 so that it will be bound by *to Yani* at a higher node of the tree. According to SLIP a MOTHER inherits *gaps* from the HEAD-DTR. Thus, for *gap* 1 to appear at the MOTHER of the *head-complement phrase* 6, it has to appear at the SLASH of the HEAD-DTR of this phrase, that is, *oti* 5. In order to allow this, *gap* 1 has to appear in the ARG-STR of the HEAD-DTR, *oti*, since a head 'picks' *gaps* from

its ARG-STR (SLAC). Thus, we are forced to assume that *oti* has an ARG-STR similar to the one of the matrix verb *ipe* in (4.17).

The second problem is an empirical one, concerning the extraction possibilities of the complement clause. If the clause is the complement (and even more so, if it appears in the ARG-STR), it is expected to undergo extraction, as complements, in general, do. Contra to this expectation though, the following examples (from English and Greek) are ungrammatical:

(4.23) *[She has met John twice $]_j$ she said that $—_j$.

(4.24) *[tha ferotan toso anoita$]_j$ den to perimena pote oti $—_j$
*[would behave-3SG that stupidly$]_j$ not it-CL expected-1SG never that $—_j$
'I never expected that s/he would behave that stupidly.'

By contrast, clauses selected by verbal heads can undergo extraction[3]:

(4.25) a. [That she has met John twice $]_j$ she denied strongly $—_j$.

b. [That [problems this difficult $]_i$ our analysis would never be able to account for $—_i$ $]_j$ almost everyone had already been convinced of $—_j$.

(4.26) [Oti tha ferotan toso anoita $]_j$ den to perimena pote $—_j$
[that would behave-3SG that stupidly $]_j$ not it-CL expected-1SG never that $—_j$
'I never expected that s/he would behave that stupidly.'

In order to sustain the hypothesis that the complementiser is a head, we need an explanation for why clausal complements selected by this head cannot undergo extraction[4].

The picture is even more complicated once long-distance extraction is considered (4.28):

(4.28) John she said that she has met twice.

The compl-as-head hypothesis leads to the awkward descriptive generalisation that complements of *that* and *oti* cannot be extracted, but their embedded complements can.

---

[3] I owe examples (4.25) to Gosse Bouma and Ivan Sag.

[4] Some adjunct clauses cannot be preposed:

(4.27) ???pu fevyis lipame
???that/because leave-2SG am-sorry-1SG
'I am sorry that/because you leave.'

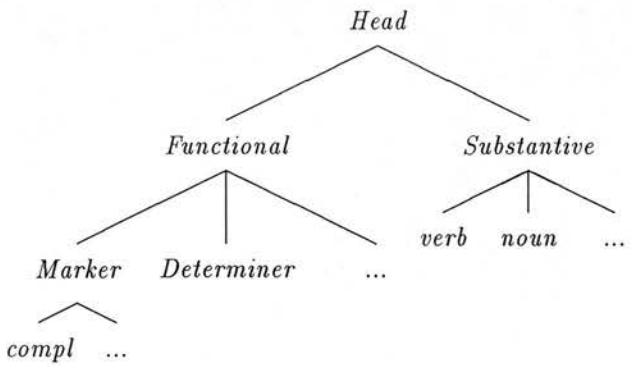
I owe this observation to Dimitra Kolliakou.

The idea that the complementiser is a phrasal head has the advantage of accounting for complementisers using an existing category (syntactic head). However, this solution is neither empirically, nor technically appealing. The complementisers considered here do not seem to share properties of other phrasal heads (e.g. verbs). Clauses selected by a verbal head can undergo extraction, but the ones selected by a complementiser cannot. As in HPSG most of the information is handled by phrasal heads, treating complementisers as heads forces them to play too powerful a role in the grammar. Cases of long-distance extraction can only be accounted for on the counterintuitive assumption that that the complementiser has some kind of ARG-STR. Though, such a solution may work for cases of long-distance extraction (4.28&4.16) it cannot block examples like (4.23&4.24). In view of this, this hypothesis is abandoned[5].

4.4.2 Compl as Marker

Alternatively, the complementiser *oti* can be analysed as a *marker*. In English, examples of *marker* are complementisers and the comparative words *than* and *as* (Pollard & Sag 1994). *Marker* is a subsort of *head/part-of-speech* (4.29) and is, therefore, an appropriate value for the attribute HEAD:

(4.29)

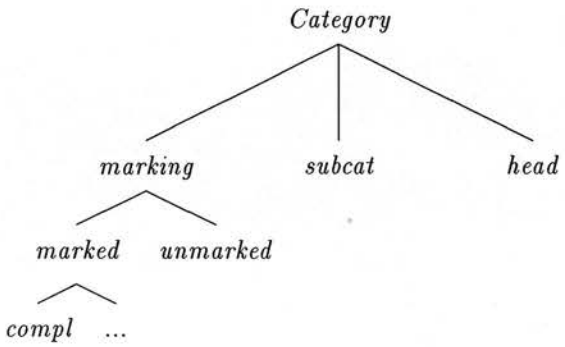


MARKING is a CATEGORY attribute with values of sort *marking*. The sort *marking* has two subsorts, *marked* and *unmarked*. *Complementiser* is a subsort of *marked*, and *oti* a subsort of

[5] Gosse Bouma and Dimitra Kolliakou pointed out to me that the ungrammaticality of (4.23) & (4.24) could be accounted for by constraining the argument of the complementiser to be of sort *canonical*.

Complementiser. The tree below illustrates the sortal hierarchy of *marking*:

(4.30)



An *oti*-clause is a phrase of sort *head-marker-phrase* with two DTRS: the HEAD-DTR and the MARKER-DTR. The HEAD-DTR is an S (HEAD:*verb*, COMPS:<>).

Though the *marker* is not the head of the *head-marker-phrase*, it resembles a head in that it selects its sister clause. This is done through the attribute SPEC in the lexical entry of the *marker*:

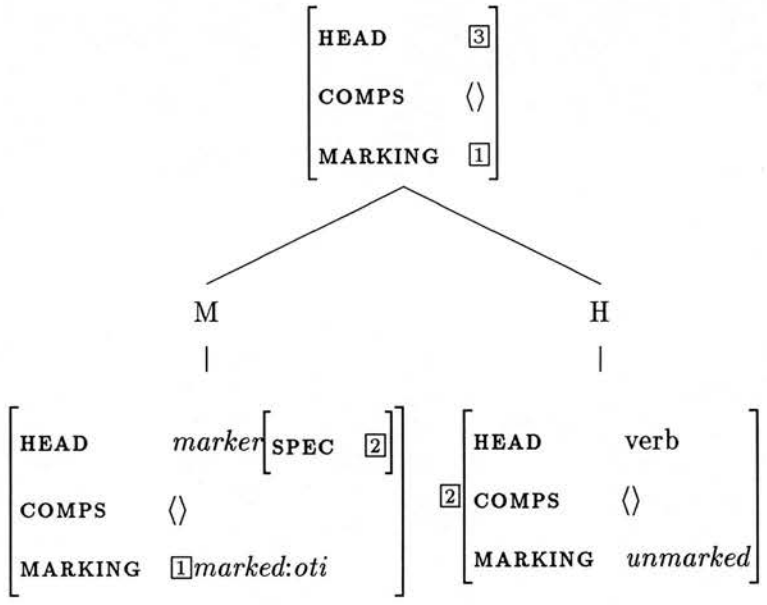
(4.31) The lexical entry of *oti*

$$\begin{bmatrix} \textsc{phon} & \textit{oti} \\ \textsc{synsem}\,|\,\textsc{loc}\,|\,\textsc{cat}\,| & \begin{bmatrix} \textsc{head} & \textit{marker}\begin{bmatrix} \textsc{spec} & \boxed{2}\begin{bmatrix} \textsc{head} & \textit{verb} \\ \textsc{comps} & \langle\rangle \\ \textsc{marking} & \textit{unmarked} \end{bmatrix} \end{bmatrix} \\ \textsc{comps} & \langle\rangle \\ \textsc{marking} & \textit{marked:oti} \end{bmatrix} \end{bmatrix}$$

Through SPEC, the marker selects an *unmarked* S (MARKING:*unmarked*). This restriction is meant to block recursive occurence of *oti*.

The *oti*-clause is a *head-marker-phrase* licensed by the HEAD-MARKER-ID-SCHEMA (4.32) (Pollard & Sag 1994):

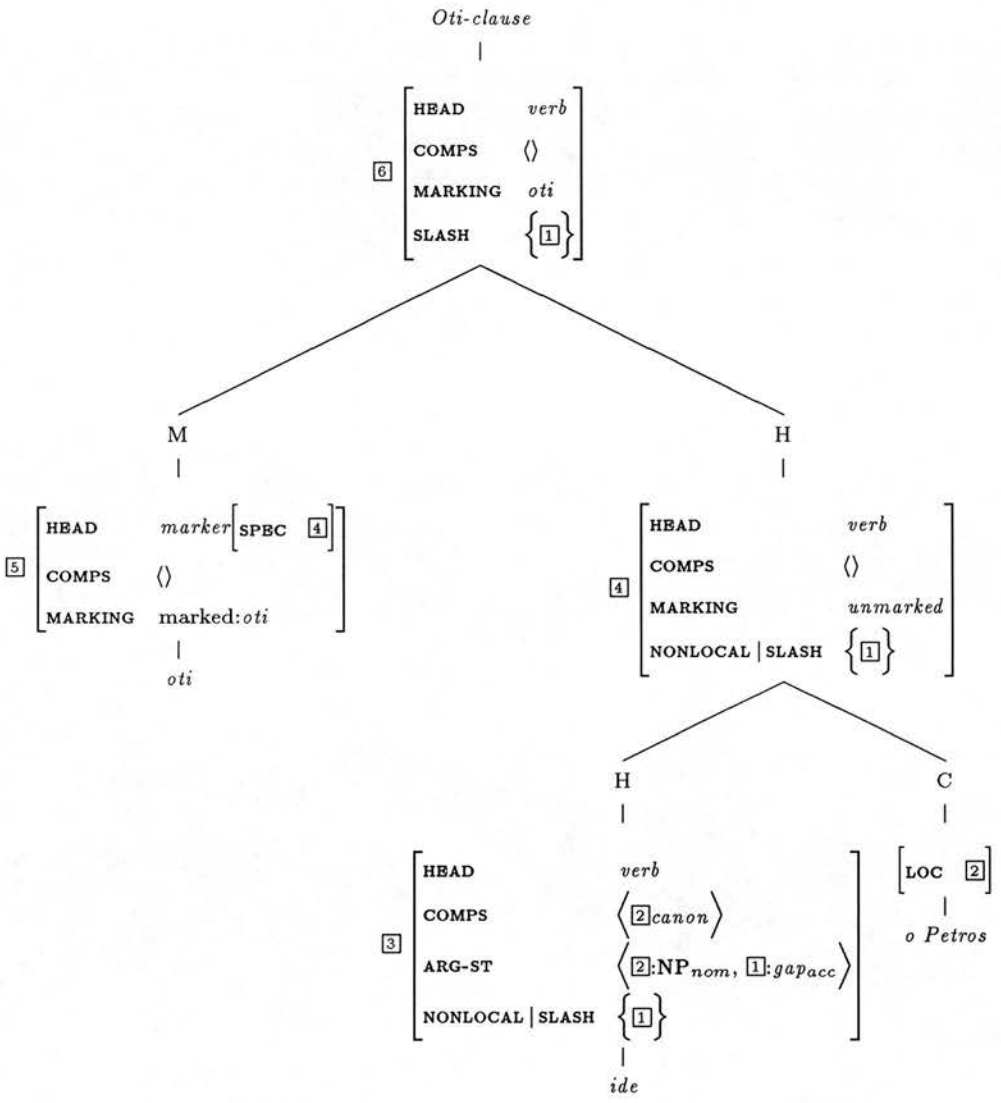
(4.32) HEAD-MARKER-ID-SCHEMA



The above schema states that the synsem value of the HEAD-DTR (2) is structure shared with the SPEC value of the MARKER-DTR. Further, it requires that the MOTHER inherits the MARKING value of the MARKER-DTR (1).

Under the marker analysis of the complementiser, the structure of (4.16) is the following:

(4.33)



The **MOTHER** of the *head-marker-phrase* 6 inherits the **SLASH** value of its **HEAD-DTR**, (**SLASH** 1). Thus, the propagation of **SLASH** values follows straightfowradly and long-distance extraction is accounted for.

Further, as a clause selected by a marker is not the complement of a head and does not appear in any **ARG-STR**, it is not expected to undergo extraction. The ungrammaticality of (4.23&4.24) is predicted. On the other hand, clausal complements selected by verbal head are expected to undergo extraction (4.25&4.26).

The idea that the complementiser is a marker and that the *oti*-clause is a constituent,

head-marker-phrase, suffers from the fact that it looks like a stipulation. The feature SPEC in the lexical entry of the complementiser is otherwise unmotivated machinery to allow the *marker* to select the clause. However, this treatment is more successful in capturing the nature of the complementiser. The propagation of SLASH values in long-distance extractions follows with no further stipulations. In addition, this approach makes the right predictions about the empirical domain. The contrast between (4.23&4.24) and (4.25&4.26) is predicted as a consequence of the fact that the clause in (4.23&4.24) is not a complement selected by a head.

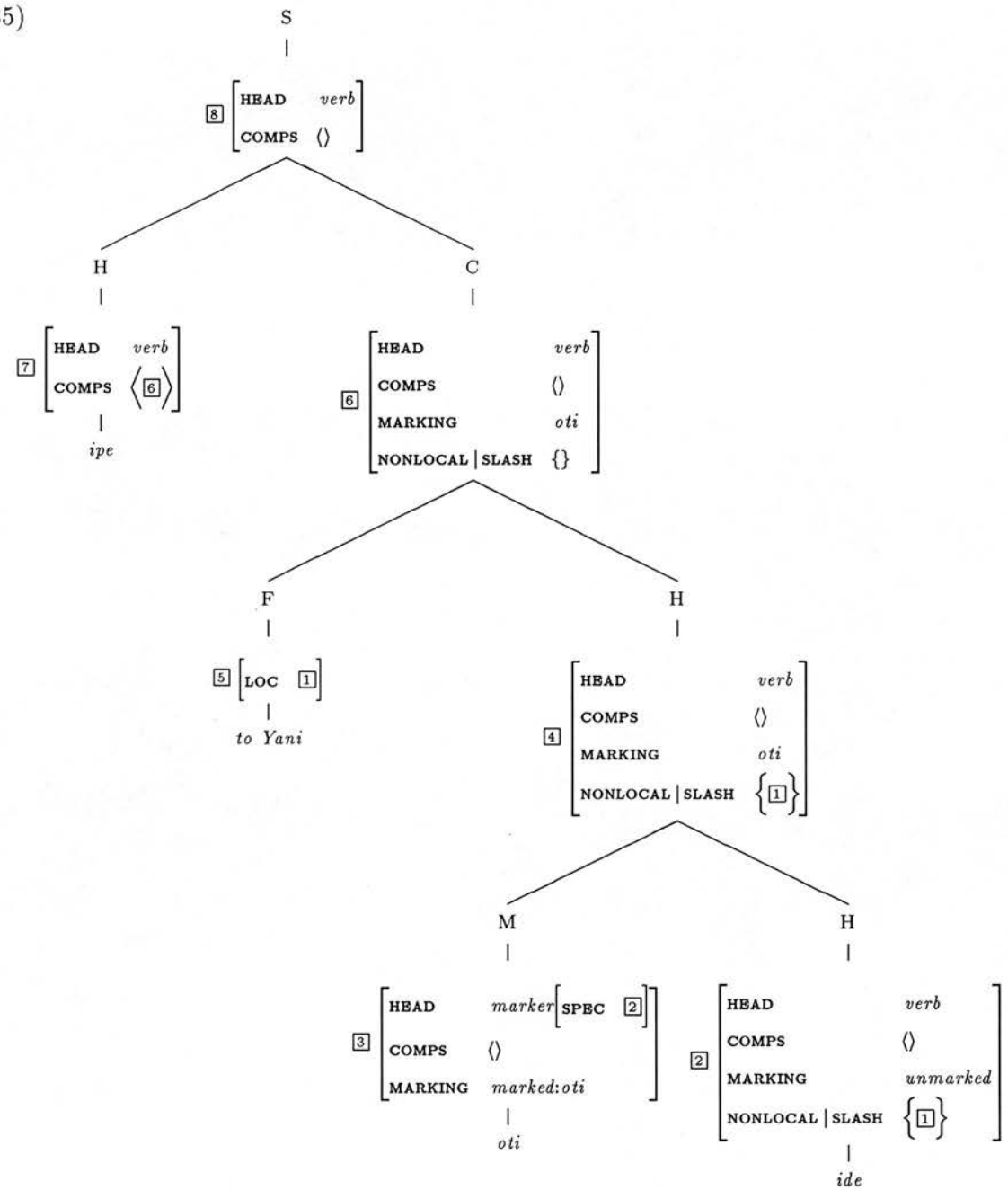
4.4.3 Long-distance extractions and complement selection

One of the properties of UDCs in Greek is that, in examples like (4.34), the *filler* does not block the selection of the lower clause by the higher verb:

(4.34) ipe to YANI **oti** ide
said-3SG the Yani-ACC that saw-3SG
'S/he said that s/he saw Yanis.'

Consider the structure of (4.34) shown in (4.35):

(4.35)



The *head-marker-phrase* 4 is licensed by the HEAD-MARKER-ID-SCHEMA (4.32). By SLIP it inherits the SLASH value of the HEAD-DTR 2 (SLASH:1) and the MARKING value of the MARKER-DTR (MARKING:*oti*). The *head-filler-phrase* 6 is licensed by the HEAD-FILLER-ID-SCHEMA. Its SLASH value is empty, as the *filler* 5 binds the *gap* 1 in the SLASH of the

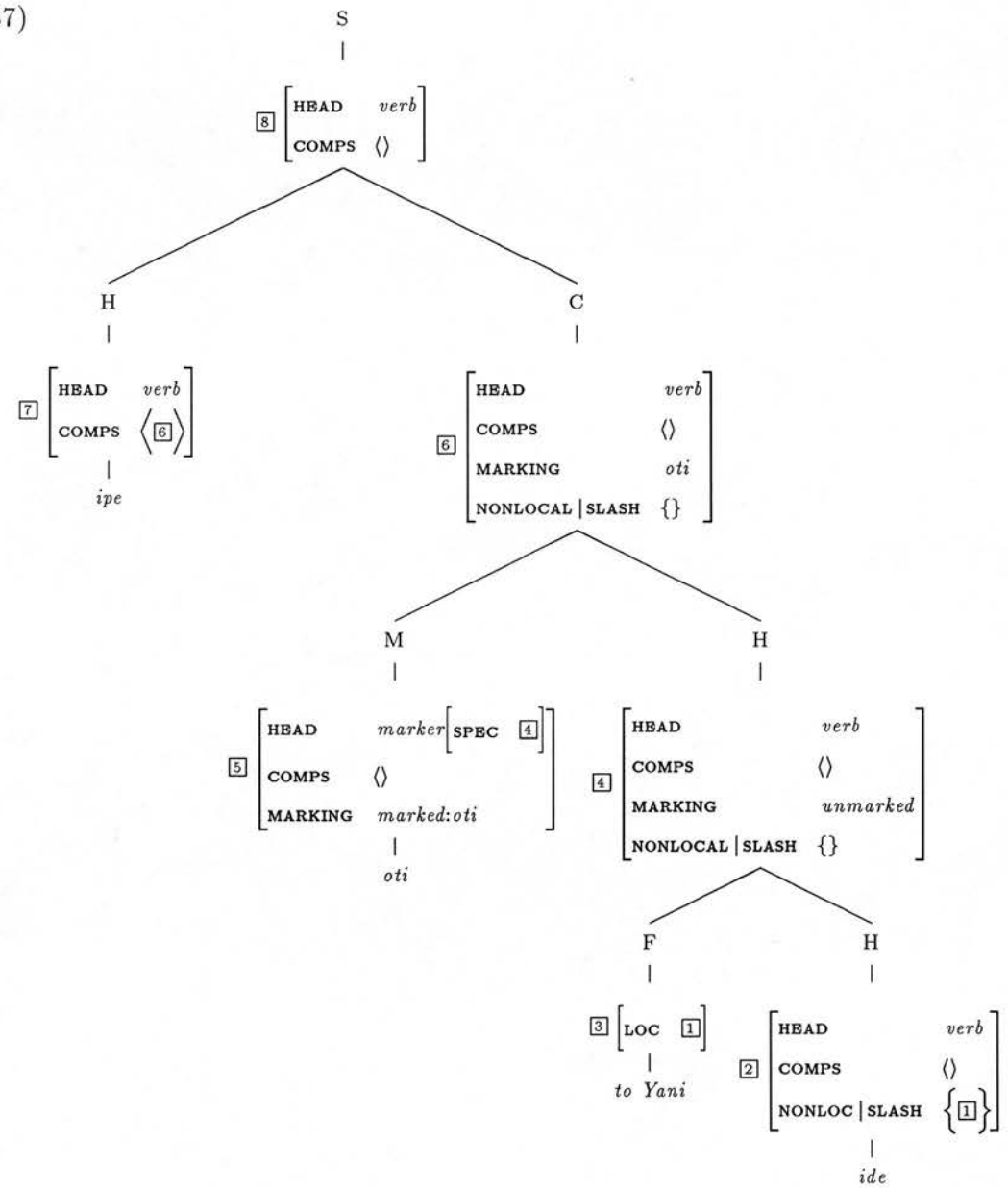
HEAD-DTR. The HEAD-FILLER-ID-SCHEMA (4.19) does not constrain the MARKING value of the HEAD-DTR. Thus, nothing prevents the *filler* from combining with a *marked* S (4). The *head-filler-phrase* inherits the MARKING value of the HEAD-DTR 4 (MARKING:*oti*). This is ensured by the MARKING PRINCIPLE which states that a MOTHER inherits the MARKING value of the HEAD-DTR in the absence of a MARKING-DTR (Pollard & Sag 1994). Thus, the *head-filler-phrase* 6 satisfies the subcategorisation requirements of the matrix verb *ipe*, which selects a saturated S, marked with *oti* (HEAD:verb, COMPS<>, MARKING: *oti*).

Consider now an example in which the extracted XP appears after the complementiser *oti*:

(4.36) ipe **oti** to YANI ide
said-3SG that the Yani-ACC saw-3SG
'S/he said that s/he saw Yanis.'

The structure of (4.36) is shown in (4.37):

(4.37)



Example (4.37) differs from (4.35) in the order between the *head-marker* and *head-filler-phrase*. The propagation of the SLASH and MARKING values follows from the same constraints.

This approach offers an elegant account of the order between *fillers* and *oti* in Greek. The head-driven nature of HPSG allows information associated with a head to be propagated to nodes higher than the node where the relevant head appears. Thus, the selection of the embedded clause by the matrix verb is not affected by intermediate *fillers*. On the other

hand, underspecified descriptions of linguistic structures allow unification at various points of the clause-structure. For example, a *filler* may combine with an S with any MARKING value, while a marker may combine with a *head-filler-phrase* (4.37) or a *head-comp-phrase* (4.35).

This analysis circumvents the problems of Discourse Configurational analyses. As discussed in Section 3.2.5, Tsimpli (1995) assumes that CP and FP are both specified for both the [f] and the [wh] feature. Tsimpli is forced to this stipulation in order to account for the fact that extracted XPs may appear on either side of the complementiser *oti*. In the analysis presented in this section *oti* or extracted XPs are not associated with a particular position in the tree specified for a specific feature. Rather, the ordering facts follow from independent assumptions about the complement of a filler and *oti* respectively.

Crosslinguistic considerations

The analysis presented in Section 4.4.3 can be extended to account for crosslinguistic variation. Unlike Greek, in English, extracted XPs can only appear after the complementiser *that*:

(4.38) a. *She said John that she has met twice.

b. She said that John she has met twice.

Examples (4.38) indicate that, in English, the *filler* combines with an *unmarked* S[6]. Thus, while in Greek the complement of a *filler* is underspecified for MARKING, the HEAD-FILLER-SCHEMA for English requires that the MARKING value of the HEAD-DTR is *unmarked*. Example (4.38a) is ruled out because it does not satisfy this restriction. On the other hand, example (4.38b) as well as examples of long-distance extraction (4.28) satisfy this restriction.[7]

4.5 Adjunct Extraction

4.5.1 Adjuncts as Complements

A thorough study of the syntax of adjuncts is beyond the aims of this thesis. However, some aspects of the syntax of adjuncts bear direct relevance to the grammar of UDCs discussed

[6] I am grateful to Robert Levine for this suggestion and discussion on these data.

[7] Italian patterns with English in that extracted XPs are not allowed immediately before 'that' (I owe this information to Ludovica Serratrice and Antonella Sorace). Unlike English though, Italian allows extracted XPs to appear immediately before an indirect question (Rizzi 1995). In this respect Italian is on a par with Greek (see Section 4.7.3).

here. In particular, the treatment of subjacency, p-gaps and adjunct extraction itself relies largely on the analysis of adjuncts. In this section, I will introduce the relevant problems and discuss recent accounts of adjuncts in HPSG.

In Pollard & Sag (1994) adjuncts select their heads by virtue of the feature MOD appearing in the lexical entry of potential adjuncts. For example, the adverb *usually* has the lexical entry in (4.39):

(4.39)
$$\left[\begin{array}{ll} \textsc{phon} & \textit{usually} \\ \textsc{synsem}\,|\,\textsc{loc}\,|\,\textsc{cat} & \left[\textsc{head}\ \textit{adverb}\ \left[\textsc{mod}\ \left[\begin{array}{ll}\textsc{head} & \textit{verb}\\ \textsc{comps} & \langle\rangle\end{array}\right]\right]\right] \end{array}\right]$$

The adverb *usually* selects a saturated VP with which it combines to form an *head-adjunct-phrase*. *Head-adjunct-phrases* are licensed by the HEAD-ADJUNCT-ID-SCHEMA which requires that the MOD value of the ADJ-DTR is token identical with the synsem value of the HEAD-DTR (Pollard & Sag 1994). Adjunct extraction is treated independently from complement extraction, by a separate lexical rule.

In a recent paper, Bouma *et al.* (1997) argue for an alternative treatment of adjuncts. They propose that adjuncts should be treated as complements and appear in COMPS. The AVM in (4.41) shows the structure of (4.40):

(4.40) It usually finds a solution.

(4.41)
$$\left[\begin{array}{ll} \textit{intrans \& 3sg} & \\ \textsc{i-form} & \textit{finds} \\ \textsc{head} & \boxed{5} \\ \textsc{subj} & \left\langle \boxed{1} \right\rangle \\ \textsc{comps} & \left\langle \boxed{2}\text{NP}, \boxed{3}\left[\begin{array}{ll}\textsc{head} & \textit{adv}\\ \textsc{mod} & \left[\begin{array}{ll}\textsc{head} & \boxed{5}\\ \textsc{cont} & \boxed{6}\end{array}\right]\end{array}\right] \right\rangle \\ \textsc{deps} & \left\langle \boxed{1}\text{NP}_{3sg}, \boxed{2}, \boxed{3} \right\rangle \\ \textsc{cont} & \boxed{6} \end{array}\right]$$

(Bouma *et al.* 1997:ex.42)

In example (4.41) the adverb appears in COMPS, as a sister of the verbal head (5) and its object (2). Adjunct-complements combine with their verbal heads like argument-complements do, (by virtue of the HEAD-COMP-ID-SCHEMA). In essence, this is a non-configurational approach to adjunction.

The introduction of adjuncts in the COMPS list has implications for the organisation of verbal signs. Bouma *et al.* (1997) propose that, in addition to ARG-STR, there is a DEPS-STR, obtained by appending ($\oplus$) the list of adjuncts to the list of arguments. The following describes the realisation of the DEPS-STR (adapted from Bouma *et al.* (1997)):

(4.42) DEPENDENTS REALISATION CONSTRAINT (*preliminary version*)

$$\textit{Word} \rightarrow \begin{bmatrix} \textsc{head} & \boxed{3} \\ \textsc{comps} & \boxed{1} \oplus \boxed{2}\textit{list}\left(\left[\textsc{mod} \quad \left[\textsc{head} \quad \boxed{3}\right]\right]\right) \\ \textsc{arg-str} & \boxed{1} \\ \textsc{deps-str} & \boxed{1} \oplus (\ \boxed{2}\ ) \end{bmatrix}$$

Note that adjuncts are optional, as indicated by the round brackets in (4.42).

DEPS-STR overlaps with ARG-STR. However, Bouma *et al.* (1997) opt to maintain both structures. This is partly motivated by the fact that the ARG-STR is the locus of binding theory in HPSG. Thus, replacing ARG-STR with DEPS-STR could possibly have implications for binding theory (in relation to this see Gill (in preparation) and Chung (1998). Since either assumption would not affect the analysis presented in this study, I will follow Bouma *et al.* (1997) and assume that both ARG-STR and DEPS-STR are present in the CAT value of *verbs*.

The non-configurational approach to adjunction is argued for in detail in Bouma *et al.* (1997). Part of their analysis is motivated by their attempt to provide a uniform analysis for argument and adjunct extraction. As adjuncts are complements, argument and adjunct-extraction are both examples of complements extraction (see Section 4.5.3).

4.5.2 Adjuncts in Greek

Adjuncts in Greek are on a par with argument-complements in the following two cases:

i) Wide Focus: as with arguments, nuclear accent on the rightmost adjunct gives rise to a wide focus reading:

(4.43) a. A: ihes nea apo ti Maria?
'Did you have any news from Maria?'

b. B: tin ide o Petros sto SINEMA
B: her-CL saw-3SG the Petros-NOM at-the cinema
'Petros met her at the cinema.'

ii) Extraction: as shown in Section 4.2, adjunct-extraction shares various properties with argument-extraction.

The extraction possibilities of adjuncts and the facts about wide focus suggest that adjuncts should be treated on a par with arguments, at least at some level of representation (e.g. DEPS-STR). This is necessary in order to avoid redundancy in the grammar of extractions and wide focus. The question of whether adjuncts in Greek are sisters of the verb or of a phrasal node (VP/S) will not be addressed here. For the purposes of the discussion, I assume a flat structure in which adjuncts are sisters of their heads (see Alexiadou (1997) for a similar view and detailed discussion on Greek adverbials). This assumption is supported by examples in which the adjunct appears between the verb and its argument:

(4.44) a. A: ti kanate me tis prosklisis
'What did you do with the invitations?'

b. B: tis stilame me to tahidromio sto YANI
B: them-CL sent-1PL with the post to-the Yani-ACC
'We sent them to Yanis by post.'

In (4.44b) nuclear accent falls on the object, *Yani*, and gives rise to a wide focus intepretation. The nuclear accent on the object shows that the adjunct is not dislocated to the right. A configurational approach to Greek adjuncts could not easily accommodate examples like (4.44b).

4.5.3 Adjunct vs. Argument Extraction in HPSG

In order to capture the similarities between adjunct and argument extraction, Bouma *et al.* (1997) modify SLAC (4.14) in the following way:

(4.45) SLASH AMALGAMATION CONSTRAINT(*preliminary version*)

$$\begin{bmatrix} \textsc{loc} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{deps-str} & \left\langle \begin{bmatrix} \textsc{slash} & \boxed{1} \end{bmatrix}, \ldots, \begin{bmatrix} \textsc{slash} & \boxed{n} \end{bmatrix} \right\rangle \end{bmatrix} \end{bmatrix} \\ \textsc{nonlocal} & \begin{bmatrix} \textsc{slash} & \boxed{1} \cup \ldots \cup \boxed{n} \end{bmatrix} \end{bmatrix}$$

The head inherits SLASH values from the DEPS-STR (rather than ARG-STR), which contains argument and adjunct-complements. In this way, the inheritance of adjunct and argument *gaps* is handled by a single constraint.

Like arguments, in cases of extraction, adjuncts are of synsem sort *gap*. The various possibilities of adjunt realisation are captured by the constraint on DEPENDENTS REALISATION (adapted from Bouma *et al.* (1997)):

(4.46) DEPENDENTS REALISATION CONSTRAINT (*final version*)

$$\textit{Word} \rightarrow \begin{bmatrix} \textsc{head} & \boxed{1} \\ \textsc{comps} & \boxed{4}:\left\langle \boxed{2} \oplus \boxed{3} \left(\begin{bmatrix} \textsc{mod} & \begin{bmatrix} \textsc{head} & \boxed{1} \end{bmatrix} \end{bmatrix} \right) \right\rangle (\textit{canon}) \\ \textsc{deps-str} & \boxed{4} \bigcirc \textit{list}(\textit{gap}) \end{bmatrix}$$

The COMPS-list [4] is obtained by appending the list of adjunct-complements ([3]) to the list of argument-complements ([2]). The members of COMPS are constrained to be of synsem sort *canonical*. The DEPS-STR is obtained by *shuffling* the members of COMPS [3] with the list of *gaps*.

The modified version of SLAC along with the DEPENDENTS REALISATION CONSTRAINT allow a uniform account for argument and adjunct extraction. SLIP and the HEAD-FILLER-ID-SCHEMA need no modification.

As noted earlier (Section 4.5.1), the analysis of adjuncts as complements is, partly, motivated by the attempt to capture the similarities between adjunct and argument extraction. However, a uniform treatment of argument and adjunct extraction is not incompatible with a configurational analysis of adjunction. Roughly, adjuncts could combine with a saturated VP and appear in an ADJ-STR. Then, the DEPS-STR, which is the locus from where heads inherit their gaps, could be obtained by appending the ADJ-STR to COMPS.

The proposal of Bouma *et al.* (1997) succeeds in capturing the similarities between adjunct and argument extraction and avoiding redundancy in the grammar. However, it cannot, as it stands, account for the differences between the two. The SLAC (4.46) allows extraction from adjunct clauses. A *gap* within an adjunct clause may be inherited by a matrix verb in the same way embedded *gaps* are inherited from argument-clauses (4.16&4.17). In other words, this analysis provides no account for subjacency effects.

In order to account for the sensitivity of UDCs to strong islands, the SLASH AMALGAMATION CONSTRAINT is modified in the following way:

(4.47) SLASH AMALGAMATION CONSTRAINT (*final version*)

$$\left[\begin{array}{ll} \textsc{head} & \boxed{1} \\ \textsc{deps-str} & \left\langle \left[\begin{array}{ll} \textsc{loc} & \boxed{2} \\ \textsc{slash} & \left\{\boxed{2} \cup \boxed{3} \cup \ldots \cup \boxed{m}\right\} \end{array}\right], \left[\begin{array}{ll} \textsc{loc} & \boxed{4}\left[\textsc{mod}\ \boxed{1}\right] \\ \textsc{slash} & \left\{\boxed{4} \cup \boxed{5} \cup \ldots \cup \boxed{l}\right\} \end{array}\right], \ldots, \left[\begin{array}{ll} \textsc{loc} & \boxed{n} \\ \textsc{slash} & \left\{\boxed{n}\right\} \end{array}\right] \right\rangle \\ \textsc{slash} & \left\{\left\{\boxed{2} \cup \boxed{3} \cup \ldots \cup \boxed{m}\right\} \cup \left\{\boxed{4}\right\} \cup \ldots \cup \boxed{n}\right\} \end{array}\right]$$

In (4.47) arguments are distinguished from adjuncts by the feature MOD in adjuncts (cf. 2 and 4). The first member of the DEPS-STR is an argument with local value 2. When this argument is of sort *gap* its SLASH value is 2. The SLASH values{...3...m} are (potentially) inherited from embedded arguments (see Section 4.3). The second member of DEPS-STR is an adjunct with SLASH values similar to 2. According to (4.47) the head inherits matrix (2) as well as embedded *gaps* ({...3...m}) from the argument (2). This is not so for adjuncts. Only 4 appears in the SLASH value of the head. Gaps contained within an clausal adjunct ({...5...l}) are not inherited by the matrix head 1.

P-gap constructions are accounted for straightforwardly. Consider the following:

(4.48) The paper they filed $-_j$ without reading $-_j$.

Example (4.48) has the following structure:

(4.49)
$$\begin{bmatrix} \textsc{head} & \boxed{1}(\textit{filed}) \\ \textsc{deps-st} & \left\langle \begin{bmatrix} \textsc{loc} & \boxed{2} \\ \textsc{slash} & \{\boxed{2}\} \end{bmatrix}, \begin{bmatrix} \textsc{loc} & \boxed{3}\begin{bmatrix}\textsc{mod} & \boxed{1}\end{bmatrix} \\ \textsc{slash} & \{\boxed{2}\} \end{bmatrix} \right\rangle \\ \textsc{slash} & \{\boxed{2}\} \end{bmatrix}$$

In (4.49) the object of the matrix verb *filed* is realised as a *gap* with SLASH value 2. The p-gap is token-identical with the matrix *gap*. Since the matrix *gap* is properly licensed, example (4.48) does not violate any constraint.

[However, this account of p-gaps may be too powerful, as it may allow example (4.50). In this example the 'p-gap' is coindexed with an in situ complement:

(4.50) *They filed the paper$_j$ without reading —$_j$.

Some restriction is needed to ensure that p-gaps are licensed when token-identical with matrix *gaps*. That is, structure-sharing with the LOC value of a 'in situ' complement should be excluded. A technical solution to this will not be pursued here. However, the constraint in (4.47) allows the possibility for a treatment of p-gap constructions.]

4.6 Clitic Left Dislocation as an Unbounded Dependency

In this section I argue that CLLD is an instance of adjunct extraction. As the analysis of CLLD is based on assumptions about the nature of the pronominal clitic and the status of the doubled NPs in Clitic Doubling, I begin the discussion with some background assumptions for the analysis of pronominal clitics and Clitic Doubling in Greek.

4.6.1 The affixal status of Greek argument clitics

Studies of Clitic Constructions in Greek treat clitics as *postlexical* elements that function as syntactic heads of distinct functional projections, Clitic Phrases (Agouraki 1993; Anagnostopoulou 1994; Schneider-Zioga 1994). In doing so, they follow standard analyses of Clitic Constructions within the Principles&Parameters framework and do not present independent evidence for the *postlexical* nature of Greek pronominal clitics. However, it has recently been

argued that certain kinds of pronominal clitics are best analysed as *affixes* rather than *postlexical* clitics (Miller 1992; Sag & Miller 1997; Monachesi 1995). The distinction between the two is a subtle one. Both cases involve phonologically deficient elements that are attached to a lexical host. Despite their close resemblance though, there are various properties that distinguish the two. Greek pronominal clitics seem to exhibit properties typical of *affixal* rather than *postlexical* clitics. In this respect, they are on a par with Romance pronominal clitics and, in particular, French and Italian ones. Some of the properties distinguishing them from *postlexical* clitics are presented below:

1) DEGREE OF SELECTION WITH RESPECT TO THE HOST:

While *postlexical* clitics tend to attach to various hosts, *affixes* appear on a single host. Greek and French pronominal clitics appear always on the verb:

(4.51) a. Il faut ne rien lui dire.
'It is necessary to tell her/him nothing.'

b. *Il faut ne lui rien dire.

(Sag & Miller 1997:ex.4)

(4.52) a. na mi tu to dosis
SUBJ.PART not him-CL.GEN it-CL.ACC give-2SG
'Don't give it to him.'

b. *na tu mi to dosis

The ungrammaticality of (4.51b) and (4.52b) shows that the pronominal clitics *lui* and *tu* can only be attached to the verb.

2) ARBITRARY GAPS IN THE SET OF COMBINATIONS

The distribution of French clitics presents arbitrary gaps in the combination possibilities with other clitics:

(4.53) a. Il le lui a présenté.
'He presented him to her.'

b. *Il me lui a présenté.
'He presented me to him.'

(Sag & Miller 1997:ex.5)

Though the direct object clitic can combine with an indirect object clitic in (4.53a), this combination is unavailable in (4.53b) where a 1st person pronoun is involved. Similarly, in Greek, though a genitive clitic may combine with an accusative one (4.54a), this sequence is ungrammatical in (4.54b):

(4.54) a. tis ton parusiase

her-CL.GEN him-CL.ACC presented-3SG
'S/he presented him to her.'

b. *tis me parusiase

her-CL.GEN me-CL.ACC presented-3SG
'S/he presented me to her.'

Arbitrary gaps of the kind displayed in (4.53&4.54) are typical of *affixal* rather than *postlexical* clitics. Of course, the fact that in both Greek and French the arbitrary gaps involve the same combinations of pronouns suggests that they might not be as arbitrary as they look at first sight.

3) RIGID AND IDIOSYNCRATIC ORDERING

Another characteristic property of *affixes* is the exhibition of rigid and indiosyncratic ordering. As shown in (4.55) genitive clitics precede accusative ones:

(4.55) a. tu to edosa

him-CL.GEN it-CL.ACC gave-1SG
'I gave it to him.'

b. *to tu edosa

4) WIDE SCOPE OVER COORDINATION

In both French and Greek, object clitics cannot have wide scope over coordinated verb-hosts:

(4.56) *Pierre les voit et écoute.

'Pierre sees and hears them.'

(Sag & Miller 1997:ex.10)

(4.57) *tous akouse ke ide

them-CL.ACC heard-3SG and saw-3SG
'S/he heard and saw them.'

Note that wide scope over coordination is possible in (4.58) where the object is a full NP:

(4.58) otan pire ke diabase ta vivlia...
when got-3SG and read-3SG the books
'When s/he got and read the books...'

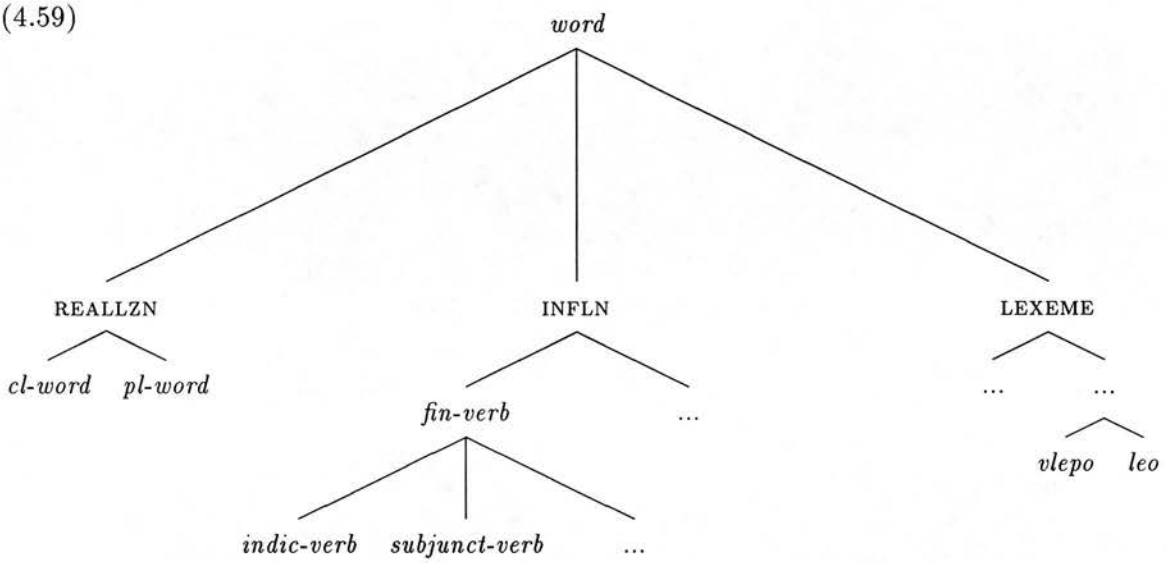
The impossibility of wide scope readings is considered strong evidence for the affixal status of pronominal clitics.

The evidence presented here supports an *affixal* treatment of Greek pronominal clitics. In essence, *affixal* clitics are argument pronouns incorporated in the verb. This is not an isolated phenomenon in Greek. As argued in Kolliakou (1997) and Kolliakou (1998a) NP internal clitics too have an *affixal* status. *Affixes* should be distinguished from inflectional agreement or tense markers. Unlike *affixes*, inflectional markers are not optional and, often, they are inseparable from their stems. [For a different view, see Tzanidaki (1994) who treats Greek clitics as agreement markers].

4.6.2 Affixal clitics in HPSG

Following Sag & Miller (1997), I assume that verbs with an attached *affixal* clitic realise a subtype of *word*, *clitic-word*. As mentioned in Section 1.3.2, each inflected *word* belongs simultaneously to two compatible types, *inflectional* type and *lexeme* type. In addition, it belongs to a third type, *(clitic) realisation* type (Sag & Miller 1997). *Clitic-words* are distinguished from *plain-words*, that is words with no *affix* attached to them. The hierarchy in (4.59) illustrates the *word* types:

(4.59)



Clitic-words differ from *plain-words* in their morphology, as well as their COMPS and ARG-STR. In a *clitic-word* the argument corresponding to the clitic does not appear in the COMPS. Thus, *clitic-words* have reduced VALENCY. Further, the relevant argument in the ARG-STR is of synsem sort *affix* (see the synsem hierarchy in Section 1.3.2, ex. 1.26). *Plain-words* and *clitic-words* are subject to the constraints in (4.60) and (4.62) respectively:

(4.60) PLAIN WORD REALISATION

$$\textit{Plain-word} \rightarrow \begin{bmatrix} \textbf{MORPH} & \begin{bmatrix} \textbf{FORM} & \boxed{0} \\ \textbf{I-FORM} & \boxed{0} \end{bmatrix} \\ \textbf{SYN} \mid \textbf{LOC} \mid \textbf{CAT} & \begin{bmatrix} \textbf{COMPS} & \boxed{1}\textit{list}(\textit{canon}) \\ \textbf{ARG-STR} & \boxed{1} \bigcirc \textit{elist}(\textit{affix}) \end{bmatrix} \end{bmatrix}$$

According to (4.60) the mophology of a *plain-word* (FORM:[0]) is its *inflectional* form (I-FORM:[0])[8].The members of COMPS are of synsem sort *canonical.* The ARG-STR may contain no *affixes*, as the list of *affixes* is the *e(mpty) list.* Consider now the constraint on the realisation of a *clitic-word*:

(4.62) CLITIC WORD REALISATION

$$\textit{Clitic-word} \rightarrow \begin{bmatrix} \text{MORPH} & \begin{bmatrix} \text{FORM} & \boxed{3}\textit{clitic-form} \\ \text{I-FORM} & \boxed{0} \end{bmatrix} \\ \text{SYN} \mid \text{LOC} \mid \text{CAT} & \begin{bmatrix} \text{COMPS} & \boxed{1}\textit{list}(\textit{canon}) \\ \text{ARG-STR} & \boxed{1} \bigcirc \textit{nelist}(\textit{affix}) \end{bmatrix} \end{bmatrix}$$

There are two differences between *clitic-words* and *plain-words.* First, unlike *plain-words*, the MORPH|FORM value of a *clitic-word* is different from its MORPH|I-FORM value. Second, the ARG-STR should contain at least one *affix*, as the list of *affixes* is the *non empty list.*

The DEPS-STR of both *clitic* and *plain-words* may contain any number of *gaps. Clitic words* inherit this possibility from their supertype, *word*, which is subject to the DEPS-REALISATION CONSTRAINT (4.46). The DEPENDENTS REALISATION CONSTRAINT allows dependents, arguments or adjuncts, to be realised as *gaps.* (The restriction that the members of COMPS are of sort *canonical* is also inherited by the DEPENDENTS REALISATION CONSTRAINT. Thus, this restriction is redundant in (4.60&4.62)).

The constraints involving the realisation of *plain* and *clitic-words* could equally make reference to the DEPS-STR rather than the ARG-STR, since the arguments are part of the DEPS-STR. In Greek, either assumption would have the same result, since clitics are available only for arguments (the direct and indirect object). However, in languages like Italian, which has clitics for adverbial PPs as well as arguments, the constraints in (4.60&4.62) should make

[8] FORM is an attribute of MORPH additional to ROOT and I-FORM. The full structure of MORPH is shown in (4.61):

(4.61)

$$\begin{bmatrix} \text{MORPH} & \begin{bmatrix} \text{ROOT} \\ \text{FORM} \\ \text{I-FORM} \end{bmatrix} \end{bmatrix}$$

reference to the DEPS-STR.

Clitic-words are related to *plain-words* through a derivational function, F_{praf}. F_{praf} is a morphological operation yielding the *clitic-form* of a verb. Roughly, F_{praf} takes the *inflectional* form of a verb with an *affixal* member in its ARG-STR as input and returns its *clitic-form* as output—for a detailed description of F_{praf} see Sag & Miller (1997).

The following illustrates the *clitic-word*, *tin-ide* ('s/he saw her'); this word belongs to *lexeme* type *vlep(o)* with *inflectional* type *3sg.indic.past*:

(4.63) *cl-word & vlep(o) & 3sg.indic.past*

$$\left[\begin{array}{ll} \textsc{morph} & \left[\begin{array}{ll} \textsc{form} & \boxed{3}\textit{tin-ide} \\ \textsc{i-form} & \textit{ide} \end{array}\right] \\ \textsc{syn}\,|\,\textsc{loc}\,|\,\textsc{cat} & \left[\begin{array}{ll} \textsc{comps} & \left\langle \boxed{1} \right\rangle \\ \textsc{arg-str} & \left\langle \boxed{1}\mathrm{NP}_{nom,3sg},\ \boxed{2}\mathrm{NP}_{aff,acc,3sg,fem} \right\rangle \end{array}\right] \end{array}\right]$$

4.6.3 Sanfilippo (1990): Quasi-Arguments and Quasi-Adjuncts

Sanfilippo (1990) provides an account of Italian CLLD in Unification Categorial Grammar. In his account, every predicate introduces a subcategorisation frame (ARGUMENTS) and entails some participant/thematic roles (ϑ-DOM). The traditional distinction of non-head constituents between arguments and adjuncts is organised along the two parameters of ARGUMENT REMOVAL, that is, satisfaction of subcategorisation requirements, and ϑ-DOM REDUCTION, that is, instantiation of participant/thematic roles. Arguments both remove elements from the subcategorisation frame (ARGUMENT REMOVAL) and reduce the ϑ-DOM. Adjuncts do none of the two.

In order to account for CLLD, Sanfilippo extends the distinction between arguments and adjuncts to include *quasi-arguments* and *quasi-adjuncts*. In cases where no clitics are involved, both the syntactic and thematic requirements of the verb are satisfied by a single (NP) phrase/constituent, the *argument*. In CLLD, these requirements are satisfied by two constituents, the clitic and the dislocated NP. The clitic is a quasi-argument; it satisfies the subcategorisation requirements of the verb (ARGUMENT REMOVAL) but does not reduce its ϑ-DOM. On the other hand, the dislocated NP is a quasi-adjunct. It reduces the verb's ϑ-DOM

by instantiating the thematic role corresponding to the argument removed by the clitic. The NP itself though, does not remove an argument from the subcategorisation frame.

Though couched in different frameworks, Sanfilippo's and Cinque's accounts (Section 3.3.3) share the intuition that no extraction/movement takes place in CLLD; the dislocated element is base-generated and combines with/enters in a binding chain with the following clause. Sanfilippo rejects the extraction analysis by noting that such an analysis implies that the object clitic functions as an agreement marker that does not reduce the valency of a verb. The optionality of the doubled phrase, though, does not support such a claim.

Kolliakou (1991) offers a GPSG account of Clitic Constructions in Greek based on Sanfilippo's analysis of CLLD in Italian. Doubled NPs are treated as quasi-adjuncts in all three constructions involving clitics, CLLD, Clitic Doubling and Clitic Right Dislocation.

4.6.4 Clitic Doubling in HPSG

The analysis of Clitic Doubling in Greek builds on the intuition of Sanfilippo (1990) and Kolliakou (1991) that the doubled NP is a syntactic adjunct. Semantically, the adjunct NP provides the semantic RESTRICTION for the referent of the coindexed pronominal clitic.

The adjunct status of doubled NPs is supported by their optionality:

(4.64) ton ide (to Yani)
him-CL saw-3SG (the Yani-ACC)
'S/he saw Yanis.'

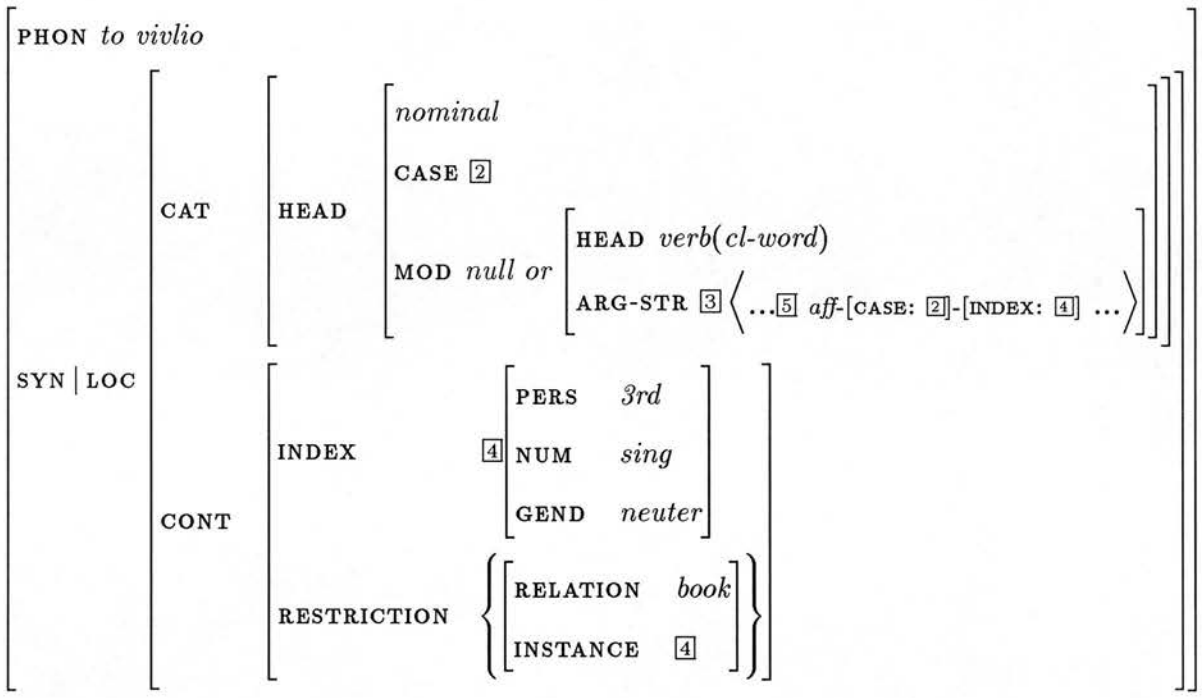
In addition, like many ordinary adjuncts (4.44b), doubled objects may appear between a verb and an argument:

(4.65) ta klidia tis ta dose tis Marias o PETROS
the keyes her-CL them-CL gave-3SG the Maria-GEN the Petros-NOM
'Petros gave the keys to Maria.'

In (4.64), the indirect doubled object, *tis Marias*, appears between the verb and the subject, *Petros*. The subject bears nuclear accent which shows that it is not dislocated to the right.

The HPSG treatment of object adjuncts is on a par with the treatment of ordinary adjuncts [see Sanfilippo (1996) for a slightly different analysis in HPSG]. Thus, object adjuncts select their head through the feature MOD in their lexical entry. The following AVM illustrates the lexical entry of the NP *to vivlio* (*'the book'*):

(4.66)



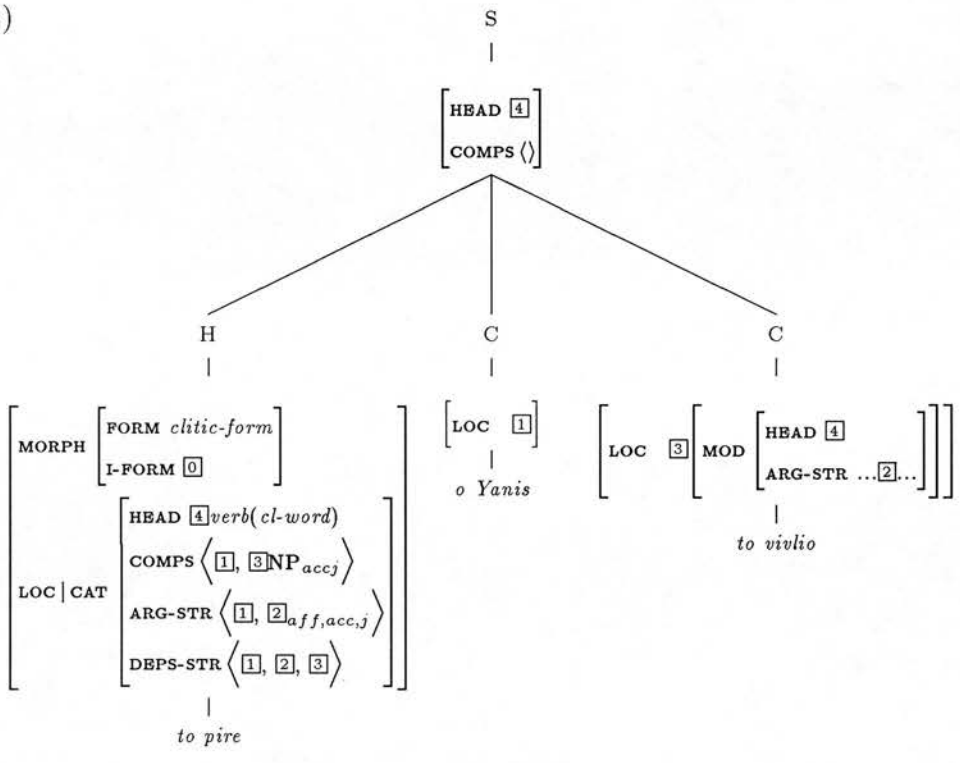
The adjunct NP selects a *verb* which is constrained to be a *clitic-word*. As already mentioned, the ARG-STR of a *clitic-word* contains an *affixal* argument. The object NP structure-shares its CASE and INDEX value (2&4) with the CASE&INDEX value of the *affix* in the ARG-STR. In this way, case, person, number and gender agreement between the doubled object and the clitic is guaranteed. (Semantically, coindexing ensures that the clitic and the NP are anchored to the same referent-see later in this section).

Example (4.67) shows a case of Clitic Doubling:

(4.67) to pire o Yanis to vivlio
it-CL took the Yanis the book
'Yanis took the book.'

Tree (4.68) illustrates the structure of (4.67):

(4.68)



In (4.68) the head is a verb of sort *clitic-word*. It takes three dependents; two arguments (1 & 2) and one adjunct (3). The subject (1) and the adjunct (3) are of synsem sort *canonical* and are realised locally as complements (COMPS: < 1,3 >). The two complements combine with their head by virtue of the HEAD-COMPS-ID-SCHEMA, giving rise to a saturated S. The second argument (2), the object, is realised as an *affix* and does not appear in COMPS. The adjunct 3 shares its CASE value (*acc*) and INDEX value (*j*) with the CASE and INDEX value of the *affixal* argument in ARG-STR (2).

The treatment of doubled objects as adjuncts presents an unusual case of syntactic selection. Typically, syntactic selection involves CATEGORY features like HEAD or MARKING. The selection of VALENCY values (COMPS, SUBJ) usually refers to whether the selected phrase is a saturated one or not (*empty* or *non-empty* COMPS/SUBJ respectively—e.g. a subject combining with a VP). In the case of doubled objects, the adjunct does not just select a head; it also selects a specific member from its ARG-STR, constrained with a particular description (*affix*). The phenomenon could be understood if more cases involving selection of ARG-STR members could be identified. However, this issue will not be pursued further here.

[Note that, if a configurational approach were adopted for adjuncts in Greek some technical amendements would be necessary. As it stands, ARG-STR is not a HEAD feature and, therefore, is not subject to the HEAD FEATURE PRINCIPLE. Thus, information concerning the members of ARG-STR is not inherited by higher nodes. If the adjunct combined with a VP/S rather than a *word*, the propagation of ARG-STR to the VP/S should be ensured. In relation to this, various participants in an informal discussion in the HPSG mailing list, have been suggesting that ARG-STR should be treated as a HEAD feature.]

Semantically, the coindexing between the pronominal clitic and the adjunct NP ensures that they are both anchored to the same referent and, in effect, realise the same semantic role in the CONTENT of the verbal head. For example, in (4.67), the clitic and the doubled NP correspond to the TAKEN (4.69) in the CONTENT value of the verb:

(4.69)
$$\begin{bmatrix} \textsc{phon} & \textit{to pire to vivlio} \\ \textsc{syn}\,|\,\textsc{loc} & \begin{bmatrix} \textsc{cat} & \begin{bmatrix} \textsc{head} & \boxed{4} \\ \textsc{deps-str} & \langle \boxed{1}_i, \boxed{2}_j, \boxed{3}_j \rangle \end{bmatrix} \\ \textsc{cont} & \begin{bmatrix} \textsc{relation} & \textit{taking} \\ \textsc{taker} & i \\ \textsc{taken} & j \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

The information that the TAKEN is a *book* is contributed by the RESTRICTION of the NP (4.66). As mentioned in Section 1.3.2, the RESTRICTION (in a NP), poses semantic conditions on the *index* of the *nominal* object. Roughly, the RESTRICTION in (4.66) states that the *index* $\boxed{4}$ should be used referentially and be anchored to a *book*. By contrast, *pronominal* objects have no restriction (4.70), as they are not used referentially. The CONTENT value of the clitic *to* in (4.67), is shown in (4.70):

(4.70)
$$\begin{bmatrix} \textsc{content} & \begin{bmatrix} \textsc{index} & \begin{bmatrix} \textsc{pers} & \textit{3rd} \\ \textsc{num} & \textit{sing} \\ \textsc{gend} & \textit{neuter} \end{bmatrix} \\ \textsc{restriction} & \{\} \end{bmatrix} \end{bmatrix}$$

Thus, the semantic contribution of the adjunct NP amounts to restricting the referent of the coindexed pronominal clitic. This view of the semantic contribution of the adjunct NP

captures Sanfilippo's intuition that the doubled NPs instantiate the participant/thematic roles of the verb.

So far, I have omitted various aspects of the internal structure of Greek NPs. In particular, as argued in Kolliakou (1995), the CONTENT value of Greek NPs has an extra feature UNIQUE, with +/— as values. UNIQUE provides information on whether the NP is anchored to a unique referent. Roughly, definite NPs have a UNIQUE: + value, whereas indefinite ones have a UNIQUE: — value. As the *affixal* clitic may combine with definite as well as indefinite NPs (Section 2.5.3), its UNIQUE value should be unspecified. The value of UNIQUE is contributed by the doubled NP along with the restriction to the index it shares with the pronominal *affix*.

4.6.5 CLLD as adjunct extraction

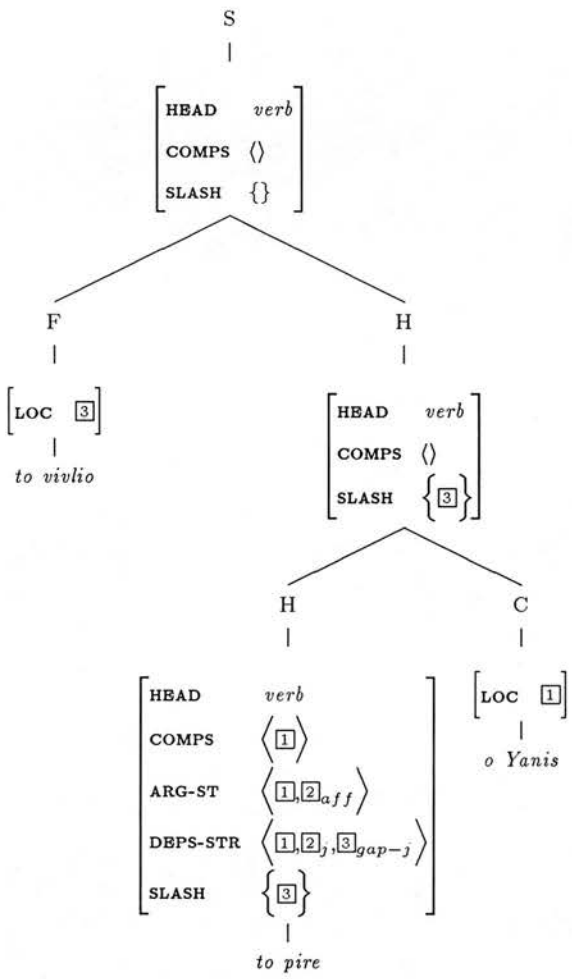
The properties of CLLD (see Sections 3.1.1&3.1.2) can be straightfowardly captured by the assumption that CLLD instantiates adjunct extraction. In particular, the availability of long-distance CLLD, its distribution (on either side of *oti* and before indirect questions), its sensitivity to strong islands, its recursive occurrence, are all general properties of UDCs in Greek. On the other hand, the unavailability of p-gaps (Section 3.1.2, example 3.11), follows from the adjunct status of the extracted doubled NP. As shown in Section 4.2, only argument extraction licenses p-gaps.

Example (4.71) illustrates a case of CLLD:

(4.71) to vivlio to pire o Yanis
the book it-CL took the Yanis
'Yanis took the book.'

The structure of (4.71) is shown in (4.72):

(4.72)



The adjunct NP[3] is coindexed with the *affixal* argument [2]. In the **ARG-STR** the adjunct is realised as a *gap*. The propagation of the *gap* and the licensing of the *head-filler-phrase* follow from general constraints on UDCs, (**SLAC**, **SLIP** and **HEAD-FILLER-ID-SCHEMA**).

Under this view, *Cinque's paradox* finds an elegant solution. Recall that Cinque (1990) and Iatridou (1995) observe that CLLD is sensitive to strong islands, an indication of movement (Sections 3.3.3&3.3.4). However, they reject the movement analysis of CLLD on the basis of the unavailability of p-gaps. Lack of p-gaps though, cannot constitute an argument that no extraction is involved, since there is a case of extraction, namely adjunct extraction, that does not license p-gaps (4.2). By contrast, the analysis proposed here captures the extraction properties of CLLD and offers an explanation for the *paradoxical* unavailability of p-gaps.

Further, this analysis captures the similarities between Topicalisation and CLLD observed

by Rizzi (1995) (Section 3.3.1). Both Topicalisation and CLLD are employed for the realisation of topics/links or, in the terms of Rizzi, both instantiate a topic-comment articulation. Rizzi accounts for this similarity by assuming that both structures involve an anaphoric operator. In Topicalisation the anaphoric operator is null whereas in CLLD it is realised by the clitic. However, as I have argued in Section 3.3.2, Rizzi's analysis is rather stipulative. Under the analysis proposed here, the fact that both Topicalisation and CLLD encode a topic-comment articulation follows from the fact that they have the same syntax (UDC) and the same phonological realisation. With respect to their syntax, CLLD is an instance of 'adjunct Topicalisation'. The CLLD in (4.73a) has the same structure with example (4.73b) which involves a topicalised adverbial PP:

(4.73) a. tin parastasi *ti* skinothetise o Dimitris **POTAMITIS**
the performance-ACC it-CL directed-3SG the Dimitris Potamitis-NOM
'Dimitris Potamitis directed the performance.'

b. me kokini mpoya tha vapsi ta **PARATHIRA** (kai me prasini tis portes)
with red paint will paint-3SG the windows (and with green the doors)
'S/he'll paint the windows with red paint (and the doors with green).'

While examples (4.73) instantiate adjunct extraction, example (4.74) instantiates argument extraction:

(4.74) tin parastasi skinothetise o Dimitris **POTAMITIS**
the performance-ACC directed-3SG the Dimitris Potamitis-NOM
'Dimitris Potamitis directed the performance.'

However, all examples above (4.73&4.74) share two properties. They all involve extraction and they all have the same phonological realisation, as the extracted XPs are not accented. Thus, in all these examples the extracted XPs are topics (irrespective of whether they are arguments or adjuncts, doubled or not)[9].

[9] Under this analysis, the cases of Empty Clitic Left Dislocation discussed in Dimitriadis (1994b) are simply instances of argument extraction.

4.7 Wh-questions as UDCs

4.7.1 Extraction Patterns

In terms of extraction possibilities, Wh-questions fall in the class of UDCs. In particular, they show the following properties, characteristic of UDCs in Greek:

i) They allow long distance extraction:

(4.75) PION ipe oti apelisan?
who-ACC said-3SG that fired-3PL
'Who did s/he say that they fired?'

ii) They obey strong islands:

(4.76) *PION efigan noris gia na prolavun?
who-ACC left-3PL early for to catch-3PL
'*Who did they leave early to catch?'

iii) In echo Wh-questions the wh-phrase may occur to either side of the complementiser *oti*:

(4.77) ipe (PION) **oti** (PION) ide?
said-3SG who-ACC that who-ACC saw-3SG
'S/he said that s/he saw who?'

iv) In echo questions the wh-phrase may appear to the left of an indirect question:

(4.78) rotise PION an idan?
asked-3SG who-ACC whether saw-3PL
'S/he asked whether they saw who?'

v) They do not create islands for extraction:

(4.79) to YANI rotise pou tha stilun
the Yani-ACC asked-3SG where FUT send-3PL
'S/he asked where they will send Yani.'

It is difficult to construct matrix questions in which the wh-phrase undergoes long distance movement from an indirect clause:

(4.80) ?PIOS rotises an tha fiyi?

who-NOM asked-2SG whether FUT leave-3SG
'(about) Who did you ask whether he'll leave?'

vi) Adjunct wh-phrases may undergo long distance extraction:

(4.81) POTE nomizis oti tha ta ehis etima?

when think-2SG that FUT them-CL have ready
'When do you think you'll have them ready?'

Finally, Focus-movement and Wh-movement pattern in that the focused phrase or the wh-phrase may be contained in a preposed PP/NP:

(4.82) a. me KOKINI mpoya ipe oti tha vapsi ta parathira

with red paint said-3SG that will paint-3SG the windows
'S/he said that s/he will paint the windows with red paint.'

b. me TI mpoya ipe oti tha vapsi ta parathira?

with what paint said-3SG that will paint-3SG the windows
'With what paint did s/he say that s/he will paint the windows?'

The above data show that extractions in Wh-questions (canonical and echo) are subject to the same constraints all extractions are in Greek. In this respect, the analysis of UDCs sketched in previous sections is straightforwardly extended to Wh-questions. Technically, Wh-questions differ from ordinary extractions in that, in Wh-questions, the *wh-gap* appears in NONLOCAL|QUE, rather than NONLOCAL|SLASH. (Here, I will remain vague as to how exactly a *wh-gap* ends up in QUE and not in SLASH. In relation to this, see Ginzburg (1992); Kathol (1995); Kathol & Pollard (1995); Pollard & Yoo (1996)).

4.7.2 Doubled wh-phrases

In Greek, wh-phrases may often be doubled, as shown in examples (3.96, 3.97 & 3.36b) repeated below:

(4.83) pia pedhia (ta) maloses

which children them scolded
'Which children did you scold?'

(Iatridou 1995:ex.48)

(4.84) [Pion apo tus dhio tus]$_i$ ipes oti dhen *ton*$_i$ simpathi i mitera
whom from the two of-them said-2,sg that not him(CLITIC) like-3,sg the mother-nom
su
of-you?
'Which of them did you say that your mother does not like?'

(Kolliakou 1991:ex.2-5a)

(4.85) pion$_i$ ton$_i$ agapai i mana tu$_i$?
who-ACC him-CL love-3SG the-NOM mother-NOM his-GEN
'Who does his (own) mother love?'

Wh-questions with doubled wh-phrases (henceforth doubled Wh-questions) are analysed as an instance of adjunct extraction. That is, the wh-phrases in examples (4.83,4.84 & 4.85) are extracted adjuncts, on a par with examples of CLLD (4.71).

As discussed in Section 3.4, various analyses of these data have been proposed in the literature. Dobrovie-Sorin (1990) and Iatridou (1995) share the view that CLLD and A-bar/Wh-movement are two distinct syntactic operations. Thus, their proposals try to explain the appearence of properties attributed to two distinct operations (CLLD, A-bar-movement) in one case. Dobrovie-Sorin (1990) proposes that doubled wh-phrases do not undergo A-bar movement and do not take scope over the clause. Iatridou (1995) suggests that doubled wh-phrases are base-generated in the Discourse-linked position and, from there, they move to [Spec,CP]. Under the analysis offered here, CLLD and doubled Wh-questions exemplify the same syntactic structure, adjunct extraction. Their similarities are captured straightforwardly without any extra assumptions for the structure of doulbed Wh-questions. In addition, there is no need to assume that the structure or the quantificational interpretation of doubled Wh-questions is different from that of non-doubled ones. Unlike the suggestions of Iatridou (1995) and Dobrovie-Sorin (1990), both cases have the same Phrase Structure. In both cases the wh-phrase appears at the same 'position' and has the same 'quantificational' interpretation.

4.7.3 Selection of indirect questions

As indirect Wh-questions are selected by verbal heads, their syntactic treatment should allow such selection to take place. I will tentatively assume that the MARKING value of a head with

a *nonempty* QUE value is *marked: wh*. The following describes this restriction:

(4.86)
$$Verb \rightarrow \begin{bmatrix} \textsc{head} & \boxed{1} \\ \textsc{marking} & \textit{marked: wh} \\ \textsc{que} & \left\{\boxed{2}\right\} \end{bmatrix}$$

A verb like *rotao* ('ask') subcategorises for an S marked with the *wh* feature (HEAD:*verb*, COMPS:<>, MARKING:*marked:wh*).

The MARKING value of Wh-questions seems to interfere with some aspects of the analysis of UDCs. Consider the following examples:

(4.87) a. *ipe pion oti ide o Petros
said-3SG who-ACC that saw-3SG the Petros-NOM
'*S/he said that who Petros saw.'

b. ipe pion oti ide o Petros?
'S/he said who that Petros saw? (echo)'

The two examples in (4.87) differ in that the first is a statement and the second an *echo* Wh-question. Example (4.87a) could be ruled out on the assumption that *wh-fillers* combine with an *unmarked* S. However, the grammaticality of (4.87b) indicates that the ungrammaticality of (4.87a) is independent of the combinatorial potential between *wh-fillers* and the MARKING value of the complement S. Recall that, the HEAD-FILLER-SCHEMA (4.19) allows a *filler* to combine with either a *marked* or *unmarked* S. Thus, example (4.87b) is predicted and is on a par with examples of pre-*oti* foci and topics (4.34). Example (4.88) in which the *wh-filler* combines with an *unmarked* S is also predicted:

(4.88) ipe oti pion ide o Petros?
'S/he said who that Petros saw? (echo)'

However, example (4.88) raises a problem for the restriction that the *marker oti* should combine with an *unmarked* S (HEAD-MARKER-ID-SCHEMA 4.32). In (4.88), the complement of *oti* has MARKING value *marked: wh*. A solution to this will not be pursued here.

Under the approach adopted here, there is no asymmetry between the clause structure of matrix and embedded Wh-questions. Recall that Tsimpli (1995) proposes that matrix

Wh-questions lack a CP projection, which embedded ones have. This stipulation is meant to account for the unavailability of Focus-movement in matrix Wh-questions, which is attributed to lack of syntactic positions. However, as discussed in Section 2.8.2, the contrast between matrix Wh-questions and indirect ones with respect to the availability of focus is independent of their syntax. Since questions do not update the hearer's information state, they are not expected to have a focus-segment. On the other hand, sentences with indirect questions are statements and are expected to have a focus segment. The focus part of the sentence may be the indirect clause itself (4.89a) or an embedded constituent (4.89b-c):

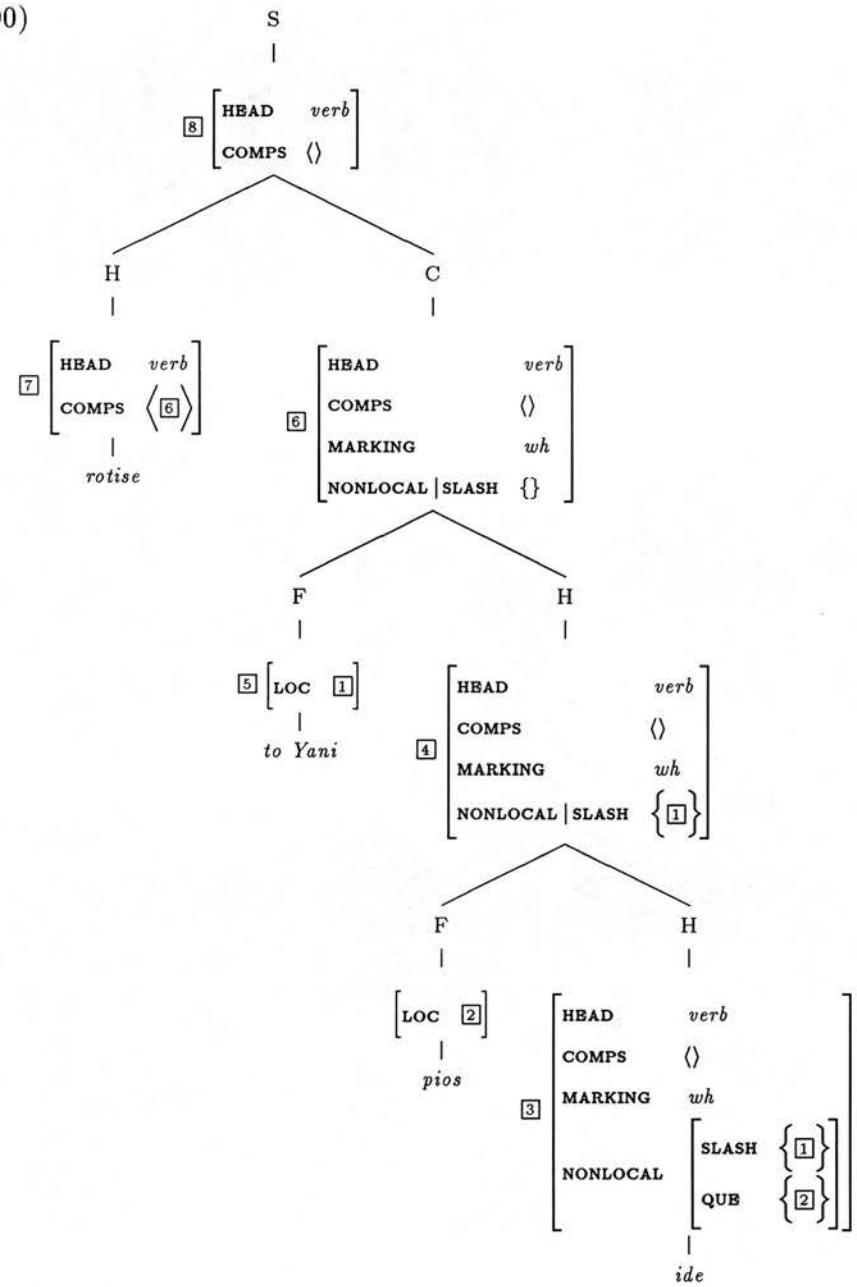
(4.89) a. rotise [$_F$ pios ide to YANI]
asked-3SG [$_F$ who-NOM saw-3SG the Yani-ACC]
'S/he asked who saw Yanis.'

b. rotise pios ide [$_F$ to YANI]
asked-3SG who-NOM saw-3SG [$_F$ the Yani-ACC]

c. rotise [$_F$ to YANI] pios ide
asked-3SG [$_F$ the Yani-ACC] who-NOM saw-3SG

Example (4.89c) involves recursive extraction (i.e. it involves two *head-filler-phrases*). The tree in (4.90) shows the structure of (4.89c):

(4.90)



The MARKING value of the indirect clause (MARKING:*wh*) is inherited by the *head-filler-phrase* in a similar way as in (4.35).

4.8 Conclusions

In this chapter I argued for a unified syntactic treatment of Focus-movement, Topicalisation and CLLD. This analysis is supported by the fact that the syntax of these constructions is not affected by the discourse import of the extracted XP. Rather, all three constructions exhibit typical properties of Unbounded Dependencies. Thus, I offered an analysis of their syntax employing the mechanism of UDCs in HPSG, which allows an elegant account of various properties of Greek extractions:

- The distribution of extracted XPs before and after *oti* was captured by the simple assumption that the complement of an extracted XP (i.e. the HEAD-DTR of a *head-filler-phrase*) has an underspecified MARKING value. Thus, a filler can combine with either an unmarked clause (in which case it appears after *oti*) or with a marked clause (in which case it appears before *oti*). On the other hand, the constraint that the complement of fillers in English is unmarked (MARKING:*unmarked*) was enough to block the ungrammatical ocurrence of a filler immediately before a *that-clause*.

- No additional assumptions were necessary to account for the fact that intermediate fillers do not block the selection of an embedded clause by the matrix verb. Already existing principles handling the propagation of MARKING values on the syntactic tree (MARKING PRINCIPLE/HEAD-MARKER-ID-SCHEMA) were enough to account for this fact.

- Since nothing in the existing analysis blocks recursive extraction, the fact that Greek extractions do not create islands for extraction is accounted for straightforwardly.

Further, I drew a distinction between argument and adjunct extraction and showed that, unlike argument extraction, adjunct extraction does not license p-gaps. This distinction is crucial for the analysis of CLLD. CLLD was analysed as an instance of adjunct extraction. This view offers an answer to *Cinque's paradox*. CLLD obeys strong islands, as all cases of extraction do, but does not license p-gaps, as all cases of adjunct extraction do not. Further, this analysis explains the interpretational similarities between Topicalisation and CLLD. Since both cases involve the same syntax and the same phonology they have the same Information Structure (see Chapter 5).

The view that CLLD instantiates extraction, allows a uniform analysis for CLLD and Clitic Doubling. In both constructions the doubled NP is an adjunct. In Clitic Doubling it

is realised as a complement whereas in CLLD as a filler. This analysis differs from the one proposed by Cinque (1990) and Iatridou (1995) who claim that the two phenomena are not related. As argued in Section 3.3.5, this claim lacks syntactic motivation and complicates the grammar of Clitic Constructions in Greek.

The uniform syntactic analysis of Focus-movement, Topicalisation and CLLD I have proposed in this chapter has implications for the Discourse-Syntax interface. Under this analysis, one construction corresponds to more than one discourse function, yielding a non-isomorphic relation between sytax and Information Structure. I will turn to this issue in the following chapter.

Chapter 5

Discourse-Syntax Interface

5.1 Introduction

The architecture of the discourse-syntax interface for a language like Greek depends on prior answers to the following two questions:

1 What is the syntax of the relevant constructions (Focus-movement, Topicalisation, CLLD)?

2 At which level of representation should discourse functions be encoded?

With respect to the first question, Discourse Configurational approaches assume distinct syntactic representations for the relevant constructions: these involve A-bar-movement, A-movement and Base-generation, respectively. As for the second question, their answer is to encode discourse functions in distinct Phrase Structure configurations: Focus Phrase, Topic Phrase. Movement to [Spec,FP] is A-bar-movement. Movement to [Spec,TP] is A-movement. CLLD involves base-generation at [Spec,TP]. Thus, the picture of the discourse-syntax interface offered by Discourse Configurational approaches involves a rather complex syntax, but a simple, isomorphic relation between discourse functions (focus, topic), Phrase Structure Configurations (FP, TP) and syntactic operations (modulo the difference between A-movement and base-generation).

The answers given in this thesis to questions 1 & 2 differ significantly from the ones proposed by Discourse Configurational accounts. With respect to the first question, it was argued in the previous chapter that Focus-movement, Topicalisation and CLLD instantiate

the same syntactic structure. This implies a simple syntax, but a non-isomorphic relation between discourse functions and syntactic operations.

In principle, a unified syntactic treatment of Focus-movement, Topicalisation and CLLD is still compatible with an analysis that maintains the view that discourse functions are encoded in Phrase Structure. Roughly, the relevant constructions would involve movement to a single preverbal position, but this would be associated with more than one discourse feature. Alternatively, they would be licensed by a discourse ID-SCHEMA. However, I have argued in Chapter 2 that discourse functions should be represented at a distinct level, Information Structure, independently of Phrase Structure configurations. This view is supported by the insensitivity of focus to syntactic constraints (subjacency) and the absence of recursive foci. Further, it allows a consistent account of languages that rely on intonation/morphology rather than syntax for the realisation of Information Packaging.

In this chapter, I will present a picture of the discourse-syntax interface for Greek that assumes a unified syntactic analysis of the relevant constructions and encodes Information Structure independently from Phrase Structure configurations. While all the constructions in question have the same syntax, discourse functions are disambiguated by nuclear accent placement. Thus, the account presented in the following sections involves mapping of both syntactic and phonological realisations to Information Structure. As will be shown, the Information Structure of Unbounded Dependencies follows from general constraints on the realisation of Information Packaging in Greek.

Further, the syntactic analysis I have adopted has implications for the account of the relative order of topics and focus. As discussed in Chapter 3, in Discourse Configurational analyses, the relative order between topics and focus is captured by the relative order of the Topic and Focus Phrase in the syntactic tree. In the syntactic analysis I proposed, topics and foci are both fillers and their relative order is not constrained. In this chapter I will argue that discourse constraints on word order should be represented independently of syntactic ones. The main argument in support of this view is the fact that the constraints on the order between links/topics, focus and tails hold irrespective of their syntactic realisation.

The chapter consists of two parts. In the first part, Section 5.2 I present an HPSG analysis of Information Packaging in Greek. In particular, Section 5.2.1 summarises some descriptive generalisations about the syntactic and phonological realisation of focus, topic and tails. In

Sections 5.2.2 & 5.2.3 I present an HPSG account of the data, essentially extending the analysis of Engdahl & Vallduví (1996)[1]. I then discuss the Information Structure of constructions involving Right Dislocation in Section 5.2.4 and offer some conclusions in Section 5.2.5. In the second part I discuss discourse and syntactic constraints on word order. In Section 5.3.1 I distinguish syntactic from discourse constraints on word order and show how Linear Precedence Rules can allow an account of discourse constraints independently of syntactic ones. In Section 5.3.2 I argue that adjacency restrictions are not syntactic in nature. In Section 5.3.3 I review previous approaches to the data. In Section 5.3.4 I present some problematic cases and conclude in Section 5.3.5.

5.2 Information Structure in Greek

5.2.1 The syntactic and phonological realisation of focus and ground in Greek

Not all syntactic constituents contribute to the Information Structure of the sentences they appear in. Weak pronouns in English and Greek clitics are such examples as discussed in Section 2.5.1. Since I have analysed Greek clitics as *affixes* and not as autonomous words or constituents, it follows that they cannot contribute to the Information Structure of the sentences they appear in. The *marker oti* is another example of such a constituent. It can neither function as ground, anchoring the new information to the hearer's information state, nor can it contribute new information. Even in a context that forces a narrow reading for *oti*, *oti* is not accented. Consider the following example:

(5.1) de mu ipe POS tha to kani; mu ipe oti/*OTI tha to KANI
not me-CL said-3SG how will it-CL do-3SG; me-CL said-3SG that will it-CL do-3SG
'S/he didn't tell me how she'll do it; s/he told me that s/he'll do it.'

In the above example, the accent may not fall on *oti*. Though the context forces a contrast between *pos* and *oti*, the contrast is established between *pos* and the embedded verb, *kani*.

On the other hand, heads (5.2a), arguments (5.3), adjuncts (5.2b), fillers (5.2b), may all bear nuclear accent.

[1] An earlier version of this analysis is presented in Alexopoulou (1998b).

(5.2) a. [$_L$o Petros][$_F$ ta EDOSE][$_T$ ta biblia sto Yani]
[$_L$the Petros-NOM][$_F$ them-CL gave-3SG][$_T$ the books to-the Yani]

b. [$_F$stin KUZINA][$_G$ afisa ta klidia]
[$_F$in-the kitchen][$_G$ left-1SG the keys]
'I left the keys in the kitchen.'

Below I summarise some phonological and syntactic constraints on constituents that do contribute to the Information Structure of their utterances (see Chapter 2).

Phonology

a) Accented constituents are obligatorily interpreted as focused. In both examples in (5.3) the accented constituent *sto Yani*[2] belongs to the focused part of the utterance.

(5.3) a. [$_L$o Petros][$_F$ edose ta biblia sto YANI]
[$_L$the Petros-NOM][$_F$ gave-3SG the books to-the Yanis]
'S/he gave the books to Yanis.'

b. [$_G$ [$_L$o Petros] edose ta biblia] [$_F$sto YANI]
[$_G$ [$_L$the Petros-NOM] gave-3SG the books] [$_F$to-the Yanis]

b) Ground elements do not bear any accent (*Petros, edose ta vivlia* in 5.3b).

c) Elements belonging to the focused part of an utterance may be accented (*Yani* in 5.3a) or not (*edose ta vivlia* in 5.3a).

Syntax

d) Any syntactic constituent contributing to the Information Structure of a sentence may belong to either part of the focus-ground partition.

That is, there is no syntactic constraint on the kind of constituents that may function as either ground or focus elements. The verbal head is focused in (5.3a&5.2a) and ground

[2] I assume here, oversimplifying matters, that the PP *sto Yani*, is accented, though accent falls only on the noun *Yani*.

in (5.3b&5.2b). The object *Yani* is focused in (5.3), whereas *ta klidia* is ground in (5.2b). Similarly for adjuncts and fillers.

However, once the distinction between links and tails is considered, there seem to be constraints on their syntactic realisation. As discussed in Section 2.4.3, links in Greek tend to appear preverbally, as fillers, while tails tend to appear postverbally. For the moment, I will assume that these facts should not be attributed to syntactic constraints. Rather, they follow from discourse constraints on the order between links, focus and tails (see Section 5.3.1).

While syntax does not constrain the kinds of constituents that are either ground or focus elements, it does play a role in the propagation of foci and links in the syntactic tree. In particular, the role of syntax is evident in the assignment of wide focus and focus in UDCs. The facts are described by (e):

e) Wide focus may arise only if the most oblique/rightmost complement is accented (5.3a); accent on the verbal head or on a preverbal XP is always interpreted as narrow focus (5.4) (Section 2.4):

(5.4) a. *[$_F$ to YANI ipe oti ide o Petros]
*[$_F$ the Yani-ACC said-3SG that saw-3SG the Petros-NOM]
'S/he said that Petros saw Yanis.'

b. *[$_L$o Petros][$_F$ ta EDOSE ta biblia sto Yani]
*[$_L$the Petros-NOM][$_F$ them-CL gave-3SG the books to-the Yani]
'Petros gave the books to Yanis.'

Examples (5.4a-b) are possible only with a narrow focus interpretation for *Yani* and *edose* respectively.

5.2.2 The representation of Information Structure in HPSG

Engdahl & Vallduví (1996) enrich the CONTEXT value of *signs* with a feature INFORMATION-STRUCTURE (INFO-STR) as shown in (5.5):

(5.5)
$$\begin{bmatrix} \text{CONTEXT} & \begin{bmatrix} \text{INFO-STR} & \begin{bmatrix} \text{FOCUS} & \\ \text{GROUND} & \begin{bmatrix} \text{LINK} \\ \text{TAIL} \end{bmatrix} \end{bmatrix} \end{bmatrix} \end{bmatrix}$$

By encoding INFO-STR in CONTEXT, they capture the independence of Information Structure from syntax and semantics (encoded in CAT and CONTENT respectively). FOCUS and GROUND take feature structures as values. Their values are structure-shared with the constituents that realise them. The various instantiation possibilities of INFO-STR values are handled by constraints on their phonological and syntactic realisation (see Manandhar (1994) for a different account in HPSG).

The phonological constraints (a-c) are captured by two constraints on *words*. In the following, *'a'* stands for *accented* and *'u'* for *non-accented/unmarked*:

(5.6) a. $$Word \rightarrow \boxed{1} \begin{bmatrix} \text{PHON} \mid \text{ACCENT} & a \\ \text{INFO-STRUCT} \mid \text{FOCUS} & \boxed{1} \end{bmatrix}$$

b. $$Word \rightarrow \boxed{1} \begin{bmatrix} \text{PHON} \mid \text{ACCENT} & u \\ \text{INFO-STRUCT} & [] \end{bmatrix}$$

(Engdahl & Vallduví 1996)

According to (5.6a) any *word* with PHON|ACCENT value of sort *a*, has its FOCUS value instantiated (Engdahl & Vallduví 1996). On the other hand, when a *word*'s PHON|ACCENT value is *u*, its INFO-STR value is not determined until it combines with other signs (5.6b). These two constraints are enough to account for (b&c). A ground element cannot be accented, since, if it were, its FOCUS value would be instantiated (5.6a). On the other hand, unaccented foci are allowed since the INFO-STR value of an unaccented word is underspecified (5.6b) and could be either focus or ground. The same constraints hold for Catalan (Engdahl & Vallduví 1996). For English, Engdahl & Vallduví (1996) assume a further constraint that associates a *word* with a B-accent with a link interpretation.

Engdahl & Vallduví (1996) capture the propagation of INFO-STR values on the syntactic tree with two principles constraining the instantiation of INFO-STR values. Their Principle I handles the propagation of links, tails and narrow focus whereas Principle II handles cases of wide focus. Their two principles are adopted for Greek and are presented below (I have modified Principle II in order to account for wide focus in sentences containing embedded clauses; I will return to these modifications later in this section):

(5.7) INFORMATION STRUCTURE INSTANTIATION PRINCIPLES FOR GREEK

- PRINCIPLE I

  if a DAUGHTER's INFO-STR is instantiated, then the MOTHER inherits this instantiation (for *narrow foci, links and tails*),

  or

- PRINCIPLE II

  the FOCUS value of the MOTHER is the sign itself (*wide focus*) just in case,

a) the FOCUS value of the most oblique COMPS-DTR is the sign itself

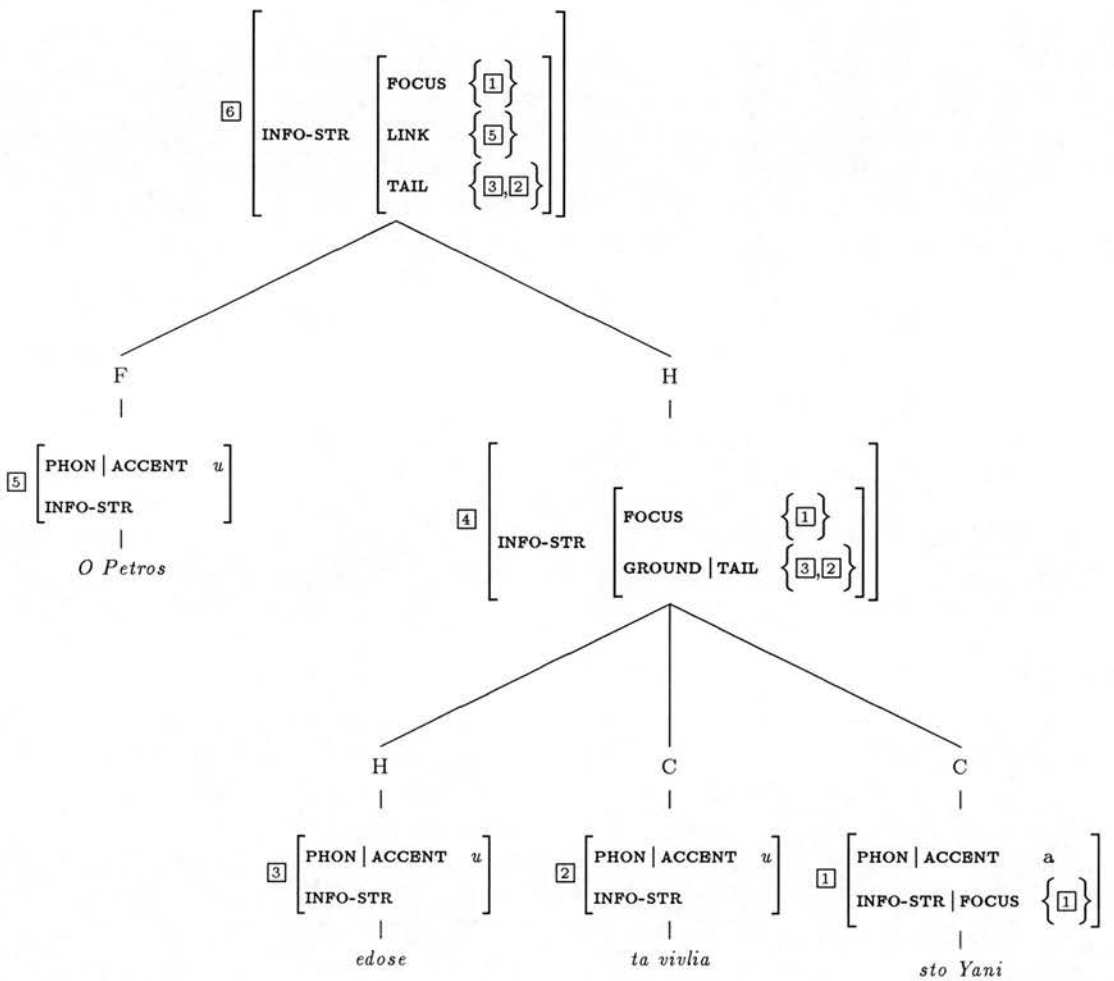
   or

b) in a *head-marker-phrase*, the FOCUS value of the HEAD-DTR is the sign itself.

Principle I allows a MOTHER to inherit the INFO-STR values of its DTRS. Consider (5.8), the structure of which is illustrated in (5.9):

(5.8) [$_G$ [$_L$o Petros] edose ta biblia] [$_F$sto YANI]
[$_G$ [$_L$the Petros-NOM] gave-3SG the books] [$_F$to-the YANI]
'Petros gave the books to Yanis.'

(5.9)

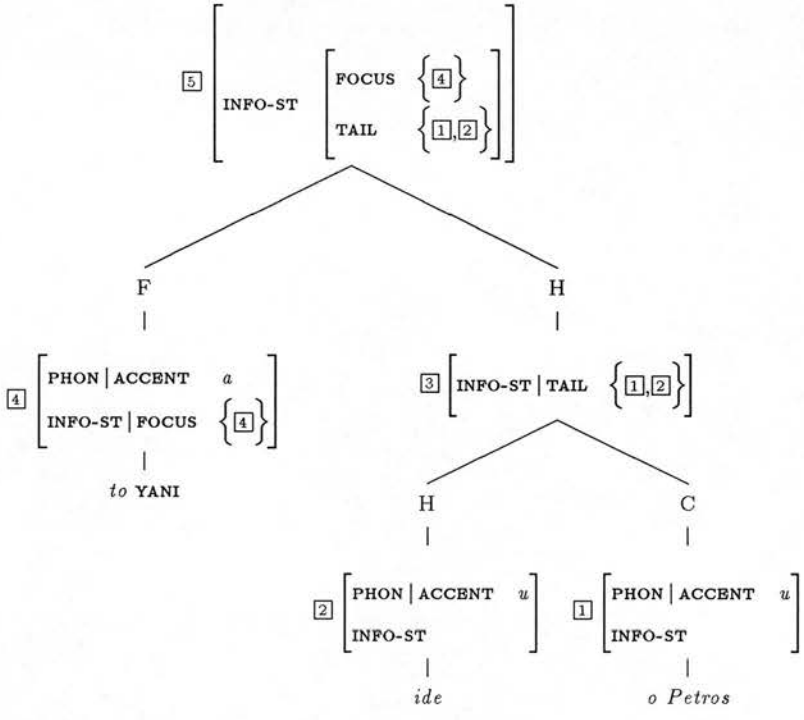


The complement PP *sto Yani* is accented (PHON|ACCENT *a*); thus, its INFO-STR|FOCUS value is instantiated and it is the sign itself [1] (5.6a). The verb and the direct object are not accented and their INFO-STR value is not instantiated. The *head-comps-phrase* inherits the FOCUS value [1] of its COMPS-DTR (Principle I). The verb and the direct object instantiate the TAIL value of their MOTHER. The *head-filler-phrase* inherits the INFO-STR value of the HEAD-DTR (FOCUS:[1], TAIL:[2],[3]). The *filler* is unaccented and its INFO-STR value is not instantiated. It appears as a link in the INFO-STR value of the *head-filler-phrase* ([6]).

Principle I also accounts for cases with preverbal focus. The *head-filler-phrase* inherits the FOCUS value of its FILLER-DTR. Tree (5.11) corresponds to example (5.10):

(5.10) [$_F$ to YANI] [$_G$ ide o Petros]
[$_F$ the Yani-ACC] [$_G$ saw-3SG the Petros-NOM]
'Petros saw Yanis.'

(5.11)



The FILLER-DTR 4 is accented and, by (5.6a), its FOCUS value 4 is instantiated. Then, the *head-filler-phrase* 5 inherits the FOCUS value 4, by Principle I.

Finally, Principle I allows examples with narrow focus within a preverbal constituent:

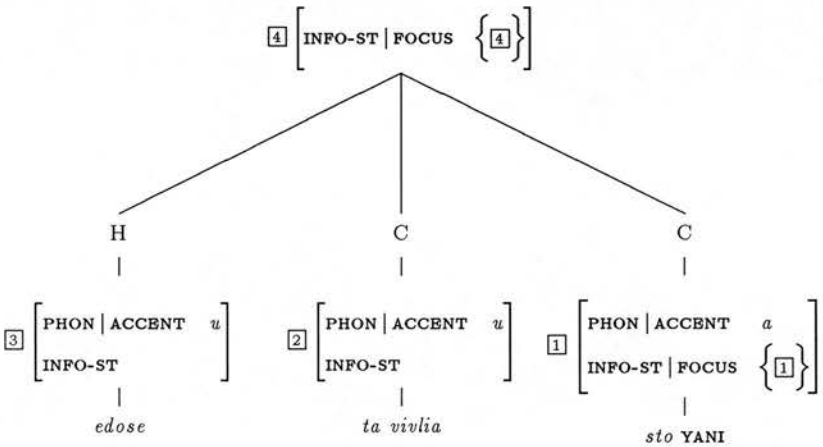
(5.12) me KOKINI mpoya evapsan ta parathira
with RED paint painted-3PL the windows
'They painted the windows with red paint.'

The NP mother inherits the FOCUS value from the ADJUNCT-DTR. The FOCUS value of the NP is iherited by the dominating PP which is the FILLER-DTR of the *head-filler-phrase*. The *head-filler-phrase* inherits the FOCUS value of the FILLER-DTR.

While Principle I accounts for narrow focus, Principle II handles wide focus data. According to Principle II.a, wide focus on a MOTHER (that is the FOCUS value of the MOTHER is the sign itself) can arise from the most oblique COMPS-DTR. Consider for example the structure of (5.13) illustrated by (5.14):

(5.13) [$_F$ edose ta biblia sto YANI]
[$_F$ gave-3SG the books to-the YANI]
'S/he gave the books to Yanis.'

(5.14)



The indirect object, *sto Yani*, is the most oblique COMPS-DTR. In addition, its FOCUS value 1 is the sign itself. Thus, the MOTHER 4 can be assigned wide focus (FOCUS:4).

The requirement that only COMPS-DTRS can give rise to wide focus blocks wide focus for *head-filler-phrases*. A *head-filler-phrase* (e.g. 5 in 5.11) cannot have wide focus, as none of its DTRS is a COMPS-DTR. In this way, the ungrammatical wide focus reading in (5.15) is ruled out:

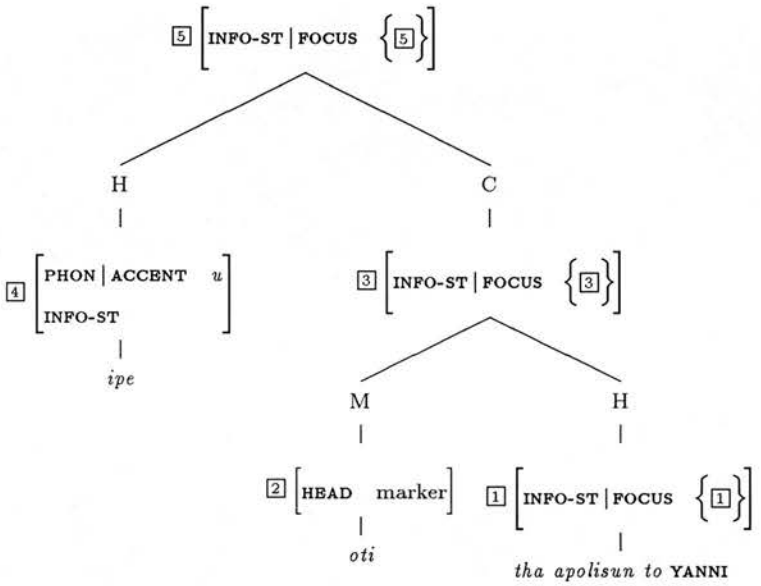
(5.15) *[$_F$ to YANI ide o Petros]
*[$_F$ the Yani-ACC saw-3SG the Petros-NOM]
'Petros saw Yanis.'

Principle II in (5.7) differs from Principle II of Engdahl & Vallduví (1996) in two ways. First, their Principle II does not contain the second clause. This clause is necessary in order to handle all focus examples in embedded contexts like (5.16):

(5.16) [$_F$ ipe oti tha apolisoun to YANI]
[$_F$ said-3SG that will fire-3SG the Yani-ACC]
'S/he said that they'll fire Yanis.'

The tree in (5.17) shows the structure of (5.16):

(5.17)



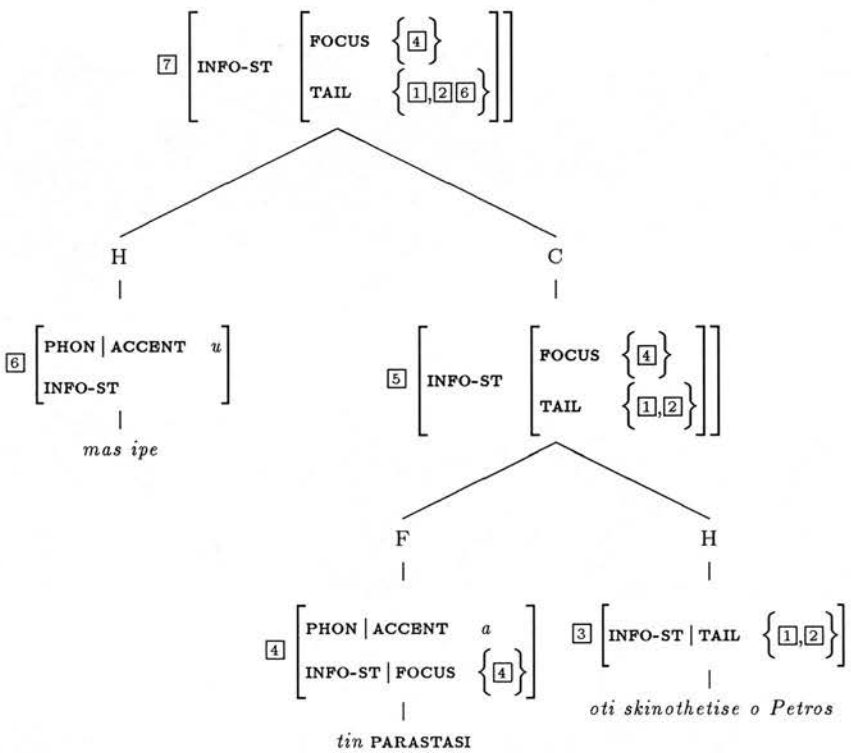
In (5.17) it is assumed that INFO-STR is not an appropriate attribute for MARKERS. This captures the fact that *oti* does not contribute to the INFO-STR value of the sign it appears in. In (5.17) the HEAD-DTR 1 has wide focus (FOCUS:1). By the second clause of Principle II the *head-marker-phrase* 3 is assigned wide focus as well. As 3 is the most oblique COMPS-DTR of the matrix verb (*ipe*) and it is widely focused the top node 5 is assigned wide focus.

The second difference between Principle II of Engdahl & Vallduví (1996) and Principle II in (5.7) involves the requirement that the COMPS-DTR (or the HEAD-DTR in a *head-marker-phrase*) is widely focused. This requirement is necessary in order to prevent a FOCUS value inherited from a FILLER-DTR from giving rise to wide focus once it appears on a COMPS-DTR. For example, consider the ungrammatical wide focus interpretation of (5.18):

(5.18) *[$_F$ mas ipe tin PARASTASI oti skinothetise o Petros]
*[$_F$ to-us said-3SG the show-ACC that directed-3SG the Petros-NOM]
'S/he said to us that Petros directed the show.'

Example (5.18) can only have a narrow focus interpretation as illustrated in (5.19):

(5.19)



In (5.19) the most oblique COMPS-DTR of *ipe* is the *head-filler-phrase* 5. However, 5 is not widely focused. Although the FOCUS value of 5 is instantiated (4), it is not the sign itself (5). Thus, Principle II cannot apply here. In effect, the ungrammatical wide focus interpretation of (5.18) is ruled out. In (5.19) the top node 7 inherits the FOCUS value 4 by Principle I.

Finally, Principle II accounts for wide focus arising from accent on an adjunct:

(5.20) a. A: ihes nea apo ti Maria?
'Did you have any news from Maria?'

b. B: tin ide o Petros sto SINEMA
B: her-CL saw-3SG the Petros-NOM at-the cinema
'Petros met her at the cinema.'

(5.21) a. A: ti eyine me tis prosklisis?
'What is happening with the invitations?'

b. B: tha ine etimes stin **ORA** tus

B: will be-**3SG** ready-**NOM** at-the time their-**CL**
'They will be ready on time.'

Since adjuncts are treated as complements (Section 4.5), Principle I can apply on examples like (5.20b&5.21b). The adjuncts realise the most oblique **COMPS-DTR**.

As discussed in Section 2.5.2, doubled objects resist accent. This restriction is expressed in the lexical entries of the relevant NPs. Their **PHON|ACCENT** value is *u*:

(5.22)
$$\left[\begin{array}{ll} \textbf{PHON} & \left[\begin{array}{ll} \textit{to vivlio} & \\ \textbf{ACCENT} & u \end{array}\right] \\ \textbf{SYN}\,|\,\textbf{LOC}\,|\,\textbf{CAT}\,|\,\textbf{HEAD} & \left[\begin{array}{ll} \textit{nominal} & \\ \textbf{CASE} & \boxed{2} \\ \textbf{MOD} & \left[\begin{array}{l} \textbf{HEAD}\ \textit{verb}(\textit{cl-word}) \\ \textbf{ARG-STR}\ \boxed{3}\left\langle \ldots \boxed{5}\ \textit{aff}\text{-}[\textbf{CASE}:\ \boxed{2}]\text{-}[\textbf{INDEX}:\ \boxed{4}]\ \ldots \right\rangle \end{array}\right] \end{array}\right] \end{array}\right]$$

Since they are not accented, their **INFO-STR** value is undetermined. They may be either ground elements (as they typically are) or part of wide focus.

In this way, a wide range of data (wide focus, narrow focus in UDCs, embedded contexts) is accounted for by two constraints mapping the phonological realisation of *words* with their **INFO-STR** value and two principles constraining the propagation of **INFO-STR** values in the syntactic tree. Note that no extra machinery is needed for the Information Structure of Focus-movement, Topicalisation, CLLD. The Information Structure of these constructions follows from general constraints.

In Catalan, the core part of the S is focus by default while the dislocated elements are ground elements (Section 2.3). Thus, Catalan uses only Principle I which allows the **FOCUS** value of the core S and the **GROUND** value of the dislocated daughters to be inherited by the top node of an utterance (a *head-disloc-phrase*; see Section 5.2.4). On the other hand, both Principles I & II are used in English. Like English, Greek uses both Principles I & II. It is remarkable that, despite significant typological differences between English and Greek, the propagation of **INFO-STR** values is subject to the same Principles in both languages.

5.2.3 More on Wide Focus

Principle II makes reference to the most oblique COMPS-DTR. In this section I will present evidence suggesting, that, at least in Greek, wide focus is sensitive to the FOCUS value of the rightmost element of an utterance rather than the most oblique COMPS-DTR. The facts presented in this section can be captured by the following generalisation:

(5.23) If the *rightmost daughter* has *wide focus*, then the *mother* may have *wide focus* as well.

Wide focus arising from post-adjunct arguments

Obliqueness is defined in terms of order in the ARG-STR or DEPS-STR. In the DEPS-STR adjunct-complements are appended to the list of argument-complements and appear as more oblique than the arguments (Section 4.5). However, as already discussed (Section 4.5), adjuncts may appear between the verb and the arguments, as illustrated by (4.44b) repeated as (5.24b):

(5.24) a. A: ti kanate me tis prosklisis
'What did you do with the invitations?'

b. B:[$_F$ tis stilame me to tahidromio sto YANI]
B:[$_F$ them-CL sent-1PL with the post to-the Yani-ACC]
'We sent them to Yanis by post.'

There is no reason to suggest that in (5.24b) the argument is more oblique than the adjunct. This example indicates that it is accent on the rightmost rather than the most oblique complement that triggers wide focus on the MOTHER.

Noun Phrases

Reference to the most oblique COMPS-DTR cannot capture the data from Greek NPs:

(5.25) a. [$_F$ agorasa ena vivlio tu CHOMSKY]
[$_F$ bought-1SG a book the-GEN Chomsky]
'I bought a book by Chomsky.'

b. [$_F$ agorasa mia kokini ZAKETA]
[$_F$ bought-1SG a red CARDIGAN]
'I bought a red cardigan.'

In (5.25a) accent falls on the genitive modifier, whereas in (5.25b) it falls on the noun. In both cases the accented contituent is the rightmost element of the NP and gives rise to a wide focus reading for the whole utterance. Principle II, as expressed in the previous section, cannot capture the similarities between NP and VP wide focus.

Links as parts of wide focus

There are cases (5.26b) in which preverbal links may appear as part of a wide focus reading:

(5.26) a. tu tilefonises;
'Did you phone him?'

b. ne, kai [$_F$ me diaveveose oti [$_L$ tis afises] tha tis ehi
yes, and [$_F$ me-CL confirmed-3SG that [$_L$ the posters-ACC] will them-CL have-3SG
etimes stin ORA-tus]
ready-ACC.PL on time-their]
'Yes, and he reassured me that he'll have the posters ready on time—that the posters will be ready on time.'

Example (5.26b) receives a wide focus interpretation despite the appearence of the embedded link (*tis afises*). The link *tis afises* both has the trappings of a link—preverbal occurrence, no accent— and it functions as link anchoring the focus part of the utterance to the hearer's information state. The analysis presented in the previous section does not account for (5.26b). The topic *tis affises* appears as a FILLER-DTR of the *head-filler-phrase*. Though the HEAD-DTR has wide focus, Principle II cannot apply as it only allows wide focus to arise from a COMPS-DTR. By contrast, 5.23 can predict (5.26b). The HEAD-DTR is the *rightmost daughter* and it is widely focused. On the other hand, *head-filler-phrases* with a focused FILLER-DTR (5.18) cannot be assigned wide focus, since the focused daughter is not the rightmost one.

In addition, 5.23 allows wide focus in cases where the verb is accented and is the rightmost daughter. Consider the following as an answer to (5.26b):

(5.27) a. A:ti eyine me tis prosklisis; tilefonises sto Yani?
A:'What happened with the invitations? Did you phone Yanis?'

b. B:ne, kai [$_F$ me diaveveose oti tis prosklisis gia ti deksiosi tha
B:Yes, and [$_F$ me-CL confirmed-3SG that the invitations for the reception will
tis TIPOSI]; gia tis prosklisis tu gamu den kseri akoma
them-CL print-3SG]; for the invitations the wedding-GEN not know-3SG yet
'Yes, and he reassured me that he'll print invitations for the reception; he doesn't know yet for the invitations for the wedding.'

Example (5.27b) receives a wide focus reading. The accent falls on the embedded verb which is the HEAD-DTR of the *head-filler-phrase*. Again, reference to the most oblique COMPS-DTR does not account for this case. The accented constituent is a head, but the rightmost daughter of the utterance.

Finally, 5.23 makes it possible to eliminate the second clause of Principle II. The HEAD-DTR of the *head-marker-phrase* is the rightmost one.

Thus, 5.23, compared with Principle II, can account for a wider range of data and do it in a more elegant way. Before its final formulation, I will consider examples of Right Dislocation and Clitic Right Dislocation.

5.2.4 Right Dislocation

As discussed in Section 2.4.3, it is assumed that elements scrambling to the right edge of the clause and crucially occurring after nuclear accent undergo Right Dislocation. In the literature, right dislocated elements are treated as distinct constituents. For example, Tsimpli (1995) proposes that right dislocated subjects (in VOS) and doubled objects are adjoined to TNSP (Section 3.2.5). However, the general freedom in the order of postverbal complements (arguments and adjuncts) in Greek, makes it hard to argue for the existence of a distinct constituent on purely structural evidence. Consider the following examples:

(5.28) a. i mama tiganize stin kuzina PATATES
the mother-nom was-frying-3SG in-the kitchen potatoes-ACC
'Mum was frying potatoes in the kitchen.'

b. i mama tiganize stin KUZINA patates
the mother-nom was-frying-3SG in-the kitchen potatoes-ACC

(5.29) a. evale sto psiyio ta **PORTOKALIA** (ke kitakse yiro yiro sa hameni)
put-3SG in-the fridge the oranges (and looked around around as-if lost)

'She put the oranges in the fridge (and looked around as if lost).'

b. evale sto **PSIYIO** ta portokalia
put-3SG in-the fridge the oranges

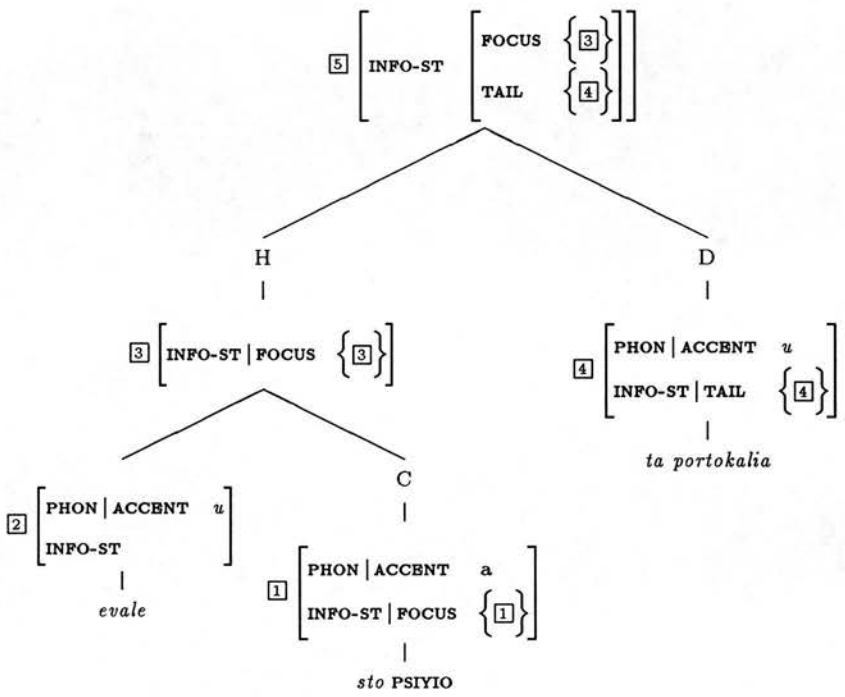
In all examples above the argument occurs after the adjunct. In the (a) examples the accent falls on the argument, which is the rightmost element of the utterance. In the (b) examples the accent falls on the adjunct. It is hard to distinguish the two cases structurally and argue that, in the (b) examples, the argument is dislocated to the right whereas in the (a) examples it is not.

The notion of a right dislocated constituent becomes useful once the Information Structure of these sentences is considered. The (a) examples may receive a wide focus reading, whereas in the (b) examples, the argument is typically interpreted as a tail. Example (5.29b) may have the following Information Structure:

(5.30) [$_F$evale sto **PSIYIO**] [$_T$ta portokalia]
[$_F$put-3SG in-the **FRIDGE**] [$_T$the oranges]
'She put the oranges in the fridge.'

Engdahl & Vallduví (1996) postulate a **HEAD-DISLOCATION SCHEMA** for Catalan that licenses a phrase (*head-disloc-phrase*) consisting of a **HEAD-DTR** and one or more **DISLOCATED-DTR**s. This **SCHEMA** is motivated in Catalan by the fact that all ground elements are dislocated out of the core S, which coincides with the focus part of utterances. If their analysis were to be adopted for Right Dislocation in Greek, (5.30) would have the structure shown in (5.31):

(5.31)

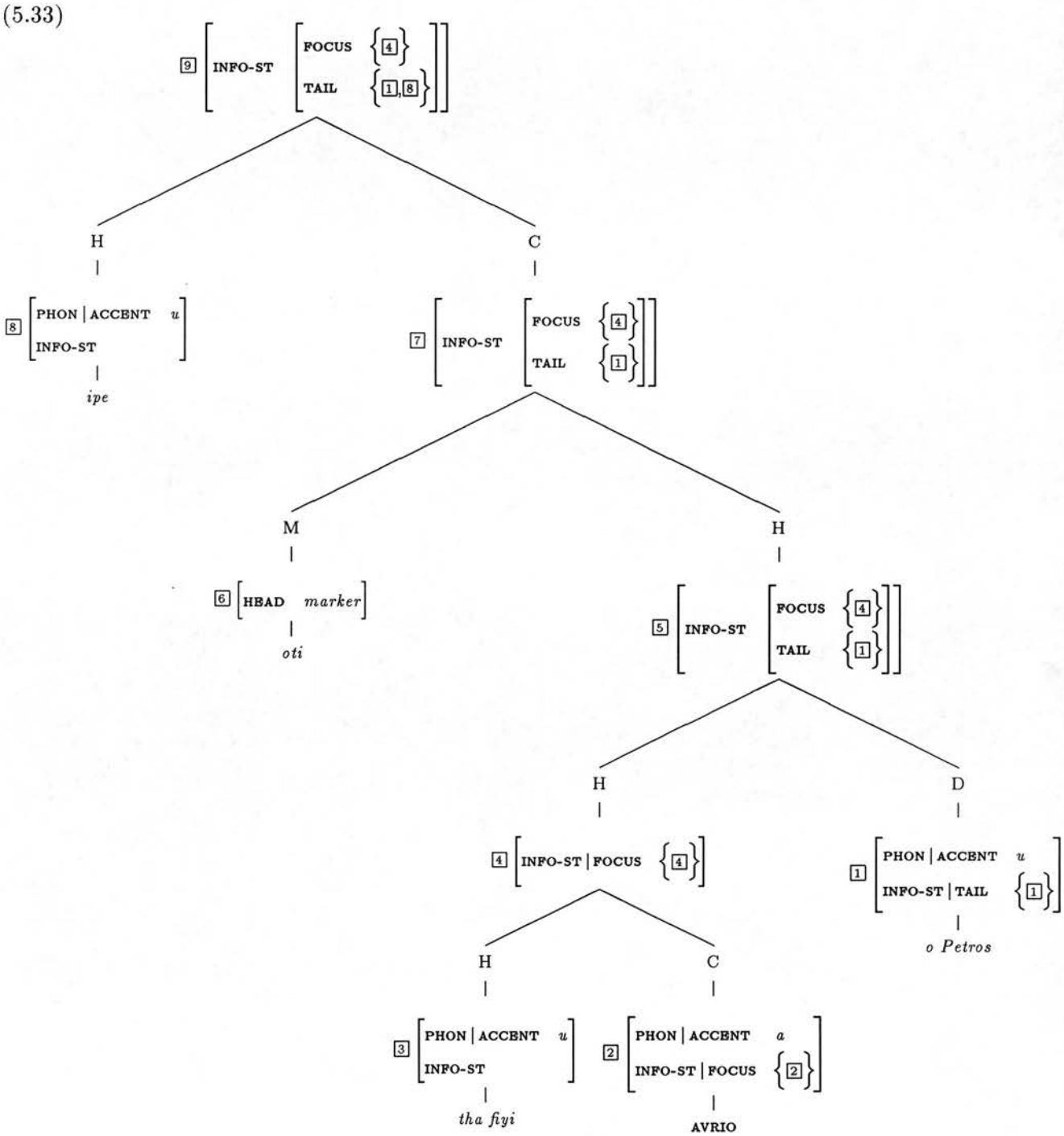


However, this analysis cannot be extended to examples like (5.32):

(5.32) [$_F$ipe oti tha fiyi **AVRIO**] [$_T$ o Petros]
[$_F$said-3SG that will leave-3SG tomorrow] [$_T$ the Petros-**NOM**]
'Petros said he'll leave tomorrow.'

The structure of (5.32) is shown in (5.33):

(5.33)



The rightmost daughter 2 of the *head-comp-phrase* 4 is accented; so, the **MOTHER** 4 is assigned wide focus. The *head-discloc-phrase* 5 inherits the **INFO-STR** values from its **DTRS**.

Principle II cannot apply here because the rightmost daughter, which is the DISLOCATED-DTR is not focused. The *head-marker-phrase* [7] inherits the INFO-STR value of its HEAD-DTR (the MARKER-DTR has no INFO-STR value). Since the rightmost daughter of [7] does not have wide focus, [7] cannot be assigned wide focus. The same holds for the *head-comp-phrase* [9] which inherits the INFO-STR value of its DTRS. Thus, wide focus 'stops' at the embedded clause.

The problem could be solved on the assumption that *o Yanis* is adjoined to the topnode [9] rather than the lower S [5]. The ID-SCHEMA postulated by Engdahl & Vallduví (1996) requires that the HEAD-DTR of a *head-disloc-phrase* is saturated (COMPS $< >$). Thus, [9] satisfies this description. In addition though, the HEAD-DISLOCATION SCHEMA requires that the DISLOCATED-DTR is in a *binding* relation with an element in the CONTENT value of the HEAD-DTR. Engdahl & Vallduví (1996) do not specify the nature of this binding relation. Whatever that is though, if *o Yanis* were to be adjoined to the higher node, some local information should be propagated from the embedded head to the matrix one. This would make Right Dislocation look like an Unbounded Dependency. However, Right Dislocation should be distinguished from UDCs. One significant difference between the two is that Right Dislocation is a local process. For example, there is no 'long distance right dislocation'. Compare the following:

(5.34) a. to Yani ipe ston Petro/s'osous ton rotisan oti ton
the Yani-acc said-3SG to-the Petros/to-whoever him-CL asked-3PS that him-CL
ide
saw-3SG
'S/he told Petros/whoever asked him that s/he saw Yanis.'

b. *ipe oti ton ide ston Petro/s'osous ton rotisan
*said-3SG that him-CL saw-3SG to-the Petros/to-whoever him-CL asked-3PS
to Yani
the Yani-acc
'S/he told Petros/whoever asked him that s/he saw Yanis.'

Example (5.34b) in which *to Yani* is extracted to the right, is ungrammatical. Thus, whatever the assumptions made for the licensing of a *head-disloc-phrase*, in (5.32) the DISLOCATED-DTR, *o Yanis*, has to be attached to the lower S [5]. It is unclear how Principle II or 5.23 could be reformulated in order to account for the wide focus reading in (5.32). In conclusion,

assuming a right dislocated constituent does not account for data involving sentences with embedded clauses.

Even if this analysis worked, the postulation of a distinct constituent is based on a spurious argument. Part of the evidence presented in Chapter 2 indicating the independence of Information Structure from syntax was based on the fact that the ground-focus partition (in Greek and English) does not correspond to syntactic constituents. It is not obvious why in the case of Right Dislocation, interpretational considerations should suffice as motivation for the existence of a right dislocated constituent.

A proper answer to the question of the existence of a right dislocated constituent requires further research. However, in the absence of solid structural evidence I assume that there is no such constituent. I will return to the issue of postverbal scrambling in Section 5.3.1.

The question that arises now is how example (5.30) can be accounted for without reference to a right dislocated constituent. Further, whether and how (5.32) could be predicted. A tentative solution is to attach a second clause to 5.23 and revise Principle II as follows:

- **Principle II** (*final version*):

  a) If the *rightmost daughter* has *wide focus*, then the MOTHER may have *wide focus* as well.

  b) In a *head-comp-phrase* with HEAD:*verb*, if any daughter is accented, then the FOCUS value of the MOTHER may have wide focus; any COMPS-DTRS occurring after the accented daughter instantiates the MOTHER's GROUND|TAIL value.

The second clause captures examples (5.30&5.32) as well as the following:

(5.35) [$_F$ipane oti tha ton DIOKSUNE] [$_T$to Yani]
[$_F$said-3PL that will him-CL fire-3PL] [$_T$the Yani]
'They said they'll fire Yanis.'

In (5.30,5.32&5.35) the accented DTR appears in a *head-comp-phrase* which is assigned wide focus. The elements occurring after the accented DTR (*ta portokalia*, *o Yanis*, *to Yani*) instantiate the TAIL value of the *head-comp-phrase*. Since the *head-comp-phrase* has wide focus, the dominating nodes in (5.32&5.35) may also have wide focus (Principle II.a.).

5.2.5 Discussion

The analysis presented in this chapter succeeds in capturing various aspects of the phenomena in question. The organisation of the HPSG sign allows a flexible account of the relation between syntax and Information Structure. The independence of discourse functions is captured by encoding Information Structure within the CONTEXT value of signs. As noted in the introduction of this chapter, the *Discourse Configurational* approach to the discourse-syntax interface, offers a rather complex syntax but a simple, isomorphic relation between discourse functions and syntactic operations. By contrast, the analysis presented in this chapter, assumes a simple syntax which yields a non-isomorphic relation between syntactic constructions and discourse functions. This, however, does not lead to a more complex architecture of the discourse-syntax interface. A small number of constraints (two constraints on the phonological realisation of *words* and two principles constraining the propagation of INFO-STR values) are enough to account for a wide range of data (propagation of INFO-STR values in *head-comp-phrases*, *head-filler-phrases* and in sentences with embedded clauses). In addition, the same set of Principles can accommodate the crosslinguistic realisation of Information Packaging in languages as different as English, Greek and Catalan.

A weakness of the presented account is that it does not account for the non-recursive nature of focus. Even though nothing forces ocurrence of recursive foci, nothing excludes it either. A constraint requiring that the FOCUS value of a *sign* is at most the singleton set would yield the correct result. Admittedely this seems an ad-hoc restriction.

The current approach raises some questions with respect to the relation of Information Structure to the rest of the grammar. It was shown that, at least for Greek, there is a descriptive advantage in referring to the rightmost-DTR of an utterance rather than the most oblique COMPS-DTR. It is not clear what is the status of the notion *rightmost* and whether it should be viewed as syntactic in nature. Reference to syntax seems necessary only because accent on some particular positions on the tree may give rise to wide focus readings. However, these positions cannot be defined in terms of familiar syntactic categories (head, complement, etc.). In addition, it is not clear how *rightmost* can be expressed in HPSG terms.

The present analysis encodes Information Structure in the *sign* and handles the wide/narrow focus distinction through constraints on the propagation of the INFO-STR values from node

to node on the syntactic tree (Principles I & II). Thus, this analysis implies that Information Structure is sensitive to the hierarchical organisation of sentences; at least, to relations of dominance between MOTHERs and DTRs. Undoubtedly Information Structure should be anchored to the *sign*, since it is the *sign* that is partitioned in a focus and a ground part. In addition, this partition is sensitive to where exactly on the tree the accent falls. However, there is no conclusive evidence that Information Structure is part of the *sign* and that it is sensitive to the full range of hierarchical structure involved in a sentence/*sign*. That is, it is not obvious that the wide/narrow focus distinction reflects sensitivity to syntactic structure of the kind implied by Principles I & II.

Right Dislocation constructions are indicative of the difficulty involved in accounting for the focus-ground partition through constraints on the propagation of INFO-STR values. The discussion of these constructions showed that neither reference to the most oblique COMPS-DTR nor to the rightmost daughter can capture the facts. Intuitively, what seems to happen in these constructions could be stated as *all the material to the left of an accented postverbal/in situ constituent may belong to focus; any material to the right is ground.* This statement does make reference to a syntactic concept, postverbal/in situ. However, the translation of this description into constraints relying on the hierarchical organisation of the syntactic tree is not straightforward. The question will not be pursued any further here. However, a better understanding of the relation between Information Structure and the rest of the grammar depends on these issues.

5.3 Syntactic and Discourse Constraints on word order

Word Order is a descriptive term referring to the linear order of words in a linguistic string. However, linguistic strings are organised hierarchically in constituents. Some ordering restrictions can be attributed to constraints on constituents. For example, illicit interleaving or extraction out of a strong island yield ungrammatical orders. This kind of constituency constraints was discussed in the analysis of Unbounded Dependencies in Greek (Chapter 4).

In the following sections, I will focus on constraints on the order of words within well-formed constituents. Two sorts of such constraints can be distinguished; syntactic and discourse ones. Syntactic constraints on word order define the order of syntactic categories; the relative order between a head and its complements, a marker and the syntactic head, a

filler and its syntactic head. Discourse constraints on word order restrict the order between discourse elements, between new and old information. I will argue that discourse constraints on word order should be represented independently of syntactic ones.

5.3.1 Linear Precedence Rules

In HPSG, ordering constraints are captured by *Linear Precedence Rules* (LPR henceforth) defining the order within a well-formed constituent (Pollard & Sag 1987; Uszkoreit 1986). For example, the following LPRs on *head-comp-phrases* states that a (verbal) head should precede its complements:

(5.36) *head-comp-phrase* → HEAD-DTR < COMPS-DTR

I assume that the order of postverbal COMPS-DTRs in Greek is unconstrained. This assumption does not express the fact that VSO is the basic order in Greek (Section 1.2). This could be captured with a LPR constraining the order of postverbal COMPS-DTRS to be that of obliqueness in the ARG-STR. However, such a constraint would require additional processes (e.g. a dislocation SCHEMA) to yield the observed variation in the order of postverbal complements. Therefore, it will not be adopted here.

Fillers and markers are ordered with respect to their heads by the following LPRs on *head-filler-phrases* and *head-marker-phrases*. The LPRs in (5.37) state that the extracted XPs and complementisers should precede their heads:

(5.37) a. *head-filler-phrase* → FILLER-DTR < HEAD-DTR

b. *head-marker-phrase* → MARKER-DTR < HEAD-DTR

The LPRs in (5.36&5.37) are examples of syntactic constraints on word order. Such constraints have three distinguishing properties. First, they apply to specific constituents (*head-comp-phrase*, *head-filler-phrase* etc.). Second, their violation results in ungrammaticality. As already shown (Section 5.2.4,ex.5.34) extraction to the right (that is, violation of 5.37a) is ungrammatical. A third feature of syntactic constraints on word order is that they do not interact with other constraints on word order. In particular, they cannot be overriden by the requirement that heavy constituents should appear at the right edge of a clause:

(5.38) *ipe oti ide ston Petro tin kopela me ta MEGALA matia
*said-3SG that saw-3SG to-the Petros the girl-ACC with the BIG eyes
'S/he told Petros that s/he saw the girl with the big eyes.'

Even though the extracted constituent *tin kopela me ta megala matia* is 'heavy', extraction to the right is not possible[3].

Discourse constraints on order define the relative order between new and given information in a well-formed constituent. In general, links tend to precede the focus part of an utterance while tails tend to follow it (in Greek as well as in Catalan). The following LPR (Engdahl & Vallduví 1996) expresses this generalisation:

(5.40) *phrase* → LINK < FOCUS < TAIL

The LPR in (5.40) is a constraint on *phrases.* Because, discourse elements—links, focus and tails— are not realised by specific syntactic categories (head, complement,etc.), (5.40) does not make any reference to such categories. For example, consider the data in (5.40):

(5.41) a. [$_L$ to party] [$_F$ i ELENI] [$_G$ to ithele] (o Yanis den
[$_L$ the party] [$_F$ the Eleni-NOM] [$_G$ it-CL wanted-3SG] (the Yanis-NOM not
ihe oreksi)
had-3SG appetite-ACC)
'Eleni wanted the party (Yanis didn't feel like it).'

b. [$_F$ ton IDE] [$_T$ to Yani]
[$_F$ him-CL saw-3SG] [$_T$ the Yani-ACC]
'S/he saw Yanis.'

In (5.41a) the link and the focused phrase are both FILLER-DTRS. In (5.41b) the focused phrase is the HEAD-DTR of the *head-comp-phrase* while the tail is realised as a COMPS-DTR.

LPR (5.40) also accounts for the following data observed by Schneider-Zioga (1994):

[3]It should be noted though, that the following, in which the object NP on the right is doubled with a clitic and bears no accent is much better:

(5.39) ?ipe oti tin IDE ston Petro tin kopela me ta megala matia
?said-3SG that her-CL saw-3SG to-the Petros the girl-ACC with the BIG eyes
'S/he told Petros that s/he saw the girl with the big eyes.'

Example (5.39) does not seem to instantiate right extraction.This is supported by the fact that when the phrase is focused/accented (5.38) the result is ungrammatical. Rather, example (5.39) could be viewed as the mirror image of Left Dislocation, which was shown to be different from extraction constructions (Section 3.1.3). It remains an open question what is the relation between (5.39) and the Right Dislocation constructions discussed in Section 5.2.4.

(5.42) a. ksero *to moro* oti o YIORGOS to-frondise
know.1s the baby that the GEORGE cl.acc.n-cared.for
'I know that GEORGE took care of the baby.'

b. *ksero o YIORGOS oti *to moro* to-frondise
know.1s the GEORGE that the baby cl.acc.n-cared.for
'I know that GEORGE took care of the baby.'

(Schneider-Zioga 1994:ex.66a,c)

Schneider-Zioga (1994) notes that the restriction that the topicalised XP precedes the focused one holds even when one of the two appears before the complementiser *oti*. Since LPR (5.40) applies on *phrases*, it applies on the *head-filler-phrase* which is the complement of the matrix verb *ksero*.

Finally, LPR (5.40) captures the obligatory preverbal appearance of Greek links. Since links are the first in the linear precedence order, they have to appear before the verb, as the verb will have to be either focused or ground/tail and follow links. (In 5.42a the link, *to moro*, appears after the matrix verb. I will return to this problem in Section 5.3.4).

Though both syntactic and discourse constraints on order are captured here by LPRs, it should be noted that the two differ in various ways. First, unlike syntactic constraints, discourse constraints on order do not apply over specific constituents (subtypes of *phrase*), but over utterances. This is not a surprising fact, given that the ground-focus partition is organised along utterances rather than specific constituents. More importantly, their violation does not give rise to strong ungrammaticality judgements. Along with Schneider-Zioga (1994) the literature on Greek has assumed that there is an adjacency restriction between focused phrases and the verbal head (Agouraki 1993; Tsimpli 1995; Tsimpli 1996; Tsiplakou 1998). Though this intuition is a valid one, I would like to claim that this constraint is not as rigid in nature as implied by various accounts in the literature. Violation of the adjacency restriction between the focused XP and the verb or the LPR in (5.40) does not give rise to strong ungrammaticality. Acceptability judgements vary from speaker to speaker with respect to the following example (in fact, some informants accept example 5.43):

(5.43) ???ti MARIA o Petros ide sto sinema
???the Maria-ACC the Petros-NOM saw-3SG at-the cinema
'Petros saw Maria at the cinema.'

It should be pointed out that there doesn't seem to be any context that would improve the acceptability of (5.43). The badness of (5.43) is due to the violation of LPR (5.40) (or the adjacency restriction). What is crucial here is the fact that the violation of LPR (5.40) in (5.43) gives rise to gradient acceptability in a way that the violation of syntactic constraints (Phrase Structure, syntactic constraints on word order) does not. This claim can only be tested by experiments on the gradience of acceptability judgments of the kind proposed in Bard *et al.* (1996). In relation to this, Keller (1998) offers some interesting conclusions. Based on experimental evidence, he distinguishes two kinds of grammatical constraints: *hard* and *soft* ones. The violation of hard constraints induces strong grammaticality judgements whereas the violation of soft constraints induces milder grammaticality judgements. Interestingly, the distinction drawn between syntactic and discourse constraints is similar to the one drawn in Keller (1998). In his experiment, syntactic constraints (on Phrase Structure, number agreement and subcategorisation requirements) appear as hard constraints whereas 'interpretational' constraints (on referentiality or definiteness) are soft ones.

Finally, the third characteristic distinguishing discourse constraints from syntactic constraints is that the former may be overriden by the restriction that heavy constituents are shifted to the right edge of clauses. In the following, ground information appears before the focus constituent; the focused constituent is a 'heavy' one:

(5.44) [$_G$ tis ipe tis Marias] [$_F$ oti tha ti di sto SINEMA]
[$_G$ her-CL said-3SG the Maria-GEN] [$_F$ that will her-CL see-3SG at-the sinema]
'S/he told Maria that s/he'll meet her at the cinema.'

The interaction with the heavy-constituent-shift together with the mild ungrammaticality caused by their violation, suggests that discourse constraints reflect strong tendencies rather than rigid, unviolable constraints.

In sum, word order arises from the combination of constraints originating from different parts of the grammar. In particular, I have identified three sources: a) constraints on constituency, b) syntactic constraints on word order and c) discourse constraints on word order. Discourse constraints on word order appear qualitatively different from syntactic ones. The first could be described as *soft/relative* constraints whereas the latter as *hard/rigid* ones. An adequate analysis of these facts should a) capture the independence of disourse constraints from syntactic ones and b) express the different nature between syntactic and non-syntactic

constraints. The LPR approach sketched here succeeds in the first goal, in that discourse constraints on word order are stated by distinct LPRs, independently of syntax. However, it does not capture the different nature of the two kinds of constraints. Rather than reflecting a tendency, the LPR in (5.40) is as 'rigid' as the LPRs in (5.36&5.37). The question of how this kind of constraint is expressed in current generative frameworks is a challenging one, which, however, remains beyond the scope of this study (see Keller (1998) with respect to this).

5.3.2 The non-syntactic nature of the adjacency restrictions

In this section I discuss two examples of adjacency restrictions between an extracted XP and a head; preverbal focused phrases and wh-phrases. I will present evidence indicating the non-syntactic nature of this restriction.

Focused phrases

Crosslinguistically focused phrases tend to appear adjacent to the verb (Brody 1990; Hoffman 1995; King 1995; Kiss 1995b). This fact has been captured as a case of Spec-Head agreement between the focused phrase and a head specified for the focus feature to which the verb moves. For Greek, it has been assumed that the focused phrase moves to [Spec,FP]. This analysis accounts for both the examples below:

(5.45) a. to YANI ide sto sinema
the Yani-ACC saw-3SG at-the cinema
'S/he saw Yanis at the cinema.'

b. rotise to YANI pios ide sto sinema
asked-3SG the Yani-ACC who-NOM saw-3SG at-the cinema
'S/he asked who saw Yanis at the cinema.'

This analysis captures simultaneously two facts. First, that there is no rightwards extraction and second, that the focused XP is adjacent to the verb or the wh-phrase. However, as discussed in the previous section, there are reasons to believe that the two constraints should be distinguished. The violation of the adjacency requirement (5.43) is not as severe as the violation of leftward extraction (5.34b). (In addition, this analysis relies on the existence of a Focus Phrase, which has been shown to be problematic for independent reasons).

Wh-phrases

It has also been assumed that wh-phrases appear adjacent to the verb. This assumption is based on examples like the following:

(5.46) *pion o Jannis idhe?
*whom the John saw?
'Who did John see?"

(Anagnostopoulou 1994:ex.43)

As with focus, the violation of the adjacency requirement in Wh-questions does not give rise to severe ungrammaticality. In addition, there are fully acceptable examples in which the wh-phrase does not appear adjacent to the verb:

(5.47) posa apo ta abstracts i epitropi tu Glow (ta) aperipse omofona?
how-many of the abstracts the committee the Glow (them-CL) rejected unanimously?
'How many of the Glow abstracts did the Glow committee reject unanimously?'

(Anagnostopoulou 1994:ex.47)

Whatever factors allow (5.47), this example indicates that the adjacency restriction is not as strict as it has been assumed to be.

Further, it appears that discourse factors affect the acceptability of adjacency violations. Unlike topicalised XPs, focused XPs cannot intervene between a wh-phrase and the verb:

(5.48) ??rotise posa apo ta abstracts i epitropi tu GLOW (ta) aperipse
asked-3SG how-many of the abstracts the committee the Glow (them-CL) rejected
omofona
unanimously
'S/he asked how many of the Glow abstracts the Glow committee rejected unanimously.'

It is unclear why the focused phrase in (5.48) is not tolerated between the wh-phrase and the verb while the topicalised one is (5.47)[4]. An adequate understanding of the nature of the

[4] Example (5.47) rendered as an indirect question is equally acceptable.

adjacency restriction warrants thorough research that is not undertaken here. However, the flexibility in the ordering patterns suggests that the differences between (5.46&5.47) as well as between (5.47&5.48) are not of syntactic nature and should not be treated in a structural way.

5.3.3 Previous approaches

Structural approaches

Discourse configurational analyses treat word order as a matter of constituency. The order between topics and preverbal focus is captured by the the relative order of the corresponding positions in the syntactic tree. This approach has, by and large, the correct results with respect to word order. However, it is unsatisfactory for two reasons. First, it is too rigid to capture readily the flexibility of the attested orders (for example the relative nature of adjacency restrictions). Second, it does not capture the independence of discourse constraints on word order from syntactic ones. In addition, it only captures the relative order between topics/links and preverbal focus. As discussed in Section 5.3.1. the relative order between topic and foci is a consequence of general constraints ordering discourse elements (LPR 5.40) that do not hold only for extracted XPs. LPR 5.40 is a constraint on *phrases*, not just *head-filler-phrases.*

Domain Union

In recent HPSG literature (Kathol & Pollard 1995; Kathol 1995; Pollard *et al.* 1994; Reape 1994) there has been a tendency to detach linear order from hierarchical structure. This is achieved through the introduction of *Order Domains* in the *sign* and the operation of *Domain Union.* The *sign* is enriched with an independent, *non-hierarchical* level of *linear* representation, the *Order Domain* (Reape 1994). A feature DOM is introduced, which takes a *list* of objects (signs or partial signs) as its value. At each level of syntactic combination the order domain of the mother (the value of DOM) is formed from the DOM values of the daughter constituents, through *Domain Union. Domain Union* allows the order domain of the DTR (which is a list) to unify/merge with the order domain of the MOTHER (which is an other list). The output is an order domain (a new list) which contains exactly those elements that appeared in the two original order domains, only arbitrarily permuted. The result is a

flat structure in which subconstituents of a DTR can enter in linearization relationships with elements of other syntactic constituents which are not their sisters. This approach has proved very useful for phenomena of *discontinuous constituency* (Reape 1994).

There are two objections to this approach with respect to the Greek data: one of theoretical and one of empirical nature.

First, Borsley (1996) observes that order domains comprise a departure from the *monostratal* nature of HPSG, a fact which might not be desirable from a theoretical point of view.

Second, as mentioned in Section 5.3, word order arises in part from the hierarchical organisation of linguistic strings into constituents. The *Domain Union* approach understates the role of constituent structure in word order. It is worth noting that, despite the word order freedom, there is no interleaving in Greek as the ungrammaticality of (5.49) shows:

(5.49) *ipe to Yani$_i$ tis Marias oti ton$_i$ ide sto SINEMA
said-3SG the Yani-ACC the Maria-GEN that him-CL saw-3SG at-the cinema
'S/he said to Mary that s/he saw Yani at the cinema.'

A *Domain Union* approach would complicate the grammar considerably, as additional Linear Precedence Rules should be stipulated to rule out ungrammatical strings with interleaving constituents. On the other hand, there is no obvious benefit in this approach for the expression of discourse constraints on word order.

In a variant of this approach, proposed by Kathol (1995), order domains are partitioned in *topological fields*. *Topological fields* constrained to contain a single element are a rough equivalent of structural positions in Principle & Parameters/Minimalist approaches. There is, however, an important difference. One member *Topological fields* are not specified for any syntactic/semantic feature. Thus, diverse elements (XPs, wh-phrases, markers) may appear at the same *topological field*. Though this notion of 'position' is more flexible, it does not seem suitable for the Greek data. Such an approach would imply a rigid organisation of the order of elements in the left periphery of the Greek clauses which is not evident in the data.

5.3.4 More on discourse constraints

Though LPR (5.40) accounts for a wide range of cases, it cannot capture examples like the following:

(5.50) [$_G$ ton ide][$_F$ sto SINEMA][$_T$ to Yani]
[$_G$ him-CL saw-3SG][$_F$ at-the CINEMA][$_T$ the Yani-ACC]
'S/he saw Yanis at the cinema.'

In (5.50) the focused phrase appears between two ground elements (the first of which cannot be interpreted as a link). A similar problem arises in embedded contexts:

(5.51) [$_L$ tis Marias] [$_G$ tis ipane] oti [$_L$ ta klidia] [$_F$ ta dosane sto
[$_L$ the Maria-GEN] [$_G$ her-CL told-3PL] that [$_L$ the keys] [$_F$ them-CL gave-3PL to-the YANI]
YANI]
'They told Maria that they gave the keys to Yani'

Example (5.51) involves two links, one belonging to the matrix clause and one to the embedded one. While they both precede the focused part of the utterance, the second link, *ta klidia* follows a part of the ground. Note as well that links may appear after the focus part of an utterance, as shown in (5.52):

(5.52) [$_L$ tis Marias] [$_F$ tis TO-PANE] oti [$_L$ ta klidia] [$_T$ ta dosane
[$_L$ the Maria-GEN] [$_F$ her-CL it-CL-told-3PL] that [$_L$ the keys] [$_T$ them-CL gave-3PL sto Yani]
to-the YANI]
'They told Maria that they gave the keys to Yanis.'

In (5.52) the stress falls on the matrix verb. The embedded clause has undergone Clitic Doubling (the clitic *to* refers to the embedded clause). The issue deserves systematic research that is not untertaken here. However, a possible generalisation capturing these data is that topics precede their verb. Here I use the term topic, in the traditional sense, as what the sentence is about, rather than link/locus-of-update. As discussed in Section 2.5.3, there is evidence suggesting that the topic of a sentence is not always the link/locus-of-update. In such a case, it would be reasonable to assume that topics are ordered with respect to their predicate/verb rather than discourse entities.

5.3.5 Discussion

Word order arises from the interaction between constraints originating from different levels of the grammar. In this chapter, I distinguished syntactic constraints from discourse ones. The two differ in various ways. Some of the differences are direct consequences of the level of the grammar they originate from; syntax and Information Structure respectively. Thus, syntactic constraints apply over specific constituents and make reference to syntactic categories (head, complement etc.). Discourse constraints apply over utterances and order new vs. given information irrespective of the syntactic realisation of focus/ground elements. The LPRs in Section 5.3.1 capture the independence of discourse constraints from syntax.

Further, discourse constraints on word order are generally weaker than syntactic ones. Unlike syntactic constraints, they do not induce strong grammaticality judgements and they can be overriden by other constraints (e.g. heavy NP shift). In previous sections, I have described the differences in rigidity between syntactic and non-syntactic/interpretational constraints in an intuitive way. The study of these differences though, raises serious methodological problems. It is unlikely that the familiar way of arguing on the basis of speakers' intuitions is enough to reveal significant differences on the grammaticality judgements induced by the two kinds of constraints. Experimental support of the claims made here is necessary. On the other hand, corpora/dialogue analysis seems of crucial importance for the understanding of the interaction between different kinds of non-syntactic constraints on word order (e.g. heavy NP shift).

Apart from the methodological problem, non-syntactic constraints pose a theoretical question. Most current generative grammar frameworks can express only the syntactic constraints, that is rigid/hard ones. It is not clear how these frameworks can capture the difference between rigid/hard constraints and relative/soft ones and the interaction beteween kinds of the latter.

5.4 Conclusions

In this chapter, I presented an analysis of the discourse-syntax interface for Greek. This analysis captures the independence of Information Structure from syntax and phonology. The Information Structure of Focus-movement, Topicalisation and CLLD were straigthforwardly

accounted for by a small set of general constraints on the realisation of Information Packaging in Greek. That is, no independent assumptions or stipulations were needed to account for the Information Structure of Unbounded Dependencies. In addition, a very small set of constraints was enough to capture the realisation of Information Packaging in languages as different as English, Greek and Catalan.

Further, I argued that the discourse constraints on word order should be captured independently of Phrase Structure positions.

Chapter 6

Concluding Remarks

6.1 Unbounded Dependencies as a syntactic primitive

Focus-movement, Topicalisation and CLLD have been considered three syntactically distinct phenomena. In this thesis I showed that their syntactic properties can be captured by a single syntactic construction/mechanism. This result allowed a reduction in the grammar of Unbounded Dependencies, since the three syntactic operations proposed in the literature were replaced by one construction.

In this study I explored the use of Unbounded Dependencies for discourse purposes and I briefly touched upon their use in Wh-questions. In Greek, Unbounded Dependencies are also central to the formation of Relative clauses and, possibly to Quantifier Raising (e.g. sentences involving *kanenas*='nobody' as in ex. 2.60b). It remains an interesting research question to which extent the Unbounded Dependencies involved in Relative clauses and QR in Greek exhibit the same properties as the ones discussed in this thesis.

Unbounded Dependencies in Greek exhibit two interesting properties: they do not create islands for extraction (also in Italian and Hungarian) and they do not block selection of an embedded clause from a matrix verb (also in Italian). In this respect, they differ from their counterparts in languages like English or Frisian (Iatridou & Kroch 1992). An explanation for these differences could prove crucial to our understanding of Unbounded Dependencies.

Under the view taken in this study, Unbounded Dependencies arise as a syntactic primitive. That is, they are constructions available in the grammar to be employed for various purposes (discourse functions, Wh-questions, Relative clauses etc.). They always involve a

dependent of the verb. For example, such a construction is not available for focusing, topicalising or questioning a verb:

(6.1) a. ????oti **EDIOKSAN** nomize ton Taki
????that fired-**3PL** thought-**3SG** the-**ACC** Takis
'S/he thought they fired Takis.'

b. *oti edioksan nomize ton **TAKI**
*that fired-**3PL** thought-**3SG** the-**ACC** Takis

Since Unbounded Dependencies involve a relation between a dependent of the verb (filler) and the clause containing the verb, they could be viewed as a case of *predication.* That is, focused, topicalised or questioned XPs realise some kind of 'subject'. In relation to this, it is worth mentioning the notion of *structural predication* proposed in Heycock (1994) and Rothstein (1983). Very roughly, these authors distinguish two operations: the *lexical function* and a relation of *predication* which is taken as a structural primitive. The *lexical function* involves the saturation of the arguments of the verb while the structural *predication* involves the relation for example between grammatical subjects in English and the VP.

As a speculation, I would like to suggest the possibility of using the notion of structural predication for Unbounded Dependencies. Very roughly, one can assume two primitive constructions in the grammar. First, a construction involving the saturation of a verbal head corresponding to the *lexical function* of Heycock (1994) and Rothstein (1983) or a *head-comp-phrase* in an HPSG style of grammar which is local in nature. Second, a construction of structural predication between a dependent of the verb and a clause, corresponding to an Unbounded Dependency (*head-filler-phrase*). The predication relation between subjects and VPs in English could then be viewed as a hybrid between the argument saturation construction (*head-comp-phrase/lexical function*) and a predication construction (*head-filler-phrase*). The local properties of the Subject-VP construction in English could be attributed to the fact that this construction is employed for the saturation of the lexical/subcategorisation requirements of the verbs.

Such an approach might prove useful in the understanding of various phenomena. For example, it might explain the idiosyncratic properties of Wh-questions involving subjects in English or multiple subject constructions in Japanese and various Semitic languages which have also been claimed to involve predication constructions (Doron & Heycock 1997).

6.2 Crosslinguistic aspects of the syntax-discourse interface

In this study I explored the way in which Unbounded Dependencies are employed in Greek for the realisation for Information Packaging. The analysis I presented in Chapter 5 differs from *Discourse Configurational* ones in two respects. First, *Discourse Configurational* approaches encode discourse functions in Phrase Structure through distinct Functional Projections. By contrast, I offered an analysis in which Information Structure is represented independently of syntax and phonology. Second, in *Discourse Configurational* analyses, the assignment of discourse functions in Unbounded Dependencies is captured by distinct syntactic mechanisms. These mechanisms do not capture the ground-focus partition of sentences that do not involve Unbounded Dependencies. By contrast, in the analysis presented in this study, the Information Structure of Unbounded Dependencies follows from general constraints on the syntactic and phonological realisation of ground and focus in Greek.

The differences between *Discourse Configurational* analyses and the approach proposed in this thesis have implications for the predictions made with respect to the crosslinguistic organisation of the grammar. According to *Discourse Configurational* analyses, the difference between Greek and English is one of Phrase Structure. Greek clauses have a more complex Phrase Structure since they contain two more functional projections, FP and TP. Under the view adopted in this thesis, English and Greek do not differ radically in their grammar. Both grammars involve Unbounded Dependencies and have, by and large, the same amount of constructions, since the clause structure of Greek is not extended with additional configurations. Further, both grammars impose essentially the same constraints on the realisation of Information Structure. No additional constraints were needed to account for the Information Structure of Unbounded Dependencies in Greek. The two languages differ in the degree of exploiting Unbounded Dependencies for discourse purposes. Greek makes extensive use of these constructions whereas English does not. Thus, this analysis allows a consistent and elegant view of the discourse-syntax interface for languages as different as Greek and English.

Bibliography

AGOURAKI, YEORYIA. 1992. Clitic Left Dislocation and Clitic Doubling: a unification. *U.C.L. Working Papers in Linguistics* 4.44–68.

——, 1993. *Spec-Head Licensing: The Scope of the Theory*. University College London dissertation.

ALEXIADOU, ARTEMIS. 1997. *Adverb Placement: a case study in antisymmetric syntax*, volume 18 of *Linguistik Aktuell*. John Benjamins Publishing Company.

ALEXOPOULOU, THEODORA. 1997. A discourse-based account of Weak Crossover Effect. In *Proceedings of the Second ESSLLI Student Session*, ed. by Alice Drewery, Geert-Jan M. Kruijff, & Richard Zuber. ESSLLI'97 Summerschool, Aix-en-Provence.

——. 1998a. Detaching Discourse Functions from Functional Projections. In *U. Penn Working Papers in Linguistics*, ed. by to appear. Proceedings of the 22d Annual Penn Linguistics Colloquium.

——. 1998b. Unbounded dependencies and the syntactic realisation of Information Packaging. In *Proceedings of the FHCG-98, Joint Conference on Formal Grammar, Head-Driven Phrase Structure Grammar and Categorial Grammar*, ed. by Gosse Bouma, Geert-Jan M. Kruijff, & Richard T. Oehrle. Saarbrücken (Germany).

ANAGNOSTOPOULOU, ELENA, 1994. *Clitic Dependencies in Modern Greek*. University of Salzburg dissertation.

BALARI, SERGIO. in press. Pronous, variables and extraction in HPSG. In *Romance in HPSG*, ed. by Sergio Balari & Luca Dini. CSLI Publications.

BARD, ELLEN GURMAN, DAN ROBERTSON, & ANTONELLA SORACE. 1996. Magnitude estimation of linguistic acceptability. *Language* 72.32–68.

BOLINGER, DWIGHT L. 1978. Asking more than one thing at a time. In *Questions*, ed. by H. Hiz, 87–106. Dordrecht: Reidel.

—— 1989. *Intonation and its uses: melody in grammar and discourse*. Stanford: Stanford Universitly Press.

BORSLEY, ROBERT D. 1996. *Modern Phrase Structure Grammar*. Blackwell.

BOUMA, GOSSE, ROBERT MALOUF, & IVAN SAG, 1997. Satisfying Constraints on Extraction and Adjunction. ftp://csli-ftp.stanford.edu/linguistics/sag/ext-adj-draft.ps.gz.

BRESNAN, JOAN, & S. MCHOMBO. 1987. Topic, Pronoun and Agreement in Chichewa. *Language* 63.741–778.

BRODY, MICHAEL. 1990. Remarks on the Order of Elements in the Hungarian Focus Field. In *Approaches to Hungarian:Structures and Arguments*, ed. by I. Kenesei, volume 3, 207–243. JATE. Szeged.

CANN, RONNIE. 1993. *Formal Semantics*. Cambridge University Press.

CATSIMALI, GEORGIA, 1990. *Case in Modern Greek: Implications for Clause Structure*. University of Reading dissertation.

CHAFE, W.L. 1976. Givenness, contrastiveness, definiteness, subjects, topics and point of view. In *Subject and Topic*. New York: Academic Press.

—— 1983. *Meaning and the Structure of Language*. Chicago: University of Chicago Press.

CHIERCHIA, G. 1992. Functional Wh and Weak Crossover. In *The Proceedings of the Tenth West Coast Conference on Formal Linguistics*, 75–90.

CHOI, HYE-WON, 1996. *Optimizing structure in context: scrambling and Information Structure*. Stanford University dissertation.

CHOMSKY, NOAM. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.

——. 1995. Bare phrase structure. In *Government and Binding Theory and the Minimalist Program*, ed. by G. Webelhuth, 381–439. Oxford: Blackwell.

——. 1996. *The Minimalist Program*. MIT Press.

CHUNG, CHAN. 1998. Case, Obliqueness and Linearization in korean. In *Proceedings of the FHCG-98, Joint Conference on Formal Grammar, Head-Driven Phrase Structure Grammar and Categorial Grammar*, ed. by Gosse Bouma, Geert-Jan M. Kruijff, & Richard T. Oehrle. Saarbrücken (Germany).

CINQUE, GUGLIELMO. 1990. *Types of A-bar Dependencies*, volume 17. MIT Press. Linguistic Inquiry Monographs.

CRUTTENDEN, A. 1986. *Intonation*. Cambridge University Press.

CRYSTAL, D. 1991. *A dictionary of linguistics and phonetics*. Blackwell.

CULICOVER, P., & M. ROCHEMONT. 1983. Stress and Focus in English. *Language* 59.

DAVIS, ANTHONY, 1997. *Lexical Semantics and Linking and the Hierarchical Lexicon*. Stanford University dissertation.

DIMITRIADIS, ALEXIS. 1994a. Clitics and island-insensitive object-drop. In *Studies in the Linguistic Sciences (Proceedings of FLSM-5)*, volume 24. Urbana-Champaign, IL.

——. 1994b. Clitics and object drop in Modern Greek. In *Proceedings of SCIL-6*, volume 23. MITWPL.

DOBROVIE-SORIN, CARMEN. 1990. Clitic Doubling, Wh-Movement, and Quantification in Romanian. *Linguistic Inquiry* 21.351–397.

DORON, EDIT, & CAROLINE HEYCOCK, 1997. Filling and Licensing Multiple Specifiers. ms.

ECKERT, MIRIAM, 1998a. *Discourse Deixis and Null Anaphora in German*. University of Edinburgh dissertation.

——. 1998b. The German Topic Position and Null Anaphora. In *Proceedings of the 2nd Colloquium on Discourse Anaphora and Anaphor Resolution*, ed. by Simon Botley & Tony McEnery.

ENGDAHL, ELISABET. 1985. Parasitic Gaps, Resumtpive Pronouns, and Subject Extractions. *Linguistics* 23.3–44.

——, & ENRIC VALLDUVÍ. 1996. Information packaging in HPSG. In *Edinburgh Working Papers in Cognitive Science, Vol. 12: Studies in HPSG*, ed. by Claire Grover & Enric Vallduví, chapter 1, 1–32. Scotland: Centre for Cognitive Science, University of Edinburgh.

ERTESCHIK-SHIR, NOMI. 1986. Wh-questions and Focus. *Linguistics and Philosophy* 9.117–149.

GAZDAR, GERALD. 1981. Unbounded dependencies and coordinate structure. *Linguistic Inquiry* 12.155–84.

GIANNAKIDOU, ANASTASIA, 1997. *The Landscape of Polarity Items*. University of Croningen dissertation.

GILL, K.H., in preparation. *Argument Structure and Long-Disatance Binding: Evidence from Korean*. University of Edinburgh dissertation.

GINZBURG, JONATHAN, 1992. *Questions, Queries, and Facts: A Semantics and Pragmatics for Interrogatives*. Stanford University dissertation.

GUNDEL, JEANETTE K. 1988. *Studies in Syntactic Typology*, chapter Universals of topic-comment structure, 209–239. Amsterdam and Philadelphia: Eenjamins.

——, NANCY HEDBERG, & RON ZACHARSKI. 1993. Cognitive status and the form of referring expressions. *Language* 69.274–307.

HAEGEMAN, LILIANE. 1991. *Introduction to Goverment and Binding Theory*. Blackwell.

HALLIDAY, MICHAEL A.K. 1967. Notes on transitivity and theme in English, part ii. *Journal of Linguistics* 3.199–244.

HEIM, IRENE, 1982. *The semantics of definite and indefinite noun phrases*. University of Massachusetts,Amherst dissertation.

——. 1983. *Meaning, Use and Interpretation of Language*, chapter File Change Semantics and the Familiarity Theory of Definiteness. Berlin and New York: de Gruyter.

HENDRIKS, HERMAN, & PAUL DEKKER. 1995. Links without locations. In *Proceedings of the Tenth Amsterdam Colloquium*, ed. by P.Dekker & M.Stokhof, University of Amsterdam. ILLC.

HEYCOCK, CAROLINE. 1993. Focus Projection in Japanese. In *Proceedings of NELS*, ed. by Merceí González, volume 24, 159–187.

——. 1994. *Layers of Predication: The Non-Lexical Syntax of Clauses*. New York & London: Garland.

HOFFMAN, BERYL. 1995. Intergrating "Free" Word Order Syntax and Information. In *Proceedings of the 1995 Conference for the European Association for Computational Linguistics*.

HOJI, H., 1985. *Logical Form Constraints and Configurational Structures in Japanese*. University of Washington dissertation.

HORROCKS, GEOFFREY. 1983. The order of constituents in Modern Greek. In *Order, concord and constituency*, ed. by Gerald Gazdar, Ewan Klein, & Geoffrey K. Pullum, 95–11. Dordrecht:Foris.

——. 1994. Subjects and configurationality:Modern Greek clause structure. *Journal of Linguistics* 30.81–109.

HORVATH, JULIA. 1995. Structural focus, structural case, and the notion of feature-assignment. In *Discourse Configurational Languages*, ed. by K.E. Kiss, Oxford Studies in Comparative Syntax, 28–64. Oxford University Press.

HUKARI, THOMAS E., & ROBERT D. LEVINE. 1995. Adjunct extraction. *Journal of Linguistics* 2.195–226.

IATRIDOU, SABINE. 1995. Clitics and Island Effects. *Penn Working Papers in Linguistics* 2.11–31.

——, & ANTONY KROCH. 1992. The Licensing of CP-recursion and its Relevance to the Germanic Verb-Second Phenomenon. *ms* .

JAEGGLI, O. 1982. *Topics in Romance Syntax*. Dortrecht: Foris.

KALLULLI, D., 1997. Clitic Doubling in Albanian and Greek. University of Durham.

KATHOL, ANDREAS, 1995. *Linearization-Based German Syntax*. Ohio State University dissertation.

——, & CARL POLLARD. 1995. On the Left Periphery of German Subordinate Clauses. *Proceedings of the Fourteenth West Coast Conference on Formal Linguistics* 14.11–31.

KELLER, FRANK. 1998. Gradient grammaticality as an effect of selective constraint re-ranking. In *Papers from the 34th Annual Meeting of the Chicago Linguistic Society*, volume 2: The Panels, Chicago.

——, & MARIA LAPATA. 1998. Object drop and discourse accessibility. In *Proceedings of the 17th West Coast Conference on Formal Linguistics*, ed. by Susan J. Blake, Eun-Sook Kim, & Kimary N. Shahin, Stanford, CA. CSLI Publications.

KING, T.H. 1995. *Configuring Topic and Focus in Russian*. CSLI Publications.

KISS, KATALIN É. (ed.) 1995a. *Discourse Configurational Languages*. Oxford University Press.

KISS, KATALIN E. 1995b. NP Movement, Operator Movement and Scrambling in Hungarian. In *Discourse Configurational Languages*, ed. by K.E. Kiss, Oxford Studies in Comparative Syntax, 207–243. Oxford University Press.

KISS, KATALIN É. 1998. Identificational focus versus information focus. *Language* 74.245–273.

KOLLIAKOU, DIMITRA, 1991. A Formal Account of Clitic Doubling Depedencies in Modern Greek. Master's thesis, University of Edinburgh.

——, 1995. *Definites and Possessives in Modern Greek: An HPSG Syntax for Noun Phrases*. University of Edinburgh dissertation.

——. 1997. Possessive affixes and complement composition. In *Computational Linguistics in the Netherlands 1996*, ed. by Landsbergen et al. IPO: Eindhoven.

——. 1998a. A composition approach to Modern Greek 'weak form' possessives. In *Lexical and Constructional Aspects of Linguistic Explanation*, ed. by Webelhuth et al. CSLI: Stanford.

——. 1998b. Linkhood and multiple definite marking. In *Proceedings of the FHCG-98, Joint Conference on Formal Grammar, Head-Driven Phrase Structure Grammar and Categorial Grammar*, ed. by Gosse Bouma, Geert-Jan M. Kruijff, & Richard T. Oehrle. Saarbrücken (Germany).

KRIFKA, MANFRED. 1991. A compositional semantics for multiple focus constructions. In *Proceedings from Semantics and Linguistics*, ed. by Steven Moore & Adam Zachary Wyner, volume 10, 127–158.

KROEGER, PAUL. 1994. *Phrase Structure and Grammatical Relations in Tagalog*. Dissertations in Linguistics. CSLI Publications.

KUHN, JONAS, 1997. Domain restriction in quantification is independent of focus marking-a DRT account of the use of adverbial quantifiers in partial answers to questions. ms.

LADD, ROBERT, AMALIA ARVANITI, & INEKE MENNEN, 1998. Wh-questions in Greek. in preperation.

LADD, ROBERT D. 1980. *The structure of intonational meaning: Evidence from English*. Bloomington: Indiana University Press.

—— 1996. *Intonational Phonology*. Cambridge University Press.

LARSON, RICHARD K., & GABRIEL SEGAL. 1995. *Knowledge of Meaning*. MIT Press.

LASCARATOU, CHRYSSOULA. 1989. *A Functional Approach to Consituent Order with particular reference to Modern Greek*, volume 5. Athens: Parousia.

LASNIK, HOWARD, & TIM STOWELL. 1991. Weakest Crossover. *Linguistic Inquiry* 22.687–720.

MANANDHAR, SURESH. 1994. Encoding Information Packaging in HPSG. In *Integrating information structure into constraint-based and categorial approaches*, ed. by Elisabeth Engdahl. Amsterdam: ILLC. DYANA-2 Report R.1.3.B.

MANNING, CHRISTOPHER D., & IVAN A. SAG, 1995. Dissociations between argument structure and grammatical relations. Unpubl. ms., Carnegie Mellon University and Stanford University.

MARANTZ, A. 1995. The minimalist program. In *Government and Binding Theory and the Minimalist Program*, 348–82. Oxford: Blackwell.

MAY, R. 1985. *Logical Form: Its Structure and Interpretation*. Cambridge,Mass: MIT Press.

MCCLOSKEY, JAMES. 1979. *Transformational Syntax and Model Theoretic Semantics*. Reidel: Dordrecht.

——. 1989. Resumptive pronouns, A-bar-binding, and levels of representations in Irish. In *Syntax and Semantics 29: The sytax of the modern Celtic languages*, ed. by Randall Hendrick, 199–248. Academin Press, New York.

MILLER, PHILIP H. 1992. *Clitics and Constituents in Phrase Structure Grammar*. New York: Garland. Published version of 1991 Doctoral dissertation, University of Utrecht, The Netherlands.

MONACHESI, PAOLA, 1995. *A grammar of Italian clitics*. University of Tilburg dissertation. ITK dissertation series,vol.3.

NIÑO, MARIA E., 1994. The Semantics of Topic NPs. Stanford University,ms.

PARTEE, BARBARA. 1991. Topic, Focus and Quantification. In *Proceedings from Semantics and Linguistics*, ed. by Steven Moore & Adam Zachary Wyner, volume 10, p. pages. Cornell Working Papers in Linguistics.

PHILIPPAKI-WARBURTON, IRENE. 1982. Problimata shetika me ti sira ton oron stis ellnvikes protasis. *Glossologia* 1.99–107.

——. 1985. Word order in Modern Greek. *Transactions of the Philological Society* 113–143.

POLLARD, CARL, ROBERT KASPER, & ROBERT LEVINE, 1994. Studies in Constituent Ordering: Toward a Theory of Linerization in Head-Driven Phrase Structure Grammar. Research Proposal to the National Science Foundation, Department of Linguistics and Center for Cognitive Science; Ohio State University.

——, & IVAN A. SAG. 1987. *Information-based Syntax and Semantics*. Stanford: CSLI.

——, & ——. 1994. *Head-Driven Phrase Structure Grammar*. University of Chicago Press, Chicago.

——, & EUN JUNG YOO, 1996. A unified theory of scope for quantifiers and wh-phrases. ms.

POSTAL, P. 1993. Remarks on Weak Crossover Effect. *Linguistic Inquiry* 24.539–556.

PRINCE, ELLEN F. 1986. On the syntactic marking of presupposed open propositions. In *CLS (Parasession)*, volume 22, 208–222.

RADFORD, ANDREW. 1997. *Syntactic Theory and the Structure of English: a mimimalist approach*. Cambridge University Press.

REAPE, MIKE. 1994. Domain Union and Word Order Variation in German. In *German in Head-Driven Phrase Structure Grammar*, ed. by C.Pollard J.Nerbonne, K.Netter, volume 46 of *Lecture Notes*, 151–198. CSLI.

REINHART, TANYA. 1982. Pragmatics and Linguistics: an Analysis of Sentence Topics. *Indiana University Linguistics Club* .

RIZZI, LUIGI, 1995. The Fine Structure of the Left Periphery. University of Geneva, ms.

ROOTH, MATS. 1996. Focus. In *The Handbook of Contemporary Semantic Theory*, ed. by Lappin Shalom, 271–297. Blackwell.

ROTHSTEIN, SUSAN, 1983. *The Syntactic Forms of Predication*. MIT dissertation.

SAG, IVAN A. 1997. English relative clause constructions. *Journal of Linguistics* 33.431–484.

——, & JANET D. FODOR. 1994. Extraction without traces. In *Proceedings of the Thirteenth West Coast Conference on Formal Linguistics*, ed. by Raul Aranovich, William Byrne, Susanne Preuss, & Martha Senturia, volume 13, Stanford University. CSLI Publications/SLA.

——, & PHILIP H. MILLER, 1997. French Clitic Movement without Clitics or Movement. ms.

SANFILIPPO, ANTONIO, 1990. *Grammatical Relations, Thematic Roles and Verb Semantics*. University of Edinburgh dissertation.

——. 1996. Thematically Bound Adjuncts. In *HPSG in Romance*, ed. by L. Balari, S. & Dini, 359–395. Stanford University.

SCHNEIDER-ZIOGA, P, 1994. *The Syntax of Clitic Doubling in Modern Greek*. University of Southern California dissertation.

STEEDMAN, MARK. 1991. Structure and Intonation. *Language* 67.260–296.

——, 1998. Information Structure and the Syntax-Phonology Interface. http://www.cogsci.ed.ac.uk/ steedman/papers.html. ms.

TSIMPLI, MARIA IANTHI. 1995. Focusing in Modern Greek. In *Discourse Configurational Languages*, ed. by K.E. Kiss, Oxford Studies in Comparative Syntax, 176–206. Oxford University Press.

——, 1996. Focus, Negation and Quantification in Greek. Cambridge University.

TSIPLAKOU, STAVROULA, 1998. *Focus in Greek: its structure and interpretation*. University of London, School of Oriental and Africal Studies dissertation.

TZANIDAKI, DIMITRA IRINI, 1994. *The syntax and pragmatics of subject and object position in Modern Greek*. University College London dissertation.

USZKOREIT, HANS. 1986. Constraints on order. *Linguistics* 24.883–906.

VALIOULI, MARIA. 1993. Post verbal subjects, reduplicated objects or retopicalisation. In *Themes in Greek Linguistics: papers from the 1st international conference in Greek Linguistics*, ed. by Irene Philippaki-Warburton & al., volume 117 of *Current Issues in Linguistic Theory*, 85–92. John Benjamins Publications.

——. 1994. Anaphora, agreement, and right dislocation in Greek. *Journal of Semantics* 11.55–82.

VALLDUVÍ, ENRIC. 1992. *The Informational Component*. Garland,New York.

——. 1995. Structural Properties of Information Packaging in Catalan. In *Discourse Configurational Languages*, ed. by K.E. Kiss, Oxford Studies in Comparative Syntax, 122–152. Oxford University Press.

——, & ELISABET ENGDAHL. 1996. The Linguistic Realization of Information Packaging. *Linguistics* 34.459–519.

——, & MARIA VILKUNA. 1997. *The limits of syntax*, chapter On rheme and kontrast. NY: Academic Press.

——, & RON ZACHARSKI. 1994. Accenting phenomena, association with focus, and the recursiveness of focus-ground. In *Proceedings of the Ninth Amsterdam Colloquium*, ed. by P.Dekker & M.Stokhof, 683–702. Amsterdam:ILLC.

VILKUNA, MARIA. 1995. Discourse configurationality in Finnish. In *Discourse Configurational Languages*, ed. by K.E. Kiss, Oxford Studies in Comparative Syntax, 244–268. Oxford University Press.

WECHSLER, STEPHEN. 1995. *The Semantic Basis of Argument Structure: A Study of the Relationship Between Word Meaning and Syntax*. Stanford University: CSLI Publications.

ZACHARSKI, RON, 1993. *Discourse Pragmatics Model of English Accent*. University of Minnesota dissertation.