



GEORGETOWN UNIVERSITY
GRADUATE SCHOOL OF ARTS AND SCIENCES

This is to certify that we have examined the doctoral dissertation of

Jorge López-Cortina, M.A.

entitled The Spanish Left Periphery: Questions and Answers

submitted to the faculty of Spanish and Portuguese

in partial fulfillment of the requirements for the degree of

Doctor of Philosophy.

This dissertation is complete and satisfactory in all respects, and any and all revisions required by the final examining committee have been made.

Héctor Campos, Ph.D.

Thesis Advisor

Signature

Date

Raffaella Zanuttini, Ph.D.

Committee Member

Signature

Date

Paul Hagstrom, Ph.D.

Ronald P. Leow, Ph.D.

Acting Director of Graduate Studies

Signature

Date

This dissertation has been accepted by the Graduate School of Arts and Sciences.

For the Dean

Date

THE SPANISH LEFT PERIPHERY:
QUESTIONS AND ANSWERS

A Dissertation
submitted to the Faculty of the
Graduate School of Arts and Sciences
of Georgetown University
in partial fulfillment of the requirements for the
degree of
Doctor of Philosophy
in Spanish Linguistics

By

Jorge López-Cortina, M.A.

Washington, DC
December 7, 2007

Copyright 2007 by Jorge López-Cortina
All Rights Reserved

**THE SPANISH LEFT PERIPHERY:
QUESTIONS AND ANSWERS**

Jorge López-Cortina, M.A.

Thesis Advisor: Héctor Campos, Ph.D.

ABSTRACT

Based on the seminal work of Rizzi (1997) and others, this dissertation explores the left periphery of Spanish sentences, with particular attention to interrogative structures and their replies.

The central claim of this work is that a sentence used as an answer has a syntactic structure specific to this function. Not every declarative sentence can be used felicitously as a reply to a question. This dissertation provides a syntax-based account of the possible structures allowed as answers to particular questions. I have argued for the differentiation of answers from identificational focus, on the grounds of the non-contrastive character of answers (including fragment answers) and their different positions in the sentence. This differentiation not only allows us to better understand the structure of answers, but also provides a new perspective for the study of the Focus layer and the left periphery as a whole.

An additional element introduced in the discussion is the Confirmation Phrase, presented here as an extended projection of the phrase requiring confirmation, which is

visible through a wh-morpheme (*qué* in Spanish). This projection allows for some syntax-based fine-tuning of the expression of the meaning of questions, distinguishing information questions from confirmation ones.

The proposal put forward in this work is that the Focus layer in the left periphery of the sentence hosts functional projections that host separate features, of which one, [+answer], hosted in Answer Phrase (AnsP), determines the character of answer of a sentence, and another, [+wh], hosted in WhP, triggers wh-movement. These features can additionally combine with a Force feature [+Q] to produce complex interrogative structures. This structure gives a unified account of regular wh-questions and split questions (interrogative structures including a proposed answer that appears as a tag linked to the question by a particular intonation). It also links split questions to the structure of sentences used as answers, by virtue of the presence of an Answer Phrase and a [+answer] feature in both structures. These elements of the analysis, together with the confirmation feature hosted in ConfP, provide tools to start a comparative analysis of the semantics of different types of questions.

ACKNOWLEDGEMENTS

This dissertation comes as the final result of a long period of my life devoted to learning, in a broad sense of the word. As such, it is but a poor reflection of all the things I have been taught and all the help and support I have received. I can only hope that in the life that I now begin, I will be able to show some of the generosity, patience, and humanity that the people I have met during these years have shown to me.

First of all I would like to thank the members of my committee. My advisor, Héctor Campos, has shown me the kind of intellectual rigor and perseverance it takes to build a solid analysis, together with the enthusiasm for language that can bring students of Syntax I to look forward to every meeting. I thank him for his dedication and patience. I am very grateful to Raffaella Zanuttini for believing in the promise of my first ideas and joining my committee in spite of knowing how illegible my first drafts could be. I always left my meetings with her with ideas that would open new directions for my analyses. And I owe thanks to Paul Hagstrom for always helping me see the big picture and linking my initial ideas to greater concepts. His flexibility and commitment to reading and commenting my work have always been a great help and encouraged me to go on.

The Department of Spanish and Portuguese and the Department of Linguistics at Georgetown have given me the opportunity to study under excellent professors. I am thankful to all of them, and particularly to Ron Leow for his continuous encouragement. Taking a course with Mario Vargas Llosa was definitely a privilege and a highlight of my graduate work, even if it had little to do with my dissertation.

During my time at Georgetown I received institutional support in the form of a University Fellowship, which made it possible for me to devote plenty of time to reading and pursuing scholarly interests beyond the Spanish Linguistics program. I would like to thank Associate Dean Jim Schaefer for, among other things, always reminding me that my job was to get out of Georgetown (with a degree), and Father Solomon Sara for knowing how to provide a word of support at exactly the right moment.

Among the experiences Georgetown provided me with was the chance to work with Congressman Bob Filner. For three years I had the privilege of discussing the news in Spanish with someone who really knew what was going on. I also thank Tony Buckles for his help while I was working there.

Among my professors at UMass, I owe special gratitude to Daphne Patai, who was instrumental not only in convincing me that I could put together an application that would get me accepted in Georgetown, but also in making me believe that I deserved it. My mentor there was Juan Zamora, who, besides Spanish Dialectology, taught me during our lunches everything there is to know about getting a job in American academia. The devotion to teaching and the patience of Kyle Johnson were also very important in my decision of continuing graduate studies. I also would like to thank Sigrid Beck, Ellen Woolford, Barbara Partee, and very specially Roger Higgins.

My interest in linguistics started in Oviedo, where I was first introduced to structuralist ideas by Emilio Alarcos Llorach and some of his disciples. Francisco García gave me my first teaching job there.

The patience and support of my colleagues and friends at Seton Hall have been a major source of strength in finalizing this dissertation. I would like to thank Daniel Zalacaín, Anne Mullen-Hohl, Gabriella Romani, and Matt Escobar for never doubting I would finish it, even when I was not so sure myself. And I would like to thank Matt also for the many conversations on grammar and for showing me that split questions might have some interest for people other than theoretical linguists. I would also like to thank Robert Kingsbury, my editor at Berlitz, for his patience and understanding.

Thanks are due to Nina Moreno, Inma Taboada, Pedro Pérez Leal and all others who spent time providing me with grammatical judgments and data about their dialects. Carlos de Cuba, Paco Fernández-Rubiera, Ángel Gallego, Francesc Roca, Guillermo Lorenzo, Eugenia Casielles, Luis Alonso-Ovalle, Jake Cáceres, and Joaquim Kuong Io-Kei helped me give shape to my ideas at different points. I appreciate their generosity. Much of this dissertation is owed to the patience, generosity, and continuous support of Susana Huidobro.

Maite Camblor, Lucas Izquierdo, Kaylea Mayer, Irina Feldman, Neal Dahan, Raúl Burneo, Haizam Amirah-Fernández, and Heshima James helped me in different ways during the difficult times, and also helped me enjoy the good ones. My warmest thanks go to all of them.

There are people whose support and friendship go well beyond what I can express with words. Paula Menéndez Benito was who first made me think of coming to the United States to study linguistics, and still is an inspiration for me to continue studying. Nate May and his wonderful family, Billy Mikades, David Soria, Alberto Pérez Pereiro, and Melissa Bowles have always provided me with a home when I felt I was far from mine. Andrea Lluch has been there for me when I needed it most, and I would not have arrived here if not for her.

My parents Sara and Jorge, and my brothers Pablo and Sergio, have supported me both as family and as occasional linguistic informants. I owe them many things beyond this dissertation.

In memoriam

Luis López López

Constantino Cortina García

However minute the employment may appear, and whatever ridicule may be incurred by a solemn deliberation upon accents and pauses, it is certain that without this petty knowledge no man can be a poet.

Dr. Samuel Johnson, *Rambler* 88

TABLE OF CONTENTS

INTRODUCTION	1
CHAPTER ONE REVIEW OF THE LITERATURE ON INFORMATION STRUCTURE: PRAGMATICS AND PHONOLOGY	4
1 Introduction	4
2 Topic and Focus. The interface between syntax and pragmatics	6
2.1 Approaches to the study of the interface	6
2.2 Delimiting information structure	8
2.3 Historical approaches	16
2.3.1 Vallduví's Informatics	18
2.4 On the notion of information	20
2.4.1 Information structure can follow from syntax	23
3 Topic and Focus. The interface between syntax and prosody	28
3.1 Underlying assumptions	28
3.1.1 Marked and unmarked Topic and Focus	29
3.1.2 Topic-Focus asymmetries. Towards a more restrictive notion of focus...	32
3.1.3 Towards a comprehensive theory	33
3.2 Underlying assumptions	35
3.2.1 Stress, intonation, and Focus	35
3.2.2 Is prosody enough?	39
3.2.3 Intonation types	51
3.3 Interface problems	53
3.3.1 The T-model issue	53
3.3.2 Focus and the T-model	55
3.4 Types of Focus: A first overview	56
4 Conclusions.....	59

CHAPTER TWO	REVIEW OF THE LITERATURE ON INFORMATION STRUCTURE: SYNTAX	60
1	Introduction	60
2	Characterization of Topic and Focus – The syntactic approach.....	62
	2.1 Types of focus	66
	2.1.1 É. Kiss’s distinction between identificational and information Focus	66
	2.1.2 Brunetti’s attempt at unification	70
	2.1.3 Domínguez’s classification for Spanish	79
3	The left periphery	80
	3.1 Rizzi’s initial proposal	82
	3.2 Benincà and Poletto’s refinements	85
4	Beyond the left periphery	90
	4.1 Kayne’s framework	92
	4.2 The “center periphery”	96
	4.2.1 Jayaseelan’s study	96
	4.2.2 Villalba’s proposal for Catalan	100
	4.2.3 Poletto’s view on DP-internal Focus	111
	4.3 Focus sightings	112
	4.3.1 Italian	112
	4.3.2 Somali	113
	4.3.3 Old Italian	114
	4.3.4 American Sign Language	115
5	Conclusions	120
	Appendix: Villalba’s classification of Catalan constructions.....	127

CHAPTER THREE	FOCUS, ANSWERS, AND QUESTIONS	132
1	Introduction.....	132
2	Focus and Interrogation as separate phenomena in the structure of answers	137
	2.1 Previous data on answers.....	139
	2.1.1 Hungarian data.....	139
	2.1.2 Italian data	143
	2.1.3 Spanish data	148
	2.2 Towards a description of (Northwestern and Central Peninsular) Spanish answers.....	154
	2.2.1 An initial classification of Spanish answers	155
	2.2.2 Fragment answers	176
	2.3 Are answers syntactic objects <i>qua</i> answers?	187
	2.3.1 Answer Phrases.....	190
	2.3.2 The structure of yes/no answers	204
	2.3.3 The hypothesis of the placement of answers in the VP periphery.....	219
	2.4 Conclusions.....	222
3	Focus and Interrogation as separate phenomena in the structure of questions.....	224
	3.1 Examples of non-correspondence between focus and wh-elements.....	225
	3.2 Conclusions.....	229
4	The semantics of questions and focus	230
	4.1 Questions as answer sets.....	230
	4.2 Focus and scope	232
	4.3 Conclusions.....	229
5	Chapter summary and conclusions	238

CHAPTER FOUR	SPLIT INTERROGATIVES	240
1	Introduction.....	240
2	Description of the construction.....	243
	2.1 Py (1971)	243
	2.2 Vígara Tauste (1992).....	247
	2.3 The syntax of split questions	250
	2.4 Follow-up questions.....	265
	2.5 Summary	270
3	Previous analyses.....	271
	3.1 Uriagereka (1988).....	272
	3.2 Lorenzo (1994)	275
	3.3 Camacho (2002).....	280
	3.4 López-Cortina (2003)	283
4	A revised analysis.....	290
	4.1 General considerations.....	290
	4.1.1 Achievements and shortcomings of previous analyses	291
	4.1.2 What is left to address.....	293
	4.1.3 The left periphery of the sentence	294
	4.2 The Focus field in Spanish	296
	4.3 The nature of <i>qué</i> and its relationship with the tag.....	301
	4.3.1 Movement of <i>qué</i>	302
	4.3.2 Position of <i>qué</i> in the structure	305
	4.3.3 There are no embedded split questions.....	314
	4.3.4 Long distance movement.....	318
	4.3.5 Negative split questions.....	323
	4.3.6 Section summary	326
	4.4 Multiple elements on the right.....	329

4.5 Dialectal variation.....	333
4.6 IP split questions.....	336
4.7 Elements to the left of the interrogative	339
4.8 Loose gerunds	346
4.9 Follow-up questions.....	347
5 The meaning of split questions	352
6 Conclusions	358
CHAPTER FIVE CONCLUSIONS AND ISSUES FOR FURTHER RESEARCH.....	361
1 General conclusions.....	361
2 Issues for further research.....	363
2.1The VP periphery.....	363
2.2 Phases	364
2.3 Question tags	365
2.4 Constructions related to split questions	366
2.5 Intonation.....	366
REFERENCES.....	368

INTRODUCTION

It has long been observed that word order, and in particular the placement of phrases towards the beginning of a sentence (its *left periphery*), has an effect on the appropriateness of that sentence in different contexts. The description and systematization of such effects has occupied linguists of all schools for over a century now. One fruitful approach has been the description of the structure of the left periphery in syntactic terms, associating the informational value of a sentence in the discourse to syntactic operations. Thus, even though many questions remain open, the notions of Focus and Topic have been integrated in syntax, linking grammar to information structure.

This dissertation takes a step forward in that line of work and explores the left periphery of Spanish interrogative constructions from the point of view of syntax. In doing that, it seeks to determine to what extent a syntactic analysis suffices to account for the particularities of structures that combine an interrogative context with other discourse functions. The study of questions is naturally complemented by the study of the structure of answers. Both questions and answers have been associated to different conceptions of focus in previous literature, most frequently by the identification of wh-movement and focus movement, and by the observation that the relevant element of an

answer must be in a relevant position within the sentence. In this work, evidence is presented that answers have a structure differentiated from sentences including focus. This is first shown by presenting a detailed summary of the structures acceptable as answers to wh-questions. The comparison of these structures leads to a hypothesis of the general structure of answers, which is subsequently confirmed by evidence from the structure of affirmative and negative answers to yes/no questions.

The syntactic operations associated to answers can also appear in other types of sentences, thus providing a syntactic account for structures that combine a question with an answer. An instance of these structures that is studied in depth is the *split question*, a combination of a question that would be ungrammatical on its own with a proposed answer, held together by a particular intonational contour, such as Spanish *¿Qué vas, a Oviedo?* ‘Where are you going, to Oviedo?’ or English *What are you, kidding?*

The dissertation is divided in five chapters. The first one considers the necessity of a relationship between information structure and syntax by reviewing the relevant literature and describing the limitations of pragmatic and phonological approaches. The second chapter reviews syntactic approaches to information structure in order to establish a framework within which the left periphery of interrogatives could be

discussed. The view is adopted (following Rizzi, 1997 and others) that part of the phenomena related to information structure is encoded in syntax in the form of functional projections in the left periphery. Taking this as a starting point, chapter 3 addresses the relationship between focus and interrogative sentences, concluding that it is not possible to assume identity between identificational focus and *wh-* in either an interrogative sentence or a sentence that is presented as an answer to an interrogative. Additionally, it is observed that answers have a special status, different from sentences involving focus, and also different from non-answer declarative sentences. On the basis of these observations, I will propose that the focus area of the left periphery of the sentence (the “Focus field” (Benincà & Poletto, 2004)) consists of at least three separate functional nodes, (identificational) Focus, Answer, and Wh. The goal of the fourth chapter is to complete this description of the left periphery of the sentence and show its usefulness in providing an explanation for constructions, the split questions mentioned above, whose analysis requires a detailed mapping of the functional projections involved. A fifth chapter summarizes the findings of the previous ones and discusses possible lines of future research.

CHAPTER 1 – REVIEW OF THE LITERATURE ON INFORMATION STRUCTURE: PRAGMATICS AND PHONOLOGY.

1 Introduction

There are grammatical phenomena whose meaning can not be easily described in terms of the semantics of a proposition. They have traditionally been described in terms of how they relate to the informational status of phrases within a sentence. This relationship with information, a hearer-based concept, has led to their classification as pragmatic issues or as an independent level of grammar. The label *information structure* is generally accepted for those phenomena since the early 1990s.

In this chapter I will review the literature on information structure. I will show how, in my view, sufficient arguments have been made to take the position 1) that certain information structure phenomena (namely, focus and topic) are encoded in grammar, since they determine parts of the grammatical structure of a sentence and they interact with pragmatic principles in the same way any other syntactic or phonological elements of the sentence do, and 2) that a better description of such phenomena is achieved when they are analyzed as independent syntactic operations rather than as a global phenomenon codified in a level of grammar separated from syntax.

The issue of the relationship between these phenomena and suprasegmental phonology presents a challenge to current grammatical theories. However, it can be shown that most phonological phenomena related to information structure are on a one-to-one relationship with syntactic phenomena. The cases where phonology is the only cue for the identification of information-related meaning are briefly discussed, but a complete analysis of such cases falls outside the scope of this dissertation.

The chapter is structured as follows: Section 2 approaches the issue of information structure from the point of view of pragmatics. In section 2.1, I give an overview of the issue, and, following mostly Vallduví (1990), in section 2.2 I address the uniqueness of information packaging (as the tradition he follows calls the set of phenomena related to topic and focus) among the issues studied within pragmatics. In section 2.3 I give a brief overview of historical approaches, including Vallduví's. Section 2.4 presents my view on information structure and why I am of the opinion that it is not an issue that can be analyzed outside of syntax. Section 3 is devoted to the phonological issues involved in the description of information structure. In section 3.1 I go over the basics of the notions of topic and focus. In section 3.2 I address their correlates in prosody. Section 3.3 explores the consequences for the theory of grammar of the interaction between information structure and prosody. Section 3.4 addresses the nature of focus and its relationship to syntax. I close this chapter with some partial conclusions.

2 Topic and Focus. The interface between syntax and pragmatics

2.1 Approaches to the study of the interface

There are a number of discourse issues that influence the shape of the sentence, which can be described in terms of informational relevance of sentence elements. That relevance can be expressed via intonational or word-order means. The puzzle presented to us is to figure out how many separate discourse functions there are, how to classify them, which ones are related to syntax, how they are expressed, and how discourse functions and means of expression interact¹.

There are two opposite possible theoretical approaches to these questions. One is to assume that an independent component of grammar is needed in order to account for discourse issues. This is Vallduví's (1990) approach. Following Chafe's information packaging (Chafe, 1976; Prince, 1986), Vallduví proposes Information Structure to be an additional level of representation, in the sense that LF and PF are levels of representation in pre-minimalist generative grammar. Vallduví does away with the T-model of grammar (Chomsky, 1981, 1986; Chomsky & Lasnik, 1977), but he does not explore the theoretical consequences of this. Rather, he uses the freedom provided by

¹ The extent to which these things can vary across languages is also an important part of the puzzle. I will, however, not address this issue at length, and largely limit my discussion to discourse functions generally assumed to be equivalent (or at least similar) in the languages discussed.

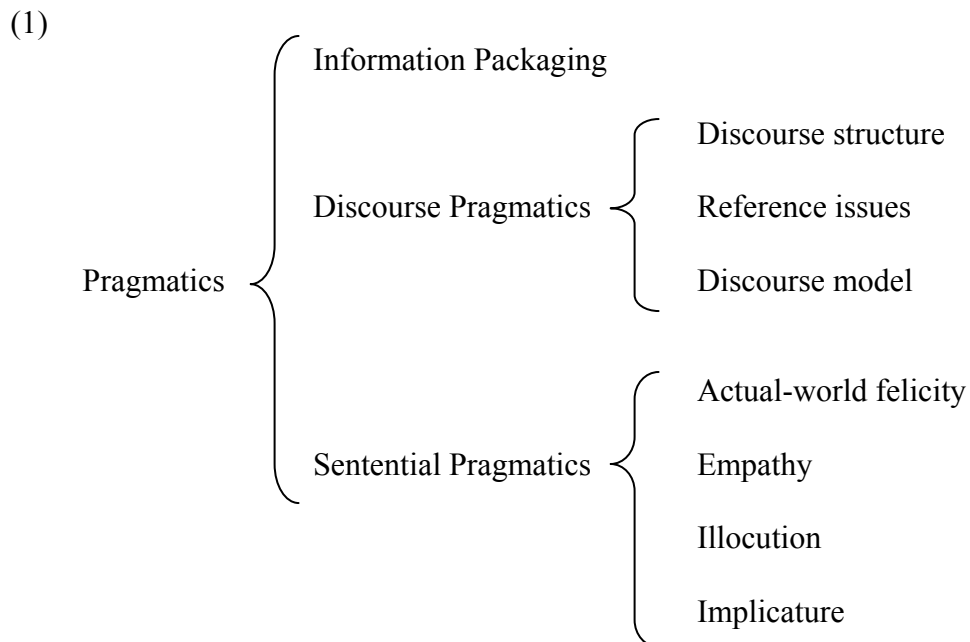
his independent module to address a number of interesting descriptive issues, for instance the different possible positions of subjects in Catalan.

The other approach, i.e. merging the informational component within the existing modules of grammar, presents a more restrictive framework for informational phenomena and, conversely, opens a new testing ground for commonly accepted generalizations, particularly in the syntax of movement.

I give an overview of the syntactic approach in chapter 2. This chapter addresses approaches from the points of view of pragmatics and phonology. In the rest of this section, I will discuss the place of information structure within pragmatics.

2.2 Delimiting information structure

Pragmatics has been portrayed as a wastebasket (Bar-Hillel, 1971) due to the number of issues that are classified as pertaining to pragmatics when other fields will not account for them. Vallduví (1990) presents a classification of the subfields contained within pragmatics that I summarize in the following chart:



I will not transcribe here the ample bibliography provided by Vallduví for each subfield, but I will provide an overview of how his classification works. Information packaging refers to Chafe's (1976) idea that some linguistic phenomena "have to do primarily with how the message is sent and only secondarily with the message itself".

The concept of information Vallduví uses is “the part of knowledge represented in a given proposition that is unknown to the hearer”(Vallduví, 1990). His concept of information packaging follows:

- (2) INFORMATION PACKAGING: a small set of instructions with which the hearer is instructed by the speaker to retrieve the information carried by the sentence and enter it into her/his knowledge store. (Vallduví, 1990, p. 66)

Note that there are other ways to understand information. A more general use of the term does not limit it to new information. Paul Hagstrom (p.c.) suggests a more general definition of information in terms of possible world semantics: A very basic unit of information would be a fact or a proposition, i.e., a specification of what the world is like (this is generally modeled by taking a proposition to be a function from possible worlds or states of affairs to truth values, such that the proposition can determine, based on the facts of each world, whether it is true in that world). The reason I transcribe Vallduví’s definitions here is that it will be necessary to understand his view of how information structure relates to previous literature where the idea of ‘new to the hearer’ is a basic concept.

Vallduví goes through the effort of developing his classification in order to clearly separate his concept of information structure from other pragmatic issues. I summarize that classification here in order to draw the limits of the phenomena I intend

to study, and give a sense of why they have been connected to pragmatics. While I agree with Vallduví's view that information structure can not be viewed as a purely pragmatic phenomenon (Brunetti, 2003 makes that case about Focus in particular), I will follow the view (adopted implicitly or explicitly in much other work, e.g. Rizzi 1997 and subsequent work) that information structure could be completely codified in syntax, and thus it would not constitute a separate, dedicated level of grammar. In other words, I will adopt the view that there are no operations specific to information structure that happens to the representation beyond what happens in the syntax. Section 3 of this chapter will deal with its relationship with phonology.

In Vallduví's classification, discourse pragmatics refers to phenomena that pertain to the linguistic context beyond the sentence level. Sentential pragmatics refers to "nonlinguistic contextual phenomena that affect the structure of the sentence regardless of linguistic content"(Vallduví, 1990, p. 19).

Within discourse pragmatics, discourse structure includes discourse segmentation and organization. Reference issues are those related to familiarity and saliency of the linguistic entities with respect to the hearer. Vallduví (1990, pp. 22-23) shows (departing from ideas by Rochemont, 1986) how information packaging (particularly the notions of focus and topic) is not determined by hearer familiarity or discourse saliency. Information packaging is a relational notion within the boundaries of

a sentence, while familiarity and saliency are absolute concepts, independent from the sentential context in which a linguistic entity appears. The last of Vallduví's subfields of discourse pragmatics, the discourse model, is more an approach to discourse description than a subset of phenomena. Discourse models are mental representations of the discourse which include the entities involved, their attributes and the links between them.

As for sentential pragmatics, with the term *actual-world felicity* Vallduví refers to the fact that otherwise well-formed sentences may not be felicitous in situations where the actual world conditions that would make the sentence true are never or rarely present. Take, for instance, Searle's (Searle, 1989) example:

(3) # I hereby fry an egg.

This sentence is not felicitous because in our world (or any worlds where egg-frying is the same as in ours) the conditions under which *hereby* can be used to perform a speech act are not met. (Those conditions would be met if eggs were fried by merely saying it.)

Empathy, or the identification of the speaker with a discourse participant, is unrelated to information packaging as defined above. Illocution "has to do with what kind of speech act is performed when a sentence is uttered" (Vallduví, 1990, p. 30)

(commanding, requesting, denying, etc.) and is not related to information packaging either. Finally, implicature occupies a central position in pragmatics since Grice (1975). Vallduví explains the difference between implicature and information packaging from the understanding that implicatures have to be obtained from logico-semantic structures, rather than from uninterpreted structures, while his assumption is that information packaging is interpreted “along with, not after, logico-semantic meaning”. This view has implications for the semantics of information structure, of course. It amounts to an assertion that information structure has an impact on the truth value of a sentence². I will come back to this issue in section 2.4.

Vallduví’s classification of pragmatic issues is a great starting point to determine the nature of focus and topic phenomena. Unlike most of the issues pragmatics deals with, the nature of focus and topic is to establish a relationship between elements within the sentence boundaries, not to carry over some sort of meaning from one part of the discourse to another. As such, topic and focus are completely encoded within the message, and they are not to be confused with their interpretations.

In order to make this clear we can give a provisional definition of focus and topic within Vallduví’s concept of information packaging. Focus would be the part of

² A summary of the impact of Focus phenomena (which are part of information structure) in truth values is discussed in chapter 3 of this dissertation. See also Rooth (1996) for an overview of these issues.

the sentence that presents information purported to be new to the hearer or in contrast with previous beliefs of the hearer. Topic would be the part of the sentence presented as the area of the hearer knowledge over which the new information has scope. Both focus and topic can be marked. When marked by being associated to the Nuclear Pitch Accent (NPA) of the sentence, most studies transcribe it using CAPITALS. Topics can be marked with a raising pitch accent, the *B-accent* (Bolinger, 1965; Jackendoff, 1972), which will be represented here using a raising slash (/) or simply by using a comma when no ambiguous reading is possible, as in traditional orthography. I address Jackendoff's description of the use of accents in more detail in section 3.2.3 of this chapter, and I also provide a summary of notational conventions at the end of this chapter, once the phenomena they describe have been characterized.

With these definitions of topic and focus in mind, we can assert that the focus of example (4a) and the topic of (4b) is *the cicadas*, regardless of whether we know anything about either the linguistic context of the sentence or the actual life-cycle of cicadas.

- (4) a. THE CICADAS died in June.
b. The cicadas/, I think they come out every seventeen years.

That is to say, the “instructions” mentioned in the definition of information packaging are there, regardless of any other considerations.

Conversely, the felicity of a sentence is a pragmatic issue, but there is no significant difference between infelicity brought up by the focalization of a particular element and infelicity brought up by lexical choice.

To see this, take the examples in (5), where all four responses can be infelicitous:

- (5) Did John see the cicadas?
- a. John has no windows.
 - b. Mary failed the exam.
 - c. I think they will never find any weapons of mass destruction in Iraq.
 - d. THE CICADAS John saw.

Grice (1975) defined a Cooperative Principle that calls on a speaker to "make your conversational contribution such as is required, at the stage at which it occurs, by the accepted purpose or direction of the talk exchange in which you are engaged" (Grice, 1989, p. 26). Note that if we assume that the speaker is cooperating we can always find ways to make (5a, b, c) felicitous. (5a) could imply that John never leaves his room and he could only have seen the cicadas from his window, if he had one. In (5b) Mary's failing would have something to do with John seeing the cicadas. Maybe Mary is John's

sister, she has to stay home to study and therefore John's family can not go on vacation to any cicada-infested areas. Even something completely unrelated, as (5c), can be made felicitous by interpreting that who is uttering it means to tell us that he does not want to talk about John, or the cicadas. The reason (5d) appears *prima facie* infelicitous is that the focus is a phrase that the hearer just referred to in her question, therefore making it part of the background. My point here is that (5d) is no more infelicitous than any of the other responses; the cooperative principle can also apply here, and there are some interpretations that make the focalization of *the cicadas* plausible, for instance cases where the speaker wants to point out that John should have been paying attention to something else. For instance, if John had just visited Washington DC for the first time, (5d) might be understood as exhaustive, i.e., as implying that the only thing that John saw in DC was the cicadas, and he ignored the museums or other attractions. In other words, the contrast expressed by focus could be interpreted to refer not to "cicadas" versus "museums and politicians" (which would be infelicitous given the question (5d) is answering) but rather to "only cicadas" versus "museums, cicadas, and politicians".

Vallduví is interested in isolating information packaging within pragmatics. In doing so he shows how information packaging is a separate linguistic device with pragmatic effects, which makes it no different from syntactic or lexical devices. The theoretical question now is whether there is an explanation for information packaging

within phonology or syntax, or it actually constitutes a separate level of grammar, but first a better characterization of the phenomena is needed. I will start with a brief review of traditional approaches.

2.3 Historical approaches

Vallduví (1990) provides a great historical overview of how the organization of information in the sentence has been described. Further clarification on the concepts of given and new information is provided by Ellen Prince (Prince, 1981). Here I will just give a brief sketch following Vallduví's work.

Theme-Rheme

The dichotomy of theme and rheme (Ammann, 1928 and many others afterwards) captures the contrast between a less informative, vehicular, part of the sentence, and a more informative one.

Topic-Comment

This approach articulates the sentence in two parts around the notion of “aboutness” (Mathesius, 1915); one part, the topic, presents what the speaker wants to talk about, and the other, the comment, what the speaker is saying about it (Hockett, 1958).

Topic-Focus

This approach articulates the sentence in terms of contextual boundedness v. contextual freeness, where contextually bound means “accessible in the hearer’s memory, i.e. salient, activated over a certain threshold in the stock of shared knowledge” (Hajicová, 1984, p. 193). The focus is the contextually free part of the sentence, and it can comprise any part of the sentence or the complete sentence. The topic is the contextually bound part of the sentence. (Sgall, Hajicová, & Panevová, 1986; Stechow, 1981).

Focus-presupposition

Generative linguistics took the term *focus* from Halliday (1967), who used it to designate the informative subset of the rheme, and used it to refer to the logical

complement to the presupposition that “denote[s] the information in the sentence that is assumed by the speaker to be shared by him and the hearer” (Jackendoff, 1972, p. 230).

Dominance

The concept of dominance differs from the other approaches above in that it does not determine a bipartite sentence structure. Instead, dominance is understood as a discourse property assigned to a particular constituent, roughly covering the cases regarded as focus in other approaches, with the difference that presupposed material is not necessarily excluded from the dominant constituent. This approach was developed by Erteschik-Shir (1973 and subsequent work).

2.3.1. Vallduví’s Informatics

The proposal by Vallduví (1990) seeks to combine the insights provided by both the focus-oriented and the topic-oriented approaches. He gives priority to the focus-oriented approach, where the focus “constitutes the only informative part of the sentence”. The rest of the sentence constitutes the *ground*, or background information. Within this background information we can find an optional, sentence-initial, topic-like element, the *link*. The rest of the *ground*, once the *link* is excluded, is called the *tail*.

Vallduví uses a hierarchical articulation in order to avoid the redundancy that would be caused by the stipulation of two different sentence divisions operating independently. He seems to give more relevance to the focus due to the lack of consistency of the presence of the link, that can be missing from the sentence, while, in other cases, a sentence can have two links. It is also possible to find all-focus sentences, where all the information is presented as new or at least without an informative contrast.

This concept of bipartite division with core elements within is picked up by Steedman (2000 for an overview), who adds an internal, intonationally marked, element within the focus, parallel to Vallduví's division within the ground between link and tail. In this work I will assume that the complexity of information structure is due to the interaction of several independent systems. Vallduví merges topic and focus into a same system, where they are opposite sides of one single articulation. The reason Vallduví chooses to do this (not unlike most of the traditional views he in other aspects surpasses) is to avoid redundancy, that he sees, for instance, in proposals such as Dahl's (1974), who includes the concepts of topic and comment as well as background and focus. An example of how Dahl's approach would be the analysis of the example in (6).

(6)	What does John drink?	topic	comment
		John	drinks beer
		background	focus

I expect to be able to show that the redundancy Vallduví vows to avoid is nothing but a mirage due to the overlap of descriptions undertaken from different theoretical and empirical grounds. In other words, there is no need to reconcile both bipartite divisions into a more articulate tripartite division; it is sufficient to accept that the operations underlying each division can appear simultaneously and independently.

2.4 On the notion of information

Vallduví (1990) takes Dretske's (1981) idea that information "is that commodity capable of yielding knowledge, and what information a signal carries is what we can learn from it" (Dretske, 1981, p. 44). Vallduví adapts this idea in the following way: "two sentences with the same propositional content may carry different information in different speaker-hearer interactions depending on how much of that propositional content is unknown by the hearer at the time of the utterance" (Vallduví, 1990, p. 14). For Vallduví, the information in a sentence is that part of the propositional content that is not previously known by the hearer. He mentions other authors (Dahl, 1976; Lyons, 1977) "that characterize information as a hearer-based contribution to sentence production and understanding" (Vallduví, 1990, p. 17).

Vallduví formalizes the idea of information as hearer-based contribution by adapting Heim's File Change Semantics (Heim, 1983), where discourse referents can be represented as file cards in a file (see also Reinhart, 1981 for an earlier approach in the same vein). The number of file cards and the contents of each of them are updated as new discourse is interpreted. Vallduví proposes to eliminate redundant entries in the file cards by interpreting the information packaging of a sentence as instructions for file updating. I repeat his definition here:

- (7) INFORMATION PACKAGING: A small set of instructions with which the hearer is instructed by the speaker to retrieve the information carried by the sentence and enter it into her/his knowledge-store. (Vallduví, 1990, p. 66)

In Vallduví's variation of File Change Semantics, the new discourse offers clues to the hearer as for what file card should be updated with the new knowledge, while the rest of discourse referents can be updated at a later stage or simply get a crossreference index. Thus, Vallduví presents a less redundant "data entry" system while capturing the intuition that information packaging serves the purpose of giving the hearer instructions about how to organize her knowledge.

The concepts of hearer-based contribution and speaker-provided instructions contradict each other to some extent. If, as Vallduví states, "Information, then, unlike

propositional content, is crucially defined with respect to the particular hearer a sentence is addressed to.” (Vallduví, 1990, p. 16), there is little reason to study information structure as an inherent property of a message. If information depends, at least partially, on the hearer and not the message, it should be expected that we study the ways different hearers break up an informatively neutral, content-oriented message in pieces of informative and less informative content. Using Chomsky’s terms, this would be the task of a theory of performance. Vallduví, however, does not work on a theory of performance. Instead, the effort of Vallduví’s work is focused in describing and analyzing the ways information is encoded within the message. Then, the hearer would be breaking up the message in order to find structural clues to identify information. These structural clues would be part of the Structural Descriptions (SDs) generated by an I-language, and would clearly fall within the scope of a theory of competence (see Chomsky, 1995).

There is no doubt that the psychological description of information packaging as a set of instructions makes more sense than assuming that there is inherent informational value to particular pieces of content, or, even more complicated, that informational value is attached to particular grammatical functions, such as the subject. It can therefore be argued that Vallduví’s work takes the correct approach. However, this approach clashes with the hearer-based concept of information expressed above. Indeed, the study of the ‘set of instructions’ suggests that there is an implicit assumption

in Vallduví's work that a) there are procedures to mark what parts of the sentence the speaker (not necessarily the hearer) understands to be informative to the hearer, and b) that such procedures are part of the knowledge of the language shared by speaker and hearer. That is to say, they are part of their linguistic competence. If we are to study the informational status of parts of the sentence in terms of encoding procedures, then the informational status of a part of a sentence is the content attached to it by each of the encoding procedures, such as word order alterations. Thus, the expression of informational status is part of the I-language and encoded in the SD.

Once this point of view is assumed there are two issues to study. The first one is what kind of content is encoded (the labels of focus and topic mentioned above would be examples of such content), which I will address mostly in chapters 2 and 3, and the second issue is where the encoding of information structure takes within the grammar of a language. The rest of this chapter deals with this issue.

2.4.1 Information structure can follow from syntax

Studying the transmission of information and the interactions between the hearer and the speaker can undoubtedly shed light on many aspects of language. However, for the purposes of this work, I will pursue an explanation to information structure phenomena (to the extent they involve word order) entirely within the bounds of syntax,

as the term is understood in the generative grammar tradition. In this, I follow the approach taken by, e.g., Rizzi (1997) and others. It is perhaps worth noting that there are other proposals outside of the generative approach. In an early attempt to study so-called free word order the Prague school of linguistics developed explanations for Czech data that are based on the newness of the referents involved (Sgall, Hajicová, & Benesova, 1973; Sgall, Hajicová, & Panevová, 1986).

I have shown above (see (4) and (5)) that the set of procedures labeled ‘information packaging’ are perfectly visible in an isolated utterance. It is not necessary to observe their effect on the hearer, or the context of the utterance, to describe them. Their meaning does not change according to the previous knowledge of the hearer. If the term ‘information’ needs to be used, it should be understood in the sense of ‘part of the discourse that the speaker understands to be informative’, regardless of the actual knowledge held by the hearer. Pragmatics deals with those aspects of linguistic communication that can be described only when the context of the utterance and its effect on the hearer are taken into account. The phenomena Vallduví labels as information packaging do not fall within pragmatics because they can be described taking the message in isolation. Information packaging takes place during the production of the message and it should be analyzed within that context.

The main contribution of Vallduví is the separation of information structure from pragmatics. He proposes a new level of linguistic representation, Informatics, which should account for information packaging within grammar, while still maintaining it separated from syntactic rules. There are other such proposals, like Erteschik-Shir's f-structure (Erteschik-Shir, 1997), placed right before Spell Out. In the rest of this chapter I will discuss whether syntax is actually sufficient to account for information packaging and therefore there is no need to postulate a separate level.

The issue under discussion here is not the merit of any of the approaches to the issue from a discourse analysis point of view. They all have some merit and can be used to clarify some points. What interests me is a) whether those analyses reflect some grammatical property or they do not (in which case they would be just useful ways to describe the manifestations of some general properties of messages or communication) and b) whether that grammatical property would belong in syntax, phonology, or both of them.

To illustrate these questions, I will reproduce an example from Vallduví (1990), who in turn takes it from Sgall et al. (1986).

(8) Beavers build DAMS.

If we adopt the Topic-Focus approach, and we define focus as the salient part of the sentence, the focus in (8) is three-way ambiguous between *dams*, *build dams*, and *beavers build dams* (compare *I was surprised to learn that dams are what beavers build*; *I was surprised to learn that building dams is what beavers do*; and *I was surprised to learn that beavers build dams*.) The sentence is ambiguous to the hearer, but not to the speaker, as long as he knows what he means to say. However, there is nothing that the speaker has put in the sentence to make it possible to disambiguate it. This kind of ambiguity phenomenon was called Focus projection by Selkirk (1984; , 1995), and, according to Brunetti (2003, p. 37), it was first noticed by Chomsky (1971), who pointed out how the Focus of a sentence can extend to larger constituents than the one immediately containing the stressed word. Vallduví points out the incompatibility, in cases like (8), of the Topic-Focus approach and the Topic-Comment articulation. In the Topic-Focus approach, as we have seen, *beavers* can be part of the Focus and therefore not a Topic, while the Topic-Comment articulation (built around the notion of “aboutness”) would see *beavers* as the sentence Topic. I believe, however, that this incompatibility is only apparent, due to terminological confusion. Beyond the obvious need to avoid using the term *topic* to refer to two different things, there is no reason to think that one of these approaches excludes the other. The actual question is whether, and how, either one of these discourse articulations reflects a grammatical property, and what the relationship between them is, if there is any. In my view, when we say that a sentential element is included in the Focus by means of Focus projection we are

referring to a different kind of Focus than in those cases where movement, or other syntactic procedure, is involved. But in order to be able to differentiate between different kinds of Foci, it is necessary first to describe the syntactic procedures associated to it. A very important breakthrough in this respect is Rizzi's (1997) article, where different informational articulations correlate with different, dedicated, functional heads in the left periphery of the sentence. I will get back to this approach in chapter 2.

Global approaches to information structure that are based on a single dichotomy fail to capture the complexity of the issues involved, and I work on the (ultimately structuralist) hypothesis, developed to some extent by Rizzi (1997), that each of those dichotomies correlates with a single linguistic feature. The next step, then, should be to look at what linguistic features have been shown to have an influence in information structure. My goal will be to present the different lines of research and the way they interact, with particular attention to syntax, since it seems to be the field that is providing the most detailed results. This should allow for a clearer view of the compatibilities and incompatibilities between results obtained in different fields of research. It will be necessary to classify different syntactic procedures, identify isolated information structure phenomena, point at those issues that require further research, and show the limitations of current phonological, syntactic and semantic theories. Of course, the availability of a particular interpretation or construction in universal grammar does not mean that it should surface in every language. Ideally, had we a complete catalogue

of semantic possibilities, or possible interpretations, as well as phonological and syntactic procedures available in universal grammar, we should find that the semantic possibilities are available in all languages, but not necessarily that they are visible (in the sense of formally distinguished from other interpretations through a syntactic or phonological procedure) in every language. It would then be possible to characterize different languages by looking at which possibilities are visible, and how they are expressed.

3 Topic and Focus. The interface between syntax and prosody

3.1 Underlying assumptions

Adopting a grammar-based approach to the treatment of information should allow us to abandon some of the assumptions that are attached to the issue of information structure when it is seen from the point of view of pragmatics. In section 3.1.1 I discuss the problems of assuming that an account of information structure must give an information-related label to every part of the sentence. In 3.1.2 I address the idea that marked Focus might not be the same thing as unmarked Focus, and therefore no explanation should be expected to cover both phenomena.

3.1.1 Marked and unmarked Topic and Focus

Going over the section on historical approaches, 2.3, it is easy to see how, in spite of the diversity of approaches, there is some agreement on that sentence content can be classified in two main categories from the perspective of information theory. These two categories can appear by default, i.e. be pragmatically inferred from the context of the utterance, or they can be explicitly encoded in the grammar of the sentence, i.e., marked.

Most approaches assume a single marked category (say Focus, Topic...) that divides the sentence in two mutually exclusive parts, labeling the non-marked category (Topic for Focus, Comment for Topic...) by default, and ignoring the possibility of marking both. Such approaches can never cover data more than partially.

A comprehensive approach would require us to assume a) that the element of the sentence intended as new information equals the Focus of the sentence, b) that the background information equals its Topic, and c) that there is nothing else in the sentence. Such an approach would lead to two problems. The first problem is that such a view would imply the impossibility of configurations where there is an overlap. As we saw in the previous section, such overlaps occur. In the example in (6), copied in (9) for convenience, it can be seen that both articulations don't match exactly, and there can be

elements, the verb in this case, that are part of the comment and of the background at the same time.

(9)	What does John drink?	topic	comment
		John	drinks beer
		background	focus

The second problem is the exact opposite, i.e., the lack of exhaustivity. The contrast between marked and unmarked categories is still unaccounted for. In cases where both Topic and Focus are marked, there will likely be something else left out. Take, for instance, (10):

- (10) a. Cicadas, I believe they spend SEVENTEEN YEARS underground.
- b. As for time underground, cicadas spend seventeen years there.
- c. As for surprising behaviors, cicadas spend seventeen years underground.

If the Topic of (10a) is *cicadas*, and its Focus is *seventeen years*, there is no way to determine whether *underground* is part of the Topic or the Focus. We have to resort to conversational context. In a conversation about animals that spend a lot of time underground, (10a) can be paraphrased as (10b) and *underground* would be a part of the Topic (Note that due to its position in the sentence *underground* would bear the default nuclear stress. This does not seem to cause it to be automatically interpreted as focus, in

spite of the usual association between focus and nuclear stress. The reason why nuclear stress and focus are disassociated in this case is the presence of a marked stress associated to the constituent *seventeen years*. The stress carried by the last syllable of the sentence is an automatic result of the end of sentence intonation.) In a conversation about strange habits of insects, however, (10c) would be a more appropriate paraphrase; *underground* would be new information, and therefore part of the Focus.

Vallduví (1990) addresses these problems by assuming that everything that is not clearly Focus is *background*, and there might be a marked part of the background, the *link*. This prevents overlaps (or, in his terms, *redundancy*), but I do not think is sufficient to describe the different usages of sentences like (10a). In fact, it could be argued that this approach precludes the (10c) interpretation, since *underground* in (10a) would be part of the background. Steedman's (2000) view of Focus and Topic is more symmetrical, with optional marked elements inside both of them. Such an approach obviously limits the usefulness of the analysis to sentences with two clearly divided parts.

In summary, descriptions based on a marked Focus or Topic leave out cases in which both seem to be marked. Descriptions based on both Focus and Topic being marked leave out cases of overlap and do not account for cases where additional elements appear in the sentence. Descriptions where Focus and Topic have a marked

core and can expand to include any additional elements leave out sentences where a two-part division of the sentence is not possible. This should suggest that attempting to comprise every element of the sentence in an information-based division might not be a good approach.

3.1.2 Topic-Focus asymmetries. Towards a more restrictive notion of Focus

At this point, two asymmetries between Topic and Focus must be noted. The first one is that Focus can be marked in situ and it would appear that the only way to mark a Topic is to have it placed at the beginning of the sentence (regardless of whether it got there by means of movement or was base-generated). I will leave the different ways Focus is marked for section 3.4 of this chapter and chapter 2. The second asymmetry is that Focus seems to always be present in a sentence (Vallduví, 1990, p. 80) while a Topic is optional.

In order to look at the second asymmetry, let us provisionally define information Focus as non-contrastive new information (this concept is introduced by É. Kiss, 1998). I will discuss the types and definitions of Focus later on, but for now let us assume that marked Focus has a contrastive value and we can take the non-contrastive Focus to be simply informational, in accordance with what was said in section 2. The Focus that a sentence always needs to have is this information Focus. Simply put, in order to say

something, *something* has to be said. I take this fact to be not a part of grammar, but rather a general property of communication. From an interpretive point of view, the Cooperative Principle (Grice, 1975) prompts us to look for, and find, a reason for a given sentence to have been uttered. For a sentence to be felicitous, it is expected to provide new information or to express a point of view. In other words, no sentence can be interpreted felicitously unless a part (or the whole) of it is interpreted as its Focus (whether it is syntactically or prosodically marked or not). The question of why Focus (in its broader, informational definition) is always present in a sentence is thus rendered moot³.

3.1.3 Towards a comprehensive theory

While there is agreement on the idea that the informational status of a given element of a sentence can be defined in terms of the sentence interpretation on the part of the hearer, there are also ways for the speaker to mark elements of the sentence

³ Paul Hagstrom (p.c.) suggests that there can be sentences without any Focus. Consider his example: — *As we all know, John left yesterday. —Indeed. John left yesterday.* He suggests that the second utterance seems to communicate something like “I agree, I share your knowledge of the truth of the proposition that John left yesterday.” I believe that the second utterance can be either an echo, where the Focus is exactly the same as in the first utterance, or an invitation to the hearer to reconsider the importance of John’s departure, in which case we would have the same focus for different reasons. My view is that as long as the Cooperative Principle prompts the hearer to wonder why the second utterance takes place, the hearer will find an information Focus. The point here is that accounts of information structure that emphasize the second asymmetry between Topic and Focus, i.e. that Focus is always present in the sentence, are using a very broad concept of Focus. Thus, even if Hagstrom’s observation is correct and there are sentences lacking both contrastive and information Foci, this would only support the idea that a more strictly defined Focus does not always have to be present in the sentence.

according to their intended discourse relevance, or informational status. Attempts to describe the way the discourse is interpreted in terms of information try to be exhaustive, i.e. to include every element of the sentence in their analysis. If we expect the same exhaustiveness from descriptions that pay attention to marked elements, we always find problems to explain the role of non-marked elements.

The sentence in (8) (*Beavers build dams*) is an example of how purely informational notions of Focus carry an implicit ambiguity that can not be solved without knowledge of the context, and do not have any syntactical or phonological properties associated to them. Also, É. Kiss (1998) and others show that contrastive Focus carries some presuppositions that cannot be cancelled without an explicit negation, while any pragmatic value carried by information Focus can be cancelled much more easily. This suggests that information Focus is not directly related to the presence and scope of contrastive Focus. I leave a more detailed presentation of É. Kiss' ideas for chapter 2.

The association between speaker-based, grammatically marked, contrastive Focus and hearer-based, unmarked, informational Focus comes naturally when the pragmatic value of a sentence is analyzed. If we try to describe the contribution to sentence meaning made by marked Focus we will probably use many of the terms used to describe the role of informational Focus. However, marked Focus is a grammatical

phenomenon, and unmarked Focus is not. This dissertation is concerned with grammatically encoded phenomena, and therefore not with informational Focus.

Section 2 of this chapter was dedicated to showing how information structure relates to pragmatics and how information-related grammatical phenomena can be separated from pragmatic effects. A first overview of the grammatical phenomena themselves will constitute the next chapter, but for now, in section 3.2, I will go over the relationship between prosody and information structure with the same idea of separating grammatical phenomena from their not grammatically encoded correlates.

3.2 The interface between syntax and prosody

3.2.1 Stress, intonation, and Focus

For the following discussion I draw from class notes by Gardent (1997) and Krifka (2001-2002), as well as Kadmon's (2001) book.

Theories about the relationship between stress and focus depend on the interpretation of stress. Here I will assume that stress is a rhythmic phenomenon (Householder, 1957; see Kadmon, 2001 for a summary; Liberman, 1975; Liberman &

Prince, 1977), independently of what acoustic phenomenon correlates with it. I will also assume Pierrehumbert's (Beckman & Pierrehumbert, 1986; Pierrehumbert, 1980) theory of intonation. A partial description of Spanish is available within the same framework (Beckman, Díaz Campos, McGory, & Morgan, 2002), but I will not discuss it here.

3.2.1.1 The Nuclear Stress Rule

It is generally accepted that intonational events, or accents, are linked to rhythmic prominence. There is disagreement about how. There is a view of stress not having a phonological existence independent of intonation (Bolinger, 1958 and subsequent work) and the opposite (Ladd, 1980; Selkirk, 1984). I will leave this issue aside. From the point of view of focus placement the question is whether stress rules can determine the position of focus. A purely prosodic account of focus was introduced in the early seventies (Chomsky, 1971; Jackendoff, 1972) but, according to Reinhart (1995, p. 55, quoted in Brunetti 2003), abandoned in favor of a semantic approach that saw focused elements undergoing Quantifier Raising (Chomsky, 1976). The purely prosodic approach reappears with Cinque (1993). He reformulates the Nuclear Stress Rule (Chomsky & Halle, 1968; Halle & Vergnaud, 1987) and proposes a relationship between the focus mark and the level of embedding within a clause.

(11) Nuclear Stress Rule (Chomsky & Halle, 1968, p. 90)

$$V \rightarrow [1 \text{ stress}] / [\# \# X \left[\overline{1 \text{ stress}} \right] Y \# \#]$$

where Y contains no vowel with the feature [1 stress]

(12) Nuclear Stress Rule (English) (Halle & Vergnaud, 1987)

- a. Parameter settings on line N ($N \geq 3$) are [-BND, +HT, right].
- b. Interpret boundaries of syntactic constituents composed of two or more stressed words as metrical boundaries.
- c. Locate the heads of line N constituents on line $N + 1$.

The English language parameters: -BND stands for not bound, +HT for having a terminal head, and *right* for the position of the head.

(13) Cinque's Null Theory

- a. Interpret boundaries of syntactic constituents as metrical boundaries.
- b. Locate the heads of line N constituents on line $N + 1$.
- c. Each rule applies to a maximal string containing no internal boundaries.
- d. An asterisk on line N must correspond to an asterisk on line $N - 1$.

The references to asterisks reflect the metrical grid notation (see, for instance, Hayes, 1995 for an explanation of the notation). An example of how Cinque's rule works would be the following grid (Brunetti, 2003):

- (14) *
- (. . *)
- (. (. *))
- (* (* (*)))

The aim behind Cinque’s modification is to avoid involving all the parameters present in the NSR. He proposes that “there is a relation between the direction of syntactic branching of a language and stress placement” (Brunetti, 2003, p. 24) Thus, the grid in the example above takes its shape due to the direction of the branching, and not due to a separate parameter specifying the position of the head, as interpreted by Halle & Vergnaud. The immediate consequence is that the Focus mark corresponds to the most embedded element in a clause. While this idea seems relatively attractive for Italian, English contrasts like the following are difficult to explain (Brunetti, 2003, p. 27):

(15) a. What did John do?

b. John LEFT.

(16) a. Who left?

b. JOHN left.

The approach is also problematic for Italian once constructions with elements in the left periphery of the sentence are included in the picture. Focus preposing (also known as Yiddish Movement) in English presents the same problem. See the following examples from Ward (1985) quoted by Domínguez (2004):

- (17) a. SIX DOLLARS it costs.
b. Butter! She's dreaming about butter! RECIPES she dreams while the world zips...

The Nuclear Stress Rule works well as a predictor of focus placement if it is applied to the focused phrase only, not to the entire sentence (Brunetti, 2003, Aritz Irurtzun, p.c.). The position of focus is not automatically determined by intonation or stress placement. The placement of the pitch accent seems, on the other hand, to follow structural accent rules within the focused phrase. Thus, "A stretch of focused elements, which doesn't necessarily form a syntactic constituent, forms a prosodic domain, which, as a unit, is marked by one accent" (Büring, forthcoming, p. 10. See this article also for an overview of the issue and additional bibliography).

3.2.2 Is prosody enough?

The facts that I have discussed so far, from pragmatics and from prosody, seem to support the idea that there is a type of Focus, informative Focus, which is a mirage that results from the interaction of sentence pragmatics and meaning-blind phonological rules. In chapter 2 I will discuss how contrastive focus can be shown to be a syntactic phenomenon. The opposite view, i.e. assuming a direct cause-effect relationship

between the informational/semantic status of a phrase and its intonational characteristics, particularly stress placement, is well represented in the literature (Neeleman & Reinhart, 1998; Reinhart, 1995; Zubizarreta, 1998) and will be discussed in this section. In section 3.2.2.1. I discuss the approach that complements the NSR with additional phonological rules. In section 3.2.2.2. I discuss the approach, presented in Zubizarreta's work, that the NSR is also complemented with movement of phrases, i.e. that there is a syntactic component of focus placement. For Zubizarreta, the motivation for this movement comes from prosody constraints, and she talks about prosodically motivated movement.

3.2.2.1 Marked rules. Stress strengthening and destressing.

The fact that Cinque's NSR is not sufficient by itself to explain the relationship between focus and stress has prompted other authors to add more rules, always pursuing a straight phonological account for focus placement. Reinhart (1995) proposes to add a marked rule, which languages would avoid whenever possible, to relocate the main stress if the NSR is not sufficient to make it coincide with the sentence focus. In Italian (18b) (taken from Brunetti, 2003, p. 32), the subject would not raise, staying in place at the end of the sentence in order to be able to receive a main stress assigned following the NSR. In contrast, assigning stress to a focused subject in English requires an extra rule, since the subject is necessarily preverbal (see 19):

- (18) a. Chi ha costruito la cupola?
 ‘Who built the dome?’
 b. [IP L’ha costruita [DP Brunelleschi]].
 it has built Brunelleschi
 ‘Brunelleschi built it’
- (19) [IP [DP MY NEIGHBOUR]_F [VP is building [DP a desk]]].

Neeleman and Reinhart (1998) go a step further and claim that there are two marked rules, and not just one, that determine stress placement. One of these rules refers to stress strengthening, as it is shown in (19). This rule would be a marked rule, and it would apply only in cases where it is necessary to change the Focus set of the sentence in order to make it unambiguous or to include a particular phrase in it (the Focus set is a concept introduced by Reinhart (1995), and defined as a set containing any sequence of constituents which includes the main stress, which is equivalent to the set of possible focus-markings a sentence might have given the surface prosody). The other rule is the rule of destressing, which applies to pronouns, due to their character of D-linked elements. This rule is independent of the informational status of the D-linked element. A linguistic item is D(iscourse)-linked when it refers to an element from a closed known set. The concept was introduced by Pesetsky (1987), who used it in reference to certain wh-phrases. An example of this (Brunetti, 2003) would be (20):

- (20) a. I SAW it.
b. * I saw IT.

What Neeleman and Reinhart are after when they separate destressing from stress strengthening is an explanation for more complicated issues, such as the contrast in (21), which Dwight Bolinger (1972) describes as stress change associated to semantically light words.

- (21) a. I have a POINT to make.
b. I have a point to EMPHASIZE.

To Neeleman and Reinhart cases like (20a) are due to destressing, and completely different from cases like (21a), that are due to the fact that the element that should bear the stress is unfit to bear focus, due to its lack of semantic content. They provide two arguments to defend the necessity of their two rules. The first argument is that if stress strengthening occurs far from the area where the NSR was insufficient to determine stress placement, like in (22) below, where the primary stress is marked in bold capitals, secondary stress placement still mimics the placement of primary stress in absence of stress strengthening. That is to say, the problem with NSR is still there independently of stress strengthening. Therefore, there must be some rule to prevent the NSR from explaining the placement of secondary stress, and an extra one to explain the

stress strengthening for the main sentence stress that takes place somewhere else in the sentence.

- (22) a. Only **MAX** can afford buying CARS.
b. Only **MAX** can afford SEEING her.

Their second argument involves Focus projection, which is not possible in cases of stress strengthening. This is not the case with destressing, where Focus projection is still a possibility.

Brunetti (2003) dismisses the argument based on secondary stress on the grounds that secondary stress is a purely rhythmic phenomenon, and thus not related to stress assignment. The premise to postulate this lack of relationship between stress assignment and a rhythmic secondary stress, of course, is that stress assignment would *not* be a rhythmic phenomenon, and therefore this counterargument appears to be inappropriate to defend the view that stress assignment is actually a rhythmic phenomenon.

More interestingly, Brunetti points out that examples like Bolinger's (21a) can be uttered out of the blue, and therefore it is difficult to group them together with cases of clear stress strengthening, like (19), which are not appropriate to be uttered out of the

blue, due, in my view, to the contrastive value of the Focus in them. For Brunetti, all cases of destressing, viewed as a violation of the NSR, are related.

Thus, an approach assuming that stress placement is a phonological process and that Focus placement takes place following stress rules seems to take us into a blind alley. The opposite approach, best summarized by Büring (forthcoming), that it is Focus placement that determines where the main stress (or nuclear pitch accent, NPA) is located, has the immediate advantage of explaining away the effects of destressing (or deaccenting) as natural consequences of applying rhythmic stress rules *after* the main stress has been assigned to the element bearing Focus.

3.2.2.2 Prosodically motivated movement

In the previous section I have discussed an approach to the relationship between the main sentence stress and the focalized phrase that involved adding extra rules to the NSR, in order to account for constructions where the stress associated to focus is not on the most embedded element of the sentence. Here I discuss Zubizarreta's (1998) work, which includes a distinction between two different stress rules and syntactic movement motivated by prosody.

Zubizarreta assumes the old basic observation in (23) that focus is linked to main stress (Chomsky, 1971; Jackendoff, 1972) and adapts it for German and Romance so it reflects the fact that the intonational nucleus is identified as the rhythmically most prominent word in those languages (24).

(23) Focus Prosody Correspondence Principle (FPCP) (Part 1)
The focused constituent (or F-marked constituent) of a phrase must contain the intonational nucleus of that phrase. (Zubizarreta, 1998, p. 38)

(24) Focus Prosody Correspondence Principle (FPCP) (adapted)
The F-marked constituent of a phrase must contain the rhythmically most prominent word in that phrase. (Zubizarreta, 1998, p. 38)

Zubizarreta also reinterprets the NSR, which she sees as “an algorithm that yields a prosodic interpretation of asymmetric c-command in the sense of Kayne (1994) and of Chomsky (Chomsky, 1994; , 1995)” (Zubizarreta, 1998, p. 40). She further assumes that it applies to the syntactic tree that is the input to Spell-Out in the sense of Chomsky (1995). The resulting NSR is presented in (25).

(25) Nuclear Stress Rule (NSR)

Given two nodes C_i and C_j that are metrical sisters, the one lower in the syntactic asymmetric c-command ordering is more prominent. (Zubizarreta, 1998, p. 40)

Rather than trying to fit contrastive stress as the result of some variation of the way the NSR is applied, Zubizarreta simply takes it to be freely assigned, with the restriction expressed in the following principle:

(26) Focus/Contrastive Stress Correspondence Principle

A word with contrastive stress must be dominated by every F-marked constituent in the phrase. (Zubizarreta, 1998, p. 45)

For non-contrastive stress, the following Focus Prominence Rule (FPR) applies:

(27) Focus Prominence Rule (FPR)

Given two sister nodes C_i (marked [+F]) and C_j (marked [-F]), C_i is more prominent than C_j .

The need for extra explanations or rules will appear in those cases of contradiction between the NSR and the FPR. Zubizarreta analyzes the differences in how such contradictions are dealt with in German, English, Spanish and Italian, as well as a few cases of other Romance languages.

Zubizarreta finds a simple way to explain the contrast between English and German, on one side, and Spanish and Italian, on the other (see section 3.2.1.1. above). She claims that “defocalized constituents are metrically invisible for the NSR in English

and German” (Zubizarreta, 1998, p. 46). While the importance of Zubizarreta’s work, in my view, lies on her proposals for Spanish and Italian, it might be interesting to ask how the rule “sees” whether a constituent is focalized or not. That is, through what interface does a phonological rule interpret informational value?

The NSR requires further tweaking to account for English and German data. I include below the versions of NSR for English, German, and Romance. The concept of selectional ordering refers to the ordered sequence of selected heads (C selects T, which selects V, which selects other V or P or D...). The actual selectional ordering presents additional complications. I give Zubizarreta’s version below her NSRs.

(28) NSR (English)

Given two sister nodes C_i and C_j , either (a) if C_i and C_j are selectionally ordered, the one lower in the selectional ordering is more prominent (the S-NSR), or (b) the one lower in the asymmetric c-command ordering is more prominent (the C-NSR). (Zubizarreta, 1998, p. 38)

(29) NSR (German)

Given two sister nodes C_i and C_j , (a) if C_i and C_j are selectionally ordered, the one lower in the selectional ordering is more prominent (the S-NSR); (b) otherwise, the one lower in the asymmetric c-command ordering is more prominent (the C-NSR). (Zubizarreta, 1998, p. 38)

(30) NSR (Romance)

Given two sister nodes C_i and C_j , the one lower in the asymmetric c-command ordering is more prominent (the C-NSR). (Zubizarreta, 1998, p. 38)

(31) Selectional ordering

$(C, T, V_1, \dots, V_i, P/V_m, D_m)$, with possibly $m=1$
 $(C, T, V_1, \dots, V_i, D_i)$, for $i = 1, 2, \dots, m-1$ (for the cases where $m > 1$)
where D_i , $i = 1, 2, \dots, m-1$ is the nominal argument of V_i (for the cases where $m > 1$) and D_m is the nominal argument of the lowest (possibly only) verb or prepositional predicate (P/V_m) in the selectional ordering.

The complexity of Zubizarreta's NSR, together with the sparing way in which it is actually applied, excluding defocalized elements, gives an idea of the extent to what prosody-based accounts of Focus rely on non-phonological elements. Zubizarreta's NSR not only has to "see" the syntactic structure (in the form of c-command relationships), which is after all the underlying assumption of this kind of rules, but it also has to apply to sentence elements according to their category and to whether they are focalized or not.

Prosodically motivated movement, or p-movement, takes place in languages where there are no metrically invisible elements. It is probably best illustrated with an example. Let us take (32), from Zubizarreta (1998, p. 125), where the main stress is on

María. The only possible interpretation of this sentence is with narrow focus on *María*, and it does not have a contrastive value.

- (32) Me regaló la botella de vino María
 To-me gave the bottle of wine María

Zubizarreta's analysis follows from the hypothesis that the underlying structure of this VOS sentence is either SVO or VSO, and the motivation for the reordering is to put the subject in a position where it can receive the nuclear stress following the Romance NSR seen above. Deriving VOS from SVO would entail moving an X' node, T', which be left-adjoined to TP:

- (33) [[T' me regaló [e_j [V₁ [la botella de vino_k [V₂ e_k]]]]]]_i [TP María_j [e_i]]]

This operation should not be possible, since X' nodes are not visible to syntactic computation. Therefore, she prefers deriving VOS from VSO, by leftward adjunction of the object (probably pied-piping the VP that immediately contains it) to the VP containing the subject. See two different possibilities for this (Zubizarreta, 1998, p. 127):

- (34) a. [TP me regaló [VP₁[DP la botella de vino]_k [VP₁ María [VP₂ V₁ [e_k [V₂ e_k]]]]]]
 b. [TP me regaló [VP₁[VP₂ la botella de vino]_k [e_k [V₂ e_k]]]_i[VP₁ María [V₁ [e_i]]]]

Italian does not have VSO order, and therefore VOS has to be derived from SVO, obtained first by raising the subject to a syntactic focus position FocP. The TP including the rest of the sentence is then left-adjoined (in a case of remnant movement) to this FocP, resulting in VOS order. Brunetti (2003, p. 46) says this analysis looks very redundant, including both syntactic and prosodic elements.

Brunetti's argument appears to be on the right track, and I would add to her criticism that we also have too many factors of each kind at play:

- a) a syntactic Focus projection that can be associated with both contrastive and non-contrastive accent, or not be used at all while Focus is assigned somewhere else.
- b) contrastive accents that are freely assigned.
- c) non-contrastive accents assigned by a prosodic rule that can detect word categories and, in certain languages, ignore non-focalized elements anyway.

Having all these elements available provides for a nuanced explanation, but hardly for a sufficiently constrained grammar, even if p-movement is interpreted as being Last Resort.

3.2.3 Intonation types

Focus is not the only element of information structure that appears related to phonology, more particularly to prosody. To see how Topic relates to prosody, note the following contrast (from Gardent, 1997):

- (35) FRED doesn't write poetry in the garden.
- a. It is Fred who doesn't write poetry in the garden.
 - b. It isn't Fred who writes poetry in the garden.

The meaning of this sentence changes depending on the intonation of *Fred*, i.e. depending on what pitch accent is associated with *Fred*. If the stressed syllable has a high pitch, it abruptly drops to a low pitch and falls at the end of the contour (H*L-L% in Pierrehumbert's (1980) notation) then the sentence is interpreted as (35a). Jackendoff (1972) calls this accent *A-accent* following Bolinger (1965). It is associated to focus interpretation. If the pitch accent is of the form H*L-H% (same as the previous one but rising instead of falling at the end) then the sentence is interpreted as in (35b). In this interpretation, *Fred* is a contrastive topic (Büring, 1997 and elsewhere). It is a topic in the sense of the topic-comment articulation, and it is contrastive in the sense that it introduces an adversative implicature. In the case of (35b), the implicature is that someone else does write poetry in the garden. This intonation is the *B-accent* (Bolinger,

1965; Jackendoff, 1972). Büring (forthcoming) points out that this correlation with intonation supports Vallduví's (1990) concept of link. A question that remains open, and poses a problem for Vallduví's system, is why sentences with no A-accent anywhere are possible.

A complete description of intonation is still in the works, particularly for Spanish (Beckman, Díaz Campos, McGory, & Morgan, 2002 ; Face, 2001 for the characterization of focus; Sosa, 1999 for a general overview). For English, Hedberg and Sosa (Hedberg, 2002; Hedberg & Sosa, 2001) have shown that we are far from a consensus on how intonational contours correlate with discourse functions in spontaneous speech. I will, however, accept that minimal pairs like the one in (35) above are an indication that some intonational marks can be meaningful when isolated. The actual realization of such marks in spontaneous speech is of course the result of the interaction of many factors that have yet to be described.

3.3 Interface problems

3.3.1 The T-model issue

The intonation-meaning line of argumentation seems to be in open contradiction with the T-model of grammar, prevalent for decades in linguistic research, which postulates no interaction between phonological rules and semantic interpretation. There is, however, a number of phenomena that suggest an interaction between syntax and phonology (Inkelas & Zec, 1995, 1990; Pullum & Zwicky, 1988). Such interactions go both ways, from syntax to phonology and vice-versa. A popular example of a phonological rule conditioned by syntax is the Italian phenomenon of Raddoppiamento Sintattico, or syntactic gemination. Take the following examples by Nespor and Vogel (1982, p. 228; , 1986, pp. 38, 170), quoted by Inkelas and Zec (1995):

- (36) Parlò [b:]ene
“He spoke well”
- (37) Devi comprare delle [_{NP} mappe [_{PP} di città [v:]ecchie]]
“You must buy some maps of old cities”
- (38) Devi comprare delle [_{NP} mappe [_{PP} di città] // vecchie]
“You must buy some old maps of cities”
- (39) Devi comprare delle [_{NP} mappe [_{PP} di città // molto vecchie]]
“You must buy some maps of very old cities”

It is easy to observe that the gemination of the initial consonant does not depend exclusively on the presence of a stressed vowel in the previous word, but it is also subject to certain other condition, which has been variously described (Chomsky & Halle, 1968; Hayes, 1989; Inkelas & Zec, 1990; Nespor & Vogel, 1982, 1986) but always taken to be a syntactic one.

More interesting for the purposes of this dissertation is the opposite case, a syntactic rule conditioned by phonology. Zec and Inkelas (1990) provide two cases where prosodic rules appear to influence syntax. In Serbo-Croatian topicalization the topic must be a branching phonological phrase. A single phonological word can not be a topic. Examples from Inkelas and Zec (1995, p. 545):

- (40) a. [NP [_ω Taj] [_ω čovek]] voleo-je Mariju
 that man loved-AUX Mary
 “Than man loved Mary”
- b. *[NP [_ω Petar]] voleo-je Mariju
 Petar loved-AUX Mary
 “Peter loved Mary”
- c. [NP [_ω Petar] [_ω Petrovic]] voleo-je Mariju
 Peter Petrovic loved-AUX Mary
 “Peter Petrovic loved Mary”

The other case Inkelas and Zec offer (Inkelas & Zec, 1995; Zec & Inkelas, 1990) is English Heavy NP Shift, that is not possible when the shifted NP consists of a single

phonological phrase (see 41). Swingle (1993) observes similar effects for Right Node Raising.

- (41) a. Mark showed to John [_{NP} [_ω some letters] [_ω from Paris]]
b. *Mark showed to John [_{NP} [_ω some letters]]

These data violate the Phonology-Free Syntax Principle postulated by Pullum and Zwicky (1988). Inkelas and Zec (1995) propose a weaker form of this principle, where “if [...] the interactions between syntax and phonology are limited to mutual, local constraints on syntactic and prosodic hierarchical configurations, then syntax will still lack access to segmental information” (Inkelas & Zec, 1995, p. 547).

3.3.2 Focus and the T-model

The issue of the phonological identification of Focus also presents difficulties for the T-model. The problem is clearly expressed by Manzini and Savoia: “Indeed even if an algorithm based on stress predicts the location of the property traditionally designated as Focus at the PF interface, this does not explain how the other interface, that of LF, knows which element undergoes the Focus interpretive process, which presumably includes the creation of a variable structure of some sort (Chomsky, 1977; Rooth, 1992, pp. 203-204). If such an interpretive process applies to phonological

representations, then we establish a direct connection between the two interfaces which effectively negates the basic reason for the existence of a computational component, namely the need for PF and LF to be connected (Chomsky, 1995)” (Manzini & Savoia, 1999).

A way to avoid the radical changes to the theory that this implies is to follow Brunetti (2003, p. 52) and say that, since Focus does not always coincide with the place where the nuclear stress falls, it is not related to prosody. Unfortunately, there is much evidence of a relationship between Focus and prosody, and it does not seem that it can be ignored completely. It would seem more appropriate to understand the mismatches between prosody and focus as exceptions to a rule, even if we have not figured out the rule yet. The real problem with prosodic accounts of Focus is the movement of focused phrases to sentence initial positions. In the next section I will briefly discuss some of the different opinions about how to classify the different manifestations of Focus.

3.4 Types of Focus: a first overview

In previous sections I have discussed a number of different views on how information structure is organized and what semantic or pragmatic content, if any, can be associated to it. Also, I have discussed a number of problems with the identification of focus using prosodic rules, even though it seems clear that there is some kind of

association between prosody and focus. It seems only logical that there is a view of the whole issue negating that all that contradictory data is a reflection of a single phenomenon. The canonical proposal for the division of focus in two different phenomena is the article by É. Kiss (1998), where two different types of foci are postulated for Hungarian (and suggested for other languages), dividing the confusing characteristics usually attributed to focus in two clearly differentiated groups. I summarize in the following table the characteristics of each type of Focus according to É. Kiss:

(42)

Identificational Focus	Information Focus
Exhaustive identification	Non-presupposed information
Universal quantifiers are excluded	No restrictions
Takes scope	Does not take scope
Movement	No movement involved
XP available for movement	Can be smaller or larger
Can be iterated	Can project

Brunetti (2003) dedicates a sizeable part of her dissertation to argue against each of these differences in Italian. I will not reproduce her arguments here. For the purposes of this work, it is worth noting that not a single one of the characteristics of Information Focus, as described by É. Kiss, seems to correlate with any actual grammatical feature,

in the sense that there seem to be no constraints on the sort of constituent that can bear Information Focus, and there are no external indications, at the sentence level, of what constituent bears it. That might be the reason it has been so difficult to pinpoint the exact relationship between any phonological (or syntactic) rule and information focus. That relationship might just not be there.

Domínguez (2004) takes a prosodic angle on the issue, but she still proposes a classification merely for descriptive purposes. She distinguishes three different types of Focus in Spanish with the following characteristics:

(43)

Focus in the left periphery	Focus in the right periphery	Focus in situ
Contrastive information	Contrastive/new information	Contrastive information
It cannot answer a wh-question	It answers a wh-question	It cannot answer a wh-question
Undergoes movement	In situ but scrambling may apply	In situ
It bears a feature	No feature if new information	It bears a feature
Marked stress	Unmarked stress	Very marked stress

These classifications, however, are partly based on syntactic arguments, so I will look at them in more detail in chapter 2.

4 Conclusions

In this chapter I have reviewed non-syntactic approaches to information structure. I have argued that information structure can not be explained as a pragmatic phenomenon, but rather requires the description of some grammatical procedure. I have discussed the major approaches to an explanation of Focus from a phonological perspective, and I have also presented current intonational descriptions of information structure phenomena.

Two important consequences can be derived from the issues considered here. One is that the realization of focus is not derivable exclusively from the application of phonological rules. The second one is there is some degree of correlation between intonation facts and semantic interpretation. This has immediate consequences for the theory of grammar, in particular for the T-model (Chomsky, 1995). Unless there is something in the computational system, i.e., in the syntax, to link the interpretation with the phonology, the model should be revised. I dedicate chapter 2 to discussing some syntactic approaches to information structure.

CHAPTER 2 – REVIEW OF THE LITERATURE ON INFORMATION

STRUCTURE: SYNTAX

1 Introduction

The syntactic approach to information structure starts from a very solid empirical ground: There are constructions where elements interpreted as sentential focus appear on the left of the sentence. Such constructions are marked. An unmarked construction would leave those elements somewhere else in the sentence, even if they can also be interpreted as focus in that other position.

- (1) a. Cicadas, I have seen.
b. I have seen cicadas.

In view of examples like (1a), it is obvious that there is an interaction between syntax and information structure, even if nothing else is completely clear. In the first chapter of this dissertation I addressed the separation of information structure phenomena from pragmatics and from prosody. This chapter will address the way they are linked to syntax.

I will argue that focus and topic are associated to functional projections, as initially proposed by Rizzi (1997). Once information focus is left out of the picture, following the line of thought of Kiss (1998) and Domínguez (2004), a functional-head approach, Rizzi style, provides a common explanation for many word order phenomena, is compatible with intonation phenomena, and does not require any major changes to the model of grammar, especially if we adopt a model like the one proposed by Kayne (1994).

The order of the argumentation will be as follows: I will first address, in section 2, several different approaches to classifying these phenomena using syntactic criteria, with particular attention to the need to distinguish between several types of focus. Once an initial description has been presented, section 3 is dedicated to the way information structure phenomena are integrated in the structure of the clause. This will be illustrated with a discussion of studies on the left periphery, based mostly on ideas developed for Italian. Section 4 to the so called “center periphery” (Camacho, 2003), as well as some other related issues, like the existence of visible Focus particles. In order to provide an appropriate theoretical framework, I include a review of Kayne’s (1994) work. After a summary of conclusions, I finish the chapter with an appendix on the classification of information structure related constructions Villalba (2000) proposes for Catalan.

2 Characterization of topic and focus – the syntactic approach

While the debate about the pragmatic and intonational characterization of information in a sentence is quite old and theories usually take a global approach, syntactic research usually concentrates on a single aspect of the issue. The obvious issue for a syntactician is word order. The main word order issues related to information structure are three: dislocations (For an overview of English constructions, see Erdmann, 1990; for Catalan, Villalba, 2000)), subject position, and clitic position¹ (Campos, 1989; Lorenzo González, 1995; Uriagereka, 1995 and others), the last two issues more relevant in Romance than in English. Here I will mostly address dislocations.

Rivero (1980) distinguishes (for Spanish, but there were not many reasons at the time to discard the possibility that her analysis would be valid for more languages) two types of constructions where an element appears at the beginning of a sentence. For one of them, left dislocation, the dislocated element is understood to be base-generated, and the construction is not subject to movement constraints, and there is an anaphoric relationship between the left-dislocated element and some element in the other part of the sentence. The other construction, topicalization, has always been considered the result of a movement operation (Chomsky, 1977; Rivero, 1978, 1980; Ross, 1967 and

¹ This issue pointed out to me by Francisco Fernández Rubiera.

many others). Chomsky's (1977) analysis, for instance, takes topicalization to be the result of wh-movement in a left-dislocation construction, where the anaphor is the wh-element and is deleted after undergoing movement. Rivero (1980) makes the point that topicalization is not subject to the same constraints as wh-movement, particularly to the wh-island constraint. Note that "topicalization" is a term associated with focus constructions, not with topic constructions. The movement was described at a time when there was not any consensus on the description of topic and focus.

An additional problem is the development of overlapping terminologies. For instance, Campos and Zampini (1990), following Suñer (1982), differentiate between contrastive focus, which introduces an adversative implicature (they say it is being emphasized as opposed to another element in the sentence) and informative focus, which corresponds to Suñer's interpretation of Halliday's (1967) focus and includes what I have been calling topics (Vallduví, 1990) but also contrastive topics (Büring, 1997), and probably also some cases of what É. Kiss (1998) calls information focus.

I will now give an example of another kind of problem that plagued early studies: notational underspecification due to a lack of a clear classification of

intonations. The following example opened an article on topicalization in Spanish² (Rivero, 1978):

(2) Dinero, dicen que tiene.

‘Money, (they) say that (he) has’

The example is presented as a clear case of topicalization. No indication about the intonation is made, but there are at least two different accents it can receive, one being the A-accent (which yields a focus reading) and the other the B-accent (which yields a topic reading) (see chapter 1 and Jackendoff, 1972 for descriptions of these accents). The difference in meaning associated to these accents is shown in (3) (Gardent, 1997). I follow the practice of marking A-accents with \ and B-accents with /. While this might seem to be irrelevant for syntactic purposes, the contrasts in (4) might show how difficult it would be for a reader who does not speak Spanish to get to conclusions of her own. The contrast in (4) shows how intonationally marked focus has to follow, never precede, left dislocated topics. While (4a) is impossible, (4b) is fine. A reader that could not count on a clear notation for intonation would determine different relationships between topicalization and the position of subjects depending on what intonation she had in mind.

² Rivero’s (1978, 1980) studies are crucial in the syntactic characterization of topicalization. By using this example I intend to point out the difficulties inherent to the field, and in no way am I resting importance to her early work.

- (3) a. Dinero\ dicen que tiene. (A-accent)
 ‘It is money they say he has’ (Focus reading)
- b. Dinero/ dicen que tiene. (B-accent)
 ‘As for money, they say he has it’ (Topic reading)
- (4) a. *Dinero\ Juan dicen que tiene. (A-accent)
 ‘It is money they say Juan has’ (Focus reading of *dinero*)
- b. Dinero/ Juan\ dicen que tiene. (B-accent, A-accent)
 ‘As for money, it is Juan who they say he has it’
 (Topic reading of *dinero*, focus reading of *Juan*)

In spite of the problems described above and having only a partial view of the number of constructions involved, Hernanz and Brucart (1987) and Zampini (1988) are able to make a number of important generalizations (listed in Campos & Zampini, 1990) that I present here in terms of topic and focus:

- (5) a. Topic precedes focus.
 b. Topic precedes interrogative words.
 c. Focus movement and wh-movement are in complementary distribution.
 d. There can be more than one topic but only one focus.

In section 3, I will show how these generalizations are captured by the split CP hypothesis (Benincà, 2001; Rizzi, 1997). But before that, in section 2, I will review the literature that defines the concept of focus I am using in this dissertation.

2.1 Types of focus

2.1.1 É. Kiss's distinction between identificational and information Focus

É. Kiss (1998) distinguishes between two types of focus, which she labels “identificational focus” and “information focus”. The claim that there is not a single kind of focus, but rather two, has been around for a long time. É. Kiss (1998) gives the examples of Halliday (1967) and Rochemont (1986) as authors that make a distinction. One kind of focus has been called contrastive, narrow, or, by É. Kiss, identificational. The other kind has been called presentational, wide, or, by É. Kiss, information focus. É. Kiss starts her article pointing out that descriptions of Greek (Tsimpli, 1994), Finnish (Vilkuna, 1994), Catalan (Vallduví, 1990), and German (Krifka, 1992) have included common syntactic explanations for both kinds of focus, even though both their perceived meaning and their syntactic behavior were different. As an illustration I include the example É. Kiss takes from Tsimpli:

- (6) a. **Ston Petro** dhanisan to vivlio
 to.the Petro lent.3pl the.acc book
 ‘It was **to Petro** that they lent the book’
- b. Dhanisan to vivlio STON PETRO
 ‘They lent the book **TO PETRO**’

In this example I keep É. Kiss’ notation, where bold typeface indicates identificational focus, and capitals indicate information focus bearing nuclear stress. Note that the Greek examples are translated differently, but Tsimpli claims that both types of focused elements end up in a left-peripheral Focus Phrase, one moving in overt syntax, the other moving covertly. As É. Kiss points out, most syntactic explanations of Focus prior to the mid-nineties have been devised trying to cover two distinct phenomena. One of those two phenomena, information focus, was syntactically ‘inert’. The search for a common explanation for both phenomena led to unnecessary covert movement explanations of information focus, which in turn obscured the actual explanation for identificational focus.

The relevant point for the purposes of this dissertation is that there is a syntactic phenomenon that has been mixed up with a non-syntactic phenomenon (loosely associated to prosody) due to the similarity of their semantic interpretation. É. Kiss takes a step further and states that the semantics of these phenomena are not actually as similar as had been claimed, and in fact, once one begins to see them as two different

types, it becomes clear that there are substantial differences in the semantics of the two types as well. The semantics of identificational focus would be expressed like this:

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

The idea that focusing is an operation on a set of comparable entities given in the context is taken from Jacobs (1986), von Stechow and Uhmann (1986), and Krifka (1992). The idea of exhaustive identification comes from Kenesei (1986) and Szabolcsi (1994).

É. Kiss provides a list of differences between the two types of focus she distinguishes. I copy it here:

- (8) Differences between identificational focus and information focus.
(É. Kiss, 1998, p. 248)
1. The identificational focus expresses exhaustive identification; information focus merely marks the nonpresupposed nature of the information it carries.
 2. Certain types of constituents, universal quantifiers, *also*-phrases, and *even*-phrases, for example, cannot function as identificational foci; but the type of constituents that can function as information focus is not restricted.

3. The identificational focus does, information focus does not, take scope.
4. The identificational focus is moved to the specifier of a functional projection; information focus, however, does not involve any movement.
5. The identificational focus is always coextensive with an XP available for operator movement, but information focus can be either smaller or larger.
6. The identificational focus can be iterated, but information focus can project.

In order to account for the difference in feature content of focus in different languages, É. Kiss proposes to adopt the features [\pm exhaustive] and [\pm contrastive], already used by Jacobs (1988), and explain the differences through the use of one or both features.

É. Kiss classification solves many of the apparent contradictions about the nature of focus presented in chapter 1, and provides a solid basis for syntax-based approaches. I will adopt the basic lines of her description as a starting point for any analysis of focus in chapters 3 and 4. The next two subsections, 2.1.2 and 2.1.3, discuss other analyses in order to point out some consequences of assuming the existence of two types of focus.

2.1.2 Brunetti's attempt at unification

Brunetti's (2003) dissertation proposes the unification of all types of focus into one³. Her position represents a return to the view against which É. Kiss was arguing. By the time Brunetti wrote her dissertation, the view that there are two different kinds of focus had become commonplace (she mentions Belletti, 2004; Zubizarreta, 1998 among others), even though the actual analyses were not always in agreement. She uses Italian data to dispute the conclusions presented by É. Kiss (1998) and Donati and Nespors (2003).

One of the arguments used to differentiate between two kinds of focus is that the phenomenon of focus projection does not take place with contrastive (identificational) focus, but only with information focus (Donati & Nespors, 2003; É. Kiss, 1998). Before addressing Brunetti's view of focus projection, I will introduce the basic facts about the phenomenon. Consider the following examples:

³ She states that she will address narrow focus only, in the sense of focus that is "only a part of the sentence, and not the whole one" (Brunetti, 2003, p. 18). The reason for this decision is not clear to me, unless it is meant to leave out the issue of focus projection (of which whole-sentence focus would be a case), and thus avoid the possible contrast between types of focus that can project (informative focus) and those that can not (contrastive focus). Brunetti, however, does discuss contrasts in the availability of focus projection.

- (9) He was warned to look out for [an ex-convict [with [a red [*shirt*]]]].
- (10) a. No, he was warned to look out for [an ex-convict [with [a red [*tie*]]]].
b. No, he was warned to look out for [an ex-convict [with [a *carnation*]].
c. No, he was warned to look out for [an *automobile* salesman].
d. No, he was warned [to expect a visit from the *FBI*].
e. No, he was [simply told to be more *cautious*].

The example in (9) is from Chomsky (1971). The sentences in (10) are possible replies depending on which of the bracketed phrases in (9) is perceived to be the focus, “depending on the context” (Brunetti, 2003, p. 106). For all the examples in this section I use Brunetti’s notation, where the word with the syllable bearing the nuclear stress appears in cursive, informative focus in bold type, and contrastive focus is marked with capital letters.

Before moving on, let me note once again how, in light of what we know now, it is easy to conclude from these examples that if we use stress placement as a basis for the identification of focused elements the results are unclear and it is necessary to resort to the context to disambiguate them.

É. Kiss (1998) and Donati and Nespó (2003) claim that contrastive (É. Kiss' identificational) focus can not project. Donati and Nespó (2003) provide the following examples, where contrast can be expressed for word elements (11a) or words (11b), but not for larger elements (11c). It is crucial to note, however, that what Donati and Nespó are claiming is that there can not be two contrastive elements active in the same sentence. That is, (11c) can not be paraphrased as *I always thought that John was YOUNG and also that he was a JOURNALIST*, where capital letters indicate contrast. Their example in (11c) is confusing because the natural reading is that the phrase *a young journalist* can be understood as a single contrastive element. I will insist on this point below in this same section.

- (11) a. I always thought John was [ANTI]_F-communist.
 b. I always thought John was [WELSH]_F.
 c. I always thought John was *[A YOUNG JOURNALIST]_F.

Brunetti puts the emphasis on the fact that Donati and Nespó say not only that a contrastive Focus cannot project, but also that it cannot be larger than a word. She provides the following examples to counter Donati and Nespó's:

- (12) a. Maria ha incontrato il Prof. Rossi.
 'Maria met Prof. Rossi'

- b. No, ha incontrato [_{DP} IL MAESTRO [_{PP} DELLA FIGLIA [_{PP} DI *PIETRO*]]].
 no (she) has met the teacher of the daughter of Pietro
 ‘No, she met Peter’s daughter’s teacher’

She notes that focus involves the whole DP in (12), which is clearly contrastive, since it contradicts the information in the previous sentence. She further notes that the stress pattern for information focus (represented in bold font) is the same as for contrastive focus (see 13), that Focus projection can also occur if the focused phrase has moved to the left (see 14) and that a contrastive interpretation of Focus does not prevent it from extending its domain to the whole clause (see 15 and 16).

- (13) a. Chi ha incontrato tua sorella?
 ‘Who did your sister meet?’
- b. Ha incontrato [_{DP} **il maestro** [_{PP} **della figlia** [_{PP} **di Pietro**]]].
 ‘She met Peter’s daughter’s teacher’
- (14) a. Tua sorella ha incontrato il Prof. Rossi?
 ‘Did your sister meet Prof. Rossi?’
- b. No, [_{DP} IL MAESTRO [_{PP} DELLA FIGLIA [_{PP} DI *PIETRO*]]] ha incontrato.
 no the teacher of the daughter of Peter (she) has met
 ‘No, it was Peter’s daughter’s teacher that she met’
- (15) a. Hanno bussato alla porta?
 ‘Did someone knock the door?’
- b. No, [STANNO [PIANTANDO [UN *CHIODO*]]].
 ‘No, they are driving a nail’

- (16) a. Gianni è ingrassato perché ha mangiato molti muffins?
'Did Gianni get fat because he ate many muffins?'
- b. No, perché [HA [BEVUTO [MOLTE [BIRRE]]]].
'No, because he drank many beers'

The examples used by Brunetti suggest that she understands the notion of Focus projection differently than Donati and Nespor did. I will now provide some common ground on the notion and revisit Brunetti's examples to see whether they sustain her claim.

Focus projection is a phenomenon by which a quasi-semantic property (status of new information), which is portrayed as associated to a prosodic feature (nuclear stress), extends beyond the element that actually bears the prosodic feature. The prosodic feature in question is constrained by phonological rules, and the domain on which these phonological rules apply seems to be constrained by syntax. What Donati and Nespor are showing in (11) is that whatever it is that is marking focus will not appear twice within a phrase. On the other hand, one sentence can express one contrast, not two (and this is not a prosodic property). Expressing two contrasts would require some kind of enumeration or coordination structure, as in *young and a journalist*. Then both elements could be focused independently as contrastive. (Another possible context is a corrective context where contrastive intonation is applied on top of a repeated or mimicked sentence, as in *I didn't give the book to Neal, I gave the NOTEBOOK to MELISSA*.) The

example in (11c) would be fine if being a “young journalist” meant, for instance, belonging to a particular association of journalists. The nuclear stress would fall on *journalist* and there would be one contrast. No ambiguity could be inferred. The analysis of Donati and Nespor can perhaps be criticized as incomplete or unclear, but the examples provided by Brunetti address a different problem. Contrastive focus is a property of a syntactic object, an XP susceptible to movement. Within it, there is one nuclear stress. This nuclear stress is associated to a certain interpretation, and it is assigned following phonological rules within the XP. Information focus is not associated to a syntactic object, in the sense that it does not limit a domain for the rule that assigns the stress, nor does it make any particular phrase susceptible to syntactic operations.

Pointing out the phenomenon of focus projection means pointing out that there is nothing that marks the boundaries of a focused element, as long as we look only at the nuclear stress. The element bearing contrastive focus, on the other hand, has clearly defined boundaries. The XP that is moved is the focused phrase. If we still want to wonder whether it projects or not, the answer is that it does not project, because there is no ambiguity as for what is focused. But not projecting does not mean that focus has to be constrained to the smallest element bearing the nuclear stress.

An additional argument against presenting this discussion in terms of “projection” comes from phonology, so I will not discuss it in depth here. I will just point out that the phonological mark of focus is not the nuclear stress itself, but an intonative contour associated to the focused phrase and anchored to the nuclear stress. That intonative contour is absent in informative focus, and that absence is what allows for ambiguity, therefore making possible a focus projection reading.

This said, it is even easier to dismiss Brunetti’s suggestion that contrastive focus can extend to the whole sentence. The fact that the meaning of a sentence contradicts some background information is not enough to establish that it carries a [+contrastive] feature. Brunetti’s example in (15) is contrastive at discourse level, but that does not tell us anything about its internal structure. Maybe there is no contrastive focus at all, and here is where the intonation might give us some clue, but in no case should we obtain an ambiguous result. The example in (16) is even more interesting, because it is clear that the so-called projection can not grow to include *perché* ‘because’ or shrink to leave out *ha bevuto* ‘drank’. So, in sum, contrastive focus is carried unambiguously by an XP that can be a whole clause, but not a whole sentence (except, perhaps, in cases where some elements have undergone ellipsis). Information focus, on the other hand, is a discourse phenomenon that might or might not be linked to the nuclear stress, but not to intonational events such as Jackendoff’s accents, nor, even more clearly, to syntax.

Brunetti also points out, contra Zubizarreta (1998), that there is no need for an Emphatic/Contrastive Stress Rule separated from the Nuclear Stress Rule, since, if we accept Brunetti's view of what a focused phrase is, they both end up assigning nuclear stress to the most embedded element of the focused phrase. The fact is that there is always a nuclear stress, and it is assigned by some kind of phonological rule. It is obvious that, in those cases when there is a contrastive focus, this rule operates taking syntax into account. However, if focus is not a prosodic phenomenon (as Brunetti herself says), we have no reason to believe that the assignment of nuclear stress in sentences that have no contrastive focus has anything to do with focus at all. Thus, even if there is in fact one single Nuclear Stress Rule, it has to include some syntax-related component.

Generally speaking, there are two main problems in the relationship between prosody and other levels of grammar. The first is the relationship between pure prosody and meaning, as perceived in wordless utterances. The second one is the relationship between suprasegmentals and syntactic operations. Information focus is related to the former problem, and contrastive focus to the latter. This dissertation is about syntax, so information focus is outside its scope. As for contrastive focus, since its prosodical marking is invariably associated to a syntactic operation, it can be assumed that no important contribution can be made to its study from the point of view of phonology.

From this point on, I will discuss prosody only when it is in correlation with syntax and can therefore serve as a clue to discern syntactic operations.

To sum up, I list here the characteristics of focus that have been discussed in this section.

- There is no more than one contrastive focus in a sentence.
- Stress rules are not sufficient to determine what constituent bears focus.
- Focus projection is not a grammatical phenomenon, but a result of structural ambiguity from the point of the view of the hearer. Arguments on whether it appears in contrastive or information focus end up making a moot point, because the phenomenon depends on the sentence having a particular prosodic structure, not on what kind of focus is actually present in, or absent of, the sentence.
- Contrastive focus is carried unambiguously by an XP that can be a whole clause, but not a whole sentence (except, perhaps, in cases where some elements have undergone ellipsis).
- Information focus is a discourse phenomenon that might or might not be linked to the nuclear stress, but not to intonational events such as Jackendoff's A-accent, nor to syntax.
- Contrastive focus can be disambiguated by movement or by intonational contour.

In chapter 4 I will analyze cases of co-occurrence of contrastive focus and interrogative features. The discussion in this section is relevant to that purpose to the extent it determines how the syntax of contrastive focus is to be understood, and what kinds of constituent can be expected to have a contrastive focus feature (such features are discussed in section 3 of this chapter).

2.1.3 Domínguez's classification for Spanish

The last contribution to the body of literature that studies Focus in Spanish is the dissertation by Laura Domínguez (2004), whose main conclusions I will not discuss in detail, but I summarize here:

- (17)
1. There are three types of focus in Spanish: contrastive in the left periphery, contrastive in situ, and information in the right periphery.
 2. Contrastive focus has a marked accent, information focus does not.
 3. Information focus does not have a focus feature
 4. Contrastive focus always moves to the left periphery, overtly or covertly

Domínguez's description of Spanish supports and builds on É. Kiss' view of focus as two separate phenomena. Since information focus is not relevant for the purposes of this dissertation, I will briefly summarize Domínguez's take on it and accept her conclusions without further discussion. When Domínguez refers to the right

periphery, she is not referring to right-dislocated elements, but to the final element of a sentence without dislocations. That is, the place where sentential stress appears. Information focus is thus not a syntactic phenomenon, but a discursive one linked to prosody, and licensed by prosody. In line with Zubizarreta's (1998) conclusions, while information focus is not a syntactic phenomenon, it can have an effect in syntax, when the need to align sentential stress and information focus produces p(rosodic)-movement. Domínguez discusses and rejects the Nuclear Stress Rule for being too dependent on syntax, and proposes a strictly phonological approach to stress placement.

3 The left periphery.

Simplification of syntactic theory (to the extent that a simpler theory is more likely to attain explanatory adequacy; for a discussion on the concept of simplicity in grammar see Chomsky (1965, pp. 37-38)) has always been a goal of modern linguistics, but a particular sense of urgency was felt among many linguists at the beginning of the 1990s. Perhaps the most celebrated results of this trend are the articles by Chomsky that were compiled as *The Minimalist Program* in 1995. The publication in 1994 of Richard Kayne's *The Antisymmetry of Syntax* represents a high point in such movement towards simplification. Rather than re-elaborate the model of grammar, Kayne focuses on the restrictions of the form of syntactic representations. Briefly put, Kayne (1994) proposes

that there is a correspondence between the structural hierarchy of sentential elements and their linear order; namely, c-commanding elements precede c-commanded ones. In Kayne's work, many traditionally accepted conditions on syntactic structures appear to be direct consequences of a general principle. At the same time, this model allows for a clearer correspondence between different levels of representation. Explanations based on structural differences, like the head-final parameter, are not possible in this framework, and cross-linguistic variation and other sources of complexity are to be explained by different instances of movement. Of course, this requires the introduction of new functional categories that will serve as landing sites for such movement. The proliferation of such functional categories has occasionally been criticized in the last few years, but it remains the cleanest approach to the problem of the interface between morphology and syntax, and it has greatly contributed to clarify the confusing field of the pragmatics-syntax interface.

Kayne's (1994) restrictive model does not leave room for right adjunction or rightward movement. Therefore, newly proposed functional categories will appear to the left of the structural representation, accumulating in a functional field that has been called left periphery. Rizzi's (1997) influential paper presented a quite comprehensive mapping of the possibilities of the field, where the clausal type or Force (Cheng, 1991; Chomsky, 1993) and the traditional informational articulations of topic/comment and

focus/presupposition were characterized as functional projections susceptible of hosting sentence elements, mostly as the result of movement.

3.1 Rizzi's initial proposal

The article by Rizzi (1997) can be said to be the point at which the syntactic research on topic and focus finds some common ground to start a coherent discussion. At least in Romance linguistics, and increasingly in other areas, the term *left periphery* does not refer to a vague left side of the clause anymore, but rather to a functional field within which concrete functional projections are postulated. The cartography of the left periphery is being developed mostly for Italian, although there is also work on Catalan (Villalba, 2000), Spanish (López-Cortina, 2003; López, 1999; Martín, 2003) and Italian dialects (Cinque, 2002; Martorana, 2004; Poletto, 2000; Rizzi, 2004).

Rizzi proposes a left periphery as in (18), where the two informational articulations of the clause (topic-comment, focus-presupposition) are represented in the syntactic structure by functional projections within a functional layer that includes all the information about the clausal type. Topic Phrases are recursive.

(18) [ForceP [TopP* [FocP [TopP* [FinP [IP...

More important than the actual number or position of the phrases are the assumptions he makes about the syntax of the left periphery. Focus, for Rizzi, is a quantificational element A'-binding a variable (for the difference between quantificational and non quantificational A'-binding, see Lasnik & Stowell, 1991), which explains the interpretive properties of contrastive focus. The Focus Phrase is also the target of wh-movement (a position nuanced in Rizzi (2001), where wh-phrases only compete for Spec, FocP in main clauses, while they have their own head in embedded clauses). Benincà (2001) reviews the structure and proposes this one instead:

(19) [DiscP [ForceP [TopP [FocP [FinP [IP...

This structure has a number of advantages. In particular, the new Disc(ourse) Phrase provides the space for a base-generated thematic element, realized in Italian as a Hanging Topic (HT). Benincà's structure also eliminates both the recurrence of the Top(ic)P(hrase) and the lack of a definite position for TopP with respect to FocP that characterized it in Rizzi's proposed structure. The problem of multiplicity of phrases in the left periphery is transferred to Spec, FocP, which Benincà considers able to host multiple elements, of which only the first one would receive focalized intonation. Benincà also identifies Spec, FocP with the wh- landing site. Benincà's work is focused on Italian data and she barely mentions other Romance languages, but it should be

noted that her proposal matches the observations for Spanish made by Hernanz and Bruccart (1987).

Subsequent proposals have expanded Focus to host several functional heads, among them an independent WhPhrase to the right of the Focus field (supported by Lee, 2001; López-Cortina, 2003; Martorana, 2004), which Rizzi (2001) confines to embedded clauses. Rizzi (2001) also proposed an Int(errogative)Phrase to the left of the FocP, that would host *se* “if” in embedded yes/no questions.

The theoretical device to justify movement in these proposals is a feature in the style of Chomsky (1995). Brunetti (2003) criticizes that feature in terms of Chomsky’s (1995) Inclusiveness Condition:

- (20) “outputs consist of nothing beyond properties of items of the lexicon (lexical features) – in other words, ... the interface levels consist of nothing more than arrangements of lexical features. To the extent that this is true, the language meets a condition of *inclusiveness*”. Chomsky (1995)

The problem is the relational character of focus, which prevents its description in lexical terms, plus the fact that contrastive focus can include more than one lexical element, or elements smaller than one. Zubizarreta (1998) proposes a modification of the condition to include focus.

The problem the Inclusiveness Condition presents for a focus feature might not be particular to focus. A wh feature triggering wh-movement could probably be criticized on similar grounds. This issue, however, is outside of the scope of this chapter, which is meant to be a presentation of the syntactic framework for the analysis of information structure phenomena. Thus, having introduced Rizzi's 1997 article, I will now talk about further developments of his proposal.

3.2 Benincà and Poletto's refinements

Paola Benincà and Cecilia Poletto have been gradually (Benincà, 2001; Benincà & Poletto, 1999, 2001, 2004) elaborating on the map of the Italian left periphery by significantly altering Rizzi's initial proposal (Rizzi, 1997). For clarity, I present as a column the structure they currently suggest (Benincà & Poletto, 2004):

- (21) *Frame*
 [Hanging Topic
 [Scene Setting
- Theme*
 [Left Dislocated elements
 [Listed Interpretation
- Focus*
 [Contrastive CP 1 (adverbials/objects)
 [Contrastive CP 2 (circumstantial and quantificational adverbs)
 [Informational CP]]]]]]]

The hypothesis that guides the work of Benincà and Poletto is that there is no recursion of functional heads, but rather a limited, ordered set. They proceed to determine, *contra* Rizzi (1997), that there are no left dislocated elements (henceforth LD) to the right of Focus. They use Topic and LD interchangeably. They provide evidence from Paduan to point out that adverbs occupying a lower position than Focus, at the IP edge, might actually be below the subject and they should not be taken into account when deciding on the location of LDs. Weak crossover (WCO) effects are observed with focalized elements but not with Topics, and any element to the right of Focus also shows WCO effects, thus supporting their claim that there is a field where several types of Focus appear in succession.

The term Informational Focus should not lead the reader to identify this position with information focus as defined by É. Kiss (1998). Benincà and Poletto claim that elements in this position do not have contrastive value, but the parallels between their Informational Focus and É. Kiss' information focus end there. The position is postulated on the basis of the distribution of the particle *pa* in the Rhaeto-Romance dialect of S. Leonardo (Poletto & Zanuttini, 2000).

The distinction between two types of contrastive focus is based on the different behavior of Focus in subordinate clauses depending on the selecting verb. The idea is that some verbs select a more complete CP layer than others, and when a limited CP

layer is selected only certain kinds of elements are allowed to appear focalized. This suggests that they occupy a different functional head than other elements.

In the Topic field, Hanging Topics (HT) on one side, and elements placed there as result of the operation of Left Dislocation (LD), on the other, are different in a number of ways noted by Cinque (1983) and Benincà herself (Benincà, Salvi, & Frison, 1988). The tests are six:

- (22)
1. LD elements maintain the preposition of the internal elements they correspond to. HTs can only be DPs.
 2. Only a single HT position per clause is available. More than one LD can appear.
 3. LDs only require a resumptive element when they correspond to direct or partitive objects. HTs always require one.
 4. LDs' resumptive elements have to be clitics. The ones for HTs can also be tonic pronouns or epithets.
 5. HTs appear before the complementizer *che* in subordinate clauses. LDs appear after the complementizer.
 6. HTs are not possible in relative clauses, neither before nor after the relative pronoun. LDs are possible after the pronoun.

This gives sufficient evidence to postulate the separate fields labeled as Frame and Theme in the summary above. Smaller divisions within the fields are suggested

after looking at the behavior of adverbs in Rhaeto-Romance dialects, and to the interaction with Right Dislocations.

As for Right Dislocation, Benincà and Poletto claim that it can be only a Theme and not a Topic. They use the term Theme for an element than can be recalled from the immediate context, and Topic for an element that must be present in the shared knowledge of speaker and hearer for the sentence to be felicitous. Regardless of the accuracy of this description, note that these definitions make the labeling of the three major fields extremely ambiguous.

Finally, a note on intonation: Benincà and Poletto choose not to use intonation as a test for distinguishing between focalized and left dislocated elements. They provide the following example to show how a left dislocated element, which would be interpreted as a Left Dislocated Topic, can receive the intonation associated to Focus:

- (23) A: Mi ha detto che il tappeto, *lo* compra l'anno prossimo
'He has told me that the carpet he will buy *it* next year'
- B: No, ti sbagli, IL DIVANO *lo* compra l'anno prossimo
'No, you are wrong, THE SOFA he will buy *it* next year'

Benincà and Poletto have used this example to support a dissociation of the intonational and syntactic levels; however, there might be other ways to interpret this example.

Focus intonation can be associated to any part of a sentence, no matter how large or small, as long as the sentence is perceived as a repetition of a previous utterance. In such sentences, the usual intonation pattern of declaratives does not apply, nor does the syntax. Rather, focus intonation is used as a means to express that the focused element has been, or has to be, changed, while the rest of the sentence must remain unchanged. Such focus intonation can be associated to any element that the speaker intends to mark as susceptible to correction, even in cases where the element is as small as a phoneme. This shows, in my view, that in these examples the intonation does not correspond to the information structure, or the syntax, of the sentence as a context-independent message. Instead, it corresponds to the sentence as a discourse object, where the usual syntactic (and even phonological) rules have been suspended.

In the particular example above, the focus intonation is not associated to a left dislocated element, because the structure of B is not one of left dislocation. *Il divano* did not get to that position in the sentence by means of dislocation, but by substitution in an echoed sentence fragment (*X lo compra l'anno prossimo*). This is obvious if one observes that a sentence like B can not be uttered without one like A immediately preceding it. Also, any differences in wording between B and A would render B infelicitous.

The main point of Benincà and Poletto, however, is not that there is no relationship between intonation and focus, but that this relationship is not one-to-one. The discussion in this section does not invalidate their point, but rather adds some additional elements to take into account when considering an analysis.

4 Beyond the left periphery.

The growing influence of Kayne's work, together with the prominence of focus as a research topic, have resulted in the development of theories of right dislocation that are closely related to those devised for left dislocation (Villalba, 2000 presents a good review of this issue). Remnant movement, i.e. the kind of movement needed to derive right-dislocation from leftward-only movement in a Kayne-style structure was originally described for German (Besten & Webelhuth, 1987, 1990; Müller, 1998, 2000; Thiers, 1985) and later for Romance (Pollock, 2002).

The elements of the left periphery are now often viewed as a layer of information-related functional categories that can be replicated within the verbal layer of the verb and not only in the actual left periphery of the sentence (Camacho, 2003; Jayaseelan, 2001; Poletto, 2004).

Today, more than twenty years after von Stechow (1981) expressed his surprise at the discovery that there was no commonly accepted definition of topic or focus, we have found many more questions than answers, which is always good news in scientific research. Rizzi's (1997) article represented a giant's leap within the purely syntactic approach to the problem, but even now there is no agreement on whether there is one focus or two (for a recent unifying account, see Brunetti, 2003), whether the often mentioned Japanese topicalization always involves topics (Kaga, 1997; Kato, 1985; Watanabe, 2003), whether there is movement in clitic left dislocation or not (besides the long discussion in Villalba's (2000) dissertation, Cecchetto & Chierchia (1998), for instance, propose moved DP and base-generated PP topics), or even how many kinds of movement are involved (Cinque, 1990; Watanabe, 2003). This last question is more general, and it shows how the study of dislocates is having an influence in general issues of syntax.

This section is dedicated to the discussion of work that goes beyond the initial hypothesis that all information structure-related materials would have their syntactic place in the left periphery. In 4.2 I discuss the ideas of IP and DP internal focus and topic, and in 4.3 I list some of the lexically overt evidence presented in favor of a Focus phrase. Before I get to that, however, I will briefly review the syntactic framework developed by Kayne (1994).

4.1 Kayne's framework

The first chapter of Villalba's (2000) dissertation thoroughly lays down the Hypothesis of the Antisymmetry of Syntax as devised by Kayne (1994) and focuses mainly on its formalization, for which purpose he resorts to the technical apparatus of Wall (1972). In this section I will give a short presentation of Kayne's framework and I will review Villalba's assessment of it (for another, important, review of Kayne see Johnson, 1997).

In Kayne (1994) the form of syntactic representations follows from three relations and one axiom. The relations are precedence, dominance, and asymmetric c-command, and they are established between the symbols (terminal and nonterminal) that constitute phrase markers. The axiom is the Linear Correspondence Axiom (LCA), which states that a linear ordering of the terminal nodes of a phrase marker can be expressed as a function of the c-commanding relationships between the nonterminals that dominate them. Under this formulation, the structures generated by the X' schema follow from a single axiom.

Perhaps more importantly, other assumptions about structure that are not codified in the X' schema, i.e. the bans on multidominance, loops, discontinuous constituents, and floating constituents, all follow from the LCA. Villalba (2000) devotes

much of its first chapter to explain these consequences of the LCA, as well as Kayne's (1994) treatment of adjunction.

Villalba (2000) adds three axioms in order to clarify Kayne's assumptions. I present them in (24).

(24) *Axiom 1—there's a node that doesn't exclude any other node.*

Given the set of nodes N forming a phrase marker, $\exists S \in N$ such that $\forall Z \in N, Z \neq S, S$ doesn't exclude Z .

Axiom 2—there's a node that asymmetrically c-commands every other node.

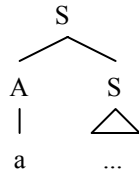
Given the set of nodes N forming a phrase marker, $\exists A \in N$ such that $\forall Z \in N, Z \neq A, A$ asymmetrically c-commands Z .

Axiom 3—there's a terminal that precedes every other terminal.

Given the set of terminals T forming a phrase marker, $\exists a \in T$ such that $\forall z \in T, z \neq a, a$ precedes z .

These axioms are to be understood as wellformedness conditions, not as generative rules, and they are a spell-out of some 'quite concise comments' by Kayne. They are necessary to provide the initial structure for the phrase marker, that Villalba represents as in (25):

(25)

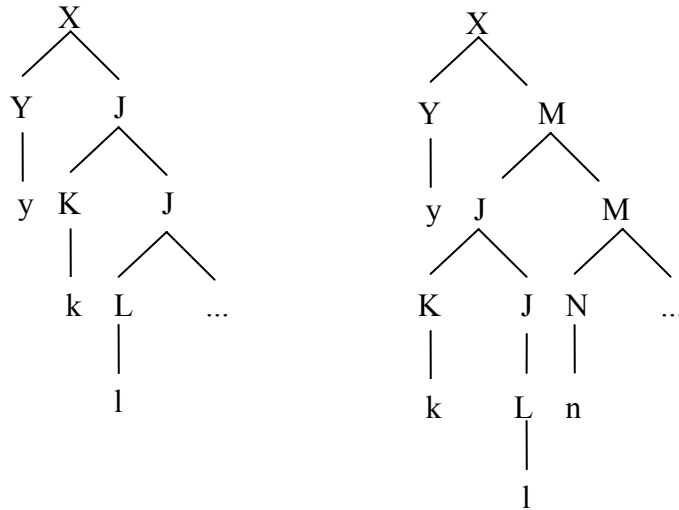


The axioms are meant to (1) state the existence of a root node S from which to establish the set of dominance relations, (2) state the existence of a nonterminal A that will asymmetrically c-command everything S dominates, and (3) state the existence of a terminal a from which to establish the precedence relation between terminals. At any rate, the nodes postulated by axioms 2 and 3 are to be understood as abstract, and adjoined to the root node (Kayne, 1994, p. 36).

Note, however, that the structure in (25) represents the adjunction of a head to a maximal projection. This configuration is not desirable, and Kayne himself (1994, p. 31), in Villalba's words, 'attempts to forbid heads in specifier position' by assuming that the highest element of a chain of heads must have a specifier. Villalba (2000, p. 27) disqualifies this assumption as *ad hoc* and states that the development of the system renders it superfluous. To see how his reasoning goes, consider (26), where K is the adjoined head. Given Axiom 2, there must be a node A asymmetrically c-commanding the whole structure. Therefore, for Villalba, the structure with the adjoined head K must be embedded in one of the two configurations presented in (26). In such configurations

there is no ordering for the pair formed by the terminal *k* of the adjoined head *K* and the terminal *y* corresponding to the node *Y* (remember that *J* does not dominate *K*).

(26)



Basically, Villalba's argument is that Kayne's stipulation of a specifier is not necessary because of a previous stipulation about a c-commanding node. However, Villalba's argument relies on the structures presented in (26) being a necessary consequence of Axiom 2, and it is difficult to see why they should be. Whether Villalba's argument is valid or we must accept the stipulation of a specifier for the highest head of a chain, the structure in (25) goes against the conclusions of both arguments. Kayne's abstract node *A* is at the same time a specifier and a head. It is beyond the scope of this work to try to

solve this contradiction, but it seems quite intriguing that the LCA can be presented as relying on an abstract node that appears to violate the LCA.

4.2 The “center periphery”

4.2.1 Jayaseelan’s study

Rizzi (1997) managed to come up with a structure that provided a common analysis for a number of previous observations and analyses of topic and focus. The common feature to the phenomena explained by Rizzi was that they were all related to the left periphery of the sentence. Jayaseelan (2001) takes the spirit of Rizzi’s conclusions and uses it as a tool to convincingly explain and link a series of facts that did not seem related at all. The big theoretical contribution that Jayaseelan makes is showing that the series of information related phrases that Rizzi found in the CP layer can also appear inside the IP, right above vP, and therefore they might be the escape hatch between phases Chomsky (1998) talks about.

Jayaseelan postulates a FocP below IP on the basis of Malayalam non-cleft questions, where “the question word must be placed immediately to the left of V in a position ‘normally’ occupied by the direct object if one is present, Malayalam being an

SOV language” (Jayaseelan, 2001, p. 39). If we accept Kayne’s (1994) framework, it is quite straightforward that Spec, FocP of a FocP right above vP would be an ideal landing site for question words, while everything else generated inside vP would move past it to the positions usually needed to create the SOV order.

Jayaseelan explores the possibility of a Topic position under FocP in Malayalam, but he can not find conclusive evidence. An iterable Topic position above FocP, however, seems more useful. Consider the following data from German:

- (27) a. ich habe meinem Bruder einen/den Brief geschickt
 I have my-dat. brother a/the letter sent
 ‘I have sent my brother a/the letter’
- b. ich habe den Brief meinem Bruder geschickt
 I have the letter my-dat. brother sent
- c. *ich habe einen Brief meinem Bruder geschickt
 I have a letter my-dat. brother sent

In German, the indirect object precedes the direct object. The order can be reversed by scrambling, but indefinite direct objects can not precede indirect objects (Abraham, 1986, p. 17; Lenerz, 1977, p. 54). If we take the scrambling to consist of the direct object being raised to a Topic position, the reason why indefinite direct objects can not be scrambled follows: “an indefinite NP which receives an existential interpretation is necessarily ‘new information’ and therefore cannot be a Topic”

(Jayaseelan, 2001, p. 54). Further evidence for TopicP is provided from German adverb placement and scrambling in Dutch and Malayalam.

The relative order of TopicP and FocP is visible with the help of adverbs. A German definite DP is usually topicalized and it occurs to the left of an adverb/neg position. If it appears to the right of that position, this can only mean that it is in FocP and we should expect it to require a contrastive focus interpretation. An indefinite DP, however, is under no pressure to raise to TopP, so a non-focused interpretation is possible. Data from Diesing (1997, p. 378) for German confirms this prediction:

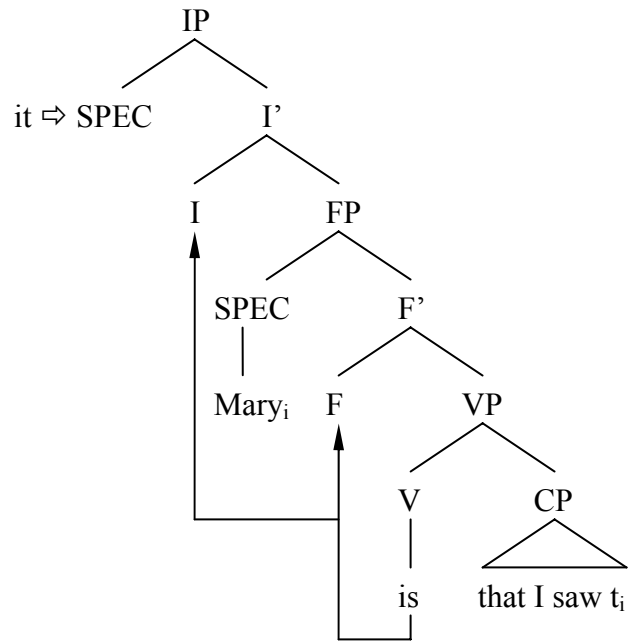
- (28) a. ... weil ich selten die Katze streichle
 since I seldom the cat pet
 ‘... since I seldom pet the CAT’

Jayaseelan provides additional evidence from Yiddish facts described by Diesing (1997), that I will not discuss here.

Postulating an IP internal FocP also provides an immediate explanation for English clefts (Jayaseelan, 2001, p. 63) and the fact that these constructions are a way to single out focalized elements:

(29) a. It is Mary that I saw.

b.



Of course, the same account appears plausible for Malayalam clefts. Thus, Jayaseelan examines data from Malayalam, German, Dutch, Yiddish, English, and a few other languages, and finds possible explanations using a structure that differs from Rizzi's proposal only in that it does not allow for Topic phrases below Focus, which is an observation made for Spanish (Campos & Zampini, 1990) and for Italian after Rizzi's data was reanalyzed (Benincà, 2001; Benincà & Poletto, 2004).

One issue open to further research would be whether at least some data of focus in-situ should be reanalyzed as involving IP-internal focus phrases, but undertaking this is well beyond the limitations of the present work.

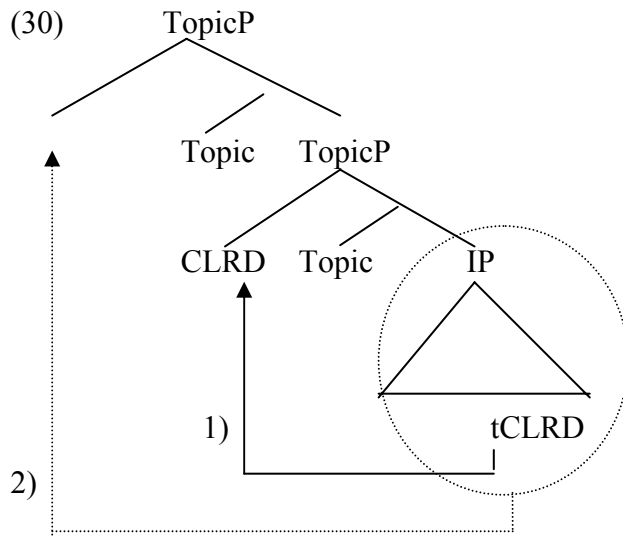
4.2.2 Villalba's proposal for Catalan

Villalba discusses and rejects three syntactic analyses of CLLD and CLRD in the fourth chapter of his dissertation (Villalba, 2000), where he brings up more syntactic properties of both constructions and carries out a more extensive comparison than he did in the descriptive section of the dissertation. Thus, the discussion against a symmetrical analysis of CLLD and CLRD (as, respectively, left and right-adjuncts) (Benincà, Salvi, & Frison, 1988; Valioui, 1994; Vallduví, 1990, 1995) serves the purpose of pointing out that, in spite of the fact that both constructions are opaque to wh-extraction, they differ in several respects: CLLD is unbounded while CLRD is upward-bounded; CLLD induces island effects, while CLRD does not (Villalba, 1996); CLLD destroys the context for licensing of Negative Polarity Items, while CLRD does not; CLLD shows antireconstruction effects, while CLRD does not; CLLD affects the bounding by quantifiers of CLLDed pronouns, while CLRD does not; CLLD is possible from left and right-dislocates, while CLRD is not possible from left-dislocates; CLLD cannot freely apply in nonfinite clauses, while CLRD can.

In sum, CLRD properties seem to suggest that what really matters about its syntax is its linear ordering with respect to the rest of the sentence. This should not be expected, were CLRDed elements right-adjuncts, since they would be in the same c-commanding position as CLLDed elements.

Kayne's (1994) own analysis obviously does not need to explain the problems of right-adjunction structures, since right-adjunction is not part of his system. Right-dislocated phrases would be in complement position, and CLRD would be an instance of covert CLLD (in Kayne's own words 'LF movement of the CLLD type'). This analysis immediately provides a solution to the differences between CLLD and CLRD with respect to upward-boundedness, island effects, and licensing of NPIs. However, this analysis does not seem to provide an obvious account for the free ordering of right dislocates, or their opacity to *wh*-extraction, or any of the other issues raised by Villalba. In addition, it does not account for the different informational roles described for CLLD and CLRD.

Cecchetto (1999) and Zubizarreta (1998), following an idea proposed in 1995 class lectures by Kayne, propose a third analysis. In this analysis, both CLLD and CLRD would consist of a movement to TopicP, which, in the case of CLRD, would be followed by the movement of everything left below TopicP to a second, higher, TopicP. I copy Villalba's schema for CLRD in (30):



This kind of ‘remnant movement’ analysis (Besten & Webelhuth, 1987, 1990; Müller, 1998, 2000; Thiers, 1985) is one of the logical consequences of Kayne’s antisymmetry, and has been extraordinarily prolific in the last ten years or so.

Villalba finds this solution quite problematic as it stands. Besides the lack of a clear motivation for IP movement (which, in any case, would be a critique to any analysis of the Antisymmetry type) the resulting structure does not explain many of the phenomena Villalba uses as tests. For ease of reference, I summarize Villalba’s conclusions in favor and against each analysis in the following table. The check mark indicates that the effect on the left column supports the analysis above. The cross indicates that it does not. The number in brackets indicates the page in his dissertation (Villalba, 2000) where the particular effect is discussed for each analysis:

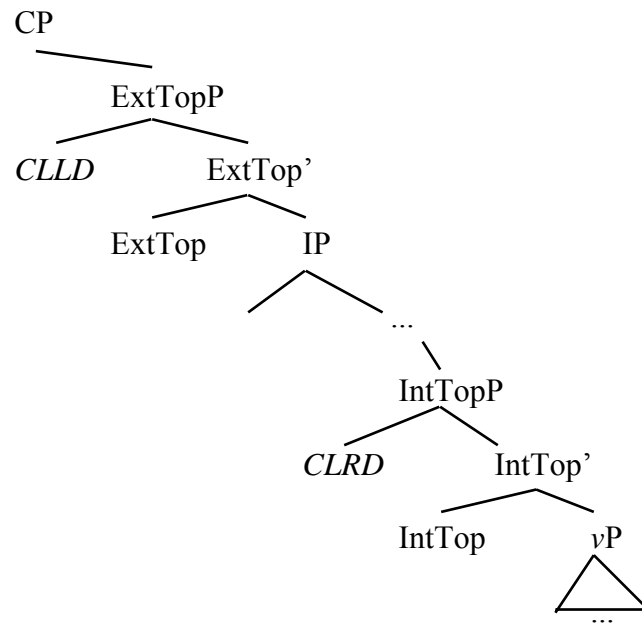
(31)

	right-adjoined CLRD (the symmetric analysis)	CLRD as covert CLLD	Double Topicalization
free dislocate ordering	✓ (183)	✗ (199)	✓ (206)
opacity	✓ (185)	✗ (201)	✗ (209)
boundedness	✗ (186)	✓ (196)	—
island effects	✗ (188)	✓ (197)	✓ (207)
licensing of NPIs	✗ (189)	✓ (197)	✗ (208)
principle C	✗ (190)	✗ (201)	✗ (209)
pronouns bound by a quantifier	✗ (191)	✗ (202)	✗ (210)
interactions between CLLD and CLRD	✗ (192)	✗ (203)	✗ (211)
non-finite clauses	✗ (193)	✗ (204)	✗ (212)
informational status	—	✗ (205)	✗ (213)

4.2.2.1 Villalba's analysis

The fifth and last chapter of Villalba's dissertation is devoted to presenting and defending his own proposal of a combined analysis for CLLD and CLRD. I will briefly summarize his analysis here. Villalba's proposal for a structure is presented in the following tree:

(32)



This proposal splits TopP in two, one of them (External Topic Phrase) above IP, and the second one (Internal Topic Phrase) just above vP. The proposal is meant to provide different structural positions for CLLD and CLRD while preserving the intuition, present in the last two analyses discussed in section 4, that there is an important shared component in CLLD and CLRD, and that they should therefore have a common position in the structure at some point in the derivation.

In Villalba's proposal, CLLD and CLRD are feature-driven movements, and CLLD consists of two successive movements: the first of them being precisely CLRD and the second one being triggered by a second feature. The idea of CLLD being CLRD plus a second movement is taken from Postal (1991). Following the interpretive

properties described for the constructions, Villalba gives the name of [background] to the feature that triggers CLRD and the first part of CLLD, and [link] to the feature that triggers the second part of CLLD. Villalba points out an interesting parallelism with focus, which by some accounts can also be described as constituted by two different features and occupying two different structural positions. This parallelism is not pursued in depth by Villalba (2000), who refers the reader to Choi (1996) for further discussion. I copy in (13) Villalba's adaptation of Choi's idea:

(33)

	-background	+background
-prominent	<i>unmarked focus</i>	<i>CLRD</i>
+prominent	<i>marked focus</i>	<i>CLLD</i>

Villalba easily fits the position of his ExtTopP within then current left periphery accounts (Benincà & Poletto, 1999; Rizzi, 1997), and it would also fit within more recent developments (Benincà, 2001; Benincà & Poletto, 2001). The position of IntTopP is, however, derived indirectly, from its relative position respect to a pre-Rizzi idea of focus immediately dominating vP (Belletti & Shlonsky, 1995). No common framework for both positions is developed, and all that can be said is that the positions, are, at most, independently plausible.

The split-topic analysis presents the advantage that it addresses most of the issues that were problematic for previous analyses. The only issue that is not completely clear is the incompatibility of CLLD with non-finite clauses. Villalba presents some examples where the presence of a *wh*-element (but not a *wh*-relative) and, therefore, probably of a richer functional structure over the nonfinite verb, seems to license CLLD. Villalba does not examine the issue any further, which is surprising, since this could be the key to justifying the topic position below IP that he needs for his IntTopP.

The bulk of the last chapter of Villalba's dissertation is devoted to counter Cinque's (1990) arguments that CLLD does not involve movement at all. As we have seen, this is a crucial point of Villalba's analysis, and he is quite successful in showing that all that is proved by Cinque's arguments is that the behavior of CLLD is incompatible with *wh*-movement. This raises the issue, of course, of what other kinds of movement are available. He adapts the following table from Borer (1995) to include the properties of CLRD and CLLD:

(34)

	A-chain	O-chain	S-chain	CLRD	clause-internal CLLD	long-distance CLLD
<i>case-marked tail</i>	no	yes	yes	yes	yes	yes
<i>new binding relations</i>	yes	no	(yes)	no	no	no
<i>licensing of parasitic gaps</i>	no	yes	yes	no	no	no
<i>WCO effects</i>	no	yes	no	no	no	yes
<i>creation of islands</i>	no	no	yes	yes	yes	yes

To understand the chart, take into account that Borer “argues that alongside with A(rgument) and O(perator)-chains —corresponding to typical A- and A’-chains—, there exist S(crambling)-chains, which share properties of both A- and O-chains. The constructions involving S-chains include XP-topics in Hebrew, short scrambled elements in Japanese, Hindi, and German, and topicalization in English.” (Villalba, 2000, p. 271). Note that Villalba has to postulate two different kinds of CLLD.

Suspensions of the existence of yet a third kind of left dislocation are raised when Villalba confronts data where CLLD is available but CLRD is not, against the predictions of his analysis. In such cases, against the norm for CLLD, a preverbal subject is banned. Villalba suggests that these cases involve a different construction and leaves the issue open.

Villalba addresses the issue of multiple, freely ordered, cases of CLLD and CLRD by proposing that dislocates form clusters, rather than a succession of independent elements. This would explain, in principle, the lack of ‘topic scope’ interpretive differences between orders, but the issue is also left open.

4.2.2.2 Orientation of further research

Villalba does not discuss what happens after a CLRDed element reaches its position above vP . The leftward movement produces a right-detachment after, we are to suppose, some kind of remnant movement takes place. Otherwise the elements left behind inside vP would surface to the right of the right-dislocate, configuration that Villalba shows to be impossible for CLRD (but not for afterthoughts) in his chapter 3.

This issue, obviously, poses further questions. While it seems plausible that CLRD results from a movement operation (see above for island effects, etc.), the scope facts that lead Villalba to place the landing site of this movement below IP are based on the idea that at some point negation or quantifiers have scope over the CLRDed element. But, if we just look at a sentence with a right dislocation, it is immediately obvious that, at least in the surface, everything has scope over a right-dislocated element. Therefore, the same scope effect would be expected to appear in the absence of movement. Villalba's strongest argument, in my view, is the one, barely developed, about the CLLD being incompatible with non-finite clauses, while CLRD is compatible with them. A careful study of the phrases involved in the remnant movement suggested by Villalba would probably provide further evidence on the issue of where CLRDed elements end up.

On the other hand, it is interesting that research completely independent from Villalba's, like Jayaseelan's (2001) on Malayalam, German, and English, points at the actual presence of IP-internal topic positions. This, added to proposals for internal focus positions (Belletti & Shlonsky, 1995; Cecchetto, 1999; Ndayiragije, 1999) seems to lend support to the idea that such positions are there and allow for different semantic interpretations than left-dislocated structures. A comparison between the interpretive properties of Villalba's (2000) CLRD and Jayaseelan's (2001) IP-internal topics would be too broad to attempt it here, but it surely would be worth the effort. Meanwhile, in the absence of sufficient syntactic evidence of the actual sentence structure, all attempts of an explanation of the interaction of syntax and semantics in these constructions will remain, to use Villalba's own words, 'quite tentative'.

To sum up, the work in Villalba's dissertation could be continued by a) explaining the implications (and the implementation) of the remnant movement included in his proposal, b) giving an independently motivated mapping of the structural positions involved in CLRD, and c) clarifying the relationship between CLLD and other very similar constructions, as well as topicalized subjects. While this last task involves enough complexity for another dissertation, the first two seem necessary to uphold Villalba's hypothesis beyond its status of mere compatibility with other proposals.

Villalba's dissertation presents a useful overview of the issues at hand in the study of dislocations. In that respect, one of the merits of Villalba's dissertation is to show the usefulness of Catalan in determining what kind of construction we are facing. The impressive array of clitics available in that language allows Villalba to reject the iteration of Hanging Topics (contra Escobar, 1995), to give a CLLD analysis of the fronting of certain non-nominal phrases (contra Cinque, 1990), or to show (using data from Bonet, 1991) that there are clear instances of CLLD that cannot be derived from clitic-doubling (contra Cecchetto, 1999; Jaeggli, 1982; Kayne, 1994).

As for its theoretical background, Villalba's dissertation shows the interaction of both trends towards restrictive accounts found in current syntactic research (Kayne's (1994) restrictions on the structure and Chomsky's (1993; , 1995; , 1998; , 1999) more general theoretical moves). However, in spite of Villalba's assertion that his hypothesis follows Rizzi's demands of a strict syntax-semantics mapping, it departs from Rizzi's (1997) line of research in a fundamental point. If Rizzi's (1997) most appealing idea is his hypothesis that there are three distinct, meaningful layers of functional projections, then Villalba's split-topic hypothesis goes against the main point of Rizzi's (1997) article.

The previous discussion is relevant for this dissertation in that it provides an exhaustive set of tests for right dislocations. Chapter 4 explores the presence of right dislocations within the context of a question.

4.2.3 Poletto's view on DP-internal Focus

The study of discourse-related positions in Old Italian syntax undertaken by Cecilia Poletto has also yielded evidence of internal Topic and Focus. The Focus in the vP phase displays the same properties noted for Focus under CP by Benincà (1995). The next logical step is to look for DP-internal Focus. She finds indirect evidence in the fact that adjectives in prenominal position can be modified, which she sees as a suggestion that the whole adjectival phrase has moved to the left of the noun.

A very interesting aspect of Poletto's hypothesis is the contrast between Old and Modern Italian. The simultaneous loss of V2 and scrambling inside the DP would be due to the dependence of both phenomena on a single hypothetical property of the Focus head, which would have changed from Old to Modern Italian.

4.3 Focus sightings

The syntactic approach to Focus presented so far implies postulating the presence of Focus-specific positions in the sentence structure. Perhaps the strongest argument to support this view would be to show not only sentence elements occupying such positions, but some kind of visible Focus particles as well. In the following sections I give an overview of several proposals that make a reference to Focus particles.

4.3.1 Italian

Particles that are correlates to focus have been proposed for Italian by Brunetti (2003), who proposes *anche* in sentences like (35) to be a realization of a Focus morpheme.

- (35) $[_{DP} \text{Bill}]_i [_{IP} t_i \text{ mangia } [_{VP} t_i t_{\text{verb}}]]$ [e $[_{DP} \text{Paolo}]_j [_{CP} \text{anche } [_{IP} t_j t_{\text{verb}}]]$ mangia $[_{VP} t_j t_{\text{verb}}]]$]
Bill eats and Paolo too eats
'Bill eats and Paolo eats, too'

In this example both subjects are at a position where they can be interpreted as Topics (even though Brunetti does not accept a specific Topic Phrase) and the deletion of the IP “does not imply deletion of the Focus morpheme” (Brunetti, 2003, p. 185).

Brunetti (2003) also mentions Donati's (2000) account of vacuous negation in comparatives. In the following example the negative word has no negative force:

- (36) Maria ha mangiato più biscotti che (non) Paolo.
Maria has eaten more cookies that (not) Paolo
'Maria has eaten more cookies than Paolo'

Donati points out that in order for this construction to be grammatical the second element of the comparison must be focused, as shown by the contrast below:

- (37) a. *Maria mangia più biscotti di quanti Paolo ne MANGI.
Maria eats more cookies of how-many Paolo of-themCL eats
- b. Maria mangia più biscotti di quanti ne mangi PAOLO
Maria eats more cookies of how-many of-themCL eats Paolo
'Maria eats more cookies than Paolo'

In this example it is necessary for the subject to bear the nuclear stress, associated to Focus, for the sentence to be grammatical. Donati concludes that the negative word *non* is the scope marker of a Focus operator.

4.3.2 Somali

Somali has an obligatory marker for Focus (Frascarelli, 2000; Frascarelli & Puglielli, in press; Puglielli, 1997):

- (38) a. Cali yuu dilay?
 Cali whom-Focus marker-3SG beat PAST
 ‘Who did Cali beat?’
- b. Cali Maryam **buu** dilay.
 Cali Maryam Focus marker-3SG beat PAST
 ‘Cali beat Maryam’

Brunetti (2003) points out that in this language, the Focus morpheme is on the right of the focused phrase, which is expected, given the fact that Somali is C final.

4.3.3 Old Italian

Poletto (2004) proposes, from an original idea of Benincà (1995), phonetic realizations of functional heads in Old Italian. In the following examples, Poletto claims that *sì* can be seen as a Spec Focus expletive particle:

- (39) La volpe andando per un bosco *sì* trovò un mulo: e il mulo *sì* li
 The fox walking through a wood *sì* found a mule and the mule *sì* to-her
 mostrò il piede dritto
 showed the foot right
 “While the fox was walking through a wood she found a mule, and the mule
 showed her right foot to her”
 (The focused material here could be *per un bosco*, *andando per un bosco*, or *la
 volpe andando per un bosco*)

- (40) Lo mperadore Federigo stando ad assedio a Melano, sì li si fuggi un
 The emperor Federico being at siege at Milan sì to him itself escaped a
 suo astore dentro a Melano (*Novellino*, p.177, r.36)
 of his hawk inside to Milan
 “While the emperor Federigo was at the siege of Milan, a hawk of his escaped
 inside Milan”
- (41) E parlandomi così, sì mi cessò la forte fantasia entro quello punto
 And talking to me so sì to me stopped the strong fantasy inside that point
 ch’io volea dicere...
 that I wanted to-say
 “And in talking to me like that, the strong fantasies about the point I wanted to
 make stopped for me”

4.3.4 American Sign Language

Colin, Hagstrom, and Neidle (2003) have identified an indefinite focus particle in American Sign Language (ASL), which serves “to widen the domain of items referred to by a focused wh-phrase or indefinite quantifier such as ‘someone’” (Conlin, Hagstrom, & Neidle, 2003, p. 26). To understand the role of this particle, consider that in ASL wh-questions can remain in situ or appear in the right periphery of the sentence. The interpretation of wh-questions is different in each case. When the wh-phrase occurs at the right periphery, the wh-phrase is focused, in the sense that it carries some kind of existential presupposition (compare with É. Kiss’ (1998) identificational focus) that is not present in wh-in situ questions (Neidle, 2002; Neidle, Kegl, Bahan, Aarons, & MacLaughlin, 1997; Neidle, Kegl, MacLaughlin, Bahan, & Lee, 2000; Neidle &

MacLaughlin, in press; Neidle, MacLaughlin, Lee, Bahan, & Kegl, 1998a, 1998b). Take the following examples (from Neidle, 2002, p. 75):

- (42)
- a. wh
 arrive who
- b. wh
 arrive who
- Who* arrived?

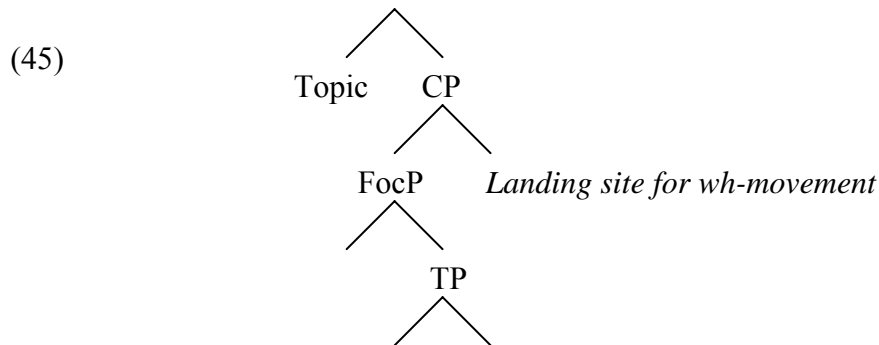
- (43)
- a. wh
 *who arrive
- b. wh
 who arrive
- Who arrived?

The line above the text represents the presence of the non-manual wh- sign, which can optionally spread to the whole utterance when the wh-phrase is in the right periphery, while it has to spread obligatorily if the wh-phrase appears in situ. The examples in (42) correspond to the focused wh-phrase, which, according to Neidle (Neidle, 2002), have a meaning that corresponds with a particular intonation of English discussed by Hajičová (Hajicová, 1976, 1983), which precludes “nobody” as an answer to the question. “Nobody” would be unremarkable as an answer to the wh-in situ questions.

The focus particle appears at the right periphery of a focused wh-question, frequently contracted with the wh-sign, but it does not appear when the wh-phrase remains in situ. It appears in the right periphery of indefinite yes-no questions and negative sentences, and it can appear immediately following or simultaneously (using the non-dominant hand) with focused constituents.

- (44)
- a. $\frac{\text{wh}}{\text{see Joan who}^{\text{part:indef}}}$
Who saw Joan?
- b. $^?*\frac{\text{wh}}{\text{who}^{\text{part:indef}} \text{ see Joan}}$
 Intended meaning: Who saw *Joan*?
- c. $\frac{\text{y/n}}{\text{IX-2p see something/one part:indef}}$
 Did you see anyone?
- d. $\frac{\text{neg}}{\text{IX-1p never seen John POSS car part:indef}}$
 I've never seen John's car
- e. $\frac{\text{y/n}}{\text{something/one}^{\text{part:indef}} \text{ see Joan}}$
 $\frac{\text{y/n}}{\text{something/one see Joan}} \quad [\text{d}]$
 $\text{part:indef -----} \quad [\text{nd}]$
 Did *anyone* see Joan?

The different distribution of the particle with wh-elements can be interpreted to mean that, in order to reach the marked interrogative position in the right periphery, the wh-element needs to reach the Focus Phrase first, from where it will move to its final position to the right in a structure like this:



This structure is of course incompatible with Kayne’s (1994) analysis. Neidle and others have argued elsewhere (Neidle, Kegl, Bahan, Aarons, & MacLaughlin, 1997; Neidle, MacLaughlin, Lee, Bahan, & Kegl, 1998a, 1998b) for rightward movement (as opposed to leftward movement followed by remnant movement) (Petronio & Lillo-Martin, 1997). Regardless of what interpretation is correct, Neidle’s work shows a) a visible focus particle, and b) a separation of focus and interrogation.

4.3.4.1 Further comments on ASL and the left periphery framework

In addition the contribution of a visible correlate of focus, it is worth noting that Neidle's conclusions about the relative order of Focus and Topic in ASL are fully compatible with the structure proposed by Benincà (2001) for Italian (see (46) below, copied from (19) above), as well as with all but one of the findings summarized in Campos and Zampini (1990) for Spanish and presented above in (5), which I also copy here for ease of reference. The observations that are compatible are (47 a, b, and d). I will get back to the distribution of wh-movement and focus in the next chapter.

(46) [DiscP [ForceP [TopP [FocP [FinP [IP...

- (47) a. Topic precedes focus.
b. Topic precedes interrogative words.
c. Focus and wh-movement are in complementary distribution.
d. There can be more than one topic but only one focus.

There are also particular, different, head-movement⁴ markings associated to Focus and Topic phrases, in a manner that appears to be exactly analogous to the way the A-accent and the B-accent described by Jackendoff (1972) (see chapter 1, section 3.2.3) are associated to Focus and Topic in many spoken languages.

⁴ 'Head' meant in the anatomical sense. That is, the person moves his or her head in a particular way to mark Focus or Topic phrases.

5 Conclusions

The idea of a direct correlation between meaning, intonation, and syntax looks more appealing as descriptive work progresses. I believe that anyone who reads Büring's (1997) book on the meaning of topic and focus after studying the syntax of dislocations will immediately come up with the idea that CLLD is contrastive topicalization; in fact, Karlos Arregi just gave that title to a paper of his (Arregi, 2003). Plenty of work is yet to be done in this direction, but there is a clear path.

As for the nature of information within a theory of grammar, it is now clear that a syntactic theory will have to account for the recent ubiquity of information-related functional projections proposed in the literature. Cartography-style proposals might be too powerful to be the explanation to phenomena that, after all, were not even considered part of the realm of syntax until relatively recently. In this respect, it is worth taking into account that the appearance of focus within IP and DP could call for a different interpretation. Brunetti (2003), for instance, thinks of focus in terms of a morpheme, which gives her more flexibility in explaining its placement. Erteschik-Shir (to appear) proposes a view of topic and focus as features that can be associated to lexical items, parallel to ϕ -features but not restricted to particular categories nor required. Proposals of this kind, however, follow from the desire to provide a common

explanation for dislocated and in-situ cases of focus (and topic) while avoiding violations of Chomsky's inclusiveness condition.

For the purposes of this dissertation I will accept the idea that focus and topic are associated to functional projections. Once information focus is left out of the picture, following the line of thought of É. Kiss (1998) and Domínguez (2004), a functional-head approach, Rizzi style, provides a common explanation for many word order phenomena, is compatible with intonation phenomena, and does not require any major changes to the model of grammar, particularly if we adopt Kayne's approach.

Appendix – Villalba’s classification of Catalan constructions

A. Left-detachments

Two chapters of Villalba’s (2000) dissertation deal, respectively, with Catalan dislocations to the left and to the right. Four types of left dislocation are described, namely Clitic Left Dislocation (CLLD), Hanging Topic Left Dislocation (HTLD), *as for* constructions, and metalinguistic topics (plus an appendix on verbal detachments, which he considers to be instances of CLLD). Of these, CLLD and HTLD have been extensively studied since Cinque (1977; , 1983; , 1990). The *as for* construction is frequently assumed to be a case of HTLD, and metalinguistic topics are described for the first time in Villalba’s dissertation.

Villalba provides a description of the syntactic and interpretive properties of each of these constructions, as well as examples from languages other than Catalan where the constructions have been attested. The following table (expanded version of the one in Villalba, 2000, p. 81) summarizes the syntactic properties for CLLD and HTLD, as well as the *as for* constructions (the numbers refer to the page of the dissertation where the properties are described):

(48)

	CLLD	HTLD	<i>as for</i>
I. category neutral	<i>yes (45)</i>	<i>NPs only (82)</i>	<i>NPs only (106)</i>
II. iterative	<i>yes (46)</i>	<i>no (85)</i>	<i>no (106)</i>
III. non-root contexts	<i>yes (47)</i>	<i>no (87)</i>	<i>yes (107)</i>
IV. free ordering of the dislocates	<i>yes (48)</i>	—	—
V. obligatory resumptive clitic	<i>yes (50)</i>	<i>resumptive, no clitic (90)</i>	<i>no resumptive (109)</i>
VI. ordering with respect to wh/C	<i>C-CLLD-wh (54)</i>	<i>C-HTLD-wh (92)</i>	<i>C-as for-wh (110)</i>
VII. connectedness	<i>yes (58)</i>	<i>no (93)</i>	<i>no (110)</i>
VIII. island sensitivity	<i>yes (60)</i>	<i>no (95)</i>	<i>no (111)</i>

The syntactic properties of HTLDs are less clear than those of CLLDs. Escobar (1995) describes Spanish HTLDs as iterative (49a), and she states that they appear in embedded contexts (49b). Villalba points out that Spanish presents particular difficulties to differentiate CLLD from HTLD, and provides examples of clearly ungrammatical iterated HTLDs (49c-e) as well as a contrasting minimal pair of Spanish (ungrammatical) HTLD (49f) and (grammatical) CLLD in an embedded context (49g). Villalba dismisses other examples provided by Escobar, as they show instances of *as for* constructions rather than HTLDs.

- (49) a. Trajes, Juan, en aquella tienda, (allí) ya no me ha comprado (él) (muchos) más.
clothes, Juan, in that shop (there) already not me has bought (he) (many) more.'
- b. No sé este libro quién lo podría reseñar para mañana.
'I don't know who could review this book for tomorrow.'
- c. *Juan, María, él habló con ella.
lit. 'Juan, Maria, he talked with her.'
- d. *El libro, Juan, María, él habló de él con ella.
lit. 'The book, Juan, Maria, he talked about it with her.'
- e. *Barcelona, Juan, María, él se divorció allí de ella.
lit. 'Barcelona, Juan, Maria, he get divorced there of her.'
- f. *No sé Juan quién lo podría examinar mañana.
'I don't know who could examine Juan tomorrow.'
- g. No sé a Juan quién lo podría examinar mañana.
'I don't know who could examine Juan tomorrow.'

Villalba's account of the interpretive properties of CLLD starts with a discussion of the properties that allow dislocability. He discards definiteness (certain definites can not be dislocated) and specificity (indefinite specific NPs are not dislocable), before taking partially ordered set relations (poset relations) as the key for dislocability, in the way proposed by Ward (1985) for English topicalization.

(50) *Discourse Condition on Preposing in Topicalization*

The entity represented by the preposed constituent must be related, via a salient partially ordered set relation, to one or more entities already evoked in the discourse model.⁵

Vallduví (1990) provides a less abstract way to express this same idea with his notion of *link*, which is ‘an instruction for updating the information conveyed by the focus under the file denoted by the CLLDed element’. Villalba presents this notion of *link* as the most appropriate way to interpret CLLDed elements, together with the more semantic view that CLLD functions as a restriction of the domain of quantification of the sentence, that he takes from Partee (1999). Villalba does not compare Vallduví’s information packaging proposal with Partee’s semantic one, and simply gives them as the ones that best reflect the actual uses of CLLD.

The interpretive properties of HTLDs have been less studied, and the best explanation still seems to be the one by Cinque (1983): ‘In HTLD the lefthand phrase is used to bring up or shift attention to a new or unexpected topic’. This view is widely shared (Bartra, 1985; Dolci, 1986), and it coincides with the description of English left dislocation (Reinhart, 1981; Rodman, 1977).

⁵ A partially ordered set is one defined by a relation that is either reflexive, antisymmetric and transitive or irreflexive, asymmetric, and transitive (Wall, 1972, p. 141). Paul Hagstrom (p.c.) points out to me a shortcoming of this definition, namely that we still need to know on what dimension things are being compared such that we can relate one to another, before we inquire as to whether the relation between them is antisymmetric, transitive, etc.

With his separate analysis for *as for* construction Villalba departs from the existing literature for Spanish (there was no literature on Catalan), which usually ignores the construction or assimilates it to HTLDs or left dislocation. The main distinctive syntactic property of the *as for* construction is that it does not require any resumptive element, unlike CLLD, which requires a clitic, or HTLD, which requires some other resumptive element. With respect to its interpretive properties, Villalba explains that ‘the *as for* construction introduces the about-topic, namely the standpoint on which the content of sentence must be evaluated’, but he does not show a clear contrast with the interpretation of HTLD. Again, like CLLD and HTLD, *as for* constructions are shown to restrict the domain of quantification of the sentence, thus having an impact on its truth conditions.

Metalinguistic topics are Villalba’s own descriptive contribution. They are only possible with interrogative intonation and in a context where the topic has been uttered in conversation right before. Two examples follow:

- (51) a. A: Demà aniré al cine amb la Maria.
‘Tomorrow I am going to the cinema with Maria.’
- B: *Amb la Maria* (diu)? Ahir parlàvem d’ella.
with the Maria say-2 yesterday talked-2PL of-her
‘With Maria (you say)? We talked about her yesterday.’

- b. A: Aquest noi és el fill de la Maria.
'This boy is Maria's son.'
- B: *De la Maria* (dius)? Doncs bé, ahir vaig parlar *amb ella*.
of the Maria say-2 so well yesterday talked-1 with-her
'Maria's (you say)? I talked with her yesterday.'

The space devoted to this construction in Villalba's dissertation is quite small, and its treatment surprisingly superficial in comparison to the systematic analysis he offers for the rest of the constructions. In my opinion, establishing that the given examples are, in fact, left detachments, is, at least, difficult. The apparent resumptive element shown in the examples in (7) might not be anything other than a case of discourse anaphora, and the separate intonational pattern might very well be telling us that the leftmost element is not associated to the rest of the utterance by anything syntactic.

B. Right-detachments

As for right-detachments, Villalba notes that they are much less studied than those to the left, due to their oral character and to the fact that they have been traditionally banned by prescriptive grammars. He finds two kinds of right detachments in Catalan. One is the Clitic Right Dislocation (CLRD), and the other is afterthoughts, purely discursive elements that have often been analyzed together with CLRD. Villalba collects instances of CLRD in many Romance languages and Greek, and he warns against possible confusion with right-detachments from other languages, like English or

Chinese Right Dislocation, where the resumptive element is a strong pronoun, or the Right Scrambling found in SOV languages like Hindi or Japanese, where no (overt) resumptive appears at all.

The syntactic properties of CLRD are shown to be quite similar to those of CLLD: it affects maximal projections of any category; is iterative; occurs in both root and embedded contexts; in cases of multiple CLRD, the ordering of the dislocates is free; there must be a resumptive element and it must be a clitic pronoun; there is obligatory connectedness between the dislocate and the resumptive element; the relation between the dislocate and the clitic is subject to (strong) island constraints. Afterthoughts, on the other hand, have a much freer status that derives from their character of discourse-repairing devices. I gather Villalba’s conclusions in the following table:

(52)

	CLLD	CLRD	<i>afterthoughts</i>
I. category neutral	<i>yes (45)</i>	<i>yes (136)</i>	<i>yes (155)</i>
II. iterative	<i>yes (46)</i>	<i>yes (137)</i>	<i>yes (156)</i>
III. non-root contexts	<i>yes (47)</i>	<i>yes (138)</i>	—
IV. free ordering of the dislocates	<i>yes (48)</i>	<i>yes (139)</i>	<i>no (158)</i>
V. obligatory resumptive clitic	<i>yes (50)</i>	<i>yes (140)</i>	<i>no (140)</i>
VI. ordering with respect to wh/C	<i>C-CLLD-wh (54)</i>	<i>N/A</i>	<i>N/A</i>
VII. connectedness	<i>yes (58)</i>	<i>yes (143)</i>	<i>no (158)</i>
VIII. island sensitivity	<i>yes (60)</i>	<i>yes (143)</i>	<i>no subjacency effects (159)</i>

French right-dislocates have been claimed to be able to surface as parentheticals, i.e. in intrasentential positions (Larsson, 1979; Postal, 1991). Villalba analyzes the Catalan counterparts of their examples and concludes that the parentheticals are in fact instances of afterthoughts, rather than right-dislocates, as it might likely be the case in French as well.

Villalba's statement that 'in cases of multiple afterthoughts, their ordering is fixed' seems at odds with the extreme freedom of placement that afterthoughts enjoy otherwise. In fact, the examples he provides (reproduced in 9) seem to portray the behavior of complements within a single afterthought. Thus, Villalba's assertion that 'the order of afterthoughts is a strict copy of the normal order of complements' should be reformulated as 'the order of complements within an afterthought is a strict copy of the normal order of complements'. This would simply reflect the fragmentary nature of afterthoughts, which produces a lack of structural positions susceptible of being the target of movement within them.

- (53) a. En Joan i en Miquel estaven barallats, però finalment li va demanar perdó, en Miquel a en Joan/*a en Joan en Miquel, vull dir.
'Joan and Miquel were fallen out, but finally he apologized, Miquel to Joan, I mean.'
- b. Sabia que en Joan i en Miquel estaven barallats, però no sabia que l'havia agredit, en Joan a en Miquel/*a en Miquel en Joan, vull dir.
'I knew that Joan and Miquel were fallen out, but I didn't know that he hit him, Joan to Miquel, I mean.'

Determining whether afterthoughts are freely ordered would require a clear criterion to determine whether we are facing one or more afterthoughts. This criterion would probably be intonational.

While CLRD appears quite parallel to CLLD from a syntactic point of view, its interpretive properties are quite different. The approach to CLLD based on poset relations doesn't extend to CLRD, and while there is a wide consensus that CLRD conveys given information (Bartra, 1985; Benincà, Salvi, & Frison, 1988; Birner & Ward, 1998; Fabra, 1956; Grosz & Ziv, 1998; Laca, 1986; Lambrecht, 1994; Larsson, 1979; Vallduví, 1990), there is not much agreement beyond that. Villalba discusses and rejects proposals a) that CLRD is an indirect focalization construction, in the sense that it allows other elements to receive focus by position, by removing itself from the final position of the sentence (Laca, 1986) and b) that 'it indicates that part of the proposition communicated is knowledge already contained under that address and that the information of the sentence must be construed in some way with that knowledge instead of merely added' (Vallduví, 1990, p. 78). His own position is that CLRD 'helps the speaker to signal to the hearer that he must shift his attention to a referent that wasn't immediately available'. This seems to sum up both contradictory traits of CLRD, namely its character of given information and its ability to bring up new information, but it is obvious that it does not go beyond an intuitive informal description of its pragmatics. Villalba goes on to describe the semantics of CLRD within Partee's (1999)

framework, to conclude that, as in the case of CLLD, it restricts the quantification domain of the sentence, with the proviso that this happens ‘either directly or as a consequence of not being part of the focus’. This leaves us wondering whether it would not be better to take focalization to be the process that actually restricts the quantification domain, and acknowledge that there is not much we can actually say about the semantics of CLRD. In fact, this criticism is parallel to the one Villalba correctly makes of Laca (1986).

The pragmatics and semantics of afterthoughts is comparatively very simple: they are mere devices to add supplementary information or solve ambiguities, not discourse organization mechanisms.

CHAPTER 3 – FOCUS, ANSWERS, AND QUESTIONS

1 Introduction

The idea that there is a relationship between focus and wh-questions is present, explicitly or implicitly, in much of the work devoted to each of those topics. This chapter aims to define the terms of that relationship.

Background concepts

Before trying to spell out the exact terms of the relationship between focus and wh-questions, it is of course necessary to determine what definition of focus will be used. I will follow É. Kiss (1998) and assume that there are two types of focus¹. The first one, *identificational focus*, appears in the left periphery of the sentence as a result of movement (Chomsky, 1976; Rivero, 1978), has a particular intonation associated to it, the A-accent² (Bolinger, 1965; Jackendoff, 1972), and changes the meaning of the sentence by introducing a contrast against a set of alternatives and/or the notion of exhaustive identification of the focused constituent (É. Kiss, 1998). The second type, *information focus*, is not syntactically marked and does not move. It does not introduce

¹ For a more detailed discussion of the idea that there are two types of focus (É. Kiss, 1998) see chapter 2, section 2.1.1.

² The A-accent starts with a high pitch on the stressed syllable, then abruptly drops to a low pitch and falls at the end of the contour (H*L-L% in Pierrehumbert's (1980) notation).

a contrast or exhaustive identification, but it is associated with new information. It is not linked to a particular accent, but it is placed towards the end of the sentence, where it bears the sentential nuclear stress. The fact that it can extend to larger constituents than the one immediately containing the stressed word (Chomsky, 1971) has been called *focus projection* (Selkirk, 1984, 1995).

Wh-questions and information focus

Perhaps the most basic concept of a relationship between focus and questions is rooted in the idea that focus can be defined as *the nonpresupposed part of the sentence* (Chomsky, 1971, 1976; Jackendoff, 1972; Zubizarreta, 1998, p. 1). The relationship between focus and the answer to a *wh*-question can be understood as follows:

To the extent that the answer to a *wh*-question has the same presupposition as the question, the focus in a statement can be identified as the part of the statement that substitutes for the *wh*-phrase in the context question.

(Zubizarreta, 1998, p. 2)

In É. Kiss's terms, this relationship pertains to information focus only, since being nonpresupposed is not a necessary and sufficient condition for being identificationaly focused³. One advantage of establishing this link between *wh*-questions and focus is

³ The set against which the identification is being made may conceivably be presupposed, but the element bearing identificational focus is not consistently either presupposed or nonpresupposed.

that it provides an immediate explanation for cases of focus projection. The ambiguity of a sentence-final focus will correspond to different possible context questions, as shown in the examples in (1)⁴.

- (1) a. —What do beavers do? —Beavers [build dams]_F
b. —What do beavers build? —Beavers build [dams]_F

Wh-questions and identificational focus

As for identificational focus, there are both syntactic and semantic facts that seem to point at some link between identificational focus and wh-questions. Syntactically, identificational focus often involves leftward movement similar to wh-movement (Chomsky, 1976), even though differences can be observed (Rivero, 1978). Wh-questions are also commonly assumed to be in complementary distribution with fronted focus (Rivero, 1978; Rizzi, 1997). Semantically, both seem to introduce sets of alternatives, possible answers in the case of wh-questions.

⁴ Note that in this dissertation I use [brackets]_F with the subscript capital F (placed either after the initial or the final bracket) to denote the extension of the pragmatic Focus (usually information Focus), which may not be in correlation with a single constituent. Whenever I use brackets to denote a syntactic Focus Phrase I use FocP instead of F. For further discussion of focus projection see chapter 1, section 2.4.1 and chapter 2, section 2.1.2 in this dissertation.

Goals and structure of the chapter

The goal of this chapter is to develop an analysis of the syntax of focus and interrogation-related structures, particularly answers, that explains the differences found between them. It will be shown:

- a) That the relationship between information focus and interrogation, to the extent that it exists, is not encoded in syntax.
- b) That identificational focus and interrogation are independent grammatical phenomena.
- c) That answer sentences have a syntactic structure different from regular declaratives.
- d) That, inside an answer sentence, the part that constitutes the answer occupies a specific position in the left periphery, adjacent to Focus Phrase. I will suggest that this position corresponds to a functional head, which I will call Answer Phrase.

On the basis of these facts, I will propose a structure of the left periphery where three different functional heads host, respectively, focus, answers, and wh-elements. These three heads are shown to be immediately adjacent. This, together with the similarities between the syntactic operations in which they are involved, suggests that they are part

of a larger functional region, a subset of projections, or Focus field, in the style of what has been proposed by Benincà and Poletto (2004).

Since the relationship between information focus and interrogation is based on the identification between focus and answer elements, it will be necessary to look at the structure of answers in order to determine the exact nature of the relationship. Section 2 addresses the structure of answers. It starts with a brief review of previous observations on Hungarian, Italian, and Spanish answers in section 2.1. These observations include the impossibility of associating identificational focus to the part of the answer that substitutes for the wh-word in the question. This impossibility leaves information focus as the only kind of focus that can be associated to the relevant part of the answer. Section 2.2 includes Spanish data that shows that not only are answers different from identificational focus, but from information focus as well. Among the conclusions of section 2.2 is the idea that the answer to a question cannot be equated to a regular declarative sentence. Section 2.3 presents the proposal that a sentence's answer character is encoded in syntax, in the form of a projection (Answer Phrase) in the left periphery, and more specifically within the series of projections referred to as the Focus field (Benincà & Poletto, 2004). The availability of such a projection provides an immediate explanation for the structural similarities between answers and focus, as well as their semantic differences.

The independence between identificational focus and interrogation is argued in section 3 by showing, in section 3.1, cases of co-occurrence of identificational focus and wh-elements within a question and, in section 3.2, the case of the gerund as an element susceptible to carry one feature, focus, but not the other, [+wh].

Section 4 discusses the structure of the left periphery of questions taking into account the conclusions of sections 2 and 3, and shows how the separation of focus and answers in the left periphery can account for differences in interpretation of simple questions. Section 5 discusses some approaches to the semantics of focus and questions and provides a sketch of what the semantics of answers could look like.

2 Focus and Interrogation as separate phenomena in the structure of answers

This section explores the semantic overlap between focus and interrogation, visible in the purported possibility to identify focalized elements with answers to questions. Such identification has been present in generative analyses at least since the early seventies (Chomsky, 1971; Jackendoff, 1972) and continues to be used as a method to identify focus in a sentence (see, for instance, Neeleman & Szendrői, 2004).

If we assume a framework where more than one type of focus is defined, it is necessary to be more specific about which focus we claim to be overlapping with answers. Here I will discuss (and reject) the possibility that either identificational or informational focus can be identified with the relevant constituent of an answer.

I will show that answers have their own structure, which is in various ways parallel to the structure of identificational focus, but nevertheless distinct from it.

In section 2.1 I will present previous observations of failures of semantic overlap between answers and identificational focus. In fact, I will show that the opposite is the case, and identificational focus is incompatible with the part of the answer that substitutes for the wh-word in the question.

Section 2.2 presents a description of Spanish⁵ answers. The data show that the association of the relevant part of the answer with information focus is not enough to make a sentence felicitous as an answer. It follows from this fact that answer sentences are different from non-answer declaratives.

⁵ The data I present come from the Spanish spoken in the center and northwest of Spain. Since, to my knowledge, there are no previous detailed descriptions of the issues I address here, it is possible that other dialects exist within that region, and also that the observations made are valid for a broader area. The dialect I study here is not perceived as markedly regional by speakers of any other variety of Peninsular Spanish, but rather acknowledged as a certain kind of standard.

Section 2.3 takes the issue a step further and advances an analysis of the structure of answers. At the same time that it shows answers to be independent from focalization, this analysis shows strong structural similarities between answers and focus movement, which explain the link traditionally established between them.

2.1 Previous data on answers

2.1.1 Hungarian data

This section presents some basic facts about Hungarian answers, together with two different analyses, proposed by É. Kiss (É. Kiss, 1998) and Szendrői (Szendrői, 2001). Their analyses are clearly incompatible (É. Kiss proposes two types of focus applying to a same constituent, while Szendrői proposes one type of focus applying to two different constituents). It is, however, not obvious from the data whether either view has any advantage over the other, since both approaches are useful in analyzing the examples given (see chart in 4). I take the analysis of focus presented by É. Kiss to be more explanatorily adequate (see chapter 2, section 2.1). Here, however, I am interested in achieving a proper description of the relationship between focus and questions, and I review other approaches in order to explore aspects of the relationship that might be obscured in an analysis that takes on focus independently from questions.

One argument to differentiate between information and identificational focus (terms from É. Kiss, 1998) is that a focalized element placed in the left periphery can only have a contrastive interpretation and therefore it can not answer a wh-question (this fact shows its inability to express new information) (Brunetti, 2003, p. 85). Let us see some examples (cited in Brunetti, 2003, pp. 89-91). É. Kiss discusses Hungarian cases where an answer containing a left-peripheral focus must be interpreted as exhaustive. An exhaustive interpretation makes (2c) infelicitous as an answer to (2a), as long as (2a) is an innocent⁶ question.

- (2) a. Hol jártál a nyáron?
 ‘Where did you go in the summer?’
- b. Jártam Olaszországban.
 went-I Italy-to
 ‘I went to Italy [among other places].’
- c. OLASZORSZÁGBAN jártam.
 Italy-to went-I
 ‘It was Italy where I went’

Szendrői (2001) rejects the idea that (2b) and (2c) are examples of two different types of focus, and argues that the difference between the examples is due to the fact that

⁶ In this work, I will use the term ‘innocent’ to refer to questions which must be understood at face value. For instance, the question *where did you go last summer?* would be an innocent question if the speaker is interested in finding out where the hearer went last summer. It would not be an innocent question if it was meant to express something else, usually to point out a purported contradiction in previous statements of the hearer. To give an example, imagine the hearer had been talking about how people should travel more and the speaker wanted to point out that the hearer himself never travels. In that case, (2c) would be a felicitous answer.

focus includes the whole VP (henceforth VP focus) in (2b), while focus includes just the DP (henceforth DP focus) in (2c). The reason this is difficult to discern in (2b), according to Szendrői, is that the verb in that example is semantically empty and incapable of bearing stress. I understand that by calling it “semantically empty” Szendrői means that the verb does not add any meaning to the sentence beyond serving as support for a complement. It would have a role similar to a copulative verb, but the notion of semantic “emptiness” is by no means clear here. In order to clarify the contrast, she provides the example in (3), where a different verb is used (one which is not semantically empty and has the capability of bearing stress).

- (3) a. Hol nyaraltál a nyáron?
‘Where did you have holidays in the summer?’
- b. ?* Nyaraltam Olaszországban.
had-holidays-I Italy-in

In Szendrői’s view, a sentence like (3b), with a non-empty verb, and without DP focus movement, becomes marginal, purportedly because the example in (3b) has a broader focus on the VP, which would render it infelicitous as an answer. The reason it would be infelicitous is that there would be no reason to include the verb in the focus, since it was already included in the question and is therefore presupposed material in the answer, so someone listening to the conversation would wonder why the person answering is repeating that information. On the other hand, a “semantically empty

verb”, like *jártam* in (2b), can be included in the focus without causing that uneasiness.

In the following chart I present a summary of the two approaches considered so far.

- (4) Comparison of analyses of focus as an answer to questions in Hungarian
(I write in *italics* my extensions of the analyses beyond the original data)

answer to wh-object verb?	interpretation	explanation	
		É. Kiss (1998)	Szendrői (2001)
(semantically empty) verb - object	non-exhaustive	information focus	<i>in sentences of this form, VP focus and DP focus are very close in meaning</i>
verb - object	infelicitous	<i>ambiguous information focus and therefore inappropriate as an answer</i>	VP focus – <i>inappropriate as an answer</i>
OBJECT - verb	exhaustive	identificational focus <i>and therefore inappropriate as an answer to an innocent question</i>	DP focus

Szendrői’s view is based on the assumption that the answer for a wh-question is a sentence with a focalized element corresponding to the wh-element of the question. (Her argument is part of a discussion on whether there are one or more types of focus, not on its relationship to questions.) She uses wh-questions elsewhere (Neeleman & Szendrői, 2004; Szendrői, 2004) as a test to determine whether an element is focalized. Thus, Szendrői’s assumptions and explanation are in contradiction with this chapter’s goal of analyzing focus and wh- as different phenomena. The distinction between

answers and focalized elements will be explicitly addressed (on the basis of Spanish data) in section 2.2.

The review of previous analyses here serves the purpose of presenting and clarifying the assumptions implicit to each point of view. Namely, if we assume that there is only one type of focus, then answers and focus are equivalent and the unacceptability of certain answers is a result of focalization of the wrong element. If, on the other hand, two different types of focus are assumed, the unacceptability of some answers is due to the incompatibility of one of the types, identificational focus, and the answer status of the focalized element. In the next subsection, these two analyses are reexamined in the light of Italian data, in order to clarify the reasons for the incompatibility of identificational focus and answers.

2.1.2 Italian data

While the Hungarian data seem to allow for either one of the explanations presented in the previous section (two types of focus v. one type), Italian data introduce a new problem. Brunetti (2003) renders in Italian (5) the Hungarian example from (2). The felicity of the answers changes somewhat. Here information focus appears in bold type, and identificational focus with contrastive value is marked with capital letters.

- (5) a. Dove sei andata quest'estate?
'Where did you go last summer?'
- b. Sono andata **in Francia**. *information focus*
(I) am gone to France
'I went to France'
- c. ?? IN FRANCIA sono andata. *identificational focus*
in France (I) am gone
'It was France where I went'

Brunetti brings up this example as part of her argument against the two different types of focus É. Kiss proposes, and in particular against É. Kiss's idea that exhaustive identification is a characteristic of left periphery focus.⁷ In Brunetti's view, the Hungarian example in (2c) shows that the assertion that identificational focus can not answer a wh-question is incorrect, at least for Hungarian (in her view, the sentence includes exhaustive identification, which may make it infelicitous in some contexts, but it is still an answer to the question). As for Italian, Brunetti claims that (5c), "to the extent that [it] is acceptable" (Brunetti, 2003, p. 91), does not denote exhaustive identification. However, other Italian speakers (Gabriella Romani, p.c.) confirm that, in order to be acceptable, (5c) has to denote a contrast (a context offered to me was *IN FRANCIA sono andata*. *Where else did you expect me to go?*). If É. Kiss's view is correct, this might be simply a difference of what semantic features ([+contrastive]

⁷ Remember (from chapter 2) that É. Kiss (1998) explains the different interpretations of identificational focus in different languages as resulting from different combinations of the semantic features [+contrastive] and [+exhaustive].

and/or [+exhaustive]) constitute the semantic contribution of identificational focus in Hungarian and Italian.

The issue, then, is whether there is one kind of focus, which is used to answer questions and subsumes both identificational and information focus (Brunetti's and Szendrői's view), or two kinds of focus, identificational and informational, of which at least identificational focus can not be used as an answer. The data show, however, that sentences including identificational focus are indeed used to reply to questions, although such replies seem to have some additional element in their meaning.

In spite of the Hungarian example in (2c), I believe that the assessment that identificational focus does not answer a *wh*-question still holds. In order to clarify this it is necessary to define what 'answering a question' means, or at least in what sense the word 'answer' will be used in this chapter. An answer to a question can be informally defined as any declarative sentence that contains the information requested, or, alternatively, it can be defined as a declarative sentence that contains the information requested *and* keeps the information structure of the question. Whenever it is necessary to make a distinction, I will limit the use of the term *answer* to this last case. I will use the term *reply* to refer to other utterances used as a response.

It is true that (2c) can be used as a ‘reply’ to the question in (2a), but it is not an ‘answer’ to the question, in the same sense that one can reply *I’d rather not talk about it* but, regardless of how felicitous it is, it can not be argued that such a reply answers the question. While the ‘reply’ in (2c) does include information that constitutes an ‘answer’ to (2a), it is not possible to ignore the fact that it is used to say something else beyond that⁸. More specifically, it introduces the idea that the answer is exhaustive (always according to É. Kiss).

In view of this, the explanation for the dubious felicity of the Italian reply in (5c) is that it is not a good ‘answer’ because it introduces new ideas that were not in the question. The difference in the reported felicity of the Hungarian and Italian examples would only be a difference in the opinion of the respective informants of what constitutes a valid answer; i.e., it would be a difference in politeness or other discourse considerations, not a difference in grammatical acceptability. To summarize, identificational focus can be used as a ‘reply’ to a question, if the discourse conditions are appropriate, but it does not constitute a proper ‘answer’ to a question, because it

⁸ The mismatch between questions and answers can be illustrated in terms of the *quaestio* of the answer. The *quaestio* of an utterance is a question that is a posteriori reconstructed from the utterance (Umbach, 2001). The problem with answers that include an identificational focus is that their *quaestiones* are not the plain *wh*-questions, but rather include additional presuppositions. For instance, the *quaestio* of (2c) would be something like *what places did you visit last summer?*, with a presupposition that more than one place was visited.

adds meaning (contrast, exhaustiveness) that was absent in the question and was not requested⁹.

At this point it does seem clear that, both in Italian and in Hungarian, there is a difference between the meaning of left periphery focus, which can be contrastive or express exhaustive identification (details depending on the language), and focus on the right of the sentence carrying nuclear stress, which appears to create no problems when a sentence containing it¹⁰ is used as an answer to a question¹¹ (see 5b). This corresponds exactly to É. Kiss's distinction between identificational focus and information focus. The data presented by Szendrői and Brunetti does not seem to invalidate this description, and in fact, when closely examined, it seems to reinforce it.

⁹ Take, for instance, *I only went to France*. It expresses exhaustive identification of the relevant part of the answer, and it would be felicitous as a response to *Did you go to all the countries you planned to go?*. As a response to (2a) or (5a), however, it is only felicitous if the person who asked is willing to accept the implicature that the person who answers had previous plans or hopes to visit other places. Identificational focus creates the same need to accept a new implicature.

The definition of answer I am proposing falls short as a formal definition, since it assumes a previous set of easily recognizable implicatures to be built in the question. Thus, a sentence like *At the Metro station* seems intuitively to be a proper answer to a question like *Where can I buy Metro tickets?*, but it would only be so in a context where the Metro tickets could not be bought anywhere else, or the question answered were something like *What's the most convenient place to buy Metro tickets?* (Hagstrom, p.c.). This leaves the distinction between answers and replies a relatively useful tool to classify replies including identificational focus or contrastive topics from replies that do not include them, but prevents its use as a formal definition more extensively.

¹⁰ Note that *all* sentences have a nuclear stress and provide new information (if they are to be felicitous); that is to say, all sentences have an information focus.

¹¹ Recall that the problems with the Hungarian information focus provided by Szendrői in (3b) were due not to the presence of the information focus, but rather to the presence of a non-semantically-empty verb.

When sentences that have an identificational focus are used to answer a question they introduce additional meaning. This additional meaning either renders them infelicitous (as in Italian according to Brunetti) or forces a reevaluation of what information speaker and hearer presume to be shared (as in Hungarian according to É. Kiss and Italian according to my informants).

The examples discussed all seem to point against the idea that identificational¹² focus can be equated with the relevant part of an answer to a wh-question, and will support my claim that identificational focus and wh-features carry a different semantic content and can co-occur in a same sentence.

2.1.3 Spanish data

Domínguez (2004) establishes a classification of three types of focus in Spanish, with the characteristics described in the table in (6). I provide corresponding examples immediately below in (7):

¹² The relationship between answers and identificational focus is left for discussion in section 3 of this chapter.

(6) Three types of Spanish focus (Domínguez, 2004, p. 92)

(a) Focus in the left-periphery	(b) Focus in the right-periphery	(c) Focus in situ
Contrastive information	Contrastive/New information	Contrastive information
It cannot answer a wh-question	It answers a wh-question	It cannot answer a wh-question
Undergoes movement	In situ but scrambling may apply	In situ
It bears a feature	No feature if new information	It bears a feature
Marked stress	Unmarked stress	Very marked stress

- (7) a. UNA HORA\ te doy.
 one hour to-you I-give
 ‘I give you one hour, no more, no less’
- b. Te doy [una hora]_F.
 to-you I-give one hour
 ‘I give you one hour’
- c. Te doy UNA HORA para terminar el trabajo.
 to-you I-give one hour for to-finish the job
 ‘I give you exactly one hour to finish the job’

The types (6a) and (6c) roughly correspond to É. Kiss’s identificational focus, and (6b) to É. Kiss’s information focus. Note that by ‘right periphery’ Domínguez means merely the rightmost part of the sentence, and she does not discuss right-dislocated elements¹³. The note Domínguez makes about (6b) having “no feature if new

¹³ The paradigm in (7) can be expanded to include right dislocation, as seen in (i):

information” refers to the fact that a focus on the right of the sentence might be either a focus in situ (6c), carrying a feature, or a new information sentence-final focus (equivalent to É. Kiss’s information focus), which would not carry a feature.

Without entering into the details of the classification, what interests me here is her observation that, in Spanish, the only kind of focus compatible with a response to a question is the one that is not syntactically¹⁴ or phonologically marked, i.e. the kind that, not being associated to syntactic movement or intonational accents, can be classified as a purely pragmatic phenomenon. At least in this respect, Spanish and Italian seem to behave similarly.

Domínguez also discusses the complex issue of subject/verb/object order and its relationship to focus. She provides the following examples as part of her classification of Spanish word orders (I provide my own glosses for clarification):

-
- (i) *Te doy para terminar el trabajo una hora.*
to-you I-give for to-finish the job one hour
'I give you, to finish the job, one hour'

The intonation required for this word order to be acceptable varies from dialect to dialect, but there is always an intonational break between the PP and the right-dislocated object.

¹⁴ When Domínguez refers to scrambling, she is talking about the possibility that *other constituents*, not the focused one, might be scrambled as a result of a prosodic requirement to leave the focalized element in sentence-final position. Thus the effect of placing the focused element in the position where nuclear stress is assigned is achieved without assigning a feature or a prosodic mark such as an A-accent.

- | | | | |
|-----|---|----------|----------------------------|
| (8) | a. Susana abrió el libro
‘Susana opened the book’ | SVO | |
| | b. Susana abrió [F el libro]
‘What Susana opened was the book’ | SVO | |
| | c. [F El libro] abrió Susana
‘It was the book, what Susana opened’ | OVS | <i>Focus Preposing</i> |
| | d. Abrió el libro [F Susana]
‘It was Susana who opened the book’ | VOS | <i>P(rosodic)-Movement</i> |
| | e. [F Susana] abrió el libro
‘It was Susana who opened the book’ | SVO | |
| | f. El libro, lo abrió [F Susana]
the book, it opened Susana
‘As for the book, Susana was who opened it’ | O-CI-V-S | <i>Dislocation</i> |

She provides additional examples of word orders and the operations involved in two tables elsewhere in her dissertation (Domínguez, 2004, pp. 133, 226). The reason this particular set is interesting for the purposes of this discussion is that in her explanation, which I quote below, she makes observations about the felicity of these sentences as answers to different questions. When Domínguez mentions unambiguous narrow focus she is referring to focus limited to a certain XP, a DP in her example, where focus cannot be interpreted to be broader as a result of focus projection. By broad focus she means a reading where the whole sentence is interpreted as non-presupposed information.

(9) In the sentences illustrated in example (8) the main stress falls on the final constituent of the sentence. However, the object, *el libro*, in example (8b), already in its canonical final position, cannot be marked with unambiguous narrow focus. Hence sentences (8a) and (8b) are ambiguous with respect to whether they show a broad and a narrow focus interpretation. In sentence (8c) the preposed object is marked with narrow focus. When focused subjects or objects appear in initial position they are obligatorily associated with a contrastive reading. Although (8b) can also be contrastive, it is ambiguous with respect to the scope of focus (narrow or broad). In the same sense, example (8e) with the subject in initial position cannot be the answer to ‘*Who opened the book?*’ which requires narrow focus on the subject, and requires the subject to appear in final position as in examples (8d) and (8f). However, (8e) would be felicitous in a context where the subject is contrastive or has a corrective interpretation, such as the answer to the question

‘*Did Juan open the book?*’ Finally, example (8f) illustrates a common strategy used in Spanish and other Romance languages (see Vallduví (1992) for Catalan). Spanish uses clitic-left dislocations very often in order to leave the presupposed information out of the sentence. This common strategy is preferred by most speakers over a sentence like (8d) where the presupposed information remains *in situ*. Notice that both sentences (8d) and (8f) are felicitous to a question such as ‘*Who opened the book?*’ and in both sentences the focused subject appears in final position. This seems to indicate that non-focal phrases have the option to remain *in situ* or move out of the intonational phrase that contains the main focus. (Domínguez, 2004, pp. 18-19)

Domínguez's brief description of Spanish answers is the most complete description I have found in the literature, in spite of the fact that they are just a few comments in a work centered on other issues. In the following sections, we will look more closely at Spanish answers, in order to expand and clarify these observations.

For instance, the purportedly optional Clitic Left Dislocation that Domínguez refers to when comparing (8d) and (8f) is associated to noticeable changes in meaning. While (8d) and (8f) do have something in common, there are also differences between them. (8d) would be roughly equivalent to "As for opening the book, it was Susana who did it", while (8f) would mean "As for the book, it was Susana who opened it". Remember that (8f) is a clitic left dislocation (CLLD), where the dislocated element has the meaning of a contrastive topic (Arregi, 2003).

Domínguez states that (8d) and (8f) are felicitous as answers to the question 'Who opened the book?' while (8e) is not¹⁵. All speakers I have consulted have judgments that agree with Domínguez's observations on (8e) but differ on (8d) and (8f). All of these sentences would be atypical as answers anyway, since the usual answer to a wh-question is limited to the XP necessary to substitute the wh-word. This fact certainly obscures judgments on longer answers. However, the clarity of the judgments provided

¹⁵ Note that the fact that (8e) (which, crucially, includes contrastive focus intonation) is not felicitous as an answer coincides with the observations made about Hungarian and Italian described in previous sections.

by speakers consulted suggests that the discrepancies are due to dialectal variation, rather than to uncertain or variable acceptability judgments.

Thus far, the Hungarian, Italian, and Spanish data discussed in section 2.1 all point to the lack of a link between identificational focus and the relevant part of an answer. In fact, the presence of identificational focus seems to prevent a sentence from functioning as a felicitous answer in a neutral context. The next task will be to identify the sentence structures that do function as felicitous answers. The Spanish results of that search, together with a detailed description of the meaning of each possible answer, are detailed in section 2.2 below.

2.2 Towards a description of (Northwestern and Central Peninsular) Spanish answers

The previous section (2.1) discussed the problems of an interpretation where answers are assumed to be equal to identificational focus. This section explores the relationship between answers and information focus, and concludes that answers are not properly described by merely associating them to information focus. This conclusion is supported by providing a general picture of the differences between answers and other declarative sentences, with particular attention to fragment answers (section 2.2.2), but including sentences with identificational focus constructions, as well as those with

information focus only. Differences in intonation are described, as well as constraints on word order (when compared to regular declaratives) which are mostly due to the incompatibility between the character of answer and other informational roles. An analysis is proposed in the next section, 2.3.

2.2.1 An initial classification of Spanish answers

The answers to the question ‘Who opened the door?’ (at least for Northern and Central Peninsular Spanish, and probably leaving out the Alicante area in the southeast of Spain), are presented in the chart in (10), followed by the answers to the question ‘What did Susana open?’ in the chart in (11), where the object, not the subject, is the answer. Two more charts, (12) and (13), present answers to questions requiring an answer containing a verb but excluding either the object or the subject. The question in (12) is ‘What happened to the door?’ and the one in (13) is ‘What did Susana do?’.

In these charts, CAPITALS indicate intonationally marked contrastive focus. When an A-accent appears as intonational mark, a downward slash (\) is added. Information focus is marked in [brackets]_F. *Italics* indicate answer intonation, which I perceive as different from A- and B- accents. I have also included the upward slash (/) to mark the B-accent intonation of topics. I add a comma to both answer intonation and topic intonation to indicate that there is an intonational event, which makes it possible

for the speaker to (optionally) insert a pause. I use the symbol * to indicate that a sentence is inappropriate as an answer, and # to indicate that, even though it might be interpreted as appropriate, it introduces additional elements that the question does not call for.

Each chart will be discussed in one or more separate subsections below. A summary of the observations made and an analysis of answers will be presented in section 2.3.

(10) Who opened the door?

Question	Answer word order	Answer	Answer description	Gloss
<p>¿Quién abrió la puerta?</p> <p>Who opened the door?</p>	unclear	a. Susana	Fragment answer, preferred	Susana
	S C I V	b. *Susana la abrió	Broad focus, clitic	Susana opened it
		c. <i>Susana</i> , la abrió	Answer intonation	Susana
	d. *SUSANA \ la abrió	Contrastive focus	It was Susana who opened it	
	S V O	e. *[Susana abrió la puerta] _F	Broad focus Domínguez (8a)	Susana opened the door
		f. <i>Susana</i> , abrió la puerta	Answer intonation	Susana
	V O S	g. *SUSANA \ abrió la puerta	Contrastive focus Domínguez (8e)	It was Susana who opened the door
		h. *Abrió la puerta [Susana] _F	Domínguez (8d)	? It was Susana who opened the door
	C I V S	i. *Abrió la puerta SUSANA	subsumed under Domínguez (8d)	It was Susana who opened the door
		j. *[Abrió la puerta Susana] _F	Broad focus, subject on the right	<i>Not available in any context</i>
	O C I V S	k. #La abrió Susana	Clitic linking to previous discourse	Susana opened it
		m. <i>La abrió Susana</i>	Clitic linking to previous discourse. Answer intonation	Susana opened it
		n. #La puerta, la abrió [Susana] _F	CLLD Domínguez (8f)	As for the door, Susana opened it

(11) What did Susana open?

Question	Answer word order	Answer	Answer description	Gloss
<p>¿Qué abrió Susana?</p> <p>What did Susana open?</p>	unclear	a. La puerta	Fragment answer, preferred	The door
	SVO	b. *[Susana abrió la puerta] _F	Broad focus Domínguez (8a)	Susana opened the door
		c. *Susana abrió [la puerta] _F	Information focus. ..(bearing nuclear stress) Domínguez (8b)	Susana opened the door
		d. (Susana) abrió <i>la puerta</i>	Answer intonation	The door
	e. *Susana abrió LA PUERTA	subsumed under Domínguez (8b)	It was the door, what Susana opened	
	OVS	f. *LA PUERTA\, abrió Susana	Contrastive focus Domínguez (8c)	The door was what Susana opened
	O C I V S	g. *La puerta, la abrió Susana	CLLD Domínguez (8f)	As for the door, it was Susana who opened it

(12) What happened to the door?

Question	Answer word order	Answer	Answer description	Gloss
<p>Qué pasó con la puerta?</p> <p>What happened to the door?</p>	SVO	a. *Susana abrió la puerta	full declarative	Susana opened the door
	S Cl V	b. #Susana la abrió	<i>Susana</i> in the left periphery	Susana, she opened it
	Cl V S	c. La abrió Susana	innocent answer	Susana opened it
	O Cl V S	d. #La puerta, la abrió Susana	CLLD	The door, Susana opened it
	C S Cl V	e. #Que Susana la abrió	subordinate clause	that Susana opened it
	C Cl V S	f. Que la abrió Susana	subordinate clause	that Susana opened it

(13) What did Susana do?

Question	Answer word order	Answer	Answer description	Gloss
<p>Qué hizo Susana?</p> <p>What did Susana do?</p>	SVO	a. #Susana abrió la puerta	subject as topic	Susana opened the door
	VO	b. Abrió la puerta	no overt subject present	She opened the door

2.2.1.1 Subject answers

This section discusses replies whose subject constitutes the answer to the posed question (see 10). These replies are found to be of the form SVO, as described in subsection 2.2.1.1.1. Other word orders are not viable, or they are associated to additional information-related meaning, and are discussed in subsections 2.2.1.1.2 and 2.2.1.1.3. The last subsection, 2.2.1.1.4, discusses subject answers in sentences with unaccusative and unergative verbs.

2.2.1.1.1 “Innocent” subject answers

In reading this subsection, it is important to keep in mind that when I use the term “innocent” I am referring to answers to the question ‘Who opened the door?’ in a context where a) the question is understood to be merely seeking information, b) the answer intends to provide information relevant to the question, and c) nothing else is implied.

When the answer to a question is the subject of the reply, the word order of an innocent answer is SVO (10f) or, in the case of the presence of an object clitic, S Cl V (10c) (the possible exception (10m) is discussed in section 2.2.1.1.3). Intonation facts must also be taken into account when determining the felicity of an SVO answer,

however. The intonation of a non-marked, broad focus sentence does not work as an answer, be it the complete sentence (10e, 8a) or a variation where the object appears as a clitic (10b). Neither does the contrastive focus intonation of (10g, 8e), which corresponds to the A-accent described by Jackendoff (1972). The actual intonation of an answer is different both from broad focus intonation and from a sentence including contrastive focus. An answer involves an accent (in the sense of an intonational event; it might be followed by a pause) different from the A-accent and the B-accent. Pending a spectrographic analysis, it is difficult to give a formal description. For the time being, and in absence of a better term, I will call it “answer intonation”¹⁶.

The presence of an accent is made noticeable by the fact that a pause can optionally be introduced between the accented phrase (the part of the answer that could serve as a fragment answer) and the rest of the sentence (10c,f).

2.2.1.1.2 Pragmatically marked subject answers

Let us now look at word orders other than SVO. The VOS word order (10h, 8d, 10i, 10j) does not seem to be acceptable at all as an answer¹⁷. As a declarative, it would require a context where the facts presented by the sentence, except for the subject, were

¹⁶ I am indebted to Susana Huidobro for useful discussion on this and many other points in this dissertation.

¹⁷ Unaccusative and unergative verbs, discussed in the next section, allow for VS order in answers.

part of a common presupposition. Say, someone opens the door every day and I am telling someone who knows this that, today, it was Susana who did it. Adding an adverb such as “today” in the role and position of a contrastive topic would improve the acceptability of the sentence¹⁸.

- (14) Hoy abrió la puerta Susana.
today opened the door Susana
'Susana was who opened the door today'

Note that this kind of context is usually viewed as appropriate as an informal description of (subject) focus, yet neither (10g, 10h, 8d) nor its more acceptable variant with a clearer topic (14) are acceptable as answers to an innocent question. The sentence in (14) can actually be used, if the subject is understood to be the focus, as the answer to a pragmatically loaded question, one where someone else is presupposed to be the one who opened the door. It would be answering a question equivalent to ‘Was it Mario who opened the door?’ in a negative way, and in that respect it would not differ from a sentence where the subject is in a contrastive focus position in the left periphery, such as (10g, 8e). Recall Domínguez’s (2004, pp. 18-19) remarks about (10g, 8e) being felicitous when interpreted as having a contrastive or corrective interpretation. Alternatively, the contrast might be understood to be expressed by the topic, and then

¹⁸ Presence of a fronted adverb is usually associated to subject inversion in Spanish. (Hernanz & Brucart, 1987, pp. 77-78).

(14) would not be equivalent to (10g, 8e) but it would rather imply that today was an exception and tomorrow someone else, not Susana, would open the door again.

2.2.1.1.3 Clitics in replies

Let us now address replies that include clitic object pronouns. The reply in (10n, 8f) presents a left dislocated object (interpreted as contrastive topic and associated with the B-accent). As a reply, this sentence has a contrastive/corrective interpretation: the addition of a contrastive topic brings in the implicature that someone else opened something else, much like the adverb in (14) helped the contrastive interpretation in the declarative.

Note that the contrast expressed by (10n, 8f) is different from the contrast expressed by (10g, 10h , 8d), discussed in the previous subsection. The sentences with a focalized subject (10g, 10h , 8d) present a contrast between the answer given and the answer the speaker presumes to be expected by the person asking the question (in the style of ‘I know you think it was Peter, but it was actually Susana who opened the door’). The contrast expressed by (10n, 8f) does not involve the answer, but other presuppositions present in the question (this reply would be appropriate, for instance, in a context where it could be paraphrased as ‘Well, to answer your question, Susana opened the door, but you are probably trying to find out who is to blame for the draft,

and that would be Mario, who opened the window. You should be asking who opened the window, not who opened the door’).

A reply with a clitic linking the sentence to previous discourse (10k,m) can be felicitous as an answer as long as it has the appropriate intonation (10m). Other cases (10k), are infelicitous in the same way CLLD constructions (10n) are. In order to clarify this let us compare the order CI-V-S to S-Cl-V. Sentences with the S-Cl-V order, with sample contexts, are listed in (15), including (15a) also listed as (10b) above.

- (15) a. ?? Susana la abrió. *broad focus*
Susana it opened
- b. Susana/, la abrió. *contrastive topic, B-accent, CLLD*
‘I don’t know who else opened the door, but Susana, she did open it’
(This would work as the answer to ‘Have any students opened the lab door lately?’)
- c. SUSANA\ la abrió. *contrastive focus, A-accent*
‘It was Susana who opened the door’
(Contrastive/corrective reply to ‘I think Mario opened the door’)
- d. Susana la ABRIÓ. *contrastive focus on the verb, prosodically marked*
‘It is true that most students rarely even look at that door, but what Susana did was open it’

I am unable to find a context where broad focus is compatible with the presence of a clitic and an initial subject. I am inclined to say that it is not possible. The CI-V-S, however, is an appropriate word order for broad focus given any context that clarifies

the reference of the clitic. Note that this is the order that appears following the dislocated element in a CLLD, where the presence of the topic makes the subject inversion mandatory (Campos & Zampini, 1990; Hernanz & Brucart, 1987). This order is not obtained, however, when the clitic is a subject clitic or non-referential.

- (16) a. Susana se bañó.
Susana CL bathed
'Susana bathed'
- b. Susana lo pasó bien.
Susana CL passed3sg well
'Susana had a great time'

These facts suggest that it is not the presence of a clitic that causes the inversion. Subject inversion in these cases would be caused by the clitic reference to a dislocated topic, whether it is visible in the sentence or not. That is to say, the word order in (10k) is caused, like in (10n), by a topic-related structure (where an explanation about the reference of the clitic is required, if not immediate available as a dislocated topic).

The issue that has to be addressed now is the availability of the CI-V-S order as a felicitous answer (10m), which appears to contradict the generalization that innocent subject answers are sentence initial. Even though it is difficult to notice in Castilian Spanish, CI-V-S sentences require a particular intonation in order to be felicitous as answers. This intonation is easier to appreciate in dialects where vowel lengthening

accompanies certain accents, such as Chilean Spanish¹⁹. The intonation of (10m) corresponds to the answer intonation mentioned above for subjects in SVO answers. Besides the difference in intonation, it is important to note that, in spite of the similarity between (10k) and (10m), the latter does not introduce a contrastive meaning. The different meaning and intonation suggest that the underlying structure of (10m) is not a CLLD, as it was proposed for (10k), but it rather involves an additional syntactic operation. Concretely, I view this structure as a variant of a fragment answer, where the elided part is smaller. In this respect, observe that *La abrió Susana* is also a good fragment answer to *qué pasó con la puerta?* ‘what happened to the door?’ (12d) (answers to such questions are discussed below, in section 2.2.1.3). An analysis for fragment answers is presented in section 2.2.2, following the description of other relevant data. I will now discuss subject answers in unaccusatives and unergatives before giving a brief summary of the sections on subject answers and moving on to object answers.

2.2.1.1.4 Unaccusatives and unergatives

Subjects of unaccusative and unergative verbs (represented, respectively, by *venir* ‘to come, to arrive’ in (17) and *hablar* ‘to speak’ in (18) below) present a pattern similar to other subjects. They can appear as fragment answers (17b, 18b), or on the

¹⁹ I owe this observation to Héctor Campos.

left, separated from the rest of the sentence by the accent I refer to as answer intonation (17d, 18d). Additionally, they can appear on the right side of the verb (17f, 18f), with the same intonation.

- (17) a. Quién vino?
who came
- b. Juan.
- c. # Juan vino.
- d. *Juan, vino. answer intonation*
- e. # Vino Juan.
- f. *Vino Juan. answer intonation*
- (18) a. Quién habló?
who spoke
- b. Juan.
- c. # Juan habló.
- d. *Juan, habló. answer intonation*
- e. # Habló Juan.
- f. *Habló Juan. answer intonation*

As seen in the previous section, subjects of transitive verbs only appear in final position in answers when there is a clitic involved. Since the sentences in (17) and (18)

have no direct object that can undergo cliticization and right dislocation, their pattern can be argued to be completely parallel to that of replies involving subjects of transitive verbs. That is, the subject can appear a) alone as a fragment answer (10a, 17b, 18b), b) preceding the rest of the sentence but separated by answer intonation (10c, 10f, 17d, 18d), or c) in postverbal position, within an intonational phrase that includes the verb (10m, 17f, 18f). The postverbal position is explained by the presence of a dislocated topic in transitive sentences, and corresponds to the usual subject position with Spanish unaccusative and unergative verbs.

To summarize section 2.2.1.1, it can be said that answer subjects appear on the left of the sentence, with a particular intonation (10a, 10c, 10f, 17b , 17d, 18b, 18d). Whenever they appear on the right, there are additional factors at play, be it the presence of additional topics on the left or other syntactic operations, such as cliticization of the object, that might have an influence in the position of the subject (10m, 17f, 18f). In any case, answers subjects are not necessarily sentence-final. This comes to show that any possible association between answers and information focus, which is always sentence-final, is purely circumstantial and not rooted in syntax, as should be expected from the assumption that information focus is not a syntactic phenomenon. The issue of the actual syntactic position of the answers is left for section 2.3, where a common analysis for subject and object answers will be presented.

2.2.1.2 Object answers

2.2.1.2.1 “Innocent” object answers

Leaving aside fragment answers, the word order for answers involving objects is SVO (11b-e, 8a,b), which leaves the object in the sentence-final position associated to information focus. There are, however, indications that object answers have a different status from both identificational and information focus. Even though, as Domínguez remarks, it is not possible to determine from the nuclear stress placement in such a sentence whether there is narrow or broad focus, I have observed that an *in situ* focus contrastive intonation such as (11e) (which would involve an intonational event, not just the nuclear stress) renders the sentence infelicitous as an answer. As for felicitous answers (11d), my perception is that, at least in Northern Castilian Spanish, their intonation is slightly different from that of information focus declaratives (11b,c). A phonetic analysis should be performed in order to determine exactly what differences there are between them. Chilean Spanish presents a much clearer difference, involving a greater vowel length in the stressed syllable of the answer and an intonational contour different from those of declaratives and contrastive focus²⁰.

²⁰ I thank Héctor Campos for all the Chilean Spanish data discussed here.

Interestingly, Chilean Spanish also allows for object initial answers when the relevant answer is the object. While in Peninsular Spanish the presence of fronted answer elements in full answers is limited to subjects, Chilean Spanish appears to allow for fronted answer indefinite objects as well.

- (19) a. A quién vio?
to who 3rd-past-see
'Who did he see?'
- b. A un chico colorín vio.
to a boy red-haired 3rd-past-see
'He saw a red-haired boy'
- c. A un exprofesor de escuela vio.
to a former-teacher of school 3rd-past-see
'He saw a former school teacher'
- (20) a. Qué compraste ayer?
what 2nd-past-buy yesterday
'What did you buy yesterday?'
- b. Un libro compré, pero luego lo devolví.
a book 1st-past-buy but later CL-it 1st-past-return
'I bought a book, but I returned it later'

Chilean Spanish allows for two different intonations in the answers presented in (19b,c) and (20b). One of those intonations is contrastive and the other one is non-contrastive. The contrastive one is of course infelicitous as an answer to an innocent question. The other one, more likely to appear in association with indefinite objects, might be the answer intonation described above for subject answers. Since focus movement is always

associated to focus intonation, this data might be independent evidence for the idea that answers have a specific position in the left periphery, different from Focus Phrase, host of identificational focus. I will discuss this issue in section 2.3.1.

2.2.1.2.2 Pragmatically marked object answers

Sentences where the object appears focalized, be it in the left periphery (11f, 8c) or in situ (11e) can only be used as replies if they are interpreted as having a contrastive/corrective interpretation, which would be possible in this case only if the question ‘What did Susana open?’ were to be interpreted as equivalent to ‘Did Susana open the wrong thing again?’ or something of the sort. The other possible order involving the subject placed on the left, the Clitic Left Dislocation in (11g, 8f), where *la puerta* ‘the door’ functions as a contrastive topic, is not a possible answer to ‘What did Susana open?’, unless the person responding assumes that the person asking knows the answer already and is trying to introduce a nuance, as in a conversation like this: —*I’ll prove to you that Susana is responsible for causing a draft in the house. What did Susana open? –La puerta, la abrió Susana, pero la ventana la abrió Mario.* ‘As for the door, it was Susana who opened it, but María opened the window.’ In such cases the dislocation of the object in the answer is related to the information structure of the answer as part of a longer argumentation, but it is not motivated by its status as answer *per se*.

2.2.1.3 VP answers whose object is part of the question

When the relevant part of the reply is a VP (see examples in 12), I have found no distinguishable answer intonation marking any phrase, unlike in the cases of subject and object answers discussed in the previous sections. To explain this, let us take into account the fact that VP internal elements present in the question do not usually appear in the answer (hence the inappropriateness of (12a)). The role of the answer intonation is to make a phrase on the left stand out from the rest of the sentence (as it does when the subject or the object are the answers). In cases where there is no division of the sentence in two parts it is not possible to have this effect. Additionally, when VP internal elements that were present in the question do appear in the answer, they are always interpreted as introducing a contrast. Thus, they introduce a new complication to the sentence structure, which is associated to its own intonational effects. This is perhaps most clearly seen in (12d), the Clitic Left Dislocation (CLLD), where the left-dislocated object serves as a contrastive topic. In such cases, the word order and the intonation are determined by the CLLD structure, and other accents that might potentially be involved (perhaps answer intonation) will be more difficult to identify.

The word order that appears in innocent VP answers (12c) is Cl V S, i.e. the order that corresponds to a CLLD, except for the dislocated element. This is to be expected, since CLLDed elements are associated to the meaning of topic, and elements

that appear in the question are topics, at least in the sense of ‘common ground information about which new information is provided’. An analysis of the structure of answers will be provided in section 2.3. For the time being, it is interesting to note the parallel between these answers and the structure of subject answers involving clitics (10m). The presence of the dislocated element in the answer (12d), on the other hand, makes the answer clearly contrastive (recall Arregi’s (2003) idea that CLLD equals contrastive topicalization), and therefore infelicitous as an answer, which makes it parallel to the structure in (10n).

The S Cl V order (12b) is somewhat problematic. I mentioned above (section 2.2.1.1.2) that this order is not compatible with broad focus (at least in any dialect I am familiar with). The meaning it has as a reply in (12) is probably best pictured with an example: In a household where children are not allowed to open a particular door, a parent finds the door open and asks what happened. The first child who answers will use this word order to accuse the other child of misbehavior. It is therefore contrastive (to the extent that it intends to present a contrast between both children; it could also be described as exhaustive, since it is used to report that it was the other child, and only him, who misbehaved), but it does not have the corrective character usually associated to contrastive focus. The accent associated to it is similar to the contrastive A-accent of sentences like (15b) above (in fact, a usual reaction to children’s accusations is a mock repetition of the intonational structure of their accusation), but it is not as clearly

marked as it is in corrective sentences, where the focus contradicts previously presented information.

Due to the lack of a formal analysis of the intonation of these examples, it is impossible to avoid an element of speculation in these remarks. It is important to note, however, that sentences like (12b) are predicted by the hypothesis that focus and *wh*- are separate syntactic phenomena. If they really are separate phenomena we should not expect to find them in complementary distribution (lest they be different versions of a broader phenomenon). We should expect to find some interaction between them. Cases like (12b) show co-occurrence of focus and whatever syntactic operation determines the structure of answers²¹. In such situations, the contrastive reading of focus holds, but not the corrective reading, which would not make sense in the absence of a previous assertion. The presence of an intonation particular to these cases might well be a result of the combination of the answer intonation and the focus accent.

It is interesting to note that this kind of answers allow for an overt complementizer to be present (12d,e), which is expected in subordinate clauses but not in matrix clauses. I will take this as an indication that these answers are indeed subordinate clauses and a matrix clause has been elided. The complete analysis is presented in section 2.3.

²¹ Co-occurrence of focus and *wh*- in separate positions in a question is discussed in section 3.

2.2.1.4 VP answers whose subject is part of the question

Nothing remarkable is to be found in cases when the question is about a VP and it includes the subject. As expected, the VP appears as the answer (13b). If the subject appears in the reply (13a), it adds an element of contrast, acting as a contrastive topic more or less intonationally marked, but definitely with an intonation different from a broad focus declarative.

Let us recapitulate the main points of the description of answers given in this section (2.2.1).

- There is a distinct intonation for answers that distinguishes them from regular declaratives.
- There is no correlation between Nuclear Stress and the position of the relevant part of an answer sentence. This is most evident in the case of subject answers. Therefore, there is no systematic correlation between the relevant part of an answer and its information focus.
- There are no cases of identificational focus serving as the relevant part of an answer. This is coherent with the findings for several languages discussed in section 2.1.

The data discussed so far gives evidence of what answers are not. Section 2.3 discusses what answers are. I will claim that an answer is a construction involving a functional head in the left periphery. This construction shares some properties with identificational focus constructions but is nevertheless distinct from them. I will place that functional projection inside a ‘focus field’ that also includes the target of the movement of wh-phrases. Before moving on to that task, however, it is necessary to address fragment answers and see how they fit in the description undertaken in this section. Fragment answers will also provide additional evidence that answers are not equivalent to identificational focus.

2.2.2 Fragment answers

2.2.2.1 Views on fragments

The use of fragments as answers to a question is a particular case of the use of nonsententials in language. There is a growing body of literature that discusses the status of nonsententials; the main point of contention is whether they are originated by derivation or by deletion/ellipsis of elements of a larger structure (see Barton & Progovac, 2005 for an overview). Among the explanations that avoid ellipsis is the proposal that the initial node of a generative grammar is not S but X^{\max} (Barton’s (1991)

X^{\max} generalization). That view allows for the generation of a fragment in isolation, without any additional sentence structure that would have to undergo ellipsis for the utterance to be limited to the fragment. The main difficulty for such explanations is to justify why a nonsentential can converge (to use the minimalist terminology) without introducing further stipulations and/or making nonsententials fundamentally different from sentences. In order to deal with this problem, it has been argued that some nonsententials are phases, and thus there is no need to have a differentiated derivation process for them (Fortin, in press). The opposite point of view, i.e. that nonsententials are originated by derivation of a sentence and subsequent ellipsis, has the advantage of not requiring any special device (besides, of course, ellipsis itself) to generate fragments, whether they are phases or not (see Merchant, 2004 for a complete development of this point of view and relevant references).

Arguments for the ellipsis interpretation of fragments are usually syntactic and based on form-identity effects. Merchant (2004) presents a number of arguments in support of the idea that fragments involve ellipsis, and fragment answers involve ellipsis after movement. The view that there is ellipsis is supported by the fact that “the morphological case form of a fragment DP is always exactly the same as the case we find on the corresponding DP in a fully sentential answer” (Merchant, 2004). He illustrates it for languages as diverse as Greek, German, Korean, English, Hebrew, Russian and Urdu. (Spanish examples of case preservation are provided in (24c,d)

below). A second argument for ellipsis is that DPs in fragments are distributed following Binding Theory principles, just as their correlates in fully sentential answers do. (Merchant, 2004 discusses this and other connectivity effects).

The view that fragment answers undergo movement before ellipsis of the rest of the answer takes place is supported by the fact that, in English and other languages, NPIs that can not be left-dislocated can not appear in fragments either (21, 22). Also, fragment answers can not be extracted from islands, as seen in (23b) (see Merchant, 2004 for these and other arguments).

(21) A: What didn't Max read?

B: *Anything.

(22) a. Max didn't read anything.

b. *Anything, Max didn't read.

(23) a. Does Abby speak the same Balkan language that *Ben* speaks?

b. *No, *Charlie*.

c. No, she speaks the same Balkan language that *Charlie* speaks.

Another argument in support of the presence of movement in fragment answers is the correlation between preposition stranding and availability of DP answers shown in (24) (Merchant, 2004 presents this argument on the light of English and Scandinavian Languages, which allow stranding, and Greek, German, Yiddish, Czech, Russian, Bulgarian, and Hebrew, which do not allow it). The mandatory presence of the preposition in the Spanish answer in (24b) suggests that the same derivational mechanisms that would prevent the separation of the preposition and its DP complement in a full sentence were also in place for the fragment. The same is true for case marking elements like the accusative personal *a* in (24c) and the dative marker *a* in (24d).

(24)	a.	Where are you from?	Jersey
	b.	De dónde eres? from where you-are	*(De) Nueva Jersey from New Jersey
	c.	A quién vio? to who he-saw 'Who did he see?'	*(A) Maia to Maia 'Maia'
	d.	A quién le escribió? to who Dat-him he-wrote 'Who did he write to?'	*(A) Francesc to Francesc 'Francesc'

Form identity between questions and answers is not always straightforward, however. There are cases in Spanish where elements that undergo wh-movement in the

question do not appear in the answer, seemingly contradicting the conclusion from (24b). Consider the following examples:

- | | | | |
|------|----|---|---|
| (25) | a. | De qué parque vives cerca?
of what park live-2sg near
'What park do you live close to?' | (*Cerca) del Parque Central
near of-the park central
'Central Park' |
| | b. | Cerca de qué parque vives?
near of what park live-2sg
'What park do you live close to?' | (Cerca) del Parque Central
near of-the park central
'Central Park' |

Spanish allows for the stranding of elements that have certain preposition-like characteristics (25a). Those elements have a feature [+N] and therefore also share characteristics with nouns and adjectives (Campos, 1991). They are relevant to this discussion because they can also undergo wh-movement, and in those cases they are optional in the answer. In the example in (25b) the word *cerca* 'near' would be expected to always appear, like the prepositions in (24b,c,d). The fact that it is optional seems to go against the idea that the answer is identical to the wh element in the question. Nevertheless, I think the form-identity can be saved by postulating an additional, separate ellipsis process for the [+N] element in the answer. This would make it parallel to the case of nouns, illustrated in (26).

- | | | |
|------|--|---|
| (26) | El piso de quién te prestaron?
the apartment of who cl lend-3pl
'Whose apartment did they let you borrow?' | El (piso) de Loli.
the apartment of Loli
'Loli's' |
|------|--|---|

The same explanation could apply to dialects where such fragments are allowed (27a) even in constructions where the [+N] element cannot be left stranded (27b,c)²². Such dialects would otherwise appear to present a challenge to the form identity generalization.

- (27) a. Salió desde detrás del sillón? %No, de la mesa.
 came-out3sg from behind of-the armchair no of the table
 ‘Did he come out from behind the armchair? No, from behind the table’
- b. * De qué salió desde detrás?
 of what came-out3sg from behind
 ‘What did he come out from behind?’
- c. * De la mesa salió desde detrás.
 of the table came-out3sg from behind
 ‘He came out from behind the table’

The question of what licenses these kinds of ellipsis inside DP, while interesting, falls outside the scope of this work.

My own argument to support the ellipsis analysis of fragments is also based on form identity, this time identity in terms of information structure. Fragments keep the informational status they would have in a full sentence (this is perhaps most noticeably when they express contrastive focus). To the extent that informational status is linked to syntactic structures, full structures must be involved in the derivation of the fragment.

²² I thank Héctor Campos for this set of data.

The next section, 2.2.2.2, argues that, precisely because of the expectation of form identity associated to ellipsis, fragment answers cannot be identified with the site of contrastive focus, FocP. Section 2.3 will put forth the idea that answers are located somewhere else in the left periphery.

2.2.2.2 Fragment answers do not move to FocP

Fragment answers to wh-questions consist of a constituent that is perceived as substituting for the wh-word in the question. Since such answers are widely regarded as the most natural way to respond to a question, it is to be expected that those accounts that equate focus with answers provide a mechanism to explain how they are derived from focus structures or otherwise linked to them.

Thus, Brunetti (2003, p. 126) proposes that “the focused constituent moves to the left periphery to a position higher than the IP, and then ellipsis of the non-focused part of the sentence applies”. The same idea is present in Merchant (2004), who makes no specific claims about the exact position targeted by the movement, but says he suspects it might be Rizzi’s FocusP. In the example below, (28b) would be derived through the ellipsis represented in (29a). The same structure without ellipsis (29b) is infelicitous, a fact that Brunetti attributes to its including redundant information.

- (28) a. Che cosa ha vinto Gianni?
 what thing has won Gianni
 ‘What did Gianni win?’
- b. [La maglietta]_F.
 ‘The T-shirt’
- (29) a. [_{DP} La maglietta]_j [_{IP} ha [_{VP} vinto [_{DP} Gianni] t_j]]
- b. #La maglietta ha vinto Gianni.
 the T-shirt has won Gianni

Similar accounts have been provided for negative fragment answers in Greek (Giannakidou, 2000) and in Spanish and Italian (Alonso-Ovalle & Guerzoni, 2004) in an attempt to explain the possibility of having negative words as fragment answers while still claiming that they are always Negative Polarity Items (30b).

- (30) a. Chi hai visto? Nessuno.
 Who did you see? Nobody.
- b. [_{FocP} Nessuno]₁ [_{Foc} Ø [_{+neg}] [_{IP} ho visto t₁.]]
 nobody have_{1sg} seen
 ‘I haven’t seen anyone’

An account for answers that calls for ellipsis after extraction of the element that suffices as an answer is appealing, because it serves as a link between the answer and the rest of the elements in the question. The idea that the answer is extracted to a focus

position, however, presents at least two serious shortcomings. Consider for instance (5) above, which I repeat here as (31) for convenience.

- (31) a. Dove sei andata quest'estate?
'Where did you go last summer?'
- b. Sono andata **in Francia**. *information focus*
(I) am gone to France
'I went to France'
- c. ?? IN FRANCIA sono andata. *identificational focus*
in France (I) am gone
'It was France where I went'

Firstly, there is a felicity argument. Assuming that the answer is extracted to a focus position goes against the fact that sentences where the answer is clearly in a focus position (31c, 29b) are precisely the ones which are infelicitous, while other sentences with no ellipsis are felicitous (31b).

The second argument against movement to focus has to do with the meaning of the sentence. There is no reason to assume that a sentence which has contrastive value when there is no ellipsis (the sentences in (29a,b) would be contrastive, just like (31c)) would lose that contrastive value when there is ellipsis (the fragment in (28b) is not contrastive). Brunetti does not accept the distinction between information and identificational focus, but the contrastive value is a descriptive fact that must be taken into account, no matter what theory we are using as a framework. If (28b) above is

derived through the ellipsis shown in (29a), a sentence equivalent to (32), the ellipsis alone can not explain why (32) is contrastive while (28b) is not.

- (32) LA MAGLIETTA\ ha vinto Gianni.
'It is the T-shirt what Gianni has won'

Therefore, I will assume that fragment answers do not move to focus. In the next section, 2.3, I propose that the relevant part of an answer does move, however, but it moves to an answer-specific functional projection, Answer Phrase.

While data from non-fragment Spanish (11d) and Italian (31b) answers does not show movement in any obvious way²³, the mere existence of fragment answers is a reason to support a movement analysis of answers. In fragment answers, there has been ellipsis. An analysis of fragment answers that postulates movement of the relevant part of the answer (i.e., the part of the answer that is actually pronounced) would require only one additional rule to explain the ellipsis of the rest of the sentence. Without the movement, it would be difficult to explain a fragment answer where constituents both to its right and its left would have been elided (this difficulty was pointed out in work by Merchant, 2004, and others).

²³ With the possible exception of non-contrastive answers with fronted indefinite objects in Chilean Spanish, as discussed in section 2.2.1.2.1.

So where is all this leading us? Let us briefly recapitulate the issues on the table. The previous sections pointed at a differentiation of answers from identificational focus from a semantic and an intonational perspective:

- Evidence from sentence interpretation, discussed in 2.1, suggests that contrastive focus and answers are different phenomena.
- Answers have specific intonational characteristics, discussed in 2.2.1, different from the A-accent associated to focus.

This section addressed the issue of fragment answers and arrived to the following conclusions:

- Ellipsis seems to be a better explanation for fragments (at least in answers) than an approach that derives the fragment without any additional structure. A number of reasons, summarized above, is compiled in Merchant (2004). To them, I add information structure form-identity between fragments and their full sentence correlates.
- An ellipsis explanation of fragment answers seems to require a movement operation that precedes ellipsis. This movement operation would extract the non-elided part of the sentence. Otherwise, obtaining the fragment might require

more than one ellipsis, in order to elide elements to the left and to the right of the remaining fragment (Merchant, 2004).

- Answer fragments do not move to the contrastive focus position, FocP.

So, if answers do not move to FocP, where do they move to? Do they move at all? An answer-specific projection, an Answer Phrase, would provide an explanation for answer-specific behavior, like the tendency to appear as a fragment and its particular intonation, and would place answers at a par with focus and topic as discourse-related syntactic phenomena. The next section explores the possibility that a phrase becomes an answer not as a result of external discourse conditions, but of a syntactic operation.

2.3 Are answers syntactic objects *qua* answers?

The previous section, 2.2, presented Spanish data that separates, within an utterance intended as a reply to a wh-question, the role of focus from the role of answer. Not only is contrastive focus incompatible with the role of an answer, but also information focus seems to be only loosely linked, if at all, to answers. Additionally, the data suggests the possibility of a syntactic structure specific to answers. This issue is well worth exploring, since it could potentially shed light on the link between the meanings of interrogation and focus, and add to our understanding of the relationship between syntax and information structure. This section addresses the syntactic character

of answers, and concludes by proposing that answers move to the left periphery, albeit not to the same position as contrastive focus, as has been shown in section 2.2.2.2. I propose the existence of an answer-specific syntactic position, which I will call Answer Phrase (AnsP), within the Focus set of projections (Benincà & Poletto, 2004), and in a fixed position within it.

There are at least three characteristics of answers that make them recognizable to the listener. First, answers have a specific intonation pattern. The second characteristic is the prevalence of ellipsis in answers and the way such ellipsis is related to the question asked. The third characteristic is that answers can include lexical items that are only used in answers and would not be appropriate as part of a regular declarative sentence (take English *yes*, for instance). All three of these characteristics are compatible with, and suggest the existence of, a syntactic operation.

Intonational events are associated to movement to the left periphery of the sentence (or, if not to movement, at least to syntactic positions there) in the cases of focalization, topic dislocation, and *wh*-movement. At least focalization and *wh*-movement have in situ correlates marked by intonation only. The similarities between the characteristics of answers and these syntactic operations are suggestive, but I will not explore the intonation issue in depth here. Some initial description of the intonation

patterns of Spanish answers was addressed in section 2.2, but a complete discussion of intonation issues falls outside the scope of this dissertation.

Section 2.3.1 looks at word order phenomena in answers, the basics of which (including the issue of fragment answers) were introduced in section 2.2, and discusses the location of seemingly non-left-peripheral answers. I will make the proposal that all answers involve movement to the left periphery of an XP including the relevant part of the answer, even in those cases where the relevant part of the answer appears in sentence-final position.

If a position on the left periphery is to be postulated, ideally it should be visible, i.e. some lexical item should be in that position under certain conditions. In section 2.3.2 I argue that English *yes* and its Spanish correlate *sí* are placed in an answer-specific position in the left periphery. Section 2.3.2.1 proposes that such elements undergo movement to arrive there. The last subsection, 2.3.3, discusses an alternative approach to subject answers (Belletti, 2005).

2.3.1 Answer Phrases

This section studies data on answers in order to test the hypothesis that answers are located in the left periphery. An indication that this might be the case is the ellipsis in fragment answers, which is easier to explain if the relevant part of the reply has moved before ellipsis applies (Alonso-Ovalle & Guerzoni, 2004; Brunetti, 2003; Merchant, 2004).

The obvious initial test for the idea that answers are located in a functional projection in the left periphery is to check whether it is the case that answers appear dislocated to the left of the sentence. As shown in section 2.2.1, that is an option for subject answers (10c,f), as well as for indefinite object answers in Chilean Spanish (19, 20). There are cases, however, of subject answers appearing at the end of the sentence (10m), as is also the case for object answers in most dialects of Spanish (11).

Subject answers where the subject appears on the left (10c,f) present no contradiction with the idea that the answer is in the left periphery. Their particular intonation, which separates them from the rest of the sentence, offers a parallel with other left periphery elements, such as focus. This might be a case where the answer element has moved to the left periphery and the ellipsis of the rest of the reply is optional.

Subject answers where the subject appears at the end of the sentence (10m) do not allow for such interpretation, nor do object answers, which must also appear at the end of the sentence. There are, however, indications that these answers have undergone some syntactic operation that differentiates their structure from that of regular declaratives. Consider the following examples:

- (33) a. Quién abrió la puerta el jueves?
 who opened the door the Thursday
 ‘Who opened the door on Thursday?’
- b. # La abrió Susana el jueves.
 it opened Susana the Thursday
 ‘Susana opened it on Thursday’
- c. El jueves, la abrió Susana.
 the Thursday, it opened Susana
 ‘On Thursday, Susana opened it’
- (34) a. Qué abrió Susana el jueves?
 what opened Susana the Thursday
 ‘What did Susana open on Thursday?’
- b. # Susana abrió la puerta el jueves.
 Susana opened the door the Thursday
 ‘Susana opened the door on Thursday’
- c. El jueves, Susana abrió la puerta.
 the Thursday, Susana opened the door
 ‘On Thursday, Susana opened the door’
- d. El jueves, la puerta.
 the Thursday, the door
 ‘On Thursday, the door’

- (35) Cuándo abrió Susana la puerta?
 when opened Susana the door
 ‘When did Susana open the door’

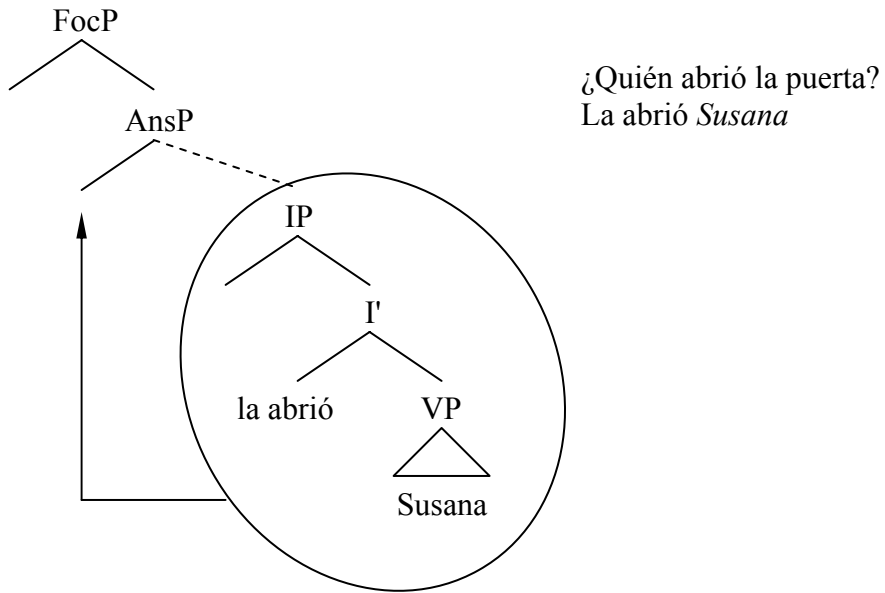
An adjunct cannot follow the answer in a reply, whether the answer is a subject or an object. Thus (33b) and (34b) are not valid answers to (33a) and (34a), although they are both good as answers to (35). This observation can be expressed as a requirement that the relevant part of the answer, if it is not left dislocated, is associated to the nuclear stress of the sentence. That is to say, the relevant part of the answer has to appear either in the left periphery or in sentence-final position²⁴.

I propose as an explanation for this set of facts that answers always move to the left periphery, where they are associated to a particular intonation. In other words, there is no difference in the syntactic position of left-dislocated and sentence-final answers. The sentence-final position of answers in certain replies is due to the possibility that other elements from the IP move together with them, by pied-piping, provided that the relevant part of the answer occupies the final position of the moved constituent. Examples of such pied-piping are provided in (36).

²⁴ Note that, in the same way there is left periphery focus and in situ focus, it would be expected that some dialects also present some sort of in situ answers, identifiable by their intonation. This seems to be the case of some varieties of Chilean Spanish. (Héctor Campos, p.c.)

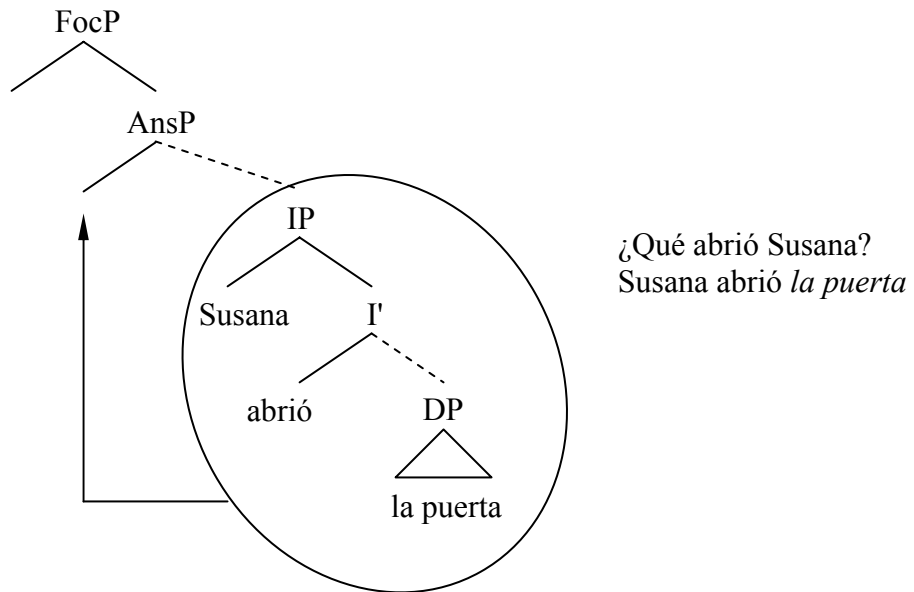
(36) PIED-PIPING IN SUBJECT ANSWERS

a.



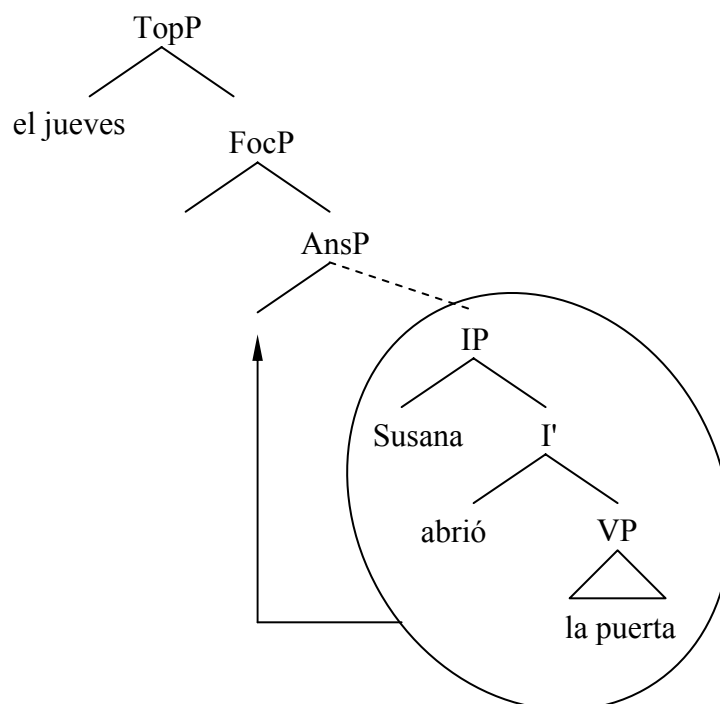
PIED-PIPING IN OBJECT ANSWERS

c.



Pied-piping does not apply to material following the relevant answer, which, if it is moved to the left periphery, must appear in a Topic position, assumed to be above FocP (Benincà, 2001). This is exemplified in (33c) and (34c,d), which are felicitous as replies to (33a) and (34a), although they are not innocent answers, in the sense that the presence of the dislocated topic brings in new presuppositions (e.g. perhaps Susana has been opening the windows, rather than the door, the rest of the week). Incidentally, the resulting word order shows that answers are placed below TopP in the structure. The structure of (34c) is presented in (37).

- (37) El jueves, Susana abrió la puerta.
 the Thursday, Susana opened the door
 ‘On Thursday, Susana opened the door’



Note that the common characteristic to all the possible answer structures is that, no matter what constituent has moved to the left periphery, the relevant answer is associated to the nuclear stress of the moved constituent. In full, non-fragment answers, the relevant part of the answer is still linked to answer intonation, which in those cases is associated to the final nuclear stress of the XP containing the answer. The infelicity of sentences like (33b) and (34b) is due to the impossibility of associating the answer intonation to the relevant answer in a structure where the relevant answer does not bear the nuclear stress²⁵. Recall that those same sentences are felicitous as replies to a question whose answer is the constituent bearing the nuclear stress, such as (35), repeated in (38) together with (33b) and (34b).

- (38) a. Cuándo abrió Susana la puerta?
 when opened Susana the door
 ‘When did Susana open the door’
- b. La abrió Susana el jueves.
 it opened Susana the Thursday
 ‘Susana opened it on Thursday’
- c. Susana abrió la puerta el jueves.
 Susana opened the door the Thursday
 ‘Susana opened the door on Thursday’

If this analysis is correct, it provides justification for the reduced inventory of full answer structures (as compared to declaratives) described in section 2.2. The

²⁵ Sentences where adjuncts seem to have ‘moved out of the way’ present an interesting parallel with Zubizarreta’s (1998) prosodic movement. Notice, however, that I am not implying a direct cause-effect relationship between the presence of an answer and the presence of sentence-initial adjuncts. I am simply suggesting that sentences with final adjuncts are prevented from being answers by prosodic reasons.

common characteristic of full answers that do not include the answer intonation split is that the relevant element of the answer appears in sentence-final position. A whole IP including a sentence-final answer could move to an Answer Phrase in the left periphery without affecting the assignment of the nuclear stress or the accent associated to answer intonation (the relevant part of the answer would receive the nuclear stress of the moved IP and the answer intonation would be associated to it).

Fragment answers, on the other hand, would result from movement of the relevant part of the answer alone, without pied-piping, and subsequent ellipsis of everything left behind. This solution is in fact quite similar to Merchant's (2004) treatment of ellipsis in sluicing and fragment answers. Data on full answers suggests that ellipsis and answers should not be completely linked; movement of answers appears to be mandatory, and ellipsis optional. Merchant (Merchant, 2001, 2004) takes the approach that ellipsis is caused by an E feature of the head whose complement is elided. Different varieties of this feature account for the different properties of ellipsis in different languages. In my account the E feature would be optionally associated to the head of Answer Phrase. There would be two varieties of these heads in the lexicon, one causing ellipsis, the other not causing it.

In section 2.2 it was noted that, while subject answers can appear fronted without elision of the rest of the sentence, object answers cannot. It was also observed

that this is not general to all dialects of Spanish; Chilean Spanish allows for indefinite object answers fronted, without ellipsis of the rest of the reply. I repeat here the relevant data, from (19) and (20).

- (39) a. A quién vio?
to who 3rd-past-see
'Who did he see?'
- b. A un chico colorín vio.
to a boy red-haired 3rd-past-see
'He saw a red-haired boy'
- c. A un exprofesor de escuela vio.
to a former-teacher of school 3rd-past-see
'He saw a former school teacher'
- (40) a. Qué compraste ayer?
what 2nd-past-buy yesterday
'What did you buy yesterday?'
- b. Un libro compré, pero luego lo devolví.
a book 1st-past-buy but later CL-it 1st-past-return
'I bought a book, but I returned it later'

The explanation for the absence of fronted object answers in Peninsular Spanish might very well be the lack of a specific output condition for those answers, such as a different intonation contour or different vowel length in the stressed syllable. There is nothing that makes them different enough from other fronted object constructions, such as regular declarative sentences or sentences including contrastive focus. Chilean Spanish, on the other hand, does present a very clear distinction between the focus intonation and

the answer intonation, which allows for both constructions to be easily identifiable. Constituents identified as answers present a longer stressed vowel and a higher pitch than they would in the corresponding declarative sentence, but they are still not quite the same as the contrastive constructions that present an A-accent.

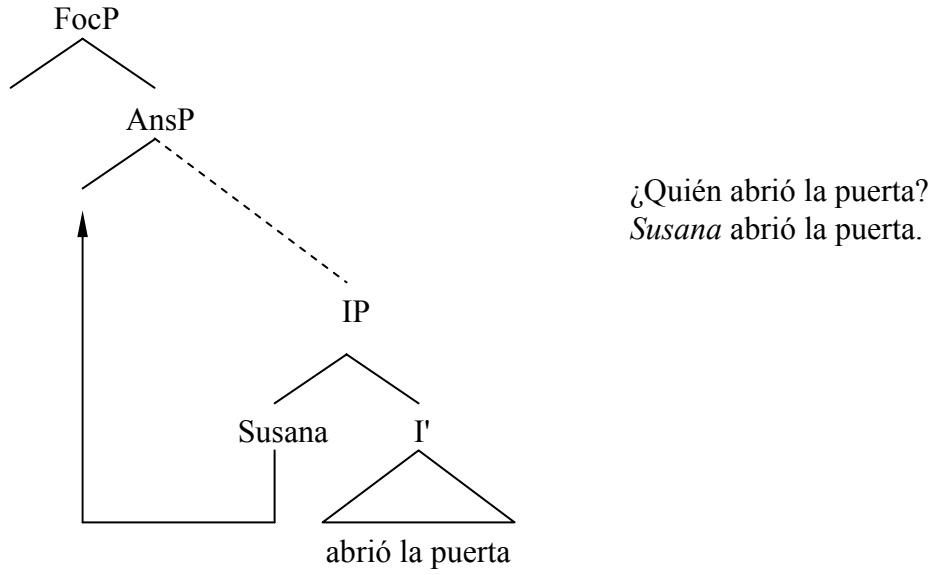
Recall from (10) that Peninsular Spanish subjects do present an intonational difference between answers, declaratives, and contrastive focus constructions. This differentiation allows for the acceptability of all three types of constructions in the appropriate contexts.

Clues to the relative position of AnsP in the left periphery are given so far by the fact that it follows the topic in sentences like (33c) and (34c), that is, below TopP. This, added to the typological similarity between answers and focus, suggests that answers are placed somewhere in the Focus set of projections (Benincà & Poletto, 2004). I have shown in section 2.2.2.2 that answers do not move to the locus of identificational focus, against what was proposed by Brunetti (2003) and suspected by Merchant (2004). Their work shows, however, that there is every indication that AnswerP must be in the same general area FocP is. I will discuss its location more specifically in chapter four of this work.

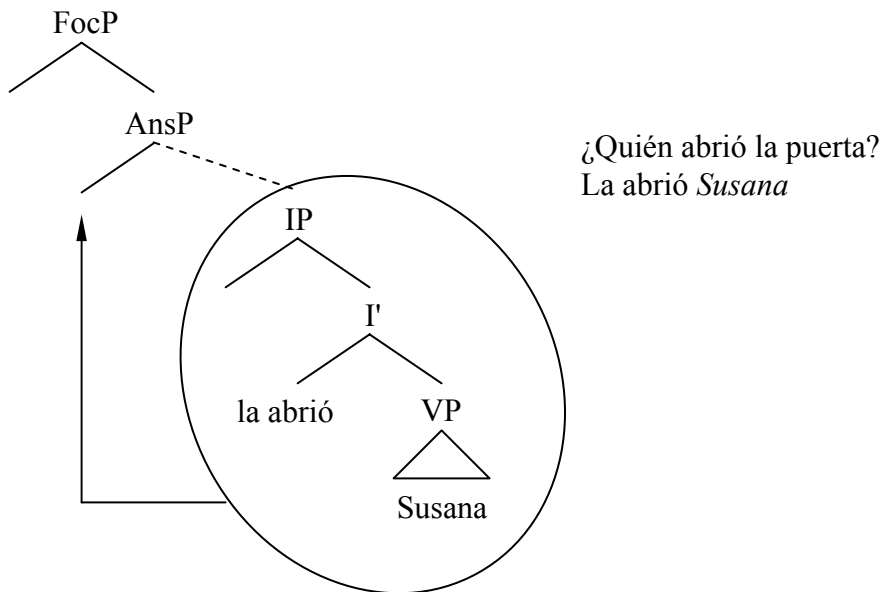
A summary of the structures corresponding to the answer typology described in the previous section follows below. In (41) I present the structures corresponding to full answers where no adjuncts are present. (41a) corresponds to (10f), the sentence-initial subject answer. (41b) corresponds to (10m), the sentence-final subject answer, and (41c) corresponds to (11d), the sentence-final object answer. The structures in (42a,b) correspond, respectively, to subject (10a) and object (11a) fragment answers. The E feature associated to ellipsis is represented associated to the head of Answer Phrase. Finally, the structure in (43) represents sentences with additional material in topic position (34c). Note also that (34d) is correctly predicted to be a good answer by this analysis. The structure would be similar to the one shown in (42b), but including a TopP.

(41) FULL ANSWERS, NO ADJUNCTS

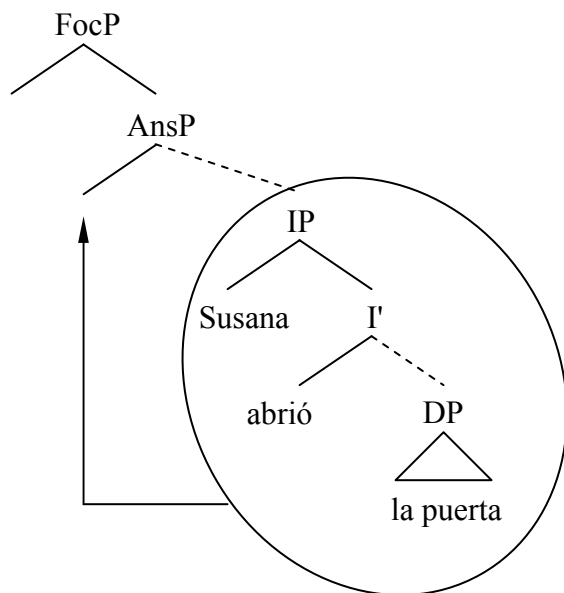
a.



b.



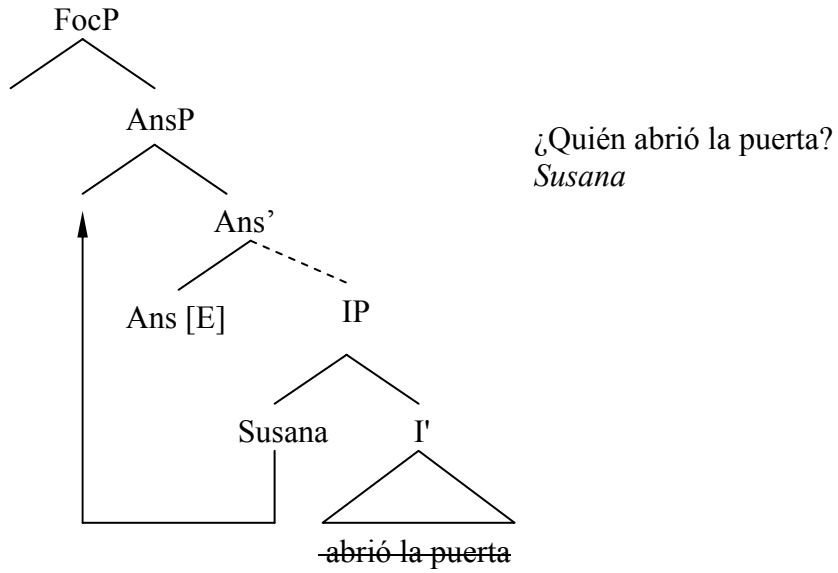
c.



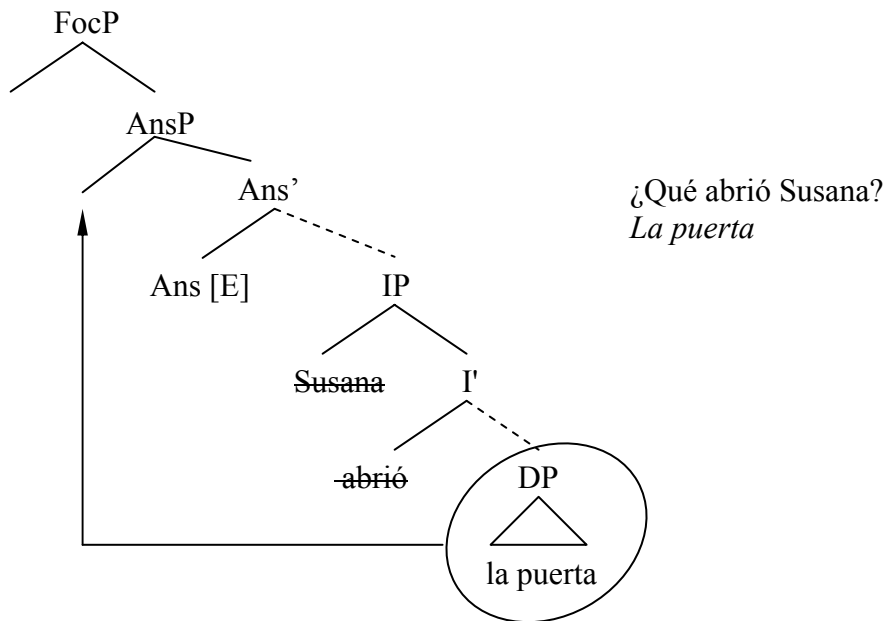
¿Qué abrió Susana?
Susana abrió *la puerta*

(42) FRAGMENT ANSWERS

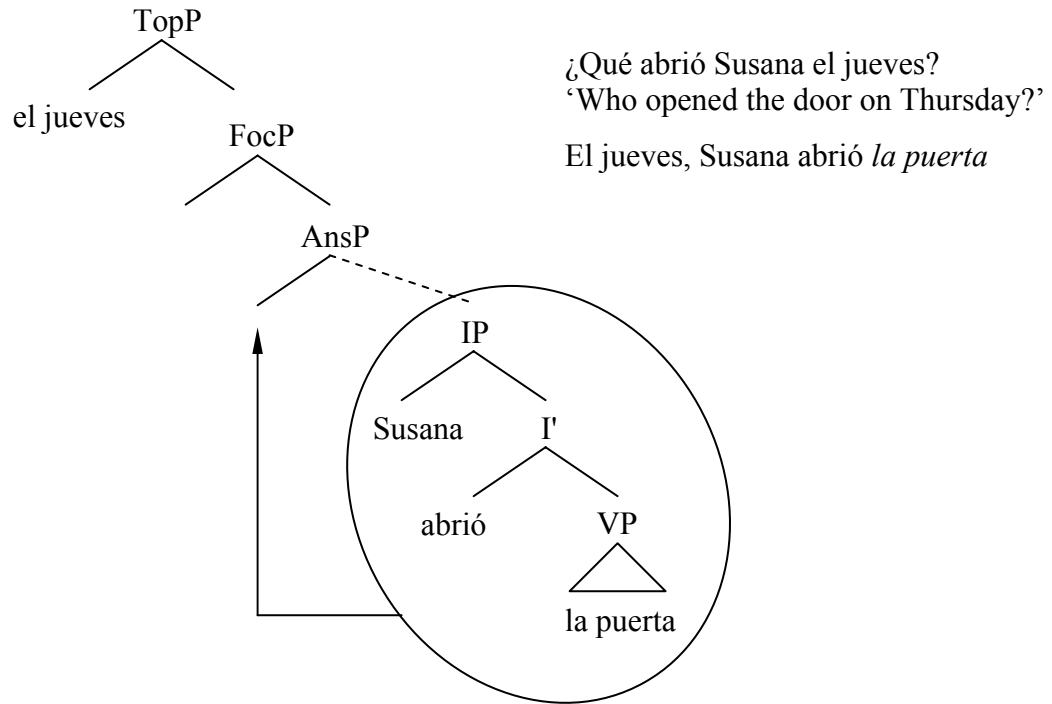
a.



b.



(43) FULL ANSWERS INCLUDING ADDITIONAL MATERIAL ON THE LEFT



2.3.2 The structure of yes/no answers

The characteristic shape of an affirmative answer to a yes/no question in Spanish, English, and a number of other languages, is a two-part structure where the two parts are separated by an intonational event –an accent– which is associated to an optional pause and is represented by a comma in writing (44b,c).

- (44) a. Do you eat pork?
b. Yes, I do.
c. Yes, I eat it.
d. Yes.
e. I do.

The first part of this two-part structure consists of one lexical item, very closely associated to the sentence-initial position it occupies in this type of structure. In addition to being the only lexical item able to occupy this position, the English version of this item, *yes*, cannot appear in contexts other than answers (leaving interjections aside).

The second part of the answer consists of a declarative sentence where a number of elements concurrent in the question may be elided.

Both parts of the answer can appear alone as answers (44d,e), though the second part, the declarative, is pragmatically charged when it appears by itself, and creates an impression of formality.

Spanish presents a similar picture, with minor differences. No auxiliary verbs like *do* are available in Spanish, and clitics can take the place of elided components²⁶.

- (45) a. Terminó Sergio la cena?
'Did Sergio finish up his dinner?'
- b. Sí, la terminó. / Sí. / La terminó.
yes it he-finished

Notice that both English and Spanish (and other languages, like Portuguese) allow for an affirmative element following, rather than preceding, the full answer (46b,d). These right dislocated phrasal adverbs might be afterthoughts (see Villalba, 2000 and chapter 2 of this work for details on afterthoughts) or completely different sentences altogether. One possible analysis is discussed in section 2.3.2.1. In any case, spectrographic analysis might shed additional light on this issue.

²⁶ Note that there is a great range of variation between Spanish dialects in terms of clitic usage. The data represented here reflects Northwestern Peninsular Spanish, where the etymological clitic system is preserved. The other extreme of the spectrum would be represented by colloquial Quiteño Spanish, where all direct object clitics have been replaced by (etymological) indirect object clitics, and there is no co-occurrence of direct and indirect object clitics. Additionally, Quiteño Spanish allows for zero realization of direct object clitics if their reference is present in left dislocated position or in the immediately preceding sentence (Suñer & Yépez, 1988). In the example in (45) the answer would be *Sí, terminó*. 'yes, he finished'.

- (46) a. Do you eat pork?
b. I do, yes.
c. Terminó Sergio la cena?
'Did Sergio finish up his dinner?'
d. (Sí) la terminó, sí.
yes it he-finished, yes

Unlike in English, the Spanish affirmative lexical item can appear in other contexts, where it is associated to the verb and adds an emphatic/confirmative element to the semantics of the sentence (47b). In Peninsular Spanish, this emphasis expresses a contradiction between the affirmative sentence and an assumed negative belief of the interlocutor. In Chilean Spanish, however, there is not necessarily an implied contradiction; the presence of *sí* may simply mark that the sentence containing it is responding to a yes/no question. In order to put these facts in context, and before continuing the analysis of affirmative answers, I will dedicate the end of this section to briefly discuss affirmative elements and determine the nature of the lexical items involved and comparing them in Spanish, English, and German.

English presents one affirmative element, *yes*, which (leaving aside indirect speech) appears only by itself as described above. Spanish has a single affirmative lexical element as well, *sí*, but it can appear both by itself and associated to a verb. German presents two separate lexical items, *ja* and *doch*, both of them able to appear by

themselves when stressed. Of these, *doch* expresses a contradiction between the affirmative answer and a negative statement implied by the question. There is also an unstressed *doch* also used in association to a verb. There is more complication to the German system, and more uses of both particles, but the facts presented will suffice for our purposes here. The presence of separate items in German, together with the absence of a lexical item for one of the positions in English, suggests that, in spite of the homophony, there might be two differentiated Spanish *sí* elements, an IP-external one and an IP-internal one^{27,28}. In terms of its placement in the structure, Spanish *sí* can

²⁷ An analysis that contemplates two different *sí* particles seems to be supported by the fact that other Spanish particles, like *ya*, can have functions similar to the internal affirmative element (Rohlf, 1970, §529). Enunciative *ya* (not to be confused with the homophonous adverb *ya* ‘already’) presents several properties that justify its classification as an inflectional element (Campos, 1986, p. 192). Among those properties are its association to a verb, its being restricted to finite clauses, and its incompatibility with negation.

- (i) a. Ves esas nubes? Ya las veo. (Campos, 1986)
 see those clouds ‘ya’ them see
 ‘Do you see those clouds? I do see them’
- b. Crees que Jorge ha trabajado mucho en su tesis? (Campos, p.c.)
 ‘Do you think Jorge has worked a lot on his dissertation?’
 Jorge ya le ha metido horas a su tesis!
 Jorge ‘ya’ IO-CI has put hours to his dissertation
 ‘Jorge has indeed spent hours on his dissertation’

Héctor Campos (p.c.) points out that Chilean Spanish presents uses of *ya* with the meaning of ‘yes’

- (ii) Vienes a la fiesta esta noche?
 Ya!

This use of *ya* seems to make it completely parallel to *sí* in Chilean Spanish. The existence of varieties differentiated by the use of a lexical item *ya* in one of the positions where we find *sí* besides varieties where *ya* is used in both positions reinforces the idea that lexical selection for each of those positions takes place separately.

The enunciative use of *ya* is common to many dialects of Spanish, although the details of its use and interpretation vary greatly from one dialect to another (for instance, (ia) above does not sound natural in Northwestern Peninsular Spanish; *ya* would be interpreted to be the adverb with the meaning of ‘already’.

The example in (ib), however, is perfectly fine in that dialect). I will therefore leave it out of this discussion, merely pointing out that its analysis as an inflectional element and the proposals put forth in this work seem to support each other. At the same time, its differences with *sí* make it a very promising element for further microcomparative work in this line of research.

²⁸ The affirmative adverb *sí* can appear followed by the complementizer *que* (see for instance Bello, 1988):

- (i) *Sí que se me está haciendo tarde.*
 yes that cl cl is making late
 ‘It is getting really late for me’

I take such adverbs to be a different element from the sentential adverbs in answers. There are several noticeable differences between *sí que* constructions and affirmative answers. In Peninsular Spanish (at least) it can be used in contexts other than answers, like as a conversation starter, expressing surprise or complaint. The function of *sí que* constructions is not to reinforce the sentence affirmation, but merely to emphasize the focused element, which usually appears preceding *sí* (Campos, 1986, 1992), and therefore they are compatible with negative statements (ii). The capitals in the examples of this section reflect the contrastive focus value of the capitalized elements.

- (ii) *Marta sí que no trabaja mucho.*
 Marta sí que not works much.
 ‘MARTA does not work much’

In addition, the *sí* in *sí que* is not an internal sentential marker, either. Campos (1992) points out clear semantic and syntactic differences between *sí que* and *sí* constructions. The *sí* construction can be interpreted to be reinforcing the affirmation of the sentence, while *sí que* does not, as seen above. Besides, *sí que* does not need to be contiguous to the verb (iii); also, *sí que* and *sí* can appear together in the same sentence (iv).

- (iii) a. *Juan sí que a veces se enfada con su novia*
 Juan sí que at times cl angers with his girlfriend
 ‘JUAN gets angry with his girlfriend sometimes’
- b. **Juan sí a veces se enfada con su novia*
 Juan sí at times cl angers with his girlfriend
 ‘Juan gets angry with his girlfriend sometimes’
- c. *Juan a veces sí se enfada con su novia*
 Juan at times sí cl angers with his girlfriend
 ‘Juan gets angry with his girlfriend sometimes’
- (iv) *Juan sí que sí se levanta temprano*
 Juan sí que sí cl gets up early
 ‘JUAN does indeed get up early’

Campos (1992) accounts for these syntactic facts, as well as for the difference in meaning the focus value by concluding that the *sí* in *sí que* is a contrastive focus marker followed by a complementizer.

appear associated to an IP (47b), where only clitics can get between *sí* and the verb. The German version of this internal affirmative element, *doch*, appears between the auxiliary and the verb.

- (47) a. I think he did (Ø/not) do it.
- b. Creo que (sí/no) lo hizo.
I-think that aff/not it he-did
'I think he did/did not do it'
- c. Ich glaube er hat es (doch/nicht) gemacht.
I believe he has it aff/not done
'I think he did/did not do it'

There is also a third Spanish *sí*, often overlooked, which is used in the same affirmative-contradictory sense as German IP-external stressed *doch*. This *sí* presents a longer vowel with a shifting intonation, and it can even be bisyllabic in some dialects. This type of *sí* is not allowed in IP-internal position.

It is easy to draw a parallel between affirmative and negative elements. English presents two negative elements *no/not*, of which the first one appears independently and can be described as a phrasal negative adverb (Merchant, 2001) and the second one appears associated to a verb. Depending on the language, this last element can be

The possibility of dialects where a *sí que* construction can function as an answer is explained by assuming that those dialects allow the *sí* element present in AnsP to take CP complements, instead of limiting this ability to the focus mark *sí*. The existence of both types of dialects supports the idea that there are two different positions for *sí* in the left periphery: one in FocP, described by Campos (1992), and one in AnsP.

interpreted either as a NegP (Zanuttini, 1990, 1991) or as a clitic; additional elements can be involved (Haegeman, 1995; Potsdam, 1997; Zanuttini, 1997). A contracted form *n't* can appear attached to the verb if not emphasis or contradiction is intended²⁹. Similar pairs appear in German (*nein/nicht*) and other languages (see Merchant, 2001 for a short overview). English has parallel syntactic structures for emphatic affirmation and negation (47a); in both cases the verb is separated from its morphological marks of person and tense, which appear attached to an auxiliary verb. This suggests that an English sentential affirmative marker is visible to syntax, even though it is not phonologically realized (or at least not realized as a sequence of phonemes. Distinct intonation and stress are certainly associated to emphatic affirmation). This idea goes back at least to Chomsky (1957). Finally, the Spanish system presents the same item *no* in all cases, with the possibility of having an emphatic intonation as a phrasal adverb.

²⁹ For a more detailed discussion of English *not* and *n't* see, for instance, Jackendoff (1972). This brief summary does not intend to present all the complexities of the English or German systems, but rather to present some facts in those systems that are parallel to the Spanish facts discussed.

(48)

		phrasal adverb		sentential marker	
			emphatic or contradicting		emphatic or contradicting
English	Affirmative	yes	yes	∅ (and presence of a modal verb)	∅ (and presence of a stressed modal verb)
	Negative	no	no	n't	not
Peninsular Spanish	Affirmative	sí	sí-i	∅	sí
	Negative	no	no, no-o	no	no
Chilean Spanish	Affirmative	sí	sí-i	sí, ∅	sí
	Negative	no	no, no-o	no	no
German	Affirmative	ja	doch	∅	doch
	Negative	nein	nein	nicht	nicht

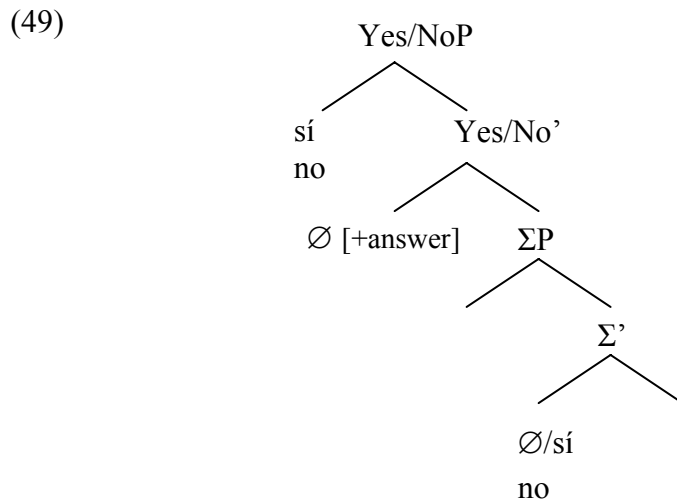
The parallel between affirmative and negative sentential markers has been noted in much of the literature. Laka (1990) proposes, on the basis of English, Basque, and Spanish data, that affirmative and negative markers are different instantiations of the same functional projection, which she calls Σ Phrase.

I will assume Laka's Σ Phrase, and I will devote the next section to explore the relationship between Σ Phrase (which includes the sentential markers) and the so-called phrasal affirmative adverbs (i.e., the affirmative adverbs that appear in sentence initial

position followed by comma intonation). I propose that these adverbs are placed in a syntactic projection in the left periphery, specifically in AnsP. The placement in the left periphery is supported by the fact that they appear on the leftmost part of the sentence and, like focus or topic, are separated from the rest of the sentence by an intonational event. Proving that this placement in the left periphery is the result of some A'-like movement, similar to focus movement, as opposed to simply being base-generated there, is more complicated. There is no obvious evidence of movement, since there is no evident link between the affirmative element on the left and the remaining material (the affirmative/negative sentence) on the right. Specifically, there is no gap where the moved adverb must be interpreted, as in wh-movement. If there is a relationship between the adverb and the rest of the sentence, it must be the kind of relationship that exists between a resumptive pronoun and its referent, or the relationship between elements in cases of doubling. The following subsection explores a possible analysis of that relationship in terms of extended projections.

2.3.2.1 The nature of affirmative/negative phrasal adverbs

This subsection explores the idea that left-peripheral affirmative/negative phrasal adverbs are generated in an extended projection of the Σ Phrase, which I will call Yes/No Phrase. The structure of the doubling would be based on Grimshaw's (1991) Extended Projection idea and can be represented as follows:

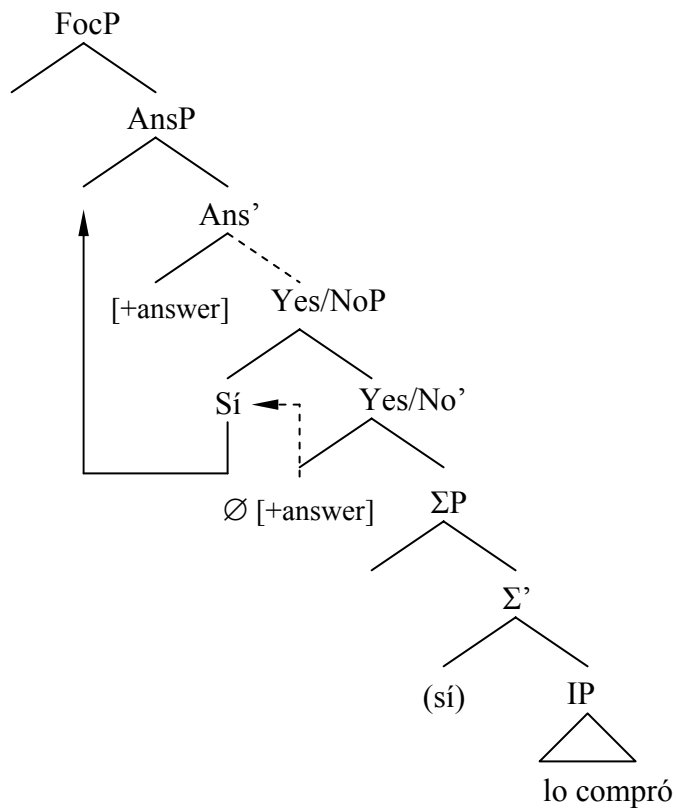


Each part of this doubled structure would correspond to one of the elements of the two-part structure described for yes/no answers in the previous section. The adjoined Yes/No phrase carries a [+answer] feature, which will trigger movement to the left periphery. The feature can associate itself to the phrasal adverb occupying the specifier position of the Yes/No Phrase, or, if there is no specifier, to the Σ Phrase to

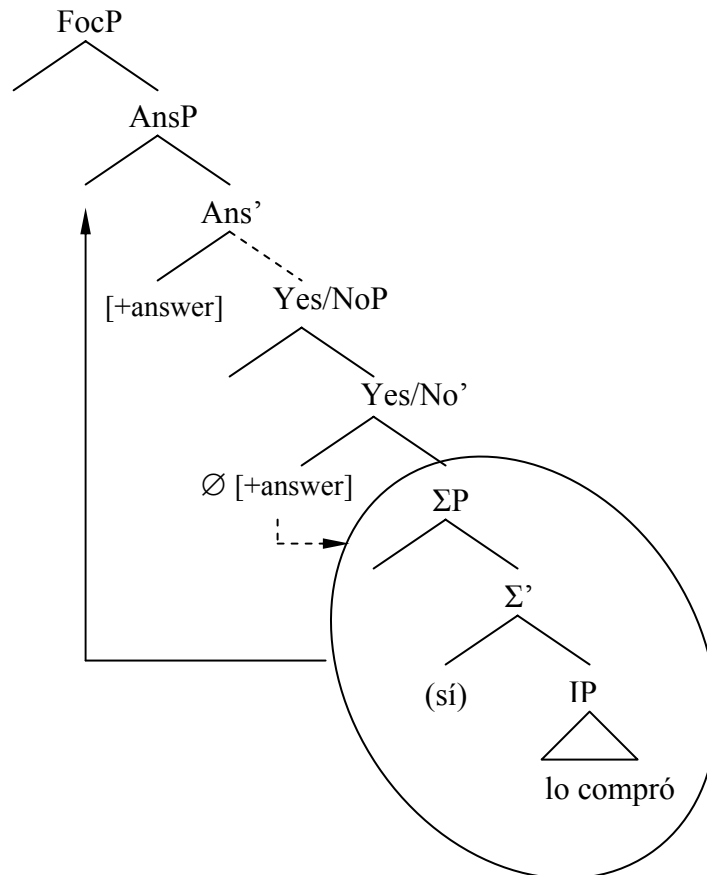
which the Yes/No Phrase is adjoined, as shown by the examples in (51), answers to the question in (50). The answer in (51a) corresponds to the case where the specifier of the adjoined element (the Yes/No Phrase) moves to AnsP while (51b) corresponds to the movement of Σ Phrase to AnsP. As mentioned in section 2.3.2, there are cases where the phrasal adverb appears right-dislocated. These cases would be explained the same way as (51b), with the only difference that a phrasal adverb would be present in Spec, Yes/No P.

(50) Lo compró?
 cl bought-3sg
 'Did he buy it?'

(51) a. Sí, (sí) lo compró.
 yes, aff cl bought-3sg
 'Yes, he bought it'



- (51) b. (Sí) lo compró.
 aff cl bought-3sg
 'He bought it'



Note that this presence of a double instance of affirmative/negative words is a unique characteristic of answers, both affirmative and negative. It does not take place in emphatic or negative declarative sentences, where there is no IP-external phrasal adverb, and only the IP-internal affirmative/negative sentential marker expresses the

polarity of the sentence. On the basis of this fact, it could be hypothesized that the reason for this double presence of an affirmative/negative word is that the answer character of these sentences requires the introduction of a [+answer] feature, which will be associated to the affirmative/negative character of the sentence. This feature would be introduced as part of an extended projection of the affirmative/negative element already present in the sentence in the Σ Phrase. The interpretable feature in this extended projection (Yes/No Phrase) would then be in agreement with its uninterpretable counterpart in the left periphery, hosted in the Answer Phrase discussed in section 2.3.1 above, which would also have an EPP feature, causing the movement to take place.

If this analysis is correct, evidence of movement of the affirmative adverb should be expected, as well as evidence of sensitivity to certain weak islands, since the moved element is being analyzed as a specifier/adjunct. Consider the differences in grammaticality of the examples of (52), taken as answers to the question in (50). Movement is possible in cases like (52a), but not possible in (52c), which would be a good answer to a question like *Does it bother you that he bought it?*, but not to *Did he buy it?*. Other cases present different degrees of acceptability (52b)³⁰.

³⁰ Answers including bridge verbs like *esperar* ‘to hope, to expect’ can also be unacceptable (see (ib) below), in spite of the possibility of movement.

- (52) Lo compró?
 cl bought-3sg
 ‘Did he buy it?’
- a. Sí, creo que (sí) lo compró.
 yes, believe-1sg that aff cl bought-3sg
 ‘Yes, I think he bought it’
- b. ?? Sí, oí el rumor de que (sí) lo compró.
 yes heard-1sg the rumor of that aff cl bought-3sg
 ‘Yes, I heard the rumor that he bought it’
- c. * Sí, me molesta que (sí) lo comprara.
 yes cl bother-3sg that aff cl bought-subj-3sg
 ‘Yes, it bothers me that he bought it’

It could be stipulated, as a way to achieve a simpler structure, that the position in the left periphery where the affirmative element appears is the same to which wh-words move in questions. If this were the case, answers would simply be a version of

-
- (i) a. Lo compró Luis?
 it bought Luis
 ‘Did Luis buy it?’
- b. *Sí, espero que lo haya comprado.
 yes, I-hope that it has-subj bought
 ‘Yes, I hope he has bought it’
- c. Cuándo esperas que lo compre?
 when you-expect that it he-buys
 ‘When do you expect him to buy it?’

In this case the unacceptability of (ib) might be derived not from obstacles to movement, but rather from the contradiction between the presence of *sí* meaning ‘yes, he bought it’ and an expression of hope which implies, at least, lack of knowledge about whether Luis bought it or not. However, it is also possible that irrealis verbs like *esperar* do not align themselves syntactically with non-factive verbs like *say* in this case, and they do constitute a weak island for answer movement, even if they are not one for wh-movement. For a recent discussion of the behavior of factive/irrealis/non-factive verbs, see de Cuba (2007). For additional observations on answer movement and weak islands, see chapter 4 of this work.

questions where the requested information is available, rather than replaced by a wh-word. This possibility is explored and rejected in chapter 4 of this work on the basis of the syntax of split questions, a construction where separate syntactic positions for answers and for wh-elements are required. The wh- position is shown to be lower than the Answer Phrase.

2.3.3 The hypothesis of the placement of answers in the VP periphery

A scenario that provides movement of an answer, but not to the CP layer, has been sketched by Belletti for Italian subject answers (Belletti, 2004, 2005). Her idea is that in answers to questions concerning the subject of the clause, the subject raises to a focus phrase in the VP periphery. The assumption is that the contrast/correction value of focus is restricted, in Italian, to the focus phrase in the external layer above the IP. The focus phrase in the internal layer, above the VP, would be associated to new information. I quote here the main lines of her analysis of subjects:

[CP ...[TP *pro*...è...partito/ha parlato...[TopP [FocP Gianni [TopP [VP]]]]]]³¹

I assume (much as in traditional accounts) that the relevant preverbal subject position (Cardinaletti (2004)) is occupied by a non overt null *pro*. I also assume (differently from traditional accounts, Belletti (2003)) a doubling derivation of “ subject inversion”

³¹ A gloss for Belletti’s example would be *is left/has spoken Gianni* ‘Gianni has left/has spoken’.

structures, with a *referential pro* moved from an original “big DP” filling the (relevant, Cardinaletti (2004)) subject position, and the lexical subject stranded in the low focus (or topic, in different discourse conditions[...]) position. (Belletti, 2005, p. 2)

The approach taken by Belletti to the issue of answers is thus very different from the position assumed here. The presence of a subject answer in a particular position is, for Belletti, a particular case of a more general syntactic phenomenon involving subject inversion. Thus, it is not to be expected that her analysis accounts for the position of other types of answers. In the previous section I have put forth a hypothesis that unifies the analysis for subject answers and object answers in Spanish. Choosing an analysis over another will, of course, depend on how accurately they can reflect actual utterances. In the absence of clearer evidence for or against either view, I see more promise in an analysis that conflates different types of answers into one model, while other researchers will be more interested in the subtleties of the position of the subject and will not perceive the lack of an account for object answers as a shortcoming. On the other hand, it is conceivable that the VP periphery also contains an Answer Phrase, since we tend to find in it the same elements we find in the left periphery. This would make both approaches compatible.

It might be worth noting here that the direct application of an analysis like Belletti's to Spanish would be precluded by the fact that full answers to *Quién abrió la puerta?* ‘Who opened the door?’ have their subjects fronted rather than following the

verb (as in (10f) *Susana, abrió la puerta* v. *#Abrió Susana la puerta*). The cases where the subject answer is postverbal, like (10m), repeated here in (53b), could however be analyzed following Belletti's idea of an answer position in the VP periphery. In fact, such an analysis would provide a syntactic distinction between the kind of replies where the answer is a fragment of the IP, but the whole IP still appears to have moved (53b) and those cases where the whole IP is the answer, like (12c), repeated in (54b).

- (53) a. Quién abrió la puerta?
 who opened-3sg the door
 'Who opened the door?'
- b. [IP La abrió [AnsP Susana [VP ~~Susana~~]]]
 it opened-3sg Susana
 'Susana opened it'
- (54) a. Qué pasó con la puerta?
 what happened-3sg with the door
 'What happened with the door?'
- b. [AnsP La abrió Susana [IP ~~La abrió Susana~~]]
 it opened-3sg Susana
 'Susana opened it'

A thorough comparison of answer positions in Italian and Spanish will be a task for future research.

2.4 Conclusions

In this section I have addressed the link between focus and answers, to show that no link can be established between the two, be it contrastive/identificational focus or information focus. Answers do show certain particularities that make them similar to identificational focus: unique intonation and fixed positions in the sentence that suggest movement to the left periphery, although the intonation of contrastive focus is different from the intonation associated to answers.

On the basis of these similarities, I have proposed that the left periphery of the sentence includes a functional projection dedicated to answers, AnsP, which hosts both the answers to yes/no questions and the answers to wh-questions. In the case of yes/no questions, the element moving to the left periphery is a phrasal adverb, which carries a [+answer] feature. This adverb is added to the answer structure by a doubling-like extension of the Σ Phrase, from where it moves to AnsP. In the case of answers to wh-questions the moved element is the co-referent of the question's wh-word. Both types of answers allow for a larger element to move to the left periphery, as long as the relevant element of the answer still receives the stress to which the answer intonation is associated. In the case of yes/no questions, this kind of movement could be the result of a structure where the doubled phrase carrying the [+answer] feature does not have a phrasal adverb as its specifier.

This analysis explains the intonation differences between contrastive focus and answer structures, still allows for a movement-plus-ellipsis analysis for fragment answers like the one developed by Merchant (2004), and explains the differences in meaning between answers that include contrastive focus and those which do not. I will now discuss, in section 3, the differences between contrastive focus and wh-movement. This will allow for an analysis of the left periphery where identificational focus operates separately from the question/answer element of the information structure.

3 Focus and Interrogation as separate phenomena in the structure of questions

This section presents arguments to separate identificational focus features from wh- features as two independent lexical properties, which cause two different phenomena, respectively focus fronting and wh-movement . Section 3.1 shows examples where, against common assumptions, a question presents an identificational/contrastive focus which is different from the element carrying the wh- feature.

3.1 Examples of non-correspondence between focus and wh-elements

It is often argued that (identificational) focus and wh-elements are in complementary distribution³². That statement holds true for left periphery contrastive focus, but not for contrastive focus in situ, as I will show in this section. I take the examples in this section from Stepanov (2005), who quotes several sources himself (Bromberger, 1992; Dretske, 1972; Geilfuss-Wolfgang, 1995). Consider the following questions:

³² A different view is held, for instance, by Erteschik-Shir (1986), who introduces the notion of *Dominance* precisely to reflect the separation between the wh-word and the element of the question to which the speaker wishes to call attention (I thank Paul Hagstrom for pointing out her work to me).

- (55) a. Why did CLYDE lease a Jaguar?
b. Why did Clyde lease a JAGUAR?
c. Why did Clyde LEASE a Jaguar?
d. Why did Clyde lease a Jaguar?

Consider now their respective answers:

- (56) a. Because he (Clyde) was the most financially careless guy around.
b. Because he is fond of Jaguars.
c. Because he couldn't afford to buy one.
d. Because he didn't know what else to do in life.

The least that can be said about these examples is that they provide clear evidence that contrastive value can be separated from the *wh*-word in a question. The difference in meaning between the questions is clear in the answers, at least when the *wh*-word is *why*. According to Stepanov (2005), other *wh*-words present a different situation, where, even if one can mark *in situ* focus, the answer does not seem to change.

- (57) a. When did CLYDE lease a Jaguar?
b. When did Clyde lease a JAGUAR?
c. When did Clyde LEASE a Jaguar?
d. When did Clyde lease a Jaguar?

- (58) At 4 p.m. on July 7, 2001.

That is, the answer to a *when* wh-question can be argued to be independent of what looks like a contrastive in situ focus present in the question (Stepanov, 2005).³³ I believe, however, that a closer look might reveal that the difference is, at least, not as clear-cut as it appears. It is easy to find an answer that covers all the questions in (55). Take, for instance, (59). Although they are perhaps not so obvious, it is also possible to find appropriately different answers to the *when* questions in (57). I list some in (60).

(59) Because he got a great deal.

- (60) a. Right after Robert did.
b. When he found out he was getting a bonus.
c. When he realized he couldn't afford more mechanic bills for his old one.
d. As soon as he got a job.

Another argument to separate *why* from other wh-words would be the following contrast (Stepanov, 2005):

- (61) a. Why was it Clyde who leased a Jaguar?
b. *When was it Clyde who leased a Jaguar?

³³ The same contrastive in situ focus found in the question could be present in a longer answer, which adds support to the idea that contrastive focus can be placed in a different element than the one that substitutes the wh-word. (Paul Hagstrom, p.c.)

Such a contrast does not appear in Spanish, as shown in (62). There might be differences in acceptability for some speakers due to pragmatic reasons, but the construction is syntactically sound, as shown by the clearly acceptable, if pragmatically marked, (63), which would be used to imply that Clyde should not be the one making decisions. This same construction can also appear in a more elaborate context such as the dialogue in (64)³⁴.

- (62) a. Por qué fue Clyde el que compró un Jaguar?
 why was Clyde the that bought a Jaguar
 ‘Why was it Clyde who bought a Jaguar?’
- b. Cuándo fue Clyde el que compró un Jaguar?
 when was Clyde the that bought a Jaguar
 ‘When was it Clyde who bought a Jaguar?’
- (63) Desde cuándo es Clyde el que toma decisiones en esta casa?
 since when is Clyde the that take decisions in this house
 ‘Since when is it Clyde who makes decisions in this household?’
- (64) a. A: —Me han dicho que editas la revista de la escuela.
 ‘I’ve been told you edit the school magazine’
- b. B: —La editamos varios estudiantes, tomando turnos.
 ‘It is edited by several students, including me, taking turns’
- c. A: —Cuándo eres tú el que la edita?
 when are you the that it edits
 ‘When are you the editor?’

³⁴ Paul Hagstrom (p.c.) points out to me that similar contexts in English can make (61b) felicitous. Therefore we have no arguments to separate *why* from other *wh*-words.

- d. B: —La segunda semana de cada mes.
'The second week of each month'

Then, besides the relative difficulty of finding an appropriate pragmatic context, it can be concluded that focus in situ can appear together with wh-elements in a question, whether the wh-element is *why* or another wh-word.

If focus in situ is the same as left periphery contrastive focus, with the only difference that it moves covertly, then the examples given show that wh- features and focus features can coexist in the same interrogative sentence. The idea that in situ focus moves covertly is commonly accepted since Chomsky (1976) showed that in situ focus presented weak crossover effects³⁵, and has been argued for Spanish on the basis of these same effects by Domínguez (2004, p. 177). I copy her example of moved focus in (65a) and in situ focus in (65b).

- (65) a. *_F[A Juan]_i ha matado su_i madre t_i
to Juan has-killed-3s his mother
‘It is Juan whom his mother has killed’
- b. *Su_i madre ha matado [_F a Juan]_i
his mother has-killed-3s to Juan

³⁵ For a different interpretation of those effects, see Rooth (1985).

Note, incidentally, that CLLD constructions, associated with the Topic position, do not present such effects and are usually assumed not to involve movement³⁶.

- (66) a. [Top A Juan]_i lo_i ha matado su_i madre t_i
to Juan cl has-killed-3s his mother
“As for Juan, his mother has killed him”

This section has presented an argument that focus features (present in focus in situ) and wh-features are not in complementary distribution, even if focus movement does not appear overtly in the presence of wh-movement. The co-occurrence of answer features and wh-features in the same sentence will be discussed in chapter 4.

3.2 Conclusions

This section has shown evidence that wh- and focus features can co-occur in an interrogative sentence and therefore are not in complementary distribution. Therefore, an analysis of the left periphery will have to take into account the existence of two different types of features, even if the movement processes associated to them show similarities.

³⁶ However, for a movement analysis of CLLD, see Cinque (1977, 1990).

4 The semantics of questions and focus

This section presents a few basic concepts of the literature on the semantics of questions and the semantics of focus. My intent here is not to formulate a coherent theory of the semantics of information structure, nor to present a summary of all, or even most, current approaches to the many related issues. My aim is rather to briefly introduce some of the background ideas in the field, only to the extent that they may be useful in the description and classification of certain interrogative structures. Such structures, involving additional informational elements such as focus, will be analyzed in chapter 4 of this dissertation.

4.1 Questions as answer sets

While a statement is defined by its truth conditions, a question is defined by its *answerhood* conditions, i.e. by our knowledge of what statements count as possible answers. The set of answers that defines a question has been portrayed in different ways in the literature. I present three of those views here, following the overview article by Hagstrom (2003).

Hamblin (1973) proposes that a question is defined by its possible answers. Thus, a question such as *Who left?* is defined by the answers where the wh-word is substituted by a person (which is what *who* stands for).

Karttunen (1977) reduces this set to the true answers to the question, in order to find a common semantics for direct and indirect questions (direct questions are those like *Is it raining? Which book did Mary read?*; indirect questions are *whether it is raining, which book Mary read*). Limiting the set of answers to the true answers would account for the intuition that *I know what Pat bought* implies only that I know the true answers to the question *What did Pat buy?* (Hagstrom, 2003, p. 189).

A third, different view of the nature of questions would come from the intuition that the answers to a question constitute an exhaustive set of mutually exclusive possibilities, a view first expressed by Hamblin (1958) and expanded by Groenendijk & Stokhof (1982). Hagstrom (2003, p. 189) characterizes this intuition in the following way: “if we know who left, we also know of people who didn’t leave that they didn’t leave (or, to put it another way, we know for each person whether that person left or not)”³⁷. This view provides a way to differentiate between complete and partial answers, as well as between partial answers and irrelevant statements.

³⁷ However, Hagstrom (p.c.) also points out that this intuition is only valid if the set of people who can leave the party is known, in which case the status of each particular person can be deduced.

4.2 Focus and scope

The problem of the semantic nature of focus is compounded by the diversity of phenomena to which the label ‘focus’ has been applied. Rooth (1996) avoids the diversity problem by limiting his analysis to intonational focus in English. The choice is of course not casual, since intonational focus in English seems to trigger all the relevant semantic phenomena. At the same time, it leaves out the contrasts in syntactic relevance and sentence interpretation that would later prompt analyses that separate those phenomena as different types of focus (É. Kiss, 1998). Rooth points out five semantic phenomena related to focus. All examples in this section are from Rooth (Rooth, 1996).

The first one is question-answer congruence, which I already discussed in section 2 of this chapter, and I see as limited to information focus. In (67a,b) the answers cannot be exchanged or they would be infelicitous.

- (67) a. Does Ede want tea or coffee? Ede wants [coffee_F].
b. Who wants coffee? [Ede_F] wants coffee.

The other four phenomena are, in one way or another, related to scope. Focus has a truth conditional effect in the context of *only*, as seen in (68). If John introduced Fred to Sue (68a) is false but (68b) does not have to be.

- (68) a. John only introduced [Bill_F] to Sue.
b. John only introduced Bill to [Sue_F].

Sentences with modals are affected as well: if a clerk escorts a ballerina that violates the rule presented in (69a) but not (69b). Sentences with adverbs or quantification are affected in the same way, as shown by (69c,d).

- (69) a. [Officers_F] must escort ballerinas.
b. Officers must escort [ballerinas_F].
c. In Saint Petersburg, [officers_F] always escorted ballerinas.
d. In Saint Petersburg, officers always escorted [ballerinas_F].

Constructions involving counterfactual reasoning (discussed by Dretske, 1972) are exemplified in (70). In a scenario where Clyde needs to get married to qualify for an inheritance and he marries Bertha, who spends most of the year abroad and had a relationship with Clyde already, (70a,b) would be true, but (70c,d) would be false.

- (70) a. The reason Clyde [married_F] Bertha was to qualify for the inheritance.
b. The reason Clyde married [Bertha_F] was to avoid making too much of a commitment.
c. The reason Clyde married [Bertha_F] was to qualify for the inheritance.

- d. The reason Clyde [married_F] Bertha was to avoid making too much of a commitment.

And lastly, focus creates conversational implicatures. The sentence in (71a), given as an answer to a question about a test I took together with other people, suggests that I did no better than passing, while (71b) suggests that the others did not pass.

- (71) a. Well, I [passed_F].
b. Well, [I_F] passed.

Rooth (1996) considers and compares two approaches to introduce such differences in the semantics of a sentence. One is the structured meaning semantics, based on the following idea:

focus has the effect of structuring the propositions denoted by sentences: the focus-influenced semantic value of a clause with a single focus is a pair consisting of i) a property obtained by abstracting the focused position and ii) the semantics of the focused phrase. The semantic values of (a) and (b) are (c) and (d) respectively:

- a. John introduced Bill_F to Sue
b. John introduced Bill to Sue_F
c. $\langle \lambda x [\text{introduce } (j, x, s), b] \rangle$
d. $\langle \lambda y [\text{introduce } (j, b, y), s] \rangle$

(Rooth, 1996)

He rejects this approach because it is not restrictive enough. As it is, it allows for the existence of focus sensitive lexical items like this one he makes up:

- (72) a. I tolfed that [he_F] resembles her \equiv I told him that he resembles her.
b. I tolfed that he resembles [her_F] \equiv I told her that he resembles her.
c. I tolfed that [he_F] resembles [he_F] \equiv I told him and her that he resembles her.

In Rooth's words, "*tolfed* φ amounts to *told the focus (or foci) of φ that φ* ". In a way, Rooth's criticism can be summarized by saying that adopting the structured meaning semantics implies equating focus with a θ -role.

The other approach he considers, based on Rooth (1985), is called 'alternative semantics' and is based on the idea that focus evokes alternatives to the sentence, which are obtained by making substitutions in the position occupied by the focus. This makes focus remarkably similar to questions, since questions determine sets of possible answers. Rooth (1992) follows this train of thought to simplify the architecture of alternative semantics to a single operator, which introduces the set of alternatives.

An advantage of structured meaning semantics is that it provides a tool to make focus available at the required syntactic level, which is not something that can be directly derived from alternative semantics. In the line of structured meaning semantics,

Chomsky analyzes focus assigning scope to focused phrases, as if they were quantifiers (Chomsky, 1976). This has also been done with *wh*-movement. Bach (1977, p. 137) states that “[t]he idea that question words like *who* and *what* are logically exactly the same as quantifiers is relatively old”. He mentions Jespersen (1924, pp. 302-305), Carnap (1937, p. 296), Reichenbach (1947, p. 340), Baker (1970), and Karttunen (1977) as authors in whose work we can find it, besides, of course, the article on *wh*-movement Bach is commenting (Chomsky, 1977).

Rooth (1996) discusses the advantages of both theories and finds that neither one can completely account for the island insensitivity of focus. On the other hand, island sensitivity is not a property of all scope-bearing operators, like, for instance, in situ *wh*- (Huang, 1982; Lasnik & Saito, 1992), so this is not a limitation particular to focus theories.

Rooth (1996) also discusses the fact that focus does not in general introduce an existential presupposition, but mentions that focus in Hungarian seems to introduce it (Szabolcsi, 1981), and concludes that maybe focus is too broad a notion to be discussed as such, and it should be replaced by separate discussions of different features. This idea is reflected in much of the literature mentioned in this dissertation and has been discussed in previous chapters.

4.3 Conclusions

Section 4 of this chapter has introduced the concepts of alternative sets as semantic analysis for questions and focus. This kind of analysis does not take into account the particular characteristics of identificational focus, which has been described as introducing a contrast or exhaustive identification (see Brunetti, 2003 for a different view and discussion), but it provides a basic framework to understand many pragmatic phenomena associated to information focus. In chapter 4 I will use these concepts as a background to develop an analysis of Spanish constructions where focus appears to be present in a question independently of the wh-element.

5 Chapter summary and conclusions

Questions require their answers to be structures containing an information focus, which is associated to the relevant part of the answer. This, however, might be just a consequence of how we choose to define information focus. That is to say, if we define information focus as information that is presented as new to the hearer, it makes sense that answering a question provides that kind of information. On the other hand, it is necessary to keep in mind that, besides being determined by contextual information, information focus is not marked syntactically or phonologically, and therefore it is fair to say that it is not in the same realm of grammar as questions, which are marked both syntactically and phonologically.

As for identificational focus, data in section 2.1 shows that it does not provide a felicitous answer to a plain question. In Hungarian, it introduces the notion of exhaustivity (É. Kiss, 1998); in Spanish, it introduces the sense of contrast or correction. In Italian it is not even acceptable as an answer (Brunetti, 2003). Since *wh*-movement does not introduce a meaning of exhaustiveness or contrast (beyond what can be expected from Grice's (1975) maxims), it can be concluded that there is no reason to identify focus movement and *wh*-movement in terms of their semantic effect. Section 2.2.2 was dedicated to show that there is no need to postulate a correlation between answers and left periphery focus in the case of fragment answers either.

The data in section 3.1 shows that, in spite of the absence of examples where both movements co-occur, it is possible to find examples of contrastive focus in situ in questions involving wh-movement. If we assume that focus in situ involves covert movement and is equivalent to focus in the left periphery, then wh- features and focus features can coexist in an interrogative sentence.

Thus, it can be stated

- 1) that focus movement introduces meanings in a sentence (contrast and/or exhaustive identification) which wh-movement does not.
- 2) that the features associated to both types of movement can coexist in a sentence, as shown by the presence of contrastive focus in situ in wh-questions.

In summary, it has been shown that arguments exist to support the idea that identificational focus and wh-movement are separated and independent from each other, both in semantic and syntactic terms. A proposal of the structure of the left periphery that takes into account these issues is developed in the analysis of split questions in the next chapter.

CHAPTER 4 – SPLIT INTERROGATIVES

1 Introduction

The previous three chapters of this dissertation discussed the relationship between syntax and information structure, with particular attention to focus. The first one showed the necessity of a relationship between information structure and syntax by reviewing the relevant literature and describing the limitations of pragmatic and phonological approaches. The second chapter reviewed syntactic approaches to information structure in order to establish a framework within which the left periphery of interrogatives could be discussed. I adopted the view (developed by Rizzi, 1997 and others) that part of the phenomena related to information structure is encoded in syntax in the form of functional projections in the left periphery. With such framework in place, chapter 3 addressed the relationship between focus and interrogative sentences, concluding that it is not possible to assume identity between identificational focus and *wh*- in either an interrogative sentence or a sentence that is presented as an answer to an interrogative. Additionally, it was observed that answers have a special status, different from sentences involving focus, and also different from non-answer declarative sentences. On the basis of these observations, I proposed that the focus area of the left periphery of the sentence (the “Focus field” (Benincà & Poletto, 2004)) consists of at least three separate functional nodes, (identificational) Focus, Answer, and *Wh*.

The goal of this last chapter is to complete this description of the left periphery of the sentence and show its usefulness in providing an explanation for constructions whose analysis requires a detailed mapping of the functional projections involved. These constructions are a type of interrogative sentences that consists of two parts separated by intonation. The first part resembles a wh-question (although it has particular constraints, described in section 2 of this chapter), and the second part (which I will call *tag*, adopting the term from Uriagereka (1988)) presents a proposed answer to the question posed by the first part. I call such constructions *split questions*¹. Such questions appear at least in English, Asturian, Catalan, some Italian dialects, and a number of Spanish varieties. Their characteristics vary from a dialect to another, but they are always used as a request for confirmation of the appropriateness of their tag as an answer to the question posed in their first part. The expected reply to a split question is a ‘yes’ or a ‘no’, and at least in that way they are closer to polar (yes/no) questions than they are to wh-questions. In the case of negative responses, however, split questions differ from polar questions in that a felicitous reply would have to include the

¹ The type of interrogative construction described here has been referred to as *interrogaciones compuestas* ‘compound interrogatives’ (Py, 1971), adjunct tags (Uriagereka, 1988), *anticipación expresiva mediante inserción de un pronombre interrogativo proléptico* ‘expressive anticipation by insertion of a proleptic interrogative pronoun’ (Vigara Tauste, 1992), *qué expletivo en preguntas dislocadas* ‘expletive *qué* in dislocated questions’ (Lorenzo González, 1994), Wh-doubling (Camacho, 2002), split interrogatives (López-Cortina, 2003) and *interrogatives escindidas* ‘split interrogatives’ (Contreras & Roca, in press). Arregi uses ‘split questions’ to designate a larger range of constructions (Arregi, 2006). The term *split question* has sometimes been used to refer to wh-subextraction, i.e. constructions where a wh-determiner moves without pied-piping its noun phrase. This construction is attested in Slavic languages and Latin, and under certain conditions in Romance (Solà & Gavarró, 2006). I adopt the term *split questions* here anyway in the belief that it is perhaps more appropriate to describe constructions where the whole sentence, rather than just the wh-phrase, is split in two parts.

actual answer to its first part. Thus, split questions have elements of a request for confirmation, but they keep the main characteristic of a wh-question, the opening of a set of alternatives. The introduction of a set of alternatives linked to a particular constituent is associated to the semantics of focus, which prompted previous analyses (López-Cortina, 2003) to interpret split questions as a result of the combination of a focus construction with a wh-question. In this chapter, a revised analysis is provided, which builds on the findings about Spanish answers previously presented in chapter 3. The combination of the structure of answers and that of questions in these constructions is the basis for an even more fine-grained analysis of the left periphery.

Section 2 of this chapter provides a description of the construction. Section 3 presents previous analyses. Section 4 takes on the analysis of split questions and its variations in different dialects. Section 5 closes the chapter with a discussion of the meaning of split questions and a sketch of their semantics.

2 Description of the construction

This section provides a description of split questions, beginning with the accounts and samples of the descriptive works by Py (1971) and Vígara Tauste (1992) in sections 2.1 and 2.2, respectively. Once the basic traits of the construction have been presented, section 2.3 discusses in more detail the characteristics and limitations of the construction, which will constitute the basis for the analysis presented in section 4. Section 2.4 discusses constructions that present certain similarities with split questions, and section 2.5 presents a brief summary of the data.

2.1 Py (1971)

The earliest reference to this type of construction that I am aware of is made by Bernard Py in his dissertation about interrogatives in spoken Madrid Spanish (Py, 1971), who describes several types of *interrogaciones compuestas* ‘compound interrogatives’, which he defines as questions performed through two or more sentences. The goal of Py’s work is to give a classification of interrogative expressions which links the possibilities derived from a conceptual model to their respective spoken forms. Thus, the concept of syntactic structure we are familiar with in generative grammar is

irrelevant for his goals, and his classification method allows grouping as a single expression two sentences that are not linked through syntax, and are sometimes separated by additional material in the discourse. I include the whole of Py's data for the purposes of comparison, but I will not discuss his conceptual model. I will instead give a provisional syntax-based classification of the sentences he collected. In (1) I transcribe the examples where the link between the two parts of the interrogation is purely conceptual, and no syntactic link is involved. The only case where there seems to be syntactic compatibility between the two segments is (1d), which I would interpret as an afterthought, as defined by Villalba (2000). In the examples in (2), (3), (4) and (5) the link between the two segments is a *wh*-word, that can come alone as a general question preceding a more specific one, as in (2), or can be a part of a longer question, of which the second segment would be a possible answer offered by the speaker, as in (3), (4) and (5). This *wh*-word can be compatible with the category and meaning of the second segment, as in (4) and (5), or be a generic *wh*-word, *qué*, as in (3).

- (1) a. ¿En Almagro? ¿Qué número?
 in Almagro what number
 ‘In Almagro Street? What number?’
- b. ¿Y esto? ¿Es sin revelar?
 and this is without develop
 ‘And this? Is it not developed?’
- c. ¿Y de aquella cuestión de tu casa? (...) ¿Cómo lo solucionaste?
 and of that issue of your house how CL3sg solved2sg
 ‘And that issue with your house? How did you solve it?’

- d. ¿Y esto es una almena? ¿Del castillo?
and this is a merlon of-the castle
'And this is a merlon? Of the castle?'
- (2) a. Bueno, ¿y qué? ¿A qué hora llegasteis allí?
well and what at what time arrived2sg there
'Well, and then? What time did you arrive there?'
- b. ¿Y qué? ¿Por qué no lo compraron?
and what for what not CL3sg bought3pl
'What then? Why didn't they buy it?'
- c. ¿Qué? ¿Te gusta España?
what CL2sg pleases3sg Spain
'Well? Do you like Spain?'
- (3) a. ¿Qué es? ¿A casete?
what is to cassette
'How does it work? With cassettes?'
- b. ¿Qué tienen? ¿Frío?
what have3pl cold
'What are they, cold?'
- c. ¿Qué lleva? ¿Mil quinientas?
what takes3sg one-thousand five-hundred
'What does it take? Fifteen hundred?'
- (4) a. ¿Qué queréis? ¿Un demonio?
what want2pl a demon
'What do you want? A demon?'
- b. Ahora, ¿qué quiere usted? ¿Cambiarlo por otro?
now what want2sg you change-CL3sg for another
'Now, what do you want? To change it for another one?'

- (5) a. Esta caja, ¿de qué es? ¿La grabadora?
 this box of what is the tape-recorder
 ‘This box, what is it for? The tape recorder?’
- b. ¿Para qué? ¿Para comer o para zumo?
 For what for eat or for juice
 ‘What for? For eating or for juice?’
- c. ¿Qué pasa? ¿Que estuvo de campo?
 what happens3sg that was3sg of field
 ‘What’s the matter? Was he in a picnic?’
- d. ¿Qué es lo que te gusta más? ¿Madrid o Badajoz?
 what is the that CL2sg likes3sg more Madrid or Badajoz
 ‘What do you like better? Madrid or Badajoz?’
- e. ¿Dónde quieres ir? ¿Al teatro?
 where want2sg go to-the theater
 ‘Where do you want to go? To the theater?’
- f. ¿Cuál quería usted? ¿Para una persona o para dos?
 which wanted2sg you for one person or for two
 ‘Which one did you want? For one person or two?’

2.2 Vigara Tauste (1992)

Vigara Tauste (1992, p. 103) observed, among other “expressive” dislocated interrogative constructions, what she called *anticipación expresiva* “expressive anticipation” in interrogative constructions, expressed by the insertion of a “proleptic” interrogative pronoun. She notes that this pronoun both anticipates the interrogative character of the sentence and explicitly links both of its parts. Vigara Tauste provides the examples in (6) and says that the construction seems to express irony, surprise, or great interest in the answer on the part of the speaker.

- (6) a. Qué estás ¿todavía desayunando?
What are2sg still having-breakfast
‘What are you doing? Are you still having breakfast?’ (surprise)
- b. Pero la postal qué es ¿para el colegio o para los abuelos de verdad?
But the postcard what is for the school or for the grandparents of truth
‘Is that postcard homework for school or is it truly for your grandparents?’
- c. La tesis cómo va, ¿bien?
The dissertation how goes good
‘How is your dissertation coming out? Good?’
- d. La música a toa leche qué, ¿para relajarte?
The music at all milk what for to-relax-CL2sg
‘What is that really loud music for? Do you find it relaxing?’ (sarcasm)

- e. Y la moral qué ¿que le den por culo?
 And the morality what, that CL3sg give3sgSUBJ through arse
 ‘And what about morality? Just screw it?’ (indignation)

I transcribe the examples with the unusual orthography Vigarà Tauste used for them, in order to point out how she noted the equally unusual intonation pattern. Spanish orthography requires an opening question mark ¿ at the beginning of a question (at least since the 1754 *Ortografía* of the Spanish Academy (Millán, 2005)). This opening question mark can sometimes appear mid-sentence if the material to its left is perceived (and pronounced) as an *as for* phrase or some other type of topic-like element. The wh-word, however, is never left outside of the question marks. The reason Vigarà Tauste leaves it out in these examples is that she perceives some kind of intonational segmentation which divides the sentence in two parts. She identifies the interrogative intonation of the second part of the sentence and places the opening question mark immediately after the end of the previous intonational event. She labels the intonation of the first part of the sentence as enunciative, somehow by default, since it is a different intonation from the interrogative one described for the second part of the sentence. This is in contradiction with descriptions of this first segment’s intonation as interrogative (Camacho, 2002) and as comma intonation (López-Cortina, 2003).

Beyond the common presence of a proleptic interrogative, there are obvious differences between the examples presented by Vigara Tauste. In three of them (6a,b,c) there is an interrogative sentence, including a verb, followed by a proposed answer, which is associated to interrogative intonation. In one of these, (6c), the proleptic wh-word is the one we would expect to find in the question if the second part (the proposed) of the sentence were not present. Comparable examples by Py were presented above in (4) and (5). In (6b) we find *qué* “what”, instead of the expected *para qué*, “what for” (compare to Py’s examples in (3)). Interestingly, there is no appropriate wh-word for the second part of (6a). I will discuss these differences, together with additional data, in sections 2.3 and 2.4.

The other two examples (6d,e) present some kind of background information followed by the proleptic interrogative and then by the proposed answer, but no verb is present in the construction. These are parallel to the construction described by Py (1971) and mentioned above in (2). A similar construction has been noticed by Lope Blanch (1972, p. 14) (quoted in Cotton & Sharp, 1988, p. 156). Lope Blanch is of the opinion that *qué* has the role of question marker, the way *est-ce que* functions in French or *do* in English. The example given by Lope Blanch is this:

- (7) Qué, no te lo dio a ti?
 what not CL2sg CL3sg gave3sg to you
 ‘Didn’t he give it to you?’

According to Lope Blanch, “such expressions, with a rising intonation on *qué*, convey surprise, disappointment, or shock” (Cotton & Sharp, 1988, p. 156)². I will get back to these constructions in section 4.6. For now, I will just note that the meaning added to the sentence by the presence of *qué* seems to be in contradiction with the role of simple question marker. Besides, Py (1971, p. 185) notes that this *qué* is never the only interrogative mark in the sentence; there is always another mark, like subject inversion or interrogative intonation.

2.3 The syntax of split questions

One of the most remarkable facts in the data described by Py and Vigara Tauste is the possibility of having the wh-word *qué* can assume the value of any wh-word (*who*, *when*, *where*, *why*, *how*, etc.) in a wh-question, as long as an appropriate answer follows the question, separated from it by an intonational event that splits the sentence in two parts. I will refer to sentences that present this structure as *split questions*. Perhaps the best way to describe what makes these constructions different from other questions, beyond the affective contexts reported by Vigara Tauste and Lope Blanch, is to look at the answers they elicit: the request for confirmation of a proposed answer makes split

² Cotton and Sharp also mention that the construction is unheard of in Spain, which is not the case, as proven by the data offered by Py (1971) and Vigara Tauste (1992).

questions yes/no questions, with the difference that a negative answer requires additional information, in the form of an actual response to the wh-question posed³.

Note that Catalan, one of the languages where split questions are found, also presents a different construction with a *que* that could be an overt complementizer or a question marker similar to Japanese *ka* (8a). This is particularly interesting because it shows very clearly the difference in intonation and meaning between split questions and other yes/no questions. While (8a) is a yes/no question for which either answer would be acceptable, a negative answer to (8b) would require further clarification as for the means of travel. Negating the travel itself would contradict a presupposition in (8b), but not in (8a), which is ambiguous in that respect.

- (8) a. Que vas en tren? (Catalan) (López-Cortina, 2003)
that go2sg in train
'Are you going by train?'
- b. Què vas , en tren?
what go-2sg, in train
'How are you going? by train?'

The main difference between split questions and similar structures where the wh-word is not *qué* (or, if it is *qué*, it is compatible with the answer provided, as in (4), (5), and (6c)) is that the wh-question of split questions can not stand alone. This is seen

³ This characterization appears for the first time in Uriagereka (1988).

in the contrast between the split questions in (9) and the ungrammatical wh-questions in (10)⁴.

- (9) a. Qué llegó , Juan?
what arrived3sg, Juan
‘Who arrived? was it Juan?’
- b. Qué vienes , el sábado?
what come2sg, the Saturday
‘When are you coming? on Saturday?’
- c. Qué vas , a Oviedo?
what go2sg, to Oviedo
‘Where are you going? to Oviedo?’
- (10) a. * Qué llegó?
what arrived3sg
- b. * Qué vienes?
what come2sg
- c. * Qué vas?
what go2sg

I will therefore draw an initial distinction between split questions, where the wh-question can not stand alone, and other structures, which I will call *follow-up questions*⁵,

⁴ I take the variety of Spanish spoken in Asturias as the canon for the examples that follow, even though different varieties present similar constructions, as discussed below. Approximate translations are given under the glosses.

⁵ Note that Arregi (2006) extends the use of the term ‘split questions’ to structures where the wh-word is other than *qué*. I have preferred to reserve the term ‘split’ to cases where there is interdependency between the two parts of the questions, as used in López-Cortina’s (2003) ‘split interrogatives’ and followed by Contreras & Roca’s *interrogatives escindidas* (Contreras & Roca, in press).

where the *wh*-question can stand alone and includes a *wh*-word that would be appropriate for a question answered with the tag. I will refer to those as *full wh-words*, as opposed to the usage of an underspecified *wh*-word in split questions.

The closeness of the relationship between the question and the tag of split questions is further illustrated by the agreement between the tag and the verb, visible when the tag is the subject. This agreement does not appear in follow-up questions.

- (11) a. Qué vienes, tú?
what come2sg, you
'Who is coming? You?'
- b. * Qué viene, tú?
what come3sg, you
- c. Quién viene? Tú?
who come3sg, you
'Who is coming? You?'

The interdependence between the question and the tag in split questions is a reason to believe that this structure requires a single-clause analysis, while the independence between question and tag in follow-up questions suggests a two-clause structure. Differences between split questions and follow-up questions are discussed in section 2.4 and analyses are provided in section 4.9, but for now I will advance that

some dialects present both single- and two-clause structures with full wh-words. In such dialects, intonation determines whether the structure is a split question or a follow-up question. The rest of section 2.3 deals with the description of split questions, and 2.4 addresses follow-up questions.

Categories and functions

The tag of a split question is not restricted to any grammatical category (12) or function (13) susceptible of answering a regular wh-question.

- (12) a. Qué lo compró, Juan / él / el viejo / un viejo?
 what it bought3sg, Juan he the old an old
 ‘Who bought it? Juan / He / The old man / An old man?’
- b. Qué lo compraron, seis?
 what it bought3pl, six
 ‘How many people bought it? Six?’
- (13) a. Qué lo compró, Juan? (subject, transitive verb)
 what it bought3sg, Juan
 ‘Who bought it? Juan?’
- b. Qué vino, Juan? (subject, unaccusative verb)
 what came3sg, Juan
 ‘Who came? Juan?’
- c. Qué compró, un libro? (direct object)
 what bought3sg, a book
 ‘What did he buy? A book?’

- d. Qué se lo compró, a María? (indirect object)
 what IOcl DOcl bought3sg, to María
 ‘Who did he buy it for? María?’
- e. Qué están hablando, de fútbol? (prepositional complement)
 what are3sg talking of soccer
 ‘What are they talking about? Soccer?’
- f. Qué lo hizo, ayer / esta tarde? (adjunct, time)
 what it did3sg yesterday this afternoon
 ‘When did he do it? Yesterday / This afternoon?’
- g. Qué lo hizo, allí / en su casa? (adjunct, place)
 what it did3sg there in his house
 ‘Where did he do it? There / In his house?’
- h. Qué lo hizo, así / muy despacio? (adjunct, manner)
 what it did3sg thus very slowly
 ‘How did he do it? Like that / Very slowly?’
- i. Qué lo hizo, con la mano? (adjunct, instrument)
 what it did3sg with the hand
 ‘How did he do it? With his hand?’
- j. Qué lo hizo, porque quiso? (adjunct, cause)
 what it did3sg because wanted3sg
 ‘Why did he do it? Because he wanted to?’

Additionally, some phrases that cannot be substituted by a wh-word in a regular wh-question can also appear as tags. Gerunds can be extracted from periphrases (14a), and tags can also include multiple constituents (14b,c,d). In such cases the interrogative word refers to everything after the comma, and a ‘yes’ answer confirms every element included in the tag.

- (14) a. Qué está, lloviendo?
 what is, raining
 ‘What is going on? Is it raining?’
- b. Qué vino, Juan a verte?
 what came3sg, Juan to see-you
 ‘Who came? Did Juan come to see you?’
- c. Qué vino, Juan ayer?
 what came3sg, Juan yesterday
 ‘Who came? Did Juan come yesterday?’
- d. Qué comió, pollo en casa?
 what ate3sg, chicken at home?
 ‘What did he eat? Chicken at home?’

Position of the subject

Subjects can not appear between the wh-word and the verb in split questions.

This is a limitation common to all wh-questions in most dialects of Spanish⁶.

- (15) a. *Qué Juan ha ido, al cine? (Uriagereka 1988, p.142)
 what Juan has3sg gone, to-the movies?
- b. *Qué Juan ha llegado, esta mañana?
 what Juan has3sg arrived, this morning
- c. *Qué Juan lo ha hecho, con las manos?
 what Juan it has3sg done, with the hands

⁶ There are Caribbean Spanish dialects that allow preverbal subjects in wh-question, particularly if the subject is a pronoun. The speakers of these dialects I have consulted do not use split questions, so it is not possible to determine whether there is any special link between split questions and subject order.

- d. *Qué Juan lo ha hecho, porque sí?
 what Juan it has3sg done, because yes

Elements between the verb and the intonation break marked by the comma are not accepted in all dialects, but even where they are not accepted there is a certain contrast between the inclusion of a subject (16b) and a complement (16a). The contrast between dialects on this point is discussed in section 4.5.

- (16) a. *Qué llegó a Oviedo, Juan? (some Asturian varieties)
 what arrived-3sg to Oviedo, Juan?
- b. ?Qué llegó Juan, a Oviedo?
 what arrived-3sg Juan, to Oviedo?
 ‘Where did Juan arrive? Oviedo?’

In situ interrogative word

Chilean and Ecuadorian Spanish, as well as some varieties of English, present a type of split question where the wh-word may remain *in situ* (17b,c). Chilean Spanish does not normally allow subjects to the right of the comma (17d), but they appear to be more acceptable if the wh-word remains *in situ* (17e), (Campos, p.c.).

- (17) a. Qué llegó Juan, a Oviedo? (some Peninsular, Chilean)
 what arrived3sg Juan, to Oviedo?

- b. Vas qué, en tren? (Ecuadorian, Chilean)
 go2sg what, in train?
 ‘You go what, by train?’
- c. You are going what, by train? (English)
- d. * Qué llegó , Juan? (Chilean)
 what arrived3sg Juan?
- e. ?? Llegó qué , Juan? (Chilean)
 arrived3sg what , Juan?
 ‘Who arrived? Juan?’

Long extraction

Speaker judgments on the grammaticality of long extraction in these constructions vary from acceptance to flat rejection (18a), while they are clearly impossible in embedded clauses, as shown in (18b).

- (18) a. % Qué dijiste que venía , Juan?
 what said2sg that came-3sg, Juan?
 ‘Who did you say that was coming? was it John?’
- b. * Dijiste que qué venía , Juan?
 said-2sg that what came3sg , Juan?

Most, but not all, of the speakers interviewed agree on the pattern presented in (19) to (21) below, where long distance movement is not allowed in split questions.

Similar constructions with pairs of juxtaposed questions⁷ (19b) (20b) (21b) do not present a problem for any speaker, unless of course one of the juxtaposed questions is ungrammatical by its own merit (22b). Structures with the same word order, but where the intonational split indicates that the answer includes the embedded IP (20c) (and therefore there is no long distance movement), are also acceptable.

- (19) a. *Qué crees que compró, un libro?
 what believe2sg that bought3sg a book
 ‘Do you think it is a book, what he bought?’
- b. Qué crees que compró? Un libro?
 what believe2sg that bought3sg a book
 ‘What do you think he bought? A book?’
- (20) a. *Qué dices que llega, mañana?
 what say2sg that arrive3sg tomorrow
 ‘Are you saying it is tomorrow when he arrives?’
- b. Cuándo dices que llega? Mañana?
 what say2sg that arrive3sg tomorrow
 ‘When are you saying he is arriving? Tomorrow?’
- c. Qué dices, que llega mañana?
 what say2sg that arrive3sg tomorrow
 ‘Are you saying he arrives tomorrow?’
- (21) a. *Qué esperas que llegue, a las cuatro?
 what expect2sg that arrive3sg at the four
 ‘Are you expecting him to arrive at four?’

⁷ The issue of juxtaposed questions, or follow-up questions, is addressed in sections 2.4 and 4.9 below.

- b. Cuándo esperas que llegue? A las cuatro?
 when expect2sg that arrive3sg at the four
 ‘When are you expecting him to arrive? At four?’
- (22) a. *Qué lamentas que llegue, a las cuatro?
 what regret2sg that arrive3sg at the four
 ‘Do you regret that it is at four when he arrives?’
- b. *Cuándo lamentas que llegue? A las cuatro?
 when regret2sg that arrive3sg at the four
 ‘When do you regret that he arrives? At four?’

The pattern of the speakers that accept long extraction is more complex⁸. For them, (19a) and (20a) are grammatical, but (21a) and (22a) are not.

Catalan data (Contreras & Roca, in press) mirrors the grammaticality judgments of the Asturian Spanish speakers that accept long extraction. Such sentences are ungrammatical (23) except with a few verbs of communication or opinion which also happen to be phonetically short (24).

- (23) a. *Què et molesta el fet que compri, molts llibres?
 what CL2sg bothers the fact that bought1sg many books
- b. *Què preguntaves qui vindria, en tren?
 what asked2sg that would-come, in train

⁸ I have found only a few speakers that use these constructions (which does not say much about their actual number, but might indicate that the phenomenon is not geographically widespread). They are all bilingual speakers of Asturian and Spanish, although not all Asturian speakers I interviewed share their judgments. An analysis of these data is presented in section 4.3.4.

- c. */?Què et preocupa que arribem, demà?
 what CL2sg worries that arrive1pl tomorrow
- (24) a. ?Què t’han dit que compraria, llibres?
 what CL2sg-have said that would-buy3sg books
 ‘What have they told you he’d buy? Books?’
- b. ?Què creus que vindrà, en tren?
 what think2sg that will-come in train
 ‘How do you think he’ll come? By train?’
- c. ?Què t’han dit que arribaríem, demà?
 what CL2g-have said that arrive1pl tomorrow
 ‘When did they tell you we’d arrive? That we’d arrive tomorrow?’

The parallelism of long extraction facts for some Asturian speakers and Catalan speakers suggests that there might be a common explanation for both sets of data. Section 4.3.4 is devoted to the characterization of the parameters that define long extraction dialects.

Negation

Split questions can not be negative, at least in the dialects of my informants in Chile and in Asturias, Spain (López-Cortina, 2007), as well as in Catalan (Contreras & Roca, in press). The following examples are all ungrammatical:

- (25) a. *Qué no quieres, el verde?
 what not want2sg, the green
 ‘What don’t you want, the green one?’

- b. *Qué no comes, cereales?
 what not eat2sg, cereals
 ‘What don’t you eat, the cereal?’
- c. *Qué no vienes, el sábado?
 what not come2sg, the Saturday
 ‘When aren’t you coming, on Saturday?’
- d. *Qué no vienes, porque te lo pidió ella?
 what not come2sg, because to-you it asked she
 ‘Why aren’t you coming, because she asked you not to?’

Note that certain cases with the same word order, but consisting of two separate questions, are grammatical (26). In the cases where the wh-question can appear by itself, the difference in intonation is key to determine whether the utterance is a split question or a juxtaposition of two questions.

- (26) Qué no comes? Cereales?
 what not you-eat cereals
 ‘What don’t you eat? The cereal?’

The left periphery of split questions

Several Spanish dialects, among them Asturian Spanish and Chilean Spanish, allow for one or more elements on the left of the interrogative in a split question, as

shown in (27a,b). Such left periphery elements have the role of establishing the sentence topic. Their exact status and position in the structure is discussed in section 4.7.

- (27) a. Pedro qué vino, en avión?
Pedro what came3sg in airplane
'How did Pedro come? by plane?'
- b. Pedro ayer qué vino, en avión?
Pedro yesterday what came3sg in airplane
'How did Pedro come yesterday? by plane?'

Distribution

It must be noted that, although they are occasionally mentioned in the literature (Py, 1971; Vígara Tauste, 1992), these constructions are not considered standard Spanish, and are rare among Castilian speakers. Due to the lack of an accepted norm, the natural variability has been preserved, and different speakers have different acceptability judgments about cases other than the canonical ones in (9). This is true even in areas where wh split interrogatives are the standard way to ask almost any question, like Asturias. The presence of split questions in the Asturian language (one of the Leonese varieties, sometimes referred to as *bable* in the literature) permeates the variety of Spanish spoken in the area, but most speakers would never use a split question in writing. Split questions are also standard in spoken Catalan, and they are

present in the Spanish spoken in Catalonia (Contreras & Roca, in press; López-Cortina, 2003).

In standard American English, split questions seem to only be possible as rhetorical questions where the tag is presented as either an obvious counterfactual or a very surprising possibility. Nouns, adjectives, gerunds, and prepositional phrases can be tags combined with the verb *be* in the *wh*-question (28). Constructions with other verbs are frequent in New York and New Jersey, but not as common (29a), and have occasionally been documented in literary texts reflecting colloquial frontier speech (29b). Some speakers from other areas where split questions are ungrammatical or dubious can ask *in situ* split questions (30).

- (28) a. What are you, Batman?
b. What are you, crazy?
c. What are you, kidding?
d. What are you, trying to get a second job here?
(Kodjo Namdi on NPR, January 9, 2004)
e. What are you, on a break?
(*Coffee and cigarettes (Twins)*, 2003, Jim Jarmusch)
f. What am I, in the morgue?
(*The bucket list*, 2007, Rob Reiner)

- (29) a. What do you got me, working?
(said by a speaker from Hudson County, New Jersey)
- b. What have you got, a cold?
(*Blood Meridian*, 1985, Cormac McCarthy)
- (30) a. When are you coming? On Saturday?
- b. ?? What are you coming, on Saturday?
- c. % You are coming what, on Saturday?

2.4 Follow-up questions

This section discusses the cases like those in (31), where the wh-word of an apparent split question is not *qué* but the one matching the possible answer.

- (31) a. Quién llegó a Oviedo , Juan?
who arrived3sg to Oviedo , Juan?
'Who arrived to Oviedo? Was it Juan?'
- b. Dónde llegó Juan , a Oviedo ?
where arrived3sg Juan , to Oviedo?
'Where did Juan arrive? Was it Oviedo?'
- c. La tesis cómo va, ¿bien? (Vigara Tauste, 1992, p. 103)
The dissertation how goes good
'How is your dissertation coming out? Good?'

The literature addressing these structures (Arregi, 2006; Camacho, 2002; López-Cortina, 2003) distinguishes two possibilities. The first one is that these questions have the same structure as split questions, the only difference being a greater flexibility in the choice of the interrogative word, which would pick up additional semantic traits from the tags. The second possibility is that these constructions are simply the juxtaposition of two questions, where the second one is what I term a *follow-up question* which has undergone ellipsis of elements previously mentioned in the discourse⁹. Dialectal variation should be expected, since the acceptability of *qué* split questions also changes drastically from one dialect to another.

I will discuss three areas where differences between follow-up questions and split questions, beyond the presence of the full wh-word, can be identified: meaning, intonation, and negation restrictions. The goal of this section is merely descriptive. A syntactic analysis is offered in section 4.3.

The difference in meaning between split questions and follow-up questions can be summarized by saying that a split question is about its second part, with the first part presenting a context assumed to be true. In contrast, a pair of juxtaposed questions poses a question and then narrows its range by adding a second, more specific one. This can be said to be the standard interpretation of the meaning of both constructions in

⁹ This is the analysis pursued by Arregi (2006), who discusses follow-up questions only.

Asturian Spanish, for example. Some speakers of other dialects, however, report being able to use both constructions interchangeably. Such speakers also report to have two possible different intonations for those questions, so it is likely that two different structures coexist in their dialect, one with split questions, and another one with two juxtaposed questions.

Regarding intonation, Arregi (2006) states that these cases have the intonation of two juxtaposed independent questions. This is the same observation Camacho (2002) makes for the cases he observes. As noted above, split questions are associated to an intonational event different from the intonation of two separate questions. In this respect, some speakers can distinguish two different intonations (Héctor Campos, p.c.). If the intonation is ascending, the construction is interpreted to consist of two juxtaposed questions. If descending, it is interpreted like a full *wh*-word split question. This second interpretation is not available in Asturian Spanish, where the split question intonation is only available for *qué* split questions¹⁰.

There are, therefore, at least four possible dialectal variations. One without split questions, a second one with *qué* split questions, another one with both *qué* and full *wh*-

¹⁰ Note that, differently from standard Spanish, Asturian Spanish question intonation is not ascending (Canellada, 1944). There is a clear difference between intonational contours for split questions and pairs of juxtaposed questions, but it can not be described as an ascending/descending contrast.

word split questions, and another one only with full wh-word split questions. I have found speakers of the first three of these possible varieties, but none of the last one.

(32)

		<i>qué</i> split question	
		yes	no
full wh-word split question	yes	Some varieties of Chilean Spanish	----
	no	Northwestern Peninsular Spanish, Asturian, Catalan	Most standard Castilian speakers

Restrictions like the impossibility of negative structures do not apply to full wh-word constructions, as shown by the contrast between the examples in (33) (repeated from (25)) and their full wh-word correlates in (34).

- (33) a. *Qué no quieres, el verde?
 what not want2sg, the green
 ‘What don’t you want, the green one?’
- b. *Qué no comes, cereales?
 what not eat2sg, cereals
 ‘What don’t you eat, the cereal?’
- c. *Qué no vienes, el sábado?
 what not come2sg, the Saturday
 ‘When aren’t you coming, on Saturday?’

- d. *Qué no vienes, porque te lo pidió ella?
 what not come2sg, because to-you it asked she
 ‘Why aren’t you coming, because she asked you not to?’
- (34) a.Cuál no quieres? el verde?
 which not want2sg, the green
 ‘Which one don’t you want, the green one?’
- b. Qué no comes? cereales?
 what not eat2sg, cereals
 ‘What don’t you eat? The cereal?’
- c. Cuándo no vienes? el sábado?
 when not come2sg, the Saturday
 ‘When aren’t you coming? on Saturday?’
- d. Por qué no vienes? Porque te lo pidió ella?
 for what not come2sg, because to-you it asked she
 ‘Why aren’t you coming? Because she asked you not to?’

Wherever intonational facts are not clear enough to determine whether a particular construction is a split question or a follow-up question, the restrictions on split questions should serve to make the distinction.

2.5 Summary

Split questions are two-part interrogative constructions, consisting of a wh-question subject to special restrictions and a tag, which corefers with the wh-word in the wh-question. In some dialects, the wh-word must be an underspecified *qué/what*, while other dialects allow for a full wh-word, matching the semantic features of the tag. Semantically, split questions are between yes/no questions and wh-questions, accepting an affirmative answer but not a negative one unless further information is supplied.

Split questions share with other wh-questions the impossibility of having a subject between the wh-word and the verb. Negative split questions are not available, nor are embedded split questions. Long extraction is generally not possible, with some exceptions, in some dialects, linked to particular lexical items. In situ questions are available in some dialects, but not all. There are also dialect-dependent restrictions on the type of additional complements available to the verb, besides the tag.

Split questions should not be confused with pairs of juxtaposed questions consisting of a wh-question and a follow-up fragment question coreferent with the first question's wh-word. Such pairs are different from split questions in intonation (they are separated by regular question intonation) and present none of the syntactic characteristics of split questions that make them different from regular wh-questions.

3 Previous analyses

Split questions are not included in traditional grammar books. In every language where they appear, they are considered non-standard, or mere oral varieties of proper interrogative constructions. The lack of an accepted standard has caused a lack of widely available descriptions, but it has also contributed to the preservation of different varieties of the construction. The combination of these factors necessarily results in analyses that start from incomplete, and often contradictory, descriptions.

In this section I review the analyses of the structure of split questions presented in the literature. Given the fact that each of these analyses starts from a partial description, strong discrepancies among them are to be expected. The main issues these analyses intend to clarify are the nature of the initial *qué*, whether its position results from movement, and what is its relationship with the tag. After López-Cortina (2003) other issues were taken into account, such as the intonational separation between the question and the tag and the variation among dialects.

3.1 Uriagereka (1988)

The first analysis of split questions within the framework of generative grammar was provided by Juan Uriagereka, on the basis of a number of examples involving adjuncts as proposed answers, which I copy in (35), with his glosses. He refers to these constructions as *adjunct tags* (Uriagereka, 1988, pp. 141-143, 202-207).

- (35) a. Qué ha ido, al cine?
what has3sg gone, to-the movies?
'What has he gone, to the movies?'
- b. Qué ha llegado, esta mañana?
what has2sg arrived, this morning
'What has he arrived, this morning?'
- c. Qué lo ha hecho, con las manos?
what it has3sg done, with the hands
'What has he done it, with his hands?'
- d. Qué lo ha hecho, porque sí?
what it has3sg done, because yes
'What has he done it, just because?'

Uriagereka notes that there is no lexical meaning associated to the initial *qué*, which takes its meaning from the tag. The tag is necessary for the question to be grammatical. He also notes that these questions can be answered in an affirmative way without adding further information, but a negative answer requires additional information for the answer to be felicitous. In his terms, *qué* is not only a scope marker, but also a

marker of “information question”. That is to say, the relationship between *qué* and the tag marks the scope of a yes/no question, but the structure retains the character of a wh-question.

Uriagereka takes adjunct tag constructions to be evidence of adjunct movement, and therefore of the presence of adjuncts in the structure prior to SS. He interprets the ungrammaticality of the examples in (36) to be a result of a subjacency violation: the initial *qué* can not cross a properly specified IP. The details of Uriagereka’s 1988 view of subjacency are not relevant here. The interesting idea is that the contrast between (35) and (36) suggests that the initial *qué* undergoes movement.

- (36) a. *Qué Juan ha ido, al cine?
what Juan has-3sg gone, to-the movies?
- b. *Qué Juan ha llegado, esta mañana?
what Juan has-2sg arrived, this morning
- c. *Qué Juan lo ha hecho, con las manos?
what Juan it has-3sg done, with the hands
- d. *Qué Juan lo ha hecho, porque sí?
what Juan it has-3sg done, because yes

The analysis Uriagereka proposes starts from the structure for adjuncts presented in (37c), a combination of his analyses of the headless nominal phrase in (37a) and the wh-phrase in (37b).

- (37) a. la sobrina del cura pateó
 the niece of-the priest kicked
 a [DP la [N' [pro] del alcalde]]
 to the of-the mayor
 'The niece of the priest kicked the one of the mayor'
- b. [DP qué [N' autora]]
 what author
 'Which author'
- c. [DP qué [N' pro]]

If the *qué* is the determiner of a *pro* as in (37c), and the *pro* is licensed by the adjunct tag, Uriagereka determines that the *qué* moves on its own, without pied-piping the *pro*, in the style of Latin "Left Branch" movement as shown in (38).

- (38) quid [flendo facies [t modi]]
 what weeping will-make limit
 'What limit will you put to your weeping?'

As an illustration, Uriagereka offers an analysis of (39). If the analysis included pied-piping, like that in (40), it would not work: regardless of what the XP is, it has to be dominated by IP so that the subject *pro* can c-command *la*¹¹. In that case, the PP would fail to c-command the *pro* in pre-IP position, and therefore that *pro* would not be licensed. The analysis where the *qué* movement does not include pied-piping is represented in (41).

¹¹ Quoting Uriagereka, "*La* in (39) is an inalienable possessor, which, as discussed in Kayne (1975), has essentially the properties of an anaphor: it has to be c-commanded by its antecedent." (Uriagereka, 1988, p. 206)

- (39) Qué lo vas a hacer, con la mano?
 what it go-2sg to do with the hand
 ‘What are you going to do it, with your hand?’
- (40) [CP [DP qué [NP pro]] C [IP pro [XP [lo vas a hacer] [PP con la ...
-
- (41) [CP qué C [IP pro [XP [lo vas a hacer [DP t [pro]]] [PP con la ...
-

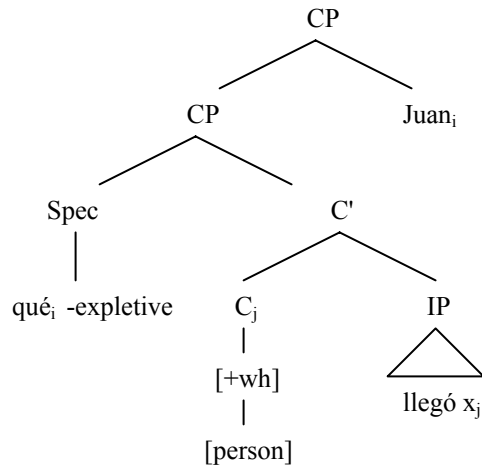
Thus, Uriagereka’s analysis captures the intuition that there is a c-command relationship between the tag and the phrase including *qué*; it does not address intonational phenomena or the exact way the coreferent adjuncts are linked in syntax.

3.2 Lorenzo (1994)

Lorenzo (1994) describes and studies cases similar to those described in (3) and (6a, b) above, where there is a *wh*-word linking both segments and the first part of the question is not viable as an independent question. He proposes the analysis illustrated in (42), where the element on the right is bearing focus and it is dislocated, and therefore base-generated (dislocated as opposed to topicalized; the terminology corresponds to that in Rivero (1980)).

- (42) a. Qué llegó , Juan?
 what arrived3sg, Juan?
 ‘Who arrived? was it Juan?’

b.



Lorenzo considers the interrogative in Spec CP to be an expletive. He discusses, as an alternative, the possibility of movement, supported by the presence of apparent subjacency effects¹² (43). He rules out movement, and attributes the ungrammaticality of (43b) to the need to establish a chain between an empty operator in C⁰ and an empty category in VP. This chain (not the movement of the expletive, which does not take place) would violate subjacency.

- (43) a. Qué debíamos alegrar, a Juan?
 what should1pl cheer-up , to Juan?
 ‘Whom were we supposed to cheer up? Juan?’

¹² I review the arguments for movement of *qué* in section 4.3. Here I present Lorenzo’s analysis without entering into detailed criticism.

- b. * Qué te empeñaste cuándo debíamos alegrar, a Juan?
 what CL2sg insist2sg when should1pl cheer-up , to Juan?
 ‘Whom did you insist when we were supposed to cheer up? Juan?’

Lorenzo’s argument in favor of dislocation is based on the contrast in grammaticality between (44a) and (44b).

- (44) a. Qué llegó e_i , Juan_i?
 what arrived3sg, Juan?
 b. * Qué llegó e_i a Oviedo , Juan_i?
 what arrived3sg to Oviedo , Juan?
 c. Quién_i e_i llegó e_i a Oviedo ?
 who arrived3sg to Oviedo ?

According to Lorenzo, (44b) is ungrammatical because the empty category is not properly governed (Lasnik & Saito, 1992). The verb *llegó* ‘arrived’ from (44b) cannot properly govern the empty category *e* because the verb is already involved in a thematic coindexation with a complement, in contrast with the grammatical (44a), in which there is no complement.

An analysis of (44b) where the element to the right had undergone movement would imply the existence of a trace of that movement in the specifier of AgrSP. This trace would be properly governed by the (sentence-final) moved element, thus licensing the chain and making (44b) grammatical. An example of this style of chain licensing

would be the grammatical (44c), where a chain is established between the trace inside VP, the trace in Spec AgrSP, and the moved interrogative in Spec CP. Therefore, Lorenzo rejects the movement analysis in favor of the dislocation analysis.

Lorenzo also describes examples where more than one element appears in the second part of the sentence, and discusses the varying grammaticality of additional elements included after the verb in the first part.

- (45) a. * Qué llegó a Oviedo , Juan?
what arrived3sg to Oviedo , Juan?
- b. ? Qué llegó Juan , a Oviedo?
what arrived3sg Juan , to Oviedo?

According to Lorenzo, subjects are more acceptable than complements in that position. In order to account for this contrast, he suggests that ‘destination’ is a type of argument that can also be interpreted as an adjunct.

In my view, the existence of such an alternative interpretation should create dubious grammaticality in both cases, not just one. The analysis does not provide any motivation for the change of interpretation of the destination as a complement or an adjunct.

Another problem with the analysis is that, in the structure Lorenzo proposes, the expletive and its associate c-command each other. Lorenzo notes this and sketches an alternative structure within Kayne's (1994) framework, where the construction would be derived by adjoining an interrogative sentence *Qué llegó?* to a "reduced" sentence structure *Juan (llegó)?* through an empty functional head X^0 , as in (46).

(46) [Qué llegó? [X^0 [Juan (llegó)?]]]

This direction of analysis eliminates the problems posed by the right dislocation, but leaves open the issue of the acceptability of (45a,b). The nature of the necessary empty head is not clear, either, and, perhaps most importantly, the left-adjoined interrogative sentence in (46), *qué llegó*, is one that would be ungrammatical on its own, and it is not clear how it would be licensed here. I address these issues in my own analysis, detailed in section 4, which fits within Kayne's framework, but postulates a stronger interdependency between the two parts of the question. Before that, in section 3.3, I will introduce Camacho's analysis, which postulates movement of the initial *qué*.

3.3 Camacho (2002)

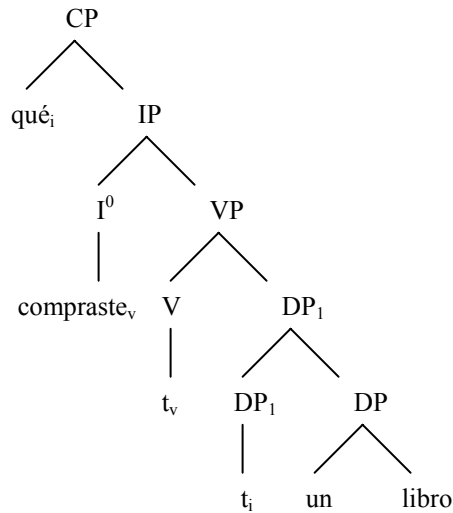
Camacho (2002) analyzes the cases where the wh-word in the first part of the question is *qué* and the second part is a DP (as *un libro* ‘a book’ in 47a) or a CP, the class presented above in (4). In such constructions no wh-word other than *qué* is expected, and therefore the first part of the question would be grammatical by itself. The wh-word and the DP or CP that constitute the second part of the question are related, and they are both interpreted as a complement of the verb *comprar* ‘to buy’ in (47a). Locality effects are also observed, as shown in (47b):

- (47) a. Qué compraste, un libro?
what bought2sg, a book?
‘What did you buy? a book?’
- b. * Qué te molesta el hecho de que comprara, un libro?
what CL2sg bothers3sg the fact of that bought3sg, a book?
‘What does it bother you the fact that he bought, a book?’

Camacho proposes that the DP or CP is adjoined to the trace of the wh-word, as in (48). This would explain the common argument position and the particular intonation of the construction.

- (48) a. Qué compraste, un libro?
 what bought2sg , a book?
 ‘What did you buy? A book?’

b.



Alternatively, Camacho considers another analysis within Kayne's (1994) framework (where there is no adjunction). In that analysis, the trace and the DP/CP would be part of a small clause complement, but he rejects it after noting that there are several problems with it, for instance the lack of other cases of small clauses inside DP in Spanish¹³.

¹³ The structure including the small clause would be something like:

- (i) [CP qué_i [IP compraste_v [VP t_v [DP₁ [DP un] [XP libro X⁰ t_i]]]]]

Here XP is a small clause complement, possibly CP; *libro* is in its specifier, and *qué* is the complement, with X⁰ null (Camacho 2002). Camacho provides the contrast between (ii a) and (ii b) to show that Spanish small clauses (*Juan presidente*) do not appear inside DPs, as in (ii b).

- (ii) a. Eligieron a Juan presidente.
 elected to Juan president

Camacho's proposal captures the intuition that both parts of the question combine to form a single structure, as well as the common assumption that the *wh*-word and the second part are somehow related. It does not explain the cases where intervening elements render the construction ungrammatical, like those in (45a), and it leaves open the issue of the nature of the *wh*-word *qué*, which would undergo *wh*-movement but would not be a [+*wh*] operator¹⁴. More importantly, it does not shed any light on the semantic contribution of the doubling in these constructions, i.e. on the difference between these constructions and regular yes/no questions.

'They elected Juan president.'

- b. *Eligieron al Juan presidente.
 elected to-the Juan president

¹⁴ Camacho takes the ungrammaticality of embedded split interrogatives as in (i) to be an indication that the *qué* in the split question does not have the [+*wh*] feature selected by the verb *preguntar* 'to wonder'.

- (i) * Se preguntan qué compraste, un libro?
 cl wonder-3pl what bought-2sg a book
 'They wonder what you bought, a book?'

3.4 López-Cortina (2003)

I will briefly present here the analysis of López-Cortina (2003). Section 4 will present an update in view of further evidence and theoretical advances.

Benincà and Poletto have suggested (Benincà & Poletto, 1999, 2001, 2004) that Focus is not a single projection but a sub-field of CP, i.e. a series of projections associated to different kinds of focus¹⁵. Then, it would be possible to separate a projection containing the interrogative *wh*- landing site. This site would be a WhP(hrase), postulated on diverse motivations elsewhere¹⁶ (Lee, 2001; Martorana, 2004). This Wh Phrase would be located below FocP. See (49): examples of a focalized element to the left of an interrogative are not obvious; in (49a) we have a case where a word is focalized by a speaker A in order to correct B's discourse. The element *hoy* 'today' is not a contrastive topic, since it is associated to a corrective A-accent (Jackendoff, 1972). Finding Focus to the right of a *wh*-word, on the other hand, is simply impossible (49b). Possible examples of focalized elements to the right of the interrogative, like the dubious example in (49c), can only be interpreted as rhetorical

¹⁵ In chapter 3 of this dissertation I have discussed a development of this idea, namely that the Focus field contains, in Spanish, the landing sites for three distinct kinds of *wh*-like movement (contrastive focus, answers, and *wh*-movement). Differences between *wh*- and focus are also discussed there.

¹⁶ Note that this Wh Phrase acts as landing site for *wh*-movement. It is a different concept from Rizzi's Int(errogative) Phrase (Rizzi, 2001, 2004), which is higher in the structure and hosts the interrogative complementizer *se* 'if'.

interrogatives. A possible analysis for them is that the wh-word would be up in an exclamative wh-position¹⁷.

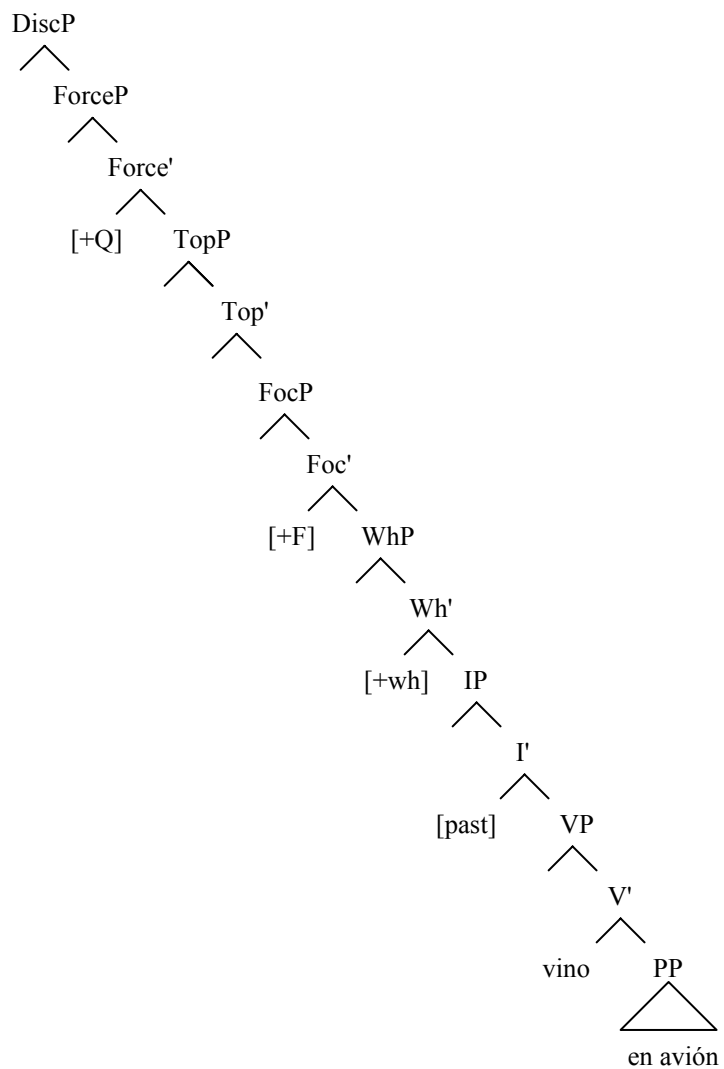
- (49) a. A: ¿Qué vas a hacer hoy?
what go2sg to do today
'What are you going to do today?'
B: Voy a estudiar mañana.
go1sg to study tomorrow
'I am going to study tomorrow'
A: HOY, qué vas a hacer?
today, what go2sg to do
'TODAY, what are you going to do?'
- b. – Voy a estudiar mañana.
go1sg to study tomorrow
'I am going to study tomorrow'
– * Qué, HOY, vas a hacer?
what , today, go2sg to do
'What, TODAY, are you going to do?'
- c. ? What, in these modern times, can we expect of the youth?

Consider now a simple sentence like (50). The sentence would be derived on the basis of the features present in the CP layer as it is represented in (50b). More specifically, I have proposed (López-Cortina, 2003) that the final output is the result of three movement operations, detailed in (51a) through (51d).

¹⁷ Alternatively, it has been proposed that English wh- landing site is higher than the Romance one (Barbosa, 2001).

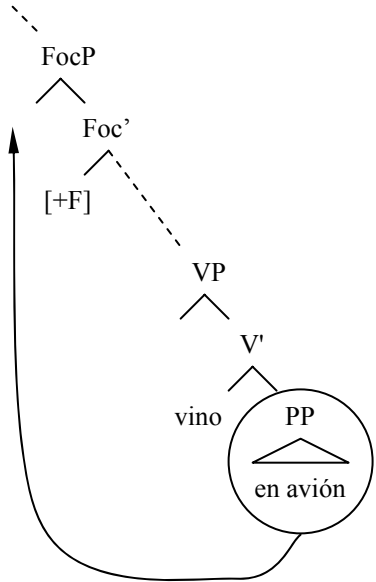
- (50) a. Qué vino , en avión?
 what came-3sg, in plane
 'How did he come? By plane?'

b.

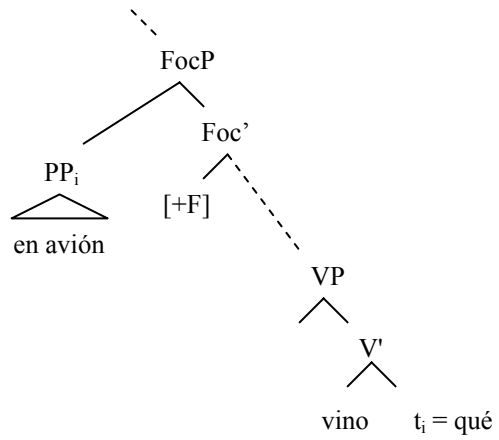


(51)

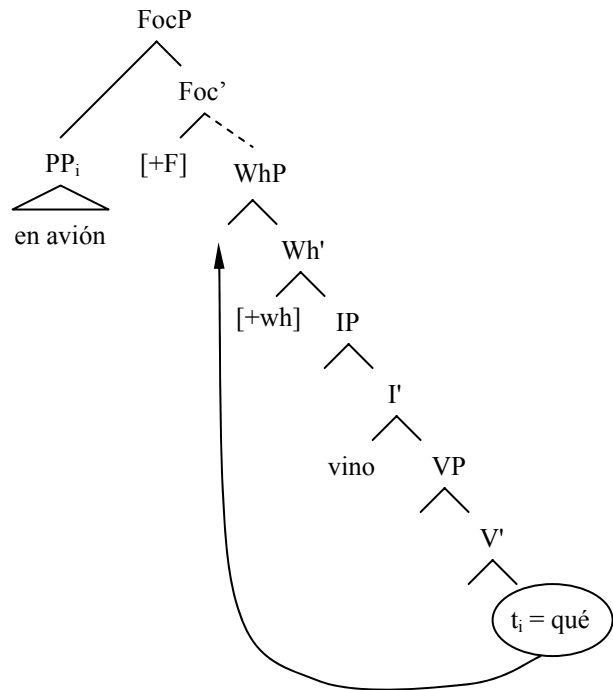
a.



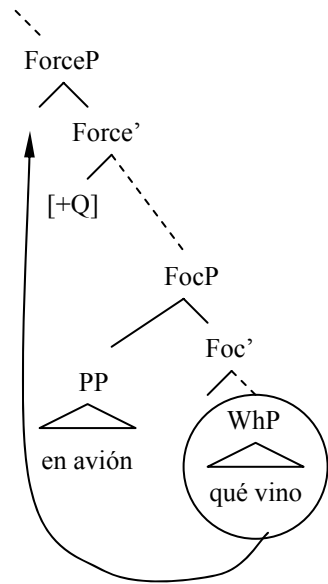
b.



c.



d.



The trees in (51a,b) illustrate the focalization of the element that will eventually appear to the right of the comma intonation. This movement leaves a trace, which will be spelled out as *qué*. (51c) shows the wh-movement of the trace left by the focalization. This movement does not take place in dialects where the *qué* is left *in situ*. In Chilean Spanish both possibilities exist¹⁸.

Finally, (51d) shows the movement of the entire WhP to Spec ForceP in order to check the [+Q] feature. This is what makes the sentence a question, as opposed to a sentence where an element has simply been focalized. Wh-words in regular wh-questions might or might not move all the way to Spec ForceP. It has been proposed that agreement with ForceP can be achieved by movement of the [+Q] feature by itself, once the element bearing it has reached some position in the left periphery (see analyses and references in Boškovič, 2001a; Hagstrom, 1998). In the case of split questions, however, the WhP would raise all the way to ForceP perhaps because otherwise its [+Q] feature would be separated from ForceP by the answer present in the Focus field¹⁹.

¹⁸ Minimalist assumptions exclude the possibility of movement being optional. It would be necessary to gather enough data from dialects that offer both possibilities (in situ *qué* and fronted *qué*) in order to determine whether there are any differences in meaning between the two constructions. Ultimately, this problem relates to the larger issue of wh- and contrastive focus in situ. The movement of *qué* is discussed in section 4.3.

¹⁹ This is one of the assumptions that will be revised in the new analysis presented in section 4.

Some of the differences between this and previous analyses follow:

- a) It provides a syntactic motivation (and, to the extent left periphery heads reflect information structure and sentence type, a semantic motivation as well) for the position of each of the elements in the sentence. The features that trigger the relevant syntactic operations are not new, and they are all independently motivated.
- b) It provides evidence for a separation between focus and wh- landing sites. The presence of an independent landing site for wh-movement is to be expected in a theory where movement-triggering features are associated to independent functional projections. If the analysis is correct, split questions provide evidence for the presence of a wh-specific functional projection. (The analysis discussed in section 4 abandons the idea that the tag of the split question is located in FocP.)
- c) It captures the intuition that the two parts of the question are constituents of a single structure.
- d) It presents the relationship between the wh-word and the proposed answer in terms of antecedent and trace. (I will show in section 4.1 that there are problems with this idea.)

There are several elements of this analysis that require a review. In the following section I reconsider the details of the analysis sketched in López-Cortina (2003) in the light of the data and proposals discussed in previous chapters of this dissertation.

4 A revised analysis

4.1 General considerations

Descriptions of split questions are growing increasingly detailed, and analyses are becoming more frequent in the last few years. This increasing relevance comes as a consequence of the fact that split questions lie at the crossroads of several current lines of linguistic research. Besides providing new data for the continuing research on wh-movement, split questions pose new challenges to the interpretation of right dislocations, fragments and tags. There are also questions about how tags relate to focus or, more generally, to the elements of information structure that are linked to the syntax of the left periphery of the sentence. Lastly, from a broader perspective, their hybrid question/answer character links their analysis to the issue of sentence typology.

Before developing a revised analysis that addresses all these questions, I will present a brief overview of the achievements of previous analyses and a summary of the theoretical framework discussed in previous chapters of this work.

4.1.1 Achievements and shortcomings of previous analyses

The analysis provided by Uriagereka (1988) addressed the role of wh-movement in split questions, showing indications that the interrogative word reaches its sentence-initial position through movement. The fact that Uriagereka only considers adjunct tags prevents him from addressing differences between wh-movement in regular wh-questions and split questions. His assumption that the tag and the interrogative word are only linked indirectly through a pro shows that he was aware of the importance of the right dislocation in these constructions. He does not address it directly, however, probably due to the lack of the necessary analytical tools within the theory at that time.

Lorenzo (1994) also gives priority to the right dislocation of the tag. His analysis, however, does not address the nature of the relationship between the initial wh-word and the tag, which are simply assumed to be coreferents.

Camacho (2002) takes the opposite approach and develops his analysis around the relationship between the wh-word and the tag. While his analysis provides an explanation for why they corefer, it does not address the dislocated character of the tag.

Finally, the analysis in López-Cortina (2003) addresses both the dislocation and the relationship between the wh-word and the tag, which are presented as different instances of a same constituent²⁰, one being a pronounced trace, or partial copy, of the

²⁰ An additional argument comes from Chilean Spanish data (Héctor Campos, p.c.), where the wh-word of split interrogatives does not move. Consider the sentences in (i):

- (i) a. Llevas leído(s) cinco libros
you-carry read-sing/plur five books
‘You have read five books’
- b. Cuántos llevas leídos, cinco?
how-many you-carry read-plur, five
‘How many have you read? Five?’
- c. Llevas leídos cuántos, cinco?
you-carry read-plur how-many, five
‘You have read how many? Five?’
- d. Llevas leídos qué, cinco?
you-carry read-plur what, five
‘You have read what, five?’
- e. *Llevas leído qué, cinco?
you-carry read-sing what, five
- f. Llevas leído(s) qué, cinco libros?
you-carry read-sing/plur what, five books
‘What have you read? Five books?’

Agreement between the past participle and its direct object is optional in a declarative (ia) but not in interrogatives (ib), even if the wh-word remains in situ (ic). This can be interpreted as evidence that there is a chain between the in situ wh-word and an invisible operator in the CP layer (see Watanabe, 2003 for the development of this idea for Japanese). If the wh-word *qué* takes the place of *cuántos* ‘how many’ (as in id), the singular form *qué* has mandatory plural agreement with the participle (see the contrast in id,e). That is to say, *qué* behaves exactly like *cuántos* regarding agreement. The fact that agreement is mandatory shows that this is not a case of *concordantia ad sensum*, which we could expect to be optional (as in (ia) or the parallel interrogative in (if)), but rather an effect of some syntactic link between the wh-word and the plural right-dislocated element. Note that, while this argument strongly supports a single clause analysis of split questions with the interrogative *qué*, it does not provide evidence in favor or against a single clause analysis of split interrogatives with full wh-words (ib,c). These might still consist of two juxtaposed questions, since the participle would be agreeing in number with the full wh-word, which is clearly within the same clause. Thus, number agreement does not show an obvious syntactic link between the participle and the possible answer. The possible analyses of full wh-word split questions are discussed in section 3.7 of this chapter.

other. This analysis does not go into the details of the syntax of split questions described in section 2 of this chapter, and it does not suffice to explain the differences between split questions and regular wh-questions.

4.1.2 What is left to address

The informational role of the tag has been assimilated to focus in previous analyses without much discussion. This is to be expected, since the concept of focus used at the time encompassed a range of phenomena, including contrastive focus and answers. Chapter 3 of this work presented answers as a sentence type structurally different from regular declaratives, the difference being the presence of an answer-specific functional projection in the left periphery. Split question tags appear to be a sort of proposed answer, so it will be necessary to establish whether there is a relationship between them and the structure of answers.

The syntax of split questions differs from that of regular wh-questions as described in section 2.3 above. The general observation may be that split questions are more restricted. It will be necessary to see what causes this difference.

4.1.3 The left periphery of the sentence

The analysis pursued here is based in the concept that the left periphery of the sentence is structured as a series of functional projections which are associated to information structure roles. That is to say, information structure is encoded in syntax, and, furthermore, it is encoded in a particular area of the sentence structure. These ideas were presented in a classic article by Luigi Rizzi (1997), who arranged previous observations about topic and focus in a single proposal of the way the left periphery is organized. The structure Rizzi proposes for the functional projections to the left of IP is this (the asterisk * is a Kleene star, indicating that there can be zero or more Topic Phrases):

(52) [ForceP [TopP* [FocP [TopP* [FinP [IP... (Rizzi 1997)

Refinements to this proposal have been formulated by other researchers as well as Rizzi himself (Benincà, 2001; Benincà & Poletto, 2004; Rizzi, 2001)²¹.

(53) [ForceP [TopP* [IntP* [TopP* [FocP [TopP* [FinP [IP... (Rizzi 2001)

²¹ Further references and a more detailed presentation of these issues can be found in chapter 2 of this work.

(54) [DiscP [ForceP [TopP [FocP [FinP [IP... (Benincà 2001)

The main lines of these proposals allow for a straightforward interpretation of previously collected data. Campos & Zampini (1990) collect and systematize previous observations for Spanish (Hernanz & Brucart, 1987; Zampini, 1988), which I present in terms of Topic and Focus in (55).

- (55) a. Topic precedes Focus.
b. Topic precedes interrogative words.
c. Focus and wh-movement are in complementary distribution.
d. There can be more than one Topic but only one Focus.

(Campos & Zampini 1990)

These observations for Italian and Spanish indicate that, in general lines, there is a Topic area or field preceding a Focus field. The elements that have usually been associated with the Focus position are of three types: contrastive Focus (identificational focus in É. Kiss terms (É. Kiss, 1998)), moved wh-words (Chomsky, 1977), and fragment answers (Brunetti, 2003; Merchant, 2004). The fact that these elements are generally perceived to be in complementary distribution has delayed the analysis of the Focus field as containing more than one projection. This step is due to Benincà and

Poletto (2004), who describe the left periphery in terms of fields rather than single projections and draw a clear distinction between “a higher Topic field hosting non-operator elements, and a lower Focus field hosting operator-like elements” (Benincà & Poletto, 2004, p. 53).

My analysis of split questions assumes that the focus field includes (at least) three independent functional projections, hosting respectively identificational focus, answers, and moved wh-words. I will summarize here the arguments for the existence of these separate functional projections.

4.2 The Focus field in Spanish

The first argument to differentiate separate elements within the Focus field is based on intonation and meaning. Left-dislocated focus, wh-questions and answers are each associated to a different intonational event that separates the left-peripheral element from the rest of the sentence. Each of them gives the sentence a different meaning, in terms of discourse relevance of the left-peripheral element. The interpretation of this element always involves a set of alternatives. In the case of focus, it is presented as a preferred option in a contrast; in the case of wh-questions, it is

presented as uncertainty relative to a set of possible alternatives. Finally, answers present the preferred alternative in a set introduced by a previous question. In chapter 3 of this work I introduced data showing that sentences used as answers were neither regular declaratives nor constructions involving (contrastive) focus movement. The difference between regular declaratives and answers, according to my analysis, is that answers involve movement to the left periphery of the material addressed by the question. Similar accounts of fragment answers have appeared in the literature (Merchant, 2004), postulating that the relevant part of the answer moves to a focus position, in order to allow for the ellipsis of the rest of the sentence. I argued that FocP cannot be the landing site of such movement, since fragment answers do not show the contrastive value associated to FocP. On the other hand, in view of intonation facts and some word order data, I proposed that all answers, fragments or not, do involve movement to the left periphery.

This argument is further illustrated by split questions. Consider now the preliminary analysis presented for split interrogatives in the previous section. In (51a) (López-Cortina, 2003), I had proposed a movement of a constituent –an answer– to the left periphery, in particular to FocP, leaving the rest of the sentence available for further operations. This is very similar to the analysis Merchant (2004) proposed for fragment answers, and presents the same problem: there is no contrastive value to this moved

constituent, at least if we understand ‘contrastive’ in the ‘corrective’ sense usually associated to the meaning of left periphery focus. The actual value of the moved element, both in fragment answers and in split questions, is that it selects one of the choices (or reduces the number of choices) of a set of alternatives created by a question (see the last section of chapter 3 for discussion and references).

Therefore, I propose that answers, be it in fragment answers, full sentence answers, or as a part of a split question, involve movement to the same answer-specific position in the left periphery, which I call Answer Phrase (AnsP). Such a phrase must be part of the Focus set of projections, with which it shares a number of characteristics, both semantic and morphological.

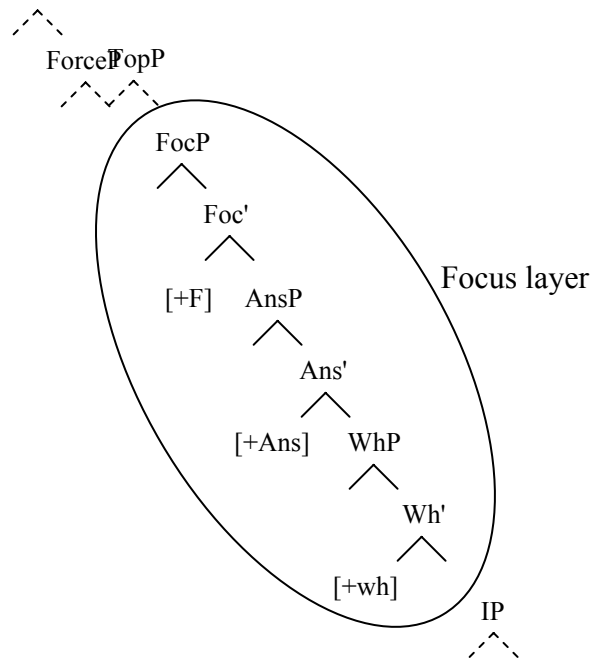
A second argument to differentiate separate projections within the Focus field comes also from split questions. Split questions provide an environment where two elements hosted in the Focus field, the *wh*-word and the answer, appear simultaneously, and thus they provide evidence not only for the existence of separate projections, but also for their ordering. In the analysis of López-Cortina (2003) the answer moves first to the left periphery, and then everything below it, including the *wh*-word, undergoes (remnant) movement to a higher position. Therefore, the position that hosts the answer must be immediately above the one targeted by *wh*-movement.

So far I have presented, on one hand, intonation and meaning facts that suggest the presence of three different projections in the Focus field, and, on the other hand, the co-occurrence of two of them, the one that hosts answers and the one that hosts wh-phrases, in split questions. The presence of in situ contrastive focus in wh-questions (56), while it shows a separation between the features of focus and those associated to wh-movement, does not suffice by itself to determine the existence of two separate functional projections.

- (56) a. Why did CLYDE lease a Jaguar? (Dretske, 1972)
b. Why did Clyde lease a JAGUAR?
c. Why did Clyde LEASE a Jaguar?
d. Why did Clyde lease a Jaguar?

The ordering of two possible separate projections for contrastive Focus and answers is not easy to determine, given the lack of clear examples where they co-occur. Since the analysis of split questions presented here does not make use of Focus features, this is not an issue that needs to be decided now, and I will leave it for later research. For the time being, I will follow the concept that each feature is associated to a separate functional head, as well as the idea that there is a Focus layer of projections, and I will propose the following structure for the focus layer in the left periphery:

(57)



In such a framework, the structure of (50) would be expressed in (58):

(58) [ForceP [WhP Qué vino] [Force' Force ... [FocP [AnsP [PP en avión] [Ans' Ans...]]]]]
what came3sg in plane
'How did he come? By plane?'

4.3 The nature of *qué* and its relationship with the tag

The analysis of the proleptic element *qué* as a phonetically realized trace²² presented in López-Cortina (2003) could be reformulated in terms compatible with a copy theory of movement, as long as an appropriate theory of partial deletion of copies is developed (McCloskey, 2006). In the minimalist framework, where the lexical array is closed to the incorporation of new elements during the derivation, movement involves copy of elements already present in the derivation, followed by internal merge and deletion of the lower copy, in order to minimize computation at the phonetic

²² Similar analyses have occasionally appeared in the literature. López-Cortina (2003) proposed that the *qué* that appears in the structure of a split question is an overt trace left by the focalization. The reason the trace was realized overtly is that, once the phrase containing it has moved to ForceP, it is left unbound, and therefore it would violate the Empty Category Principle. Phonetically realized traces were also proposed by Koopman for Kru languages (Koopman, 1982, 1984; see also McCloskey, 2006 for references). They have been proposed for Swedish by Engdahl (1985), whose examples I copy below. In Swedish, resumptive pronouns (*det* in (a); *han* in (b)) appear in wh-questions and can license parasitic gaps (*p* in (b)).

- (a) Vilket ord_i visste ingen hur_j det_i/*t_i stavas t_j ?
which word knew no-one how it is-spelled
“Which word did no one know how it is spelled?”
- (b) Vilken fånge_i var det läkarna inte kunde avgöra om han_i verkligen var sjuk utan att tala med *p*
personligen?
“Which prisoner was it the doctors couldn’t determine if he really was ill without talking to in person?”

However, Swedish resumptive pronouns that are variables in wh-questions are not spelled out as wh-, in spite of being traces of wh-phrases (in Engdahl’s analysis). Here I present a different situation, where the trace retains the interrogative character. It might be plausible that a [+wh] resumptive is spelled out as wh- in such cases. In fact, the idea of an wh-element being in an antecedent-trace relation with a fully spelled out wh-phrase has been proposed before, as a possible analysis for German wh-scope marking, by Jürgen Pafel and Gereon Müller in personal communication to Beck & Berman (Beck & Berman, 1996). Finding out whether there is some kind of parallelism between the phenomena for which pronounced traces have been proposed will be left for future research.

interface (Chomsky, 2005). In such accounts of movement, there are no “traces” to be pronounced, but, under certain special conditions, pronunciation of a residue of the lowest copy could be required, to satisfy interface conditions (Abels, 2001; Boškovič, 2001b; Chomsky, 2005; Hiraiwa; Landau, 2004).

Terminological questions aside, however, detailed review of the data reveals inconsistencies between the movement of the proleptic *qué* in split questions and the movement of fully spelled out wh-words. This suggests that the proleptic *qué* is subject to different movement restrictions than its coreferent is, and would therefore require reconsidering the issue of the antecedent-trace (or copy) relationship between them. In this section I discuss the characteristics of *qué* and propose a new analysis that takes into account the empirical difficulties posed by the previous one. Namely, I will propose that *qué* in split interrogatives is an interrogative pronoun that always occupies an adjunct position.

4.3.1 Movement of *qué*

The analysis in López-Cortina (2003) postulates movement of *qué*. Another possible interpretation would be to take *qué* to be an expletive marking the scope of the dislocated possible answer (Lorenzo González, 1994). However, as shown above in (17),

there are some dialects in which *qué* appears to have moved and some in which *qué* stays in situ. The occasional presence of *qué* in situ should be enough to rule out an interpretation as mere scope marker, if a common analysis is sought for both constructions.

In fact, the relationship between the two structures seems to present a clear parallel with regular wh-questions, where fronted wh-elements, in situ wh-elements, or both structures can be found, depending on the language or the dialect²³ (see Jiménez, 1997 for a review of the literature on the issue and an analysis of Spanish).

If the position of *qué* follows from wh-movement (51c), some restrictions to the construction should be expected, and they have in fact been noted in previous work. However, constraints on movement seem to be stronger in the case of split questions than they are in normal wh-movement, most clearly in dialects where no long extraction is allowed, but also in other dialects where long extraction does take place. Consider the examples (59c) and (60c) below. In both cases it can be shown that, in spite of having the same reference as an argument, *qué* does not behave like an argument, which could be extracted from a wh-island (59a)²⁴ or a complex NP (60a), but rather like an adjunct,

²³ Dialectal variation is discussed in section 4.5.

²⁴ Compare this example with the classic one from Torrego (1984):

(i) Quién no sabes cuánto pesa?
who not know2sg how-much weighs3sg
'Who don't you know how much he weighs?'

which could not (59b and 60b, respectively)²⁵. These facts are all immediately explained if we assume that *qué* is always an adjunct, independently of the status of the phrase it stands for. (For comparison, I repeat here Lorenzo González's (1994) subjacency example from (43) and Camacho's (2002) locality effects example from (47) in (59d) and (60d) respectively).

- (59) a. A quién sabes cuándo visitaron?
to who know2sg when visited3pl
'Who do you know when they visited (him)?'
- b. * Dónde sabes cuándo visitaron a Pedro?
where know2sg when visited3pl to Pedro
'Where do you know when they visited Pedro?'
- c. * Qué sabes cuándo visitaron , a Pedro?
what know2sg when visited3pl to Pedro
'Who do you know when they visited (him)? Pedro?'
- d. * Qué te empeñaste cuándo debíamos alegrar, a Juan?
what CL2sg insist2sg when should 1pl cheer-up , to Juan?
'Whom did you insist when we were supposed to cheer up? Juan?'
(Lorenzo González 1994)
- (60) a. Con quién oíste el rumor de que se quiere casar?
with who heard2sg the rumor of that CL3sg want3sg to-marry
'Whom did you hear the rumor that he wants to marry (her)?'
- b. * Cuándo oíste el rumor de que se quiere casar?
when heard2sg the rumor of that CL3sg want3sg to-marry
'When did you hear the rumor that he wants to marry (then)?'

²⁵ I thank Héctor Campos for pointing out these contrasts to me.

- c. * Qué oíste el rumor de que se quiere casar, con Mónica?
 what heard2sg the rumor of that CL3sg want3sg to-marry with Mónica
 ‘Who did you hear the rumor that he wants to marry (her)? Mónica?’
- d. * Qué te molesta el hecho de que comprara , un libro?
 what CL2sg bothers the fact of that bought3sg, a book?
 ‘What does it bother you the fact that he bought, a book?’
 (Camacho 2002)

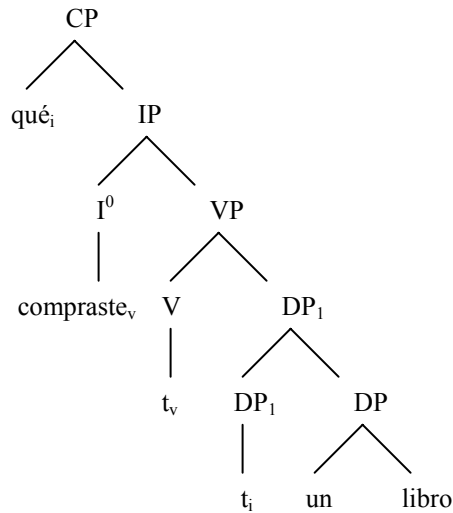
4.3.2 Position of *qué* in the structure

The idea that *qué* is always an adjunct, rather than a copy of the argument or adjunct it stands for, would require that *qué* appears in the initial numeration, which would be expected within a minimalist framework. The *qué* would be adjoined to the answer and carry the wh feature, while the answer would only carry the feature associated to the AnsP. The motivation for the doubling would be to enable the phrase to carry two different features: the one corresponding to a question and the one corresponding to an answer²⁶. This structure would be somewhat similar to the one proposed by Camacho (2002), which I presented above in (48) and copy here for convenience:

²⁶ The remaining puzzle, of course, would be why a sentence would require both question and answer features. In short, split questions are a sort of confirmation question, where the constituent to be confirmed carries a feature identifying it as the answer. The details of this, plus the differences with other types of questions, are discussed in sections 4.3 and 5.

- (61) a. Qué compraste, un libro?
 what buy-past-2sg, a book?
 ‘What did you buy? A book?’

b.



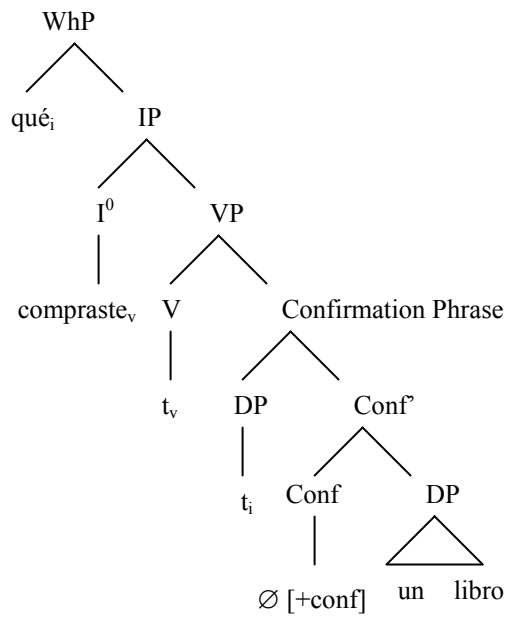
Such an adjunction structure is incompatible with Kayne’s Linear Correspondence Axiom. I will propose instead a doubling structure based on Grimshaw’s Extended Projection concept²⁷ (Grimshaw, 1991, 2005), where a whole functional projection acts as a functional extension of a lexical projection, and on Belletti’s (2003) extended doubling. See the structure of the same sentence in (62). Note that, for clarity, the movement of the answer to the left periphery and the subsequent remnant movement are not represented in this tree. A step-by-step derivation is presented in (63) for ease of reference, although some of the operations included in that derivation will be further

²⁷ I owe Héctor Campos the suggestion of looking into extended projections as a possible analysis. Alternative concepts of doubling and numerous references are presented in Belletti (2003).

discussed in the remainder of this section, including, for instance, the final landing site for the WhP (63c). Additionally, a tree presenting the final result can be seen in (64d).

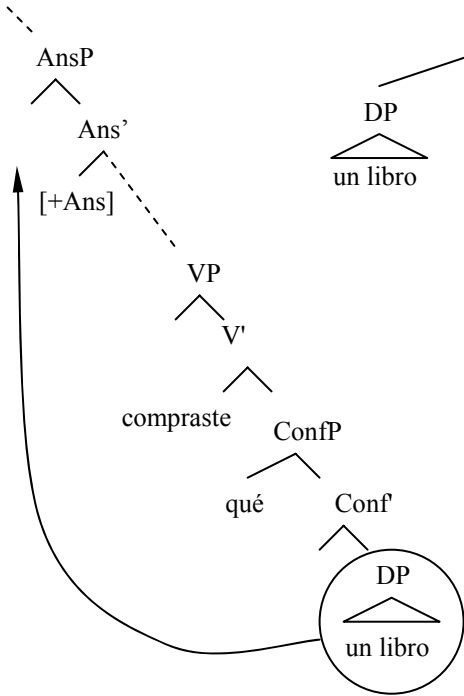
- (62) a. Qué compraste, un libro?
 what buy-past-2sg, a book?
 ‘What did you buy? A book?’

b.

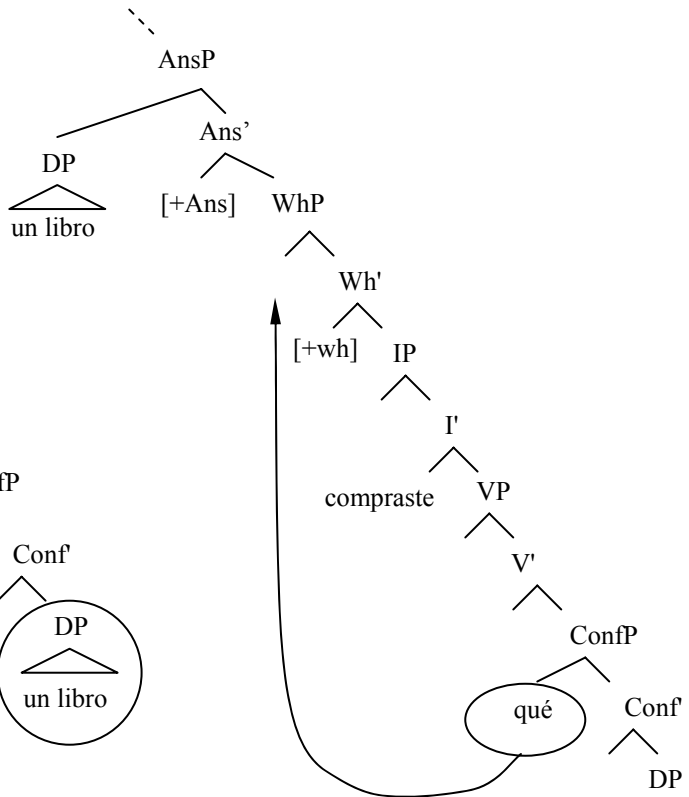


(63)

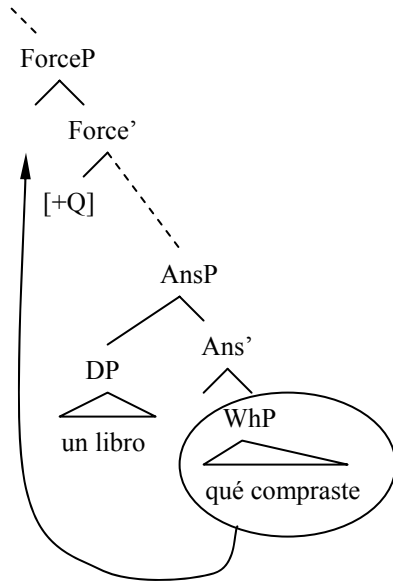
a.



b.²⁸



c.²⁹



I take *qué* to be in the Spec position of an extended projection, which would account for its adjunct-like behavior. This projection introduces the notion of need for confirmation, which I will represent as a feature [+conf], and it is optionally associated to a [+wh] feature. I call this phrase Confirmation Phrase (ConfP).

The semantic value introduced by ConfP is similar to that of Muysken's (1995) Non-Factual Phrase (NFP). Muysken proposes this phrase, whose head (in Quechua) would be the enclitic particle *-chu*, as part of a group of functional projections above IP whose combination accounts for the relationship between focus and negation in Quechua. The particle *-chu* appears both in interrogative and negative sentences, but is limited to main tense contexts and does not appear in wh-questions. There is, therefore, no obvious syntactic parallel between Quechua NFP and the ConfP I am proposing. The conceptual link, however, is clear: Both Quechua *-chu* and the *qué* of Spanish split questions are particles that appear associated to negative or interrogative sentence types to express that a sentence (in the case of NFP) or a constituent (in the case of ConfP) have not been confirmed as factual³⁰.

²⁸ Assuming that VP is a phase, probably both *un libro* and *qué* would adjoin to the periphery of VP on their way to the left periphery.

²⁹ The actual landing site and other details of this movement will be discussed in section 4.7.

³⁰ The example provided in (62) illustrates the doubling for a DP. The same doubling takes place at least with CP, IP, vP, PP, and adverbs, as illustrated by the following questions:

- (i) Qué dices, que compró una boina?
what say2sg that bought3sg a beret
'What are you saying, that he bought a beret?'

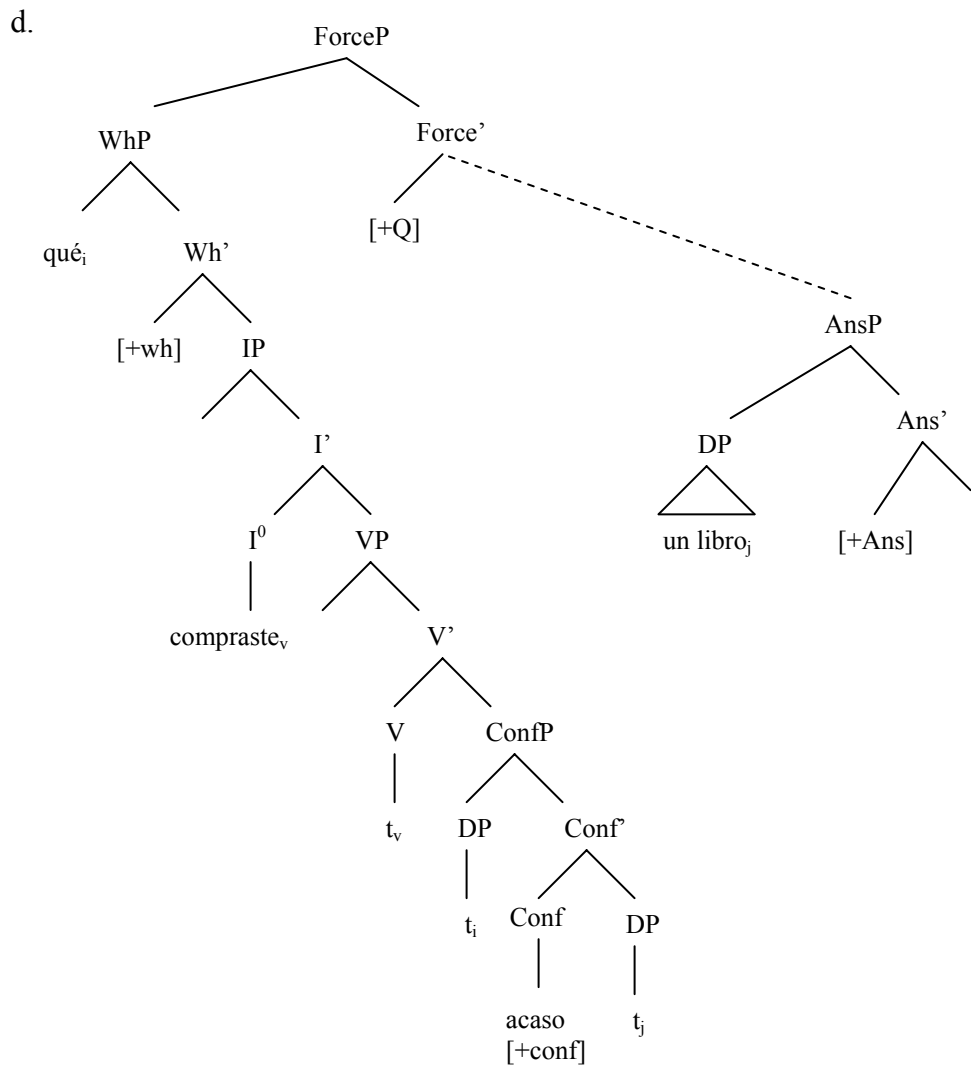
The head of ConfP is generally not pronounced in Spanish, although the interrogative marker *acaso* ‘perhaps, perchance’ seems to occupy that position in Chilean Spanish³¹ (64). Note how it is positioned before the intonational break in (64a,b), and immediately following *qué* in split questions where the wh-word is left in situ (64c). The structure proposed for (64a) is presented in (64d), where for the moment I am assuming movement to ForceP, but see section 4.7. Besides the introduction of ConfP, his structure is similar to the one proposed in López-Cortina (2003). The other difference is that moves to AnsP, not FocP. The same structure, except for the wh-movement, would apply for (64c).

-
- (ii) Qué, compró una boina?
 what bought3sg a beret
 ‘What, he bought a beret?’
 - (iii) Qué compró, una boina ayer?
 what buy-past3sg a beret yesterday
 ‘What did he buy, a beret yesterday?’
 - (iv) Qué la compró, para el frío?
 what cl-DO buy-past3sg for the cold
 ‘What did he buy it for? For the cold?’
 - (v) Qué la compró, ayer?
 what cl-DO buy-past3sg yesterday
 ‘When did he buy it? Yesterday?’

In all these cases the Confirmation Phrase is proposed to be present right above the node that ends up being the proposed answer after the comma.

³¹ Héctor Campos (p.c.).

- (64) a. Qué compraste acaso, un libro?
 what bought2sg *acaso* a book?
 ‘What did you buy? Perhaps a book?’
- b. Qué te vas acaso, mañana?
 what CL go2sg *acaso* tomorrow?
 ‘When are you leaving? Perhaps tomorrow?’
- c. Compraste qué acaso, un libro?
 bought2sg what *acaso* a book?
 ‘What did you buy? Perhaps a book?’



Acaso also serves as an interrogative marker in regular (65a) and embedded (65b,c,d) yes/no questions (Campos, 1993). In embedded questions in Chilean Spanish, *acaso* alternates with *si* ‘if’ as the element introducing the embedded clause (65b,c), but both can also appear simultaneously with *acaso* following *si* (65d).

- (65) a. *Acaso* conoces a María? (Chilean Spanish)
acaso know2sg to María
 ‘Do you know María?’
- b. Yo no sé si conoce a María. (Chilean Spanish)
 I not know1sg if know3sg to María
 ‘I don’t know whether he knows María’
- c. Yo no sé *acaso* conoce a María. (Chilean Spanish)
 I not know1sg *acaso* know3sg to María
 ‘I don’t know whether he knows María’
- d. Yo no sé si *acaso* conoce a María. (Chilean Spanish)
 I not know1sg if *acaso* know3sg to María
 ‘I don’t know whether he knows María’

The meaning of *acaso* in some Peninsular Spanish varieties matches its association to a confirmation role in Chilean Spanish split questions. Consider the contrast between (66a) and (66b). The question including *acaso* (66a) is not felicitous in neutral contexts, where (66b) would be uttered instead. Including *acaso* introduces the presupposition that the person asking is expecting either a *yes* or either a *no*. An example where a negative answer is expected would be a situation where the person asking wants to point out that the person asked is acting or speaking as if he knew

María, but he actually does not know her. Examples where a positive answer is expected would be situations where the person asked and María are acting as if they in fact had met before, and the person asking wishes to confirm this. The use of *acaso* with positive expectations sounds somewhat archaic in Peninsular Spanish, but it is very common in cases where the expected answer is negative.

- (66) a. *Acaso conoces a María?* (some Peninsular Spanish)
acaso know2sg to María
 ‘Do you happen to know María?’
- b. *Conoces a María?*
 know2sg to María
 ‘Do you know María?’
- c. *Qué conoces, a María?* (Asturian Spanish)
 what know2sg to María
 ‘So you happen to know *María*?’ or ‘Who do you know? María?’
- d. *Qué ye, que conoces a María?* (Asturian)
 what is that know2sg to María
 ‘So you happen to know María?’

In varieties where a word like *acaso* is not available or popular, but split questions are present, split questions can be used to introduce this sort of confirmation value in the question. Thus, Asturian Spanish (66c,d) can be used with a meaning similar to standard Spanish (66a), with either a positive or a negative expectation. The confirmation value extends to the whole question if the whole IP is placed in an answer

position (see Asturian/Asturian Spanish (66d)). The contrast between (66c) and (66d) is the contrast between the person asking being surprised that the person asked knows María (66d) (corresponding to *conoces a María* being the information focus in (66a)) or surprised that the person asked knows *precisely* María (66c) (corresponding to *María* being the information focus in (66a)).

The link between the meaning introduced by *acaso* and that introduced by split questions is therefore clear. In Chilean Spanish, where *acaso* can serve to define an interrogative sentence type, as it does in the indirect polar question of (65c), it also appears in to split questions, where it is associated to the same yes/no need-for-confirmation value (64).

4.3.3 There are no embedded split questions

Note that, in my proposal, (against Camacho, 2002) *qué* still is a [+wh] operator, carrying a [+wh] feature (see 64d). Camacho argues that it cannot be. He presents as evidence the impossibility of sentences like (67), which he takes to be ruled out by selectional restrictions of the verb *preguntar* ‘to wonder’:

- (67) * Se preguntan qué compraste, un libro?
CL wonder3pl what bought2sg , a book?
‘They wonder what did you buy, a book?’

In my view, the reason for the impossibility of (67), and in general the impossibility of embedded split questions, is that split questions involve additional features ([+answer], [+conf]) which are not present in a regular embedded (or indirect) question and are incompatible with the sort of verbs that select embedded questions.

The issue of selection of embedded questions in Spanish presents certain complexity (Plann, 1982; Suñer, 1991, 1993, 1994); in order to provide an appropriate context to the hypothesis presented here, I will start by noting that verbs like *preguntarse* ‘to wonder’ select for structures that present an optional complementizer in addition to the wh-word (see 68), while other verbs like *saber* ‘to know’ select for structures with the wh-word only.

- (68) Preguntan (que) qué compraste.
ask3pl that what bought2sg
‘They ask what you bought’

Suñer (1991) distinguishes between true indirect questions (including the complementizer or not) and semi-questions or pseudo-questions (obligatorily missing the complementizer). She refers to the co-occurrence of the complementizer and the wh-word in true indirect questions as “doubly-filled Comp”. Thus, true indirect questions would have a double CP.

Suñer’s analysis attributes a “semantic +WH” feature to the higher CP of true indirect questions, which would be missing in semi-questions, while syntactic features as [+wh] are associated to the lower CP. Such an analysis is straightforwardly compatible with the view of the left periphery assumed in this dissertation: Suñer’s higher CP, which determines the sentence type, corresponds to Rizzi’s ForceP, while the lower CP hosting the wh-word corresponds to the WhP proposed in López-Cortina (2003).

In order to capture Suñer’s insights in the framework I am assuming, I now present a classification of interrogative sentences on the basis of two features, one of them, which I will term [+Q], associated to ForceP (Suñer’s “higher CP”) and the other one, [+wh], associated to WhP (Suñer’s “lower CP”)³². True questions would carry a “question” feature in ForceP, which would roughly correspond to the feature proposed by Baker (1970) and to Suñer’s semantic +WH. The issue left to clarify is the difference between direct and indirect questions. Following Suñer, it can be assumed that direct and indirect true interrogatives share a common feature [+Q] which is missing in semi-

³² Suñer proposes that *si* ‘if’ appears in the lower CP and it might be associated to a [+Q] in the style of Baker (1970). Following Rizzi (2001), I will assume that *si* ‘if’ is associated to a separate functional head, below the projection that hosts the complementizer *que*, as in (i), thus dodging the question of whether it carries the same [+wh] feature as wh-words.

- (i) Preguntan (que) si vienes.
 ask3pl that if come2sg
 ‘They ask whether you are coming’.

questions. The difference between direct and indirect true interrogatives is the presence of an optional *que* complementizer in indirect interrogatives, which is not allowed in direct ones. In some, but not all, dialects, the intonational contour is also different. There would be, thus, two different types of complementizer for interrogative sentences. The first type might be visible in the utterance through the presence of a suprasegmental, and has no other overt markings³³. It is associated to direct questions. The second type has no link to suprasegmentals, but appears instead as an (optional) segment *que*, and is associated to indirect questions. Semi-questions are lacking an interrogative [+Q] ForceP, as they lack the higher CP in Suñer's analysis³⁴. Verbs that select for indirect questions, like *preguntar* 'to ask', select for phrases carrying the [+Q] feature, but only those with the indirect question complementizer. This would be a sort of lexical selection on top of a semantic one.

In terms of features, an embedded semi-question has a [+wh] feature but not a [+Q] feature. Polar questions have a [+Q] feature but not a [+wh] one. Direct wh-questions, as well as embedded true questions, must have both features, the difference between them being the sort of complementizer they can have, and in some dialects

³³ Except perhaps in Spanish dialects that allow for an interrogative marker *y* to precede a direct question.

(i) Y cuándo vienes? (Chilean Spanish)
 int when come2sg
 'When are you coming?'

³⁴ It might be the case that these constructions are free relatives, rather than just CPs. I will leave this open for further research.

their intonational contour. Split questions have a [+Q] feature, which makes them questions, and a [+wh] feature, that accounts for the movement of *qué*. The [+answer] feature associated to the tag, and the [+conf] feature associated to its coreferent *qué* allow for the particular interpretation of these structures and prevent them from being selected by verbs of interrogation. This is summarized in the chart in (69).

(69)

Features	Sentence type	Example
[+wh]	Suñer's semi-question	Saben (*que) qué compraste. 'They know what you bought'
[+Q]	Direct yes/no question	Compraste un libro? 'Did you buy a book?'
[+Q], [+wh]	Suñer's true indirect question	Preguntan (que) qué compraste. 'They ask what you bought'
[+Q], [+wh]	Direct wh-question	Qué compraste? 'What did you buy?'
[+answer]	Answer	Un libro (, compré). 'A book (I bought)'
[+Q], [+conf], [+answer], [+wh]	Split question	Qué compraste, un libro? 'What did you buy, a book?'

4.3.4 Long distance movement

Let us look now at the problem of long distance movement in split questions. Acceptability of long distance movement varies from dialect to dialect. I will present here the three patterns I have found, starting from the most constrained dialect, and

summarize the differences between them in a chart in (74). Recall from section 2 that most of the speakers interviewed agree on the pattern presented in (70) to (72) below, copied from (19) to (21) for convenience. These data show that long distance movement is not allowed in the corresponding dialects. Similar constructions with pairs of juxtaposed questions (70b) (71b) (72b), do not present a problem, unless of course one of the juxtaposed questions is ungrammatical by its own merit (73b). Structures with the same word order, but where the intonational split indicates that the answer includes the embedded IP (71c) (and therefore there is no long distance movement), are also acceptable.

- (70) a. *Qué crees que compró, un libro?
 what believe2sg that bought3sg a book
 ‘Do you think it is a book, what he bought?’
- b. Qué crees que compró? Un libro?
 what believe2sg that bought3sg a book
 ‘What do you think he bought? A book?’
- (71) a. *Qué dices que llega, mañana?
 what say2sg that arrive3sg tomorrow
 ‘Are you saying it is tomorrow when he arrives?’
- b. Cuándo dices que llega? Mañana?
 when say2sg that arrive3sg tomorrow
 ‘When are you saying he is arriving? Tomorrow?’
- c. Qué dices, que llega mañana?
 what say2sg that arrive3sg tomorrow
 ‘Are you saying he arrives tomorrow?’

- (72) a. *Qué esperas que llegue, a las cuatro?
 what expect2sg that arrive3sg at the four
 ‘Are you expecting him to arrive at four?’
- b. Cuándo esperas que llegue? A las cuatro?
 when expect2sg that arrive3sg at the four
 ‘When are you expecting him to arrive? At four?’
- (73) a. *Qué lamentas que llegue, a las cuatro?
 what regret2sg that arrive3sg at the four
 ‘Do you regret that it is at four when he arrives?’
- b. *Cuándo lamentas que llegue? A las cuatro?
 when regret2sg that arrive3sg at the four
 ‘When do you regret that he arrives? At four?’

Recall, also, that the pattern of some of the speakers that accept long distance movement is more complex. For them, (70a) and (71a) are grammatical, but (72a) and (73a) are not. The same pattern appears in Catalan (Contreras & Roca, in press). Lastly, there is a third group of speakers (of Chilean Spanish) whose speech appears to present no restrictions to long distance movement in split questions except for factive verbs like *lamentar* ‘to regret’. The chart in (74) summarizes the differences between dialects:

(74)

	Dialect A Most dialects		Dialect B some Asturian Spanish speakers, Catalan		Dialect C Chilean Spanish	
	regular question	split question	regular question	split question	regular question	split question
<i>decir, creer</i> (non-factive verbs) argument	(70b) ✓	(70a) *	(70b) ✓	(70a) ✓	(70b) ✓	(70a) ✓
<i>decir, creer</i> (non-factive verbs) adjunct	(71b) ✓	(71a) *	(71b) ✓	(71a) ✓	(71b) ✓	(71a) ✓
<i>esperar</i> (irrealis verb) adjunct	(72b) ✓	(72a) *	(72b) ✓	(72a) *	(72b) ✓	(72a) ✓
<i>lamentar, molestar</i> (factive verbs) adjunct	(73b) *	(73a) *	(73b) *	(73a) *	(73b) *	(73a) *

The data show no correlation between the acceptability of long distance split questions and the quality of argument or adjunct of the moved element. In both dialects that allow for long distance split questions there are adjuncts that can move and other adjuncts that can not move. There is also no obvious correlation between the acceptability of long distance movement in regular questions and in split questions. This suggests that the variation between dialects is not a result of general constraints on wh-movement, but rather of lexical particularities of the verbs involved. There are, then, two issues here. One, what determines differences between dialects. The differences between dialects seem to run along the lines of familiar classifications of verbs. Differences between factive, non-factive, and irrealis verbs are well documented in the

literature, from the classic analysis by Kiparsky & Kiparsky (1971) to the more recent one by Carlos de Cuba (2007).

The second issue is the lack of correlation between the behavior of regular wh-questions and split questions with respect to long distance movement across the same verbs. In the dialects where there is a difference, long distance movement is more constrained in split questions than it is in regular questions. The one element that is present in all split questions but is not present in any regular question is the answer tag. It is possible that it is the long distance movement of the answer to the left periphery that is blocked (in the dialects where it is always blocked), rather than the wh-movement of the *qué*. Consider the following structure for the sentence in (71a), repeated here:

- (75) a. *Qué dices que llega, mañana?
 what say2sg that arrive3sg tomorrow
 ‘Are you saying it is tomorrow when he arrives?’
- b. [ForceP [AnsP [WhP [IP dices [ForceP que [AnsP [WhP [IP llega [ConfP qué mañana]]]]]]]]]]]

The answer *mañana* would need to cross a whole left periphery in order to reach its final position in the leftmost AnsP. It is possible that answer movement is bound in Dialect A and unbound in Dialects B and C. The types of verbs that do not allow for extraction in Dialects B and C present consistently a greater obstacle to movement in

other languages, so it is to be expected that they do the same here. The obstacle could consist of these verbs selecting sentence complements with a defective left periphery, where no AnsP would be available as an intermediate step for answer movement.

Alternatively, there might be no difference in the characteristics of answer movement between the dialects. In that case, Dialect A would have no verbs selecting complements with complete left peripheries, while Dialects B and C would have one or more types of verbs selecting them.

4.3.5 Negative split questions

Further differences between regular questions and split questions are visible in the presence of negation. Wh split questions cannot be negative, at least in the dialects of my informants in Chile and in Asturias, Spain. The status of negation as a weak island is well known; it allows for movement of arguments but not adjuncts (Ross, 1983).

- (76) a. Who don't you think that John talked to?
b. *Why don't you think that John talked to Mary?

In Spanish, there is also a contrast between arguments and adjuncts in short extraction. Arguments extraction is allowed (77a,b), but adjunct extraction is not (77c), except in some cases where the adjunct is d-linked, which in some ways resemble an echo-question (77d)³⁵.

- (77) a. Qué no quieres?
 what not you-want
 ‘What don’t you want?’
- b. Qué no comes?
 what not you-eat
 ‘What don’t you eat?’
- c. *Cuándo no vienes?
 when not you-come
 ‘When aren’t you coming?’
- d. Cuando no vienes? (d-linked)
 when not you-come
 ‘When did you say you weren’t coming?’

In the case of split questions, however, it doesn’t seem to matter whether the dislocated element is an argument or an adjunct³⁶, as seen in the ungrammatical examples in (78) (copied from (25) for convenience).

³⁵ Movement of different types of adjuncts through weak islands presents different degrees of acceptability. In Spanish, it seems that *dónde* ‘where’ is more acceptable than *cuándo* ‘when’, which in turn would be more acceptable than *cómo* ‘how’. See Rizzi (1990).

³⁶ The same is true for German partial wh-movement (examples from Stechow, 1996).

- a. *Was glaubt niemand wen Karl gesehen hat?
 What believes nobody who, Karl seen has?

- (78) a. *Qué no quieres, el verde?
 what not want2sg, the green
 ‘What don’t you want, the green one?’
- b. *Qué no comes, cereales?
 what not eat2sg, cereals
 ‘What don’t you eat, the cereal?’
- c. *Qué no vienes, el sábado?
 what not come2sg, the Saturday
 ‘When aren’t you coming, on Saturday?’
- d. *Qué no vienes, porque te lo pidió ella?
 what not come2sg, because to-you it asked she
 ‘Why aren’t you coming, because she asked you not to?’

These facts are all immediately explained if we assume that *qué* is always an adjunct, independently of the status of the phrase it stands for. This would lend additional support to the idea that *qué* originates as the specifier of an extended projection of the phrase carrying the answer feature.

-
- b. Wen glaubt niemand daß Karl gesehen hat?
 Who believes nobody that Karl seen has?

Beck (1996) proposes that negation and quantifiers block movement at LF, but not overt movement. This kind of filter would explain the facts for split questions as well, if we assume that the element on the right in split questions has to move to the [+wh] position at LF.

4.3.6 Section summary

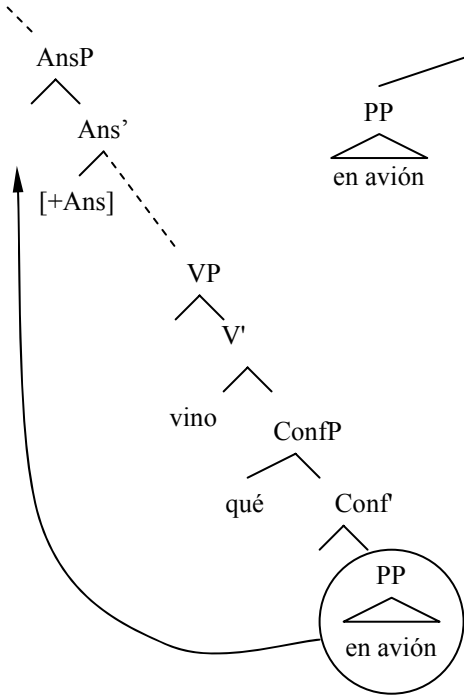
In the last two sections I have presented a more detailed structure of the left periphery, and a more appropriate description of *qué*. Taking these elements into account, I present here a more elaborate proposal for the syntax of split questions, illustrated in (79) below. It maintains the conceptual basis of López-Cortina's (2003) proposal, but recovers Camacho's view of *qué* being adjoined (Camacho, 2002), abandoning the assumption of *qué* being a trace of the answer, which was one of the premises of the previous analysis. This view (that *qué* in split constructions has different properties than a regular interrogative word) is now supported with data on islands presented in section 4.3.1 above, as well as for the impossibility of negative split questions discussed in section 4.3.5.

This analysis is also compatible with the idea that answers are separated from focus, and this, in turn, allows for a better account of the semantic properties of the construction, which line up with those of questions and answers, rather than with contrastive left periphery focus. Those dialects that include a confirmation adverb *acaso* provide a clear example of how the confirmation value is associated to split questions by means of additional structure adjacent to the position where the answer tag is generated. Additionally, the idea that there is a landing site for wh-movement separate

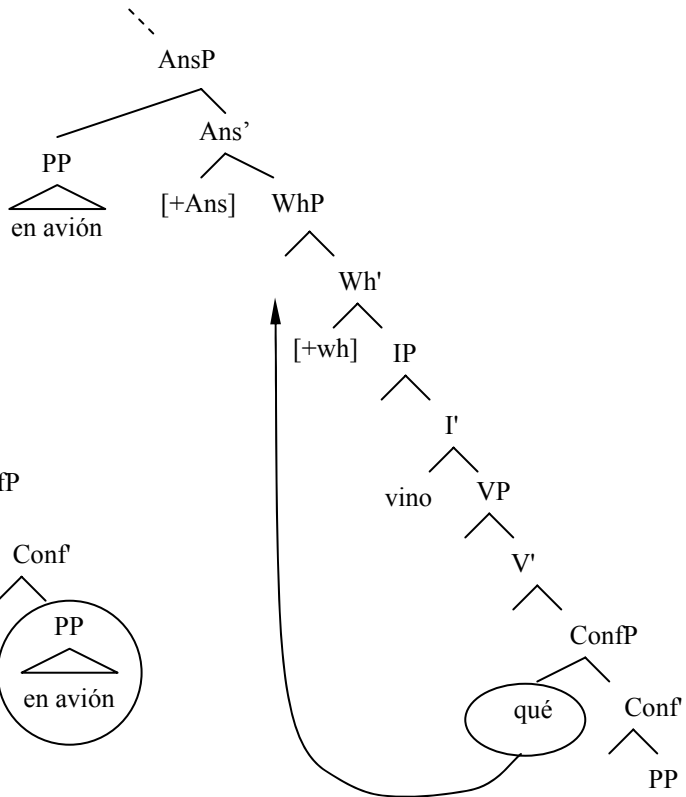
from the locus of interrogative force, crucial to this analysis, has been shown to be supported by previous observations and analyses of the structure of Spanish embedded questions (Suñer, 1991). The analysis, thus, reflects the intuition that split questions are the result of the presence of four features, three of which cause three separate syntactic effects, while a fourth one, purely semantic, introduces the confirmation character of the question. The first syntactic operation realizes the answer character of a constituent, associated to the [+answer] feature, which raises to the left periphery of the sentence, as illustrated in (79a). The second feature, [+wh], causes the movement of the adjunct doubled element of the answer. This movement is illustrated in (79b). The third feature is associated to the interrogative force of the sentence, and causes a third movement, illustrated in (79c), which allows the interrogative part of the split question to take precedence over the answer and provides the characteristic interrogative contour. Although for the time being I am assuming this movement to target ForceP, its actual landing site will be discussed in detail in section 4.7. The fourth feature, [+conf], causes no syntactic effect, and can be described as the semantic value of a sort of confirmation morpheme.

(79)

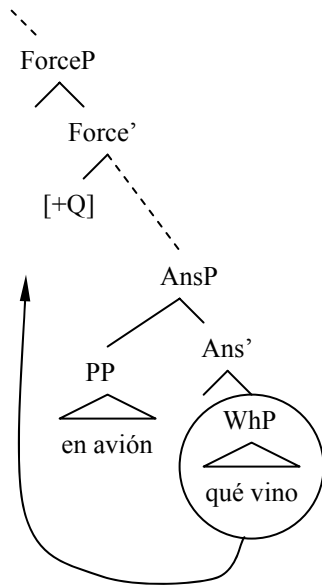
a.



b.



c.



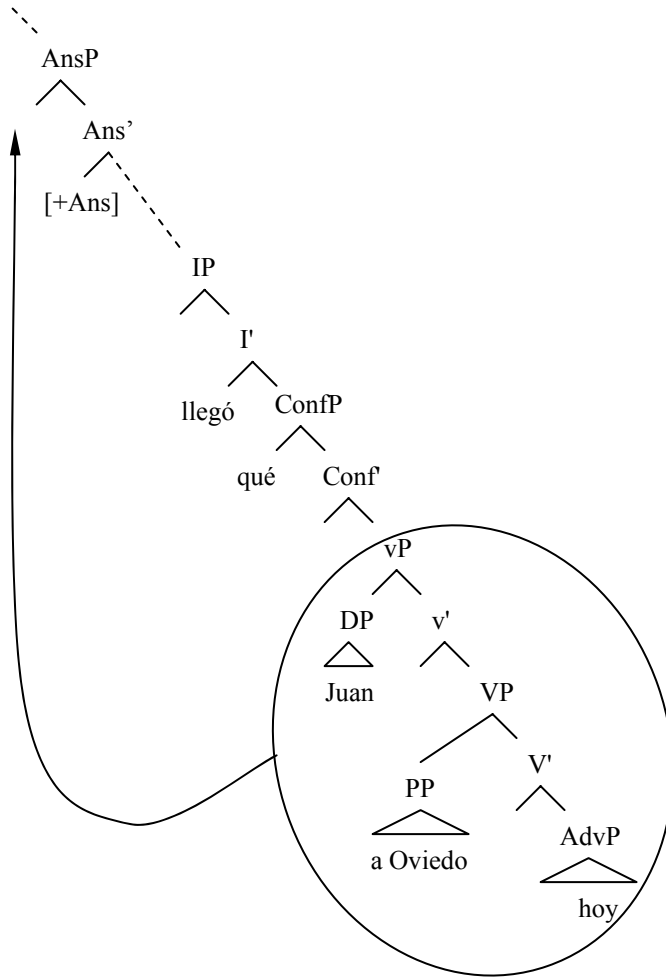
The next sections discuss how this analysis applies to a variety of instances of split questions. Section 4.4 discusses those cases where the tag appears to include more than one constituent. Section 4.5 deals with dialectal variation, showing in particular what sorts of constituent can be associated to ConfP and the effects this has in the acceptability of sentences with constituents between the verb and the tag. Section 4.6 continues this topic with discussion of those cases where ConfP appears to be associated to IP. Section 4.7 discusses the presence of elements to the left of *qué* in a split question. Section 4.8 briefly discusses split questions involving gerunds, and shows how they support a single-clause analysis. Finally, section 4.9 returns to the issue of two-clause structures resembling split questions, which I refer to as follow-up questions.

4.4 Multiple elements on the right

The analysis of multiple elements to the right of the comma intonation is straightforward once Larsonian style VP shells (Larson, 1988) are assumed: the entire vP is moved to the left periphery once the verb has raised to I. The doubled element in Spec, ConfP stands for the whole vP³⁷.

³⁷ Not all dialects allow for this sort of doubling. Differences are discussed in section 4.5.

- (80) Qué llegó , Juan a Oviedo hoy?
 what arrived-3sg , Juan to Oviedo today?



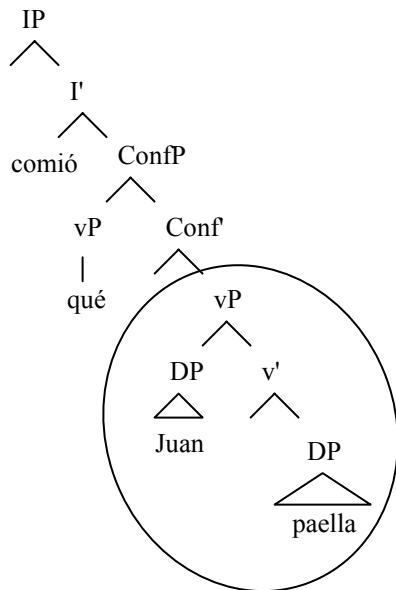
A fact that supports this interpretation is that the order of the elements on the right is not free. I take the examples in (81) from Lorenzo (1994):

- (81) a. * Que comió, paella Juan?
 what ate rice Juan
 ‘What did who eat, paella Juan’
- b. * Que comió paella, Juan en el desayuno?
 what ate rice Juan in the breakfast
 ‘When did who eat rice, Juan for breakfast?’
- c. Qué llegó, Juan en bicicleta?
 what arrived Juan in bicycle
 ‘Who arrived how, Juan on a bike?’
- d. ?Qué llegó, Juan a Oviedo?
 what arrived, Juan to Oviedo
 ‘Who arrived where, Juan to Oviedo?’

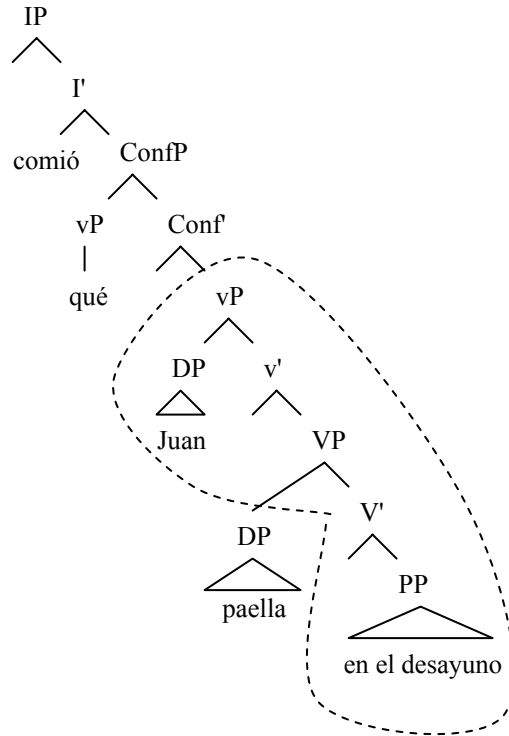
Lorenzo perceives (81d) to be dubious. I find it fine, but regardless of the quality of that particular example, there is no doubt that grammaticality disappears as soon as there is a change in the canonical order of the subject and the complements or adjuncts (81a) or the two elements are not contiguous in the structure (81b). This shows that the element on the right of the comma in these sentences is a single constituent, not several, and that no movement operations have taken place inside it. The trees in (82a,b) show, respectively, how the elements involved in (81a) are in the opposite order they would be expected to be, and how there is no single phrase containing the elements extracted in (81b). The ungrammaticality of these sentences is therefore predicted by this analysis.

(82)

a.



b.



Notice that the order of constituents follows from the order in which they are generated, assuming a Larsonian structure including vP shells. This is also consistent with the observation that an affirmative or negative answer to the question confirms or denies the whole constituent, not individual elements within it.

4.5 Dialectal variation

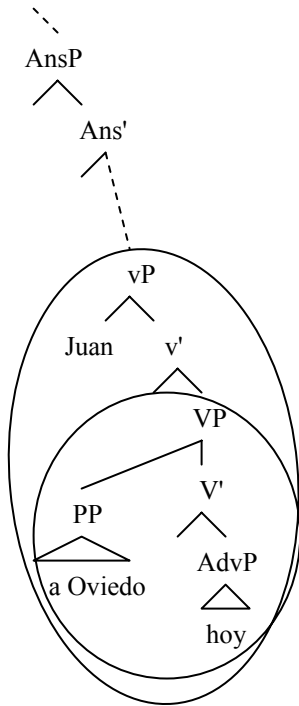
Dialectal variation would follow from the different structural elements subject to Answer movement in each variety. Asturian Spanish allows this kind of movement only in the case of verbal projections, either including or excluding the subject³⁸. Other northern Peninsular Spanish varieties allow it for any given projection below IP, and Chilean Spanish allows the movement of any projection below the one that contains the subject. ConfP will appear in the structure as an extended projection of the one that moves in each case. A comparison of the answer extractions allowed in each dialect follows (83). The trees shown do not include ConfP, in order to be able to represent several possible structures in a single tree for each dialect.

³⁸ Galician appears to present the same pattern (María del Carmen Parafita-Couto, p.c.).

(83)

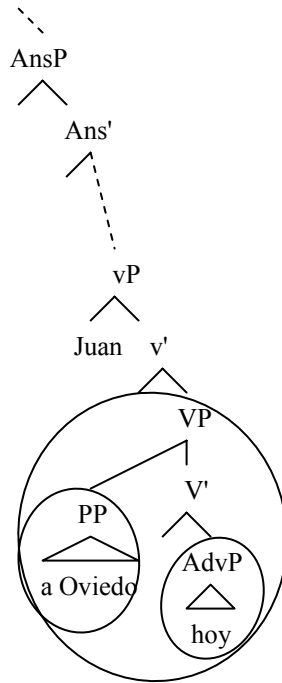
qué llegó Juan a Oviedo hoy
what arrive-past-3rd Juan to Oviedo today

'qué' refers to the element or elements to the right of the comma in each case



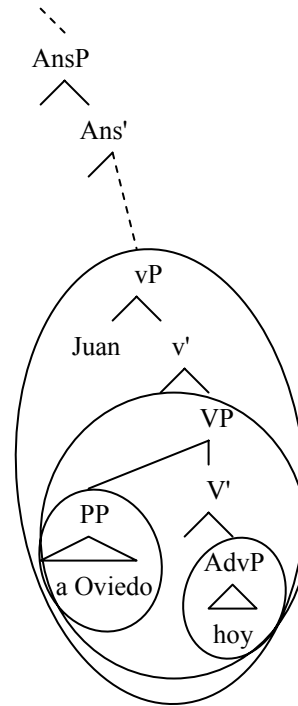
Asturian

- Qué llegó, Juan a Oviedo hoy?
- Qué llegó Juan, a Oviedo hoy?
- * Qué llegó Juan hoy, a Oviedo?
- * Qué llegó Juan a Oviedo, hoy?
- * Qué llegó a Oviedo hoy, Juan?



Chilean

- * Qué llegó, Juan a Oviedo hoy?
- Qué llegó Juan, a Oviedo hoy?
- Qué llegó Juan hoy, a Oviedo?
- Qué llegó Juan a Oviedo, hoy?
- * Qué llegó a Oviedo hoy, Juan?



Some Peninsular varieties

- Qué llegó, Juan a Oviedo hoy?
- Qué llegó Juan, a Oviedo hoy?
- Qué llegó Juan hoy, a Oviedo?
- Qué llegó Juan a Oviedo, hoy?
- Qué llegó a Oviedo hoy, Juan?

This analysis of dialect variation is supported by the behavior of the interrogative particle *acaso* in Chilean Spanish. Recall (from section 4.3) the proposal that *acaso* is the head of ConfP, the extended projection whose specifier is occupied by the doubling element *qué*. Data on the position of *acaso* in (84) shows where the answer was extracted from in each case:

- (84) a. Qué llegó Juan acaso, a Oviedo hoy?
 what arrived3sg Juan *acaso* to Oviedo today
 ‘Where/when did Juan arrive? Oviedo today?’
 [ForceP [WhP Qué_j llegó Juan [ConfP t_j acaso t_i]], [AnsP [VP_i a Oviedo hoy?]]]
- b. Qué llegó Juan acaso hoy, a Oviedo?
 what arrived3sg Juan *acaso* today to Oviedo
 ‘Where did Juan arrive today? Oviedo today?’
 [ForceP [WhP Qué_j llegó Juan [ConfP t_j acaso t_i] hoy], [AnsP [VP_i a Oviedo?]]]
- c. *Qué llegó Juan hoy acaso, a Oviedo?
 what arrived3sg Juan today *acaso* to Oviedo
- d. Qué llegó Juan a Oviedo acaso, hoy?
 what arrived3sg Juan to Oviedo *acaso* today
 ‘When did Juan arrive to Oviedo? Today?’
 [ForceP [WhP Qué_j llegó Juan a Oviedo [ConfP t_j acaso t_i]], [AnsP [VP_i hoy?]]]
- e. *Qué llegó Juan acaso a Oviedo, hoy?
 what arrived3sg Juan *acaso* to Oviedo today

4.6 IP split questions

I have proposed in chapter 3 of this work that the IP might also be susceptible to carry an answer feature and raise to AnsP, in the case of affirmative answers with *sí* ‘yes’. It could be hypothesized that IPs can also take that answer role in a split question and appear as its answer tag, the whole IP being coreferent of a confirmation *qué*. In fact, examples like standard Spanish in (85) seem to support that hypothesis:

(85) Qué, llegó Juan?
what arrived3sg Juan
‘Well, did Juan arrive?’

(86) Llegó Juan?
arrived3sg Juan
‘Did Juan arrive?’

The propositional meaning of this sentence presents no difference from an absolute interrogative like (86). The difference between them is that a speaker uttering (85) would not be interested in alternative answers, while (86) leaves that possibility open. Note that (87) is a felicitous answer to (86) in most circumstances, but it is only felicitous as an answer to (85) if there is a known relationship between both events (for instance, Raúl is arriving to cover Juan’s shift and therefore Juan is not coming). The speaker of (85) does not care whether someone else arrived or any other valid answers to (86) beyond yes or no.

- (87) No, llegó Raúl.
no, arrived Raúl
'No, Raúl arrived'

In sum, the difference between (86) and its split counterpart (85) is the difference between a question and the same question expressed as in (88).

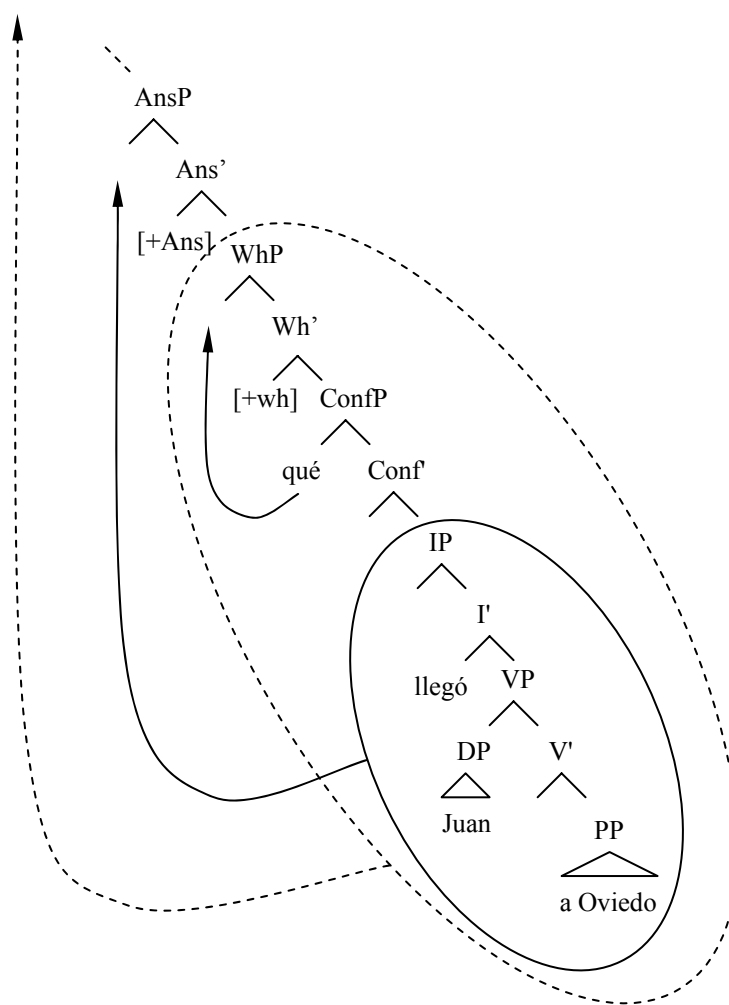
- (88) Is it the case that Juan arrived?

A question like (88) is not about who arrived, or about what Juan did, but about the event of Juan arriving³⁹. This seems to support the idea that split intonation marks answerhood (whose relationship to focalization was discussed in chapter 3) of the element on the right. By marking that answerhood, the speaker determines the set of felicitous answers. The only felicitous answers will be those compatible with the same informational structure implied by the proposed answer expressed in the tag. In a sense the tag, i.e. the answer, is what the question is about, and other interpretations of the intentions of the question are excluded.

³⁹ That is, of course, the reason this type of sentences is used as a way to spell out statements in propositional logic.

The syntactic analysis of these structures as split questions is straightforward: everything is included in Answer Phrase except for the *qué* in ConfP, which undergoes remnant movement alone, being the only constituent left in the WhP that moves past AnsP.

- (89) Qué, llegó Juan a Oviedo?
 what arrived-3sg Juan to Oviedo?
 ‘Well, did Juan arrive to Oviedo?’



4.7 Elements to the left of the interrogative

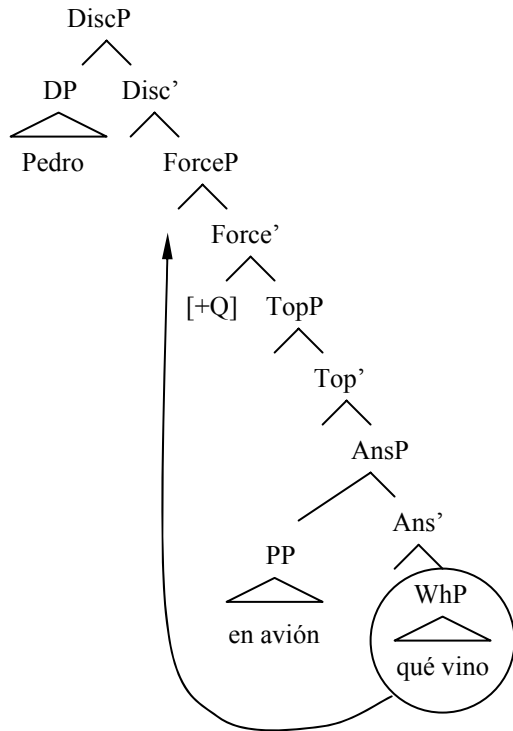
Elements on the left of the interrogative in a split question, as shown in (90), are found in several Spanish dialects, among them Asturian Spanish and Chilean Spanish, which I will discuss here.

- (90) Pedro qué vino, en avión?
Pedro what came-3sg in airplane
'Now, about Pedro, how did he come? by plane?'

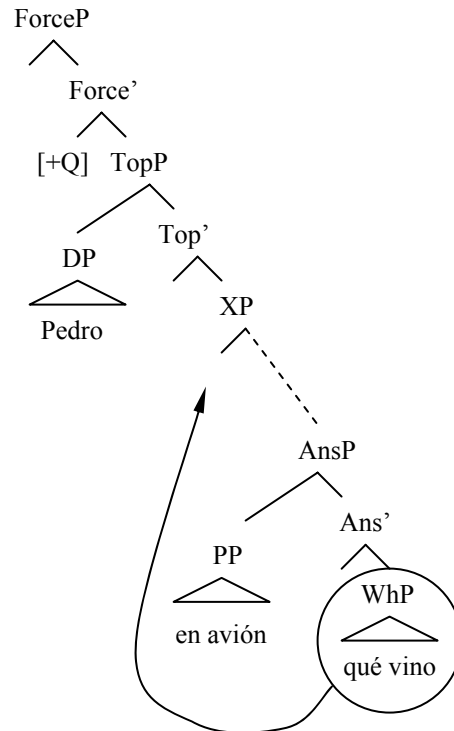
The structure proposed in (50b) includes one phrase on the far left, Discourse Phrase, which was proposed for Italian by Benincà (2001). This projection has the role of hosting Hanging Topics, thematic elements that are clearly differentiated from other left dislocated elements. These elements are outside of ForceP, the projection that includes information about clause typing. In this section, I will show how this can not be the location of the elements to the left of *qué*, and provide an alternate analysis. For ease of reference, I provide a graphic here including the relevant projections and showing the two possible analyses of (90). The tree in (91a) shows an analysis where the leftmost element is in DiscP, and the one in (91b) places that element in TopP, requiring an additional landing site for the movement motivated by the [+Q] feature.

(91)

a.



b.



If the analysis in (91a) is correct, the elements on the left should have the characteristics of Hanging Topics. There are available descriptions of Hanging Topics for Italian (Benincà, 2001), where they have to be DPs and can not be iterated, but are otherwise not distinguishable from Topics, and Catalan (Villalba, 2000), where they

have to be NPs and can not be iterated. A list of tests is provided by Benincà and Poletto (2004) to differentiate Hanging Topics (HT) on one side, and elements placed there as result of the operation of Left Dislocation (LD), on the other. Some of the tests were pointed out by Cinque (1983) and Benincà herself (Benincà, Salvi, & Frison, 1988). The tests are six:

- (92)
1. LD elements maintain the preposition of the internal elements they correspond to. HTs can only be DPs.
 2. Only a single HT position per clause is available. More than one LD can appear.
 3. LDs only require a resumptive element when they correspond to direct or partitive objects. HTs always require one.
 4. LDs' resumptive elements have to be clitics. The ones for HTs can also be tonic pronouns or epithets.
 5. HTs appear before the complementizer *che* in subordinate clauses. LDs appear after the complementizer.
 6. HTs are not possible in relative clauses, neither before nor after the relative pronoun. LDs are possible after the pronoun.

For our purposes here, tests involving embedded clauses (5 and 6) are irrelevant, since split questions cannot appear embedded. Elements on the left in a split question can be iterated (93a) (test 2). Additionally, they can appear in different orders, proving that they are not part of a single constituent (93b).

- (93) a. Pedro ayer qué vino, en avión?
 Pedro yesterday what came3sg in airplane
 ‘How did Pedro come yesterday? by plane?’
- b. Ayer Pedro qué vino, en avión?
 yesterday Pedro what came3sg in airplane
 ‘How did Pedro come yesterday? by plane?’

Elements on the left do not seem to be limited to NPs (94a); they can include a preposition (94a) (test 1) and appear without an internal coreferent (test 3) (94a):

- (94) a. Con Pedro qué habló, Pablo?
 with Pedro what spoke3sg Pablo
- b. Ayer qué vino, en avión?
 yesterday what came3sg in airplane
 ‘How did he come yesterday? by plane?’

Test 4 calls for the availability of non-clitic resumptive elements for HTs. This does not seem to be a possibility with split questions, as shown in (95b), whose grammatical equivalent with a clitic is shown in (95a).

- (95) a. Pedro qué lo trajo, Pablo?
 Pedro what cl brought-3sg Pablo
 ‘As for Pedro, who brought him? Was it Pablo?’
- b. *Pedro qué trajo a ese, Pablo?
 Pedro what brought-3sg to that Pablo
 ‘As for Pedro, who brought that guy? Was it Pablo?’

In summary, from the tests used to differentiate HTs from LDs, we can conclude that elements on the left of the *qué* in split questions are LDs. Two tests do not apply (5 and 6), and four (1, 2, 3, 4) support the LD analysis.

Additional evidence against the placing of elements to the left of ForceP would be the presence of distinct enunciative elements in Chilean Spanish. The particle *y* can appear preceding a sentence without adding any meaning to it, simply indicating the interrogative character of the sentence that follows (Campos, 1986)⁴⁰. The examples in (96) are taken from Campos (1986, pp. 188-189), who borrows (96a) from Rohlfs (1970).

- (96) a. Y no bebas?
 'y' not drink-2sg
 'Aren't you drinking?'
 b. Y/*e hiciste la tarea
 'y'/and did-2sg the homework
 'Did you do your homework?'

This enunciative element is homophonous with the conjunction *y* 'and' in most cases, but there is no doubt that it is not the conjunction because, besides not adding any meaning to the sentence, it is not affected by the rule that changes the preposition to *e* when it precedes a word beginning with the same sound [i] (Campos, 1986).

⁴⁰ A similar structure, involving the homophonous conjunction *y* 'and', is available in Peninsular Spanish; such constructions do add meaning to the sentence beyond merely marking its interrogative character.

When there are elements to the left of a split question in Chilean Spanish, the particle *y* appears (if it does) to the left of those elements.

- (97) *Y Pedro ayer qué vino, en avión?*
 y Pedro yesterday what came_{3sg} in airplane
 ‘How did Pedro come yesterday? by plane?’

If the particle *y*, as it appears, is occupying a position in ForceP, it may very well be the case that WhP does not raise to all the way to ForceP, but rather stays in a lower position in the CP layer, below the thematic layer that hosts left dislocated topics.

This analysis would be compatible with some proposals for wh-elements that do not carry the whole weight of marking the interrogative character of the sentence. Concretely, Grohmann (2006) proposes that wh-elements in German multiple wh constructions might involve wh-movement to Spec, TopP in a structure that provides several such positions. The XP of (91b) would then be an additional TopP. There are, however, problems with this solution. In terms of agreement, there is no Topic feature associated to WhP that can agree with a feature in TopP. Also, since Topics are generally assumed to be base-generated and not a landing site for movement, they would not be associated to an EPP feature, so it is not clear what would trigger a movement to the specifier of a TopP.

Assuming, then, a structure like that of (91b), and rejecting a movement to an additional TopP, the pending issue is to clarify the details of the movement of WhP, which is carrying the [+Q] feature. It has been determined that it does not raise all the way to ForceP. Perhaps it is simply not necessary for it to move to that node, as long as there is no intervening element preventing long-distance feature-checking. Recall that it has been proposed that agreement with ForceP can be achieved by movement of the [+Q] feature by itself, once the element bearing it has reached some position in the left periphery (see analyses and references in Bošković, 2001a; Hagstrom, 1998). Topics would not be intervening elements. The WhP bearing the [+Q] feature would only need to raise enough to be closer to ForceP than any other element in the Focus field. It is possible that this is a consequence of either AnsP or FocusP being a phase, and then the target of the movement would be the edge of that phase, probably the edge of the Focus field. Recall my assumption that there is a Focus layer including contrastive focus, answers, and the landing site for wh-movement. The movement of WhP would reach the periphery of this Focus layer, from where the [+Q] feature could be checked by the corresponding feature in ForceP. Further research on the phase character of the Focus layer will be necessary in order to confirm this speculation.

4.8 Loose gerunds

There are sentences, like (98a) (from López-Cortina, 1997), where the focalized element includes everything below IP (but not IP) and there is a difference in meaning with (98b), where only the subject appears dislocated. Similar constructions (like (6a), repeated here as (99a)) have been noted by Vigara Tauste (1992) and are also found in English (99b).

- (98) a. Qué está, llegando el tren?
what is, arriving the train?
'What is going on? Is the train arriving?'
- b. Qué está llegando, el tren?
what is arriving, the train?
'What is arriving? Is it the train?'
- c. *Qué está?
what is
'What is it (doing)?'
- (99) a. Qué estás ¿todavía desayunando?
What are2sg still having-breakfast
'What are you doing? Are you still having breakfast?' (surprise)
- b. What are you, sleeping?

Dislocated gerunds are common, in both Topic and Focus positions. These split questions are particular in that they have a wh-word as coreferent with the gerund. The

first part of the question is ungrammatical by itself (98c) (at least with this intended answer), but there is also no possible alternate question that would take the second part of the split question as an answer⁴¹. This would represent a problem for an analysis of split questions as a juxtaposition of two questions, since it clearly shows that the first question requires the tag to be grammatical. These examples, therefore, constitute additional evidence that the formation of split questions must be explained within a single syntactic structure including both elements (question and proposed answer tag).

4.9 Follow-up questions

The analysis of split questions presented so far in section 4 provides us with a basis to start examining other constructions involving interrogative elements associated to dislocations.

As an example of this, I will now consider Arregi's analysis of follow-up questions, another type of bipartite interrogative structure (Arregi calls them *split questions* as well. For the sake of clarity, I will continue to use my own terminology here). The goal of the section is to determine whether there is any particular element of

⁴¹ At least in Peninsular Spanish. Chilean Spanish gerunds can be part of a PP, and that PP can undergo wh-movement.

the structure of follow-up questions that requires revision of the analysis of split questions presented here, and to establish whether there is any incompatibility between the presence of split questions and follow-up questions in a given dialect. A description of follow-up questions and their differences with split questions was presented in section 2.3. I briefly summarize it here before discussing follow-up questions.

(100)

	Structure	Movement restrictions	Wh-word
Split questions	Single clause	Some dialects do not allow long distance movement. Incompatible with negation.	Generic <i>qué</i> . (Some varieties of Chilean Spanish allow for full wh-words)
Follow-up questions	Two clauses	No different from regular wh-questions.	Full wh-word matching the expected answer.

Arregi (2006) notes that follow-up question constructions can be embedded, but only if the tag presents an explicit alternative, as in (101).

(101) Juan no sabe qué compraste, si un libro o una revista. (Arregi, 2006).
 Juan not knows what bought2sg if a book or a magazine
 ‘Juan doesn’t know what you bought, whether a book or a magazine’

Note the presence of the complementizer *si* ‘whether’ in (101). This *si* is necessary for the sentence to be viable, and it would not be present in non-embedded versions of this question.

- (102) a. *Juan no sabe qué compraste, si un libro.
Juan not knows what bought2sg if a book
- b. *Juan no sabe qué compraste, un libro o una revista.
Juan not knows what bought2sg a book or a magazine
- b. *Juan no sabe qué compraste, un libro.
Juan not knows what bought2sg a book
- (103) a. *Qué compraste, si un libro o una revista?
what bought2sg if a book or a magazine
- b. *Qué compraste? Si un libro o una revista?
what bought2sg if a book or a magazine

In my view, this shows that the structure above includes two separate questions, which require separate complementizers to be embedded. Compare with the examples in (104).

- (104) a. Su madre siempre pregunta si comió, si va a llover,
his mom always asks3sg whether ate3sg whether go3sg to rain
- si lleva dinero...
whether carries3sg money
- ‘His mother is always asking whether he ate, whether it’s going to rain, whether he is carrying cash...’

- b. *Su madre siempre pregunta si comió, va a llover,
 his mom always asks3sg whether ate3sg it-goes to rain
 lleva dinero...
 carries3sg money

This does not contradict Arregi's (2006)⁴² view, since he posits the analysis in (105), where the construction is understood to consist of two juxtaposed, independent questions, a wh-question followed by a non-wh-question.

- (105) [_{CP} [Qué libro]_i leyó Juan t_i], [_{CP} ~~Juan leyó~~ [Don Quijote]_F] ?
 what book read3sg Juan Juan read3sg Don Quijote

The non-wh question undergoes ellipsis of every element except for its focus, which is assumed to be equivalent to the answer to the 'question under discussion' in the semantics of questions developed by Roberts (1998). Arregi leaves it open whether the focus of the second question (which he calls the *tag*) moves to the left periphery before ellipsis takes place (as in Hankamer, 1979; Merchant, 2004). Ellipsis is licensed by the presence of the previous question as an immediate antecedent. This analysis also leaves open, implicitly, the issue of what makes the tag a question, as opposed to a declarative with a focalized element.

⁴² Arregi refers to these constructions as 'split questions' and mentions the constructions analyzed by Lorenzo (1994) and López-Cortina (2003) to say that they are a particular type of split question he does not extend his analysis to.

Arregi shows that the data he presents is better analyzed as juxtapositions of two clauses than as single clauses, as it is in analyses like the one proposed by Camacho (2002). The key argument in his analysis is that the question on the left behaves exactly like a regular wh-question, unlike the *qué* of split questions analyzed here. The observations about the differences in meaning made in this dissertation support the idea that there should not be a common analysis for both constructions, and therefore they can be added to his arguments. These arguments, of course, support the presence of pairs of questions in the language, but do not preclude the presence of single clause split questions. In the same way, the constructions described by Camacho (2002), where *qué* is the expected wh-word, are ambiguous and could be read as single clauses or as two juxtaposed clauses. There is in principle no reason, given the data provided by other work, to take both analyses to be incompatible, even within the same dialect⁴³. The data from speakers who differentiate between two intonational contours, and meanings, supports this coexistence.

⁴³ I can get both intonations and readings myself. The existence of dialects that allow for only one of them seems to be suggested by the “either or” approach of Camacho and Arregi, although they do not explicitly reject the possibility of getting the other reading.

5 The meaning of split questions

Let us go back now to the question about the reason for the existence of the split question construction, i.e., what this construction expresses that other interrogative constructions do not. The next two subsections discuss the differences in meaning between split questions and, respectively, polar questions and clefts. The main goal of this section is to make the meaning of split questions as explicit as possible and contrast it with the meanings of similar interrogative structures. I will not attempt to devise a full semantic account of how these meanings are derived compositionally.

Split questions and polar questions

One of the differences between a split question and a regular polar question is that a regular polar question can have more than one interpretation due to the ambiguity of its information focus⁴⁴, which may be clear to the speaker, but not necessarily to the hearer. Both (106b) and (106c) are felicitous answers to (106a), since there is no external indication in the question of whether it has to be interpreted with narrow (*a Oviedo* ‘to Oviedo’) or broad (*vas a Oviedo* ‘you go to Oviedo’) information focus. The question in (107), however, presupposes that the person to whom it is addressed is

⁴⁴ Information focus is inherently ambiguous due to the phenomenon of focus projection, i.e., the possibility that focus can “extend to larger constituents than the one immediately containing the stressed word” (Chomsky, 1971).

going somewhere, and therefore can not take (106c) as an answer without further explanation, since (106c) contradicts the presupposition.

- (106) a. Vas a Oviedo?
go2sg to Oviedo
'Are you going to Oviedo?'
- b. No, a Gijón. (*a Oviedo* interpreted as information focus)
'No, to Gijón'
- c. No, me quedo en casa. (broad focus)
'No, I'm staying home'
- (107) Qué vas , a Oviedo?
what go2sg, to Oviedo?

The narrow interpretation of a split question is achieved by the singling out of the proposed answer. In section 4.3 above I proposed that this is done by means of a “confirmation morpheme”, consisting of an extended projection, the Confirmation Phrase. This ConfP contains what I called a [+conf] feature in order to capture the idea that the complement of ConfP is the part of the question that has to be confirmed or denied by a felicitous response. In a way, this ConfP, together with the [+answer] feature that causes movement to the left periphery, brings the proposed answer (the tag) into focus. It is worth noting here that the focus associated with the proposed answer is

not contrastive, left peripheral focus, but rather information focus⁴⁵. In fact, the difference between the two interpretations of (106a) can be expressed as a difference in information focus, which adds no contrastive value to the sentence (see É. Kiss, 1998 and previous chapters of this work). In split questions, the right dislocation of the proposed answer has the effect of restricting potential information focus projection to the dislocated material, the tag.

Split questions and clefts

Consider the structure in (108a). It resembles a cleft, but it is not equivalent to the English cleft in (108c). The Spanish equivalent of (108c) would be (108b). While (108b,c) have a meaning similar to (109b), the meaning of (108a) is more similar to (109a). It is probably the case that (108b) does make perfect sense in certain registers of formal/academic Spanish, but the fact of the matter is that it is extremely difficult to find in normal conversation in any dialect I am familiar with. What one does find in conversation is the structure in (108a)⁴⁶.

⁴⁵ Note also that elements carrying the [+answer] feature in sentences used as responses to questions are not necessarily associated to information focus. See chapter 3 of this dissertation for discussion of this point.

⁴⁶ Additionally, some dialects allow for constructions like *Es a Oviedo donde/que vas?* 'Is to Oviedo what you-go?' (Héctor Campos, p.c.). While these constructions show that the properties or relatives vary from one dialect to another, they still present the same meaning contrast with the cleft relevant to the discussion that follows in this section.

- (108) a. Es Oviedo donde vas?
is Oviedo where go-2sg
- b. Es a Oviedo a donde vas?
is to Oviedo to where go2sg
- c. Is it Oviedo where you are going?
- (109) a. ‘So, this Oviedo you are talking about, is it the place you are going?’
- b. ‘So, this place where you are going, is it Oviedo?’

The sentence topic in (108a) is *Oviedo*. The counterpart of that sentence where Oviedo does not assume the topic position would be (110a). This construction, however, sounds extremely awkward, and is perhaps ungrammatical.

- (110) a. ?? Es donde vas Oviedo?
is where go2sg Oviedo
‘The place where you are going, is it Oviedo?’
- b. Qué vas, a Oviedo?
what go2sg to Oviedo

Note that the meaning of this construction, were it grammatical, would be very similar to the English cleft in (108c). What we have in (110a) is a question which includes a wh-word and a constituent that is presented as the answer to the question the wh-phrase would pose. The other structure that contains those three things (question, wh-word, proposed answer), but is actually grammatical, is a split question like the one in (110b).

Towards a formalization of the semantics of split questions

The semantics of a split question seems to combine characteristics of polar questions and wh-questions. If we represent wh- and polar questions as in, respectively, (111) and (112), then we can represent split questions as in (113), taking into account the topic status (topic in the sense of “issue under discussion”) of the wh-question as presented in the previous section.

- (111) a. Quién llegó?
who arrived-3sg
‘Who arrived?’
b. ?x llegó(x)
- (112) a. Llegó Sergio?
arrived-3sg Sergio
‘Did Sergio arrive?’
b. ? [llegó (Sergio)]
- (113) a. Qué llegó, Sergio?
what arrived-3sg Sergio
‘Who arrived, was it Sergio?’
b. ? [∂ [$\exists x$ [llegó (x)]] \wedge llegó (Sergio)]

The ∂ operator in (113b) is modeled on the one discussed by Beaver (1995) within the framework of dynamic semantics. Here, however, I am leaving out the context change potentials associated to this operator in a dynamic semantics, and I am writing the meanings in a non-dynamic way. That is to say, the operator ∂ , as it is used here, is meant to require its argument to hold of the context, but it is not necessarily introducing a referent that can be picked up later as an antecedent. Basically, what this representation does is equating being out of information focus with being a sentential topic, or at least assumed to exist in the prior context. This might require some tweaking in a complete analysis, but it should suffice for the modest purposes of this section, i.e. sketching out the contrasts in meaning between these constructions as clearly as possible.

In that spirit, this representation can be used to capture the contrast between a split question and a cleft. Both structures put the same constituent in a prominent, focus-like, position. The difference between the cleft-like structure and the split question presented here is which constituent's existence is presupposed.

- (114) a. Es Oviedo donde vas?
 is Oviedo where go-2sg
 $?\partial[\exists x[x=Oviedo]] \wedge go(\text{you}, Oviedo)$
- b. Qué vas, a Oviedo?
 what go-2sg to Oviedo
 $?\partial[\exists x[go(\text{you}, x)]] \wedge go(\text{you}, Oviedo)$

6 Conclusions

The analysis of split questions presented in this chapter shows that a syntactic approach is able not only to provide an account of their structure, but also of their particular meaning and their relationship to answers.

In chapter 3 of this work it was proposed that sentences used as answers resulted of a syntactic operation involving an answer-specific functional projection. One of the predictions this approach makes is that the answer character of a phrase within a sentence should be available to be a part of the syntax of other constructions, not only those which are used as a reply in the discourse. Split questions are constructions that include an answer, even though they are not used to reply to a question. Thus, they provide additional support to the idea that the answer character of a sentence is encoded in syntax, rather than being one of the ways a regular declarative is used in the discourse.

From the point of view of the syntax of the left periphery, split questions provide a context where a syntactic explanation requires a landing site for wh-movement that is separated from the Focus projection in the left periphery, and thus these structures become supporting evidence for an interpretation of the left periphery

as a series of layers consisting of several functional projections.

In summary, split questions can be analyzed as the result of three separate syntactic operations. The first one realizes the answer character of a constituent, associated to the [+answer] feature, which raises to the left periphery of the sentence, concretely to Answer Phrase, an answer-specific functional projection, in the same way the affirmative/negative adverbs reach the left periphery in yes/no answers. The second operation is triggered by a [+wh] feature, and causes the movement of the adjunct doubled element of the answer, *qué* in Spanish, to the landing site of wh-movement, WhP, the lowest projection within the Focus layer. The third operation, an instance of remnant movement that affects the whole sentence except for the constituent previously moved to Answer Phrase, is linked to the interrogative force of the sentence, and is associated to the interrogative feature [+Q]. Its landing site is at the leftmost edge of the Focus layer, from where the [+Q] feature in ForceP can check the corresponding feature associated to the moved phrase. This movement has the interrogative part of the split question take precedence over the answer and provides the characteristic word order and the intonational split. A fourth feature, [+conf], is not linked to movement, but rather to the presence of an extended projection. It can be described as the semantic value of a sort of confirmation morpheme.

The analysis proposed here introduces a number of elements that should be able to appear independently from each other as part of other constructions. The identification and analysis of those constructions, as well as a number of other questions, are left to further research. Chapter 5 discusses some directions such research could take.

CHAPTER 5 – CONCLUSIONS AND ISSUES FOR FURTHER RESEARCH

1 General conclusions

Perhaps the main contribution of this dissertation is the concept that a sentence used as an answer has a syntactic structure specific to this function. Not every declarative sentence can be used felicitously as a reply to a question. This dissertation provides a syntax-based account of the possible structures allowed as answers to particular questions. Additionally, a syntactic approach to answers provides the grounds for the analysis of interrogative constructions which seem to include an answer for confirmation.

I have argued for the differentiation of answers from identificational focus, on the grounds of the non-contrastive character of answers and their different positions in the sentence. This differentiation not only allows us to better understand the structure of answers, but also provides a new perspective to the study of the Focus layer and the left periphery as a whole.

An additional element introduced in the discussion is the Confirmation Phrase, presented here as an extended projection of the phrase requiring confirmation. This extended projection is visible through a wh-morpheme (*qué* in Spanish). This projection

allows for some fine-tuning of the expression of the meaning of questions, distinguishing information questions from confirmation ones, and integrating this distinction in syntax.

The proposal put forward in this dissertation, in brief, is that the Focus layer in the left periphery of the sentence hosts functional projections that host separate features, of which one, [+answer], hosted in AnsP, determines the character of answer of a sentence, and another, [+wh], hosted in WhP, triggers wh-movement. These features can additionally combine with a Force feature [+Q] to produce complex interrogative structures, such as split questions. This structure gives a unified account of regular wh-questions and split questions. It also links split questions (which include answers) to the structure of sentences used as answers, by virtue of the presence of an Answer Phrase and a [+answer] feature in both structures. These elements of the analysis, together with the confirmation feature hosted in ConfP, provide tools to start a comparative analysis of the semantics of different types of questions. The comparison of the meaning of split questions with that of clefts sketched in chapter 4 is a first step in that direction.

2 Issues for further research

2.1 The VP periphery

Belletti (2004, 2005) has argued for the presence of a left periphery of the VP which replicates the periphery of CP and which contains post-verbal subjects in Italian. If my proposal is on the right track, it is worth exploring whether the VP periphery contains an Answer Phrase as well, and what effects its presence would have. It might very well be the case that the presence of an Answer Phrase in the VP periphery accounts for the distinction between sentences like (1b), with a subject answer, and others where the whole IP is the answer, like (2b).

- (1) a. Quién abrió la puerta?
who opened-3sg the door
'Who opened the door?'
- b. La abrió Susana.
it opened-3sg Susana
'Susana opened it'
- (2) a. Qué pasó con la puerta?
what happened-3sg with the door
'What happened with the door?'
- b. [_{AnsP} La abrió Susana [_{IP} ~~La abrió Susana~~]]
it opened-3sg Susana
'Susana opened it'

Besides the left periphery of IP and VP, some data seems to indicate that this type of projections could also appear higher in the structure, to the left of TopP. Thus, the presence of what appears to be a topic to the right of the confirmation particle *qué* in sentences like (3) seems to indicate that ConfP appears to the left of TopP. The the analysis presented in chapter 4 does not necessarily provide a landing site for any answers above that, unless a left-periphery style set of projections is available to the left of TopP.

- (3) a. [ForceP [XP Qué, [TopP a Juan [IP lo invitaste ya?]]]]
 what to Juan cl invited-2sg already
 ‘Well, what happened with Juan, did you invite him already?’

2.2 Phases

The notion of phase has been largely ignored in this dissertation. Both vP and CP have been argued to be phases, and this makes it necessary to pursue further research into phases and these constructions. The analysis of split questions might be useful in determining, for instance, whether the Focus layer constitutes a phase, but a reexamination of the current analysis taking phases into account should be a priority of further work on these constructions.

2.3 Question tags

One construction that presents an assertion and a question about that assertion is the question tag. It is easy to sketch an analysis of question tags using the features employed in the analysis of split questions. Question tags in sentences such as (4) seem to be a mirror image of split questions, including a non-interrogative sentence followed by what appears to be a yes/no question.

- (4) Estás escribiendo una ponencia, no?
'You are writing a presentation, aren't you?'

Sentences like this could be interpreted as a subtype of split questions where there is no wh-element. In such questions, the tag has the same answer value it has in wh-split questions, and the first part of the sentence is in fact a yes/no question. This would allow for a classification of question tags in two groups, answer tags and proper question tags, which would explain at least some of the differences in availability of tag types in Spanish and other languages like English (López-Cortina, 2006).

2.4 Constructions related to split questions

There are constructions very similar to split questions that are not straightforwardly explained by the analysis presented. An instance would be the structures that include disjunction in the tag, such as (5).

- (5) Qué vienes , el sábado o el domingo?
what come2sg, the Saturday or the Sunday
'When are you coming? on Saturday or on Sunday?'

These structures, in the dialects where they are available, are used to present either/or questions, not merely confirmation ones. In that respect, they might be an interesting ground to test theories on disjunction.

2.5 Intonation

The clear relevance of the role of intonation in the issues discussed in the last two chapters of this dissertation begs the question of why a phonetic analysis has not been included in this work.

It is true that, at this point in the research, a phonetic description of the constructions studied does seem promising as a source of additional information to improve their classification and refine the analysis. However, it must be noted that this is only the case because the analysis presented here (on the basis of the meaning and syntax of the constructions) has provided the basic guidelines of what a phonetic analysis should look at. It is dubious that the presentation of bare phonetic data could have helped in developing an analysis. Indeed, the intonation facts are so confusing that they have led some linguists to ignore them altogether and claim they are not related to the syntax of the left periphery, in spite of all the evidence supporting a relationship (evidence discussed, for instance, in Büring, 1997; Jackendoff, 1972). Thus, intonational facts have for a long time been the trees that prevented us from seeing the forest.

From this perspective, this dissertation should provide a starting point for phonetic analysis of split questions and other related constructions, since it presents an account of the left periphery that distinguishes where intonation facts are relevant to syntax and where they are linked to other phenomena, either purely phonetic or at discourse level. The phonetic description of these intonations should be a central goal of further work.

References

- Abels, K. (2001). The predicate cleft constructions in Russian. In S. Franks (Ed.), *Proceedings of FASL 9*.
- Abraham, W. (1986). Word order in the middle field of the German sentence. In W. Abraham & S. de Meij (Eds.), *Topic, focus and configurationality (Papers from the 6th Groningen Grammar Talks, Groningen, 1984)* (pp. 15-38). Amsterdam: John Benjamins.
- Alonso-Ovalle, L., & Guerzoni, E. (2004). Double Negation, Negative Concord and Metalinguistic Negation. In M. Andronis, E. Debenport, A. Pycha & K. Yoshimura (Eds.), *Proceedings of the Chicago Linguistics Society Meeting, 38* (Vol. 1 The Main Session, pp. 15-31). Chicago, Illinois: CLS Publications.
- Ammann, H. (1928). *Die menschliche Rede 2. Der Satz*. Darmstadt: Lahr.
- Arregi, K. (2003). Clitic Left Dislocation is Contrastive Topicalization. *Proceedings of the 26th Annual Penn Linguistics Colloquium, U. Penn Working Papers in Linguistics, 9*(1), 31-44.
- Arregi, K. (2006). *Syntax and Semantics of Split Questions*. Paper presented at the Linguistic Symposium on Romance Languages 36, Rutgers University.
- Bach, E. (1977). Comments on Chomsky. In P. W. Culicover, T. Wasow & A. Akmajian (Eds.), *Formal Syntax* (pp. 133-155). New York: Academic Press.
- Baker, C. L. (1970). Notes on the Description of English Questions: The Role of an Abstract Question Morpheme. *Foundations of Language, 6*, 197-219.
- Bar-Hillel, Y. (1971). Out of the pragmatic wastebasket. *Linguistic Inquiry, 2*, 401-406.
- Barbosa. (2001). On inversion in wh-questions in Romance. In A. Hulk & J.-Y. Pollock (Eds.), *Subject inversion in Romance and the theory of universal grammar*. Oxford; New York: Oxford University Press.
- Barton, E. (1991). Nonsentential Constituents and Theories of Phrase Structure. In K. Leffel & D. Bouchard (Eds.), *Views on Phrase Structure* (pp. 193-214). Dordrecht: Kluwer.
- Barton, E., & Progovac, L. (2005). Nonsententials in Minimalism. In R. Elugardo & R. J. Stainton (Eds.), *Ellipsis and non-sentential speech* (pp. 71-93). New York: Springer.
- Bartra, A. (1985). *Qüestions de la sintaxi d'ordre en català*. Unpublished doctoral dissertation, Universitat Autònoma de Barcelona.
- Beaver, D. (1995). *Presupposition and Assertion in Dynamic Semantics*. Unpublished doctoral dissertation, CCS, Edinburgh.
- Beck, S. (1996). *Wh-constructions and transparent Logical Form*. Unpublished doctoral dissertation, University of Tübingen.

- Beck, S., & Berman, S. (1996). Wh-Scope Marking: Direct vs. Indirect Dependency. In U. Lutz & G. Müller (Eds.), *Papers on Wh-Scope Marking*. Stuttgart: Universität Stuttgart.
- Beckman, M. E., Díaz Campos, M., McGory, J. T., & Morgan, T. A. (2002). Intonation across Spanish, in the Tones and Break Indices framework. *Probus*, 14(1), 9-36.
- Beckman, M. E., & Pierrehumbert, J. (1986). Intonational structure in English and Japanese. *Phonology Yearbook*, 3, 255-310.
- Belletti, A. (2003). *Extended doubling and the VP periphery*. Unpublished manuscript, University of Siena.
- Belletti, A. (2004). Aspects of the low IP area. In L. Rizzi (Ed.), *The structure of CP and IP. The Cartography of Syntactic Structures* (Vol. 2). New York: Oxford University Press.
- Belletti, A. (2005). *Answering with a "cleft": The role of the null subject parameter and the VP periphery*. Paper presented at the 30 Incontro di Grammatica Generativa, Venezia.
- Belletti, A., & Shlonsky, U. (1995). The order of verbal complements: A comparative study. *Natural Language and Linguistic Theory*, 13, 489-526.
- Bello, A. (1988). *Gramática de la lengua castellana destinada al uso de los americanos; con las notas de Rufino José Cuervo. Estudio y edición de Ramón Trujillo*. Madrid: Arco Libros.
- Benincà, P. (1995). Complement clitics in Medieval Romance: the Tobler-Mussafia Law. In A. Battye & I. Roberts (Eds.), *Clause structure and language change* (pp. 296-325). Oxford: Oxford University Press.
- Benincà, P. (2001). The position of Topic and Focus in the left periphery. In G. Cinque & G. Salvi (Eds.), *Current Studies in Italian Syntax offered to Lorenzo Renzi* (pp. 39-64). Amsterdam; New York: Elsevier.
- Benincà, P., & Poletto, C. (1999). Topic, Focus and V2: defining the CP sublayers. Certosa di Pontignano: talk delivered at the Workshop on the cartography of syntactic positions and semantic types.
- Benincà, P., & Poletto, C. (2001). Topic, Focus and V2: defining the CP sublayers. Talk at Georgetown University.
- Benincà, P., & Poletto, C. (2004). Topic, Focus and V2: defining the CP sublayers. In L. Rizzi (Ed.), *The Structure of CP and IP. The Cartography of Syntactic Structures* (Vol. 2). Oxford: Oxford University Press.
- Benincà, P., Salvi, G., & Frison, L. (1988). L'ordine delle parole e le costruzioni marcate. In L. Renzi (Ed.), *Grande grammatica italiana di consultazione* (Vol. I, pp. 129-194). Bologna: Il Mulino.
- Besten, H. d., & Webelhuth, G. (1987). *Adjunction and Remnant Topicalization in the Germanic SOV-Languages*. Paper presented at the GLOW, Venice.

- Besten, H. d., & Webelhuth, G. (1990). Stranding. In G. Grewendorf & W. Sternefeld (Eds.), *Scrambling and Barriers* (pp. 77-92). Amsterdam/Philadelphia: John Benjamins.
- Birner, B., & Ward, G. (1998). *Information status and noncanonical word order in English*. Amsterdam/Philadelphia: John Benjamins.
- Bolinger, D. (1958). A theory of pitch accent in English. *Word*, 14, 109-149.
- Bolinger, D. (1965). *Forms of English: Accent, Morpheme, Order* (I. Abe and T. Kanekiyo ed.). Cambridge: Harvard University Press.
- Bolinger, D. (1972). Accent is predictable (if you are a mind reader). *Language*, 48, 633-644.
- Bonet, E. (1991). *Morphology after syntax: Pronominal clitics in Romance*. Unpublished doctoral dissertation, MIT.
- Borer, H. (1995). The ups and downs of Hebrew verb movement. *Natural Language and Linguistic Theory*, 13, 527-606.
- Boškovič, Ž. (2001a). On the interpretation of multiple questions. *Linguistic Variation Yearbook*, 1, 1-15.
- Boškovič, Ž. (2001b). *On the Nature of the Syntax-Phonology Interface*: Elsevier.
- Bromberger, S. (1992). What we don't know when we don't know why. In S. Bromberger (Ed.), *On what we know we don't know* (pp. 145-169). Chicago: The University of Chicago Press.
- Brunetti, L. (2003). *A Unification of Focus*. Unpublished doctoral dissertation, Università degli Studi di Firenze.
- Büring, D. (1997). *The meaning of topic and focus: the 59th Street Bridge accent*. London/New York: Routledge.
- Büring, D. (forthcoming). Semantics, Intonation and Information Structure. In G. Ramchand & C. Reiss (Eds.), *Handbook on Interface Research in Linguistics*. Oxford: Oxford University Press.
- Camacho, J. (2002). Wh-Doubling: Implications for the Syntax of Wh-Movement. *Linguistic Inquiry*, 33(1), 157-164.
- Camacho, J. (2003). *The coarse structure of the center periphery*. Unpublished manuscript, Rutgers University.
- Campos, H. (1986). *Inflectional elements in Romance*. Unpublished doctoral dissertation, UCLA.
- Campos, H. (1989). Clitic Position in Modern Gallegan. *Lingua*, 77(1), 13-36.
- Campos, H. (1991). Preposition Stranding in Spanish? *Linguistic Inquiry*, 22(4), 741-750.
- Campos, H. (1992). Los bearneses que quequean, ¿y nosotros, qué? *Hispanic Linguistics*, 4(2), 329-349.
- Campos, H. (1993). *De la oración simple a la oración compuesta: Curso superior de gramática española*. Washington, D.C.: Georgetown University Press.

- Campos, H., & Zampini, M. (1990). Focalization Strategies in Spanish. *Probus*, 2(1), 47-64.
- Canellada, M. J. (1944). *El bable de Cabranes*. Madrid: RFE.
- Cardinaletti, A. (2004). Towards a Cartography of Subject Positions. In L. Rizzi (Ed.), *The structure of CP and IP. The Cartography of Syntactic Structures* (Vol. 2, pp. 115-165). New York: Oxford University Press.
- Carnap, R. (1937). *The Logical Syntax of Language*: Kegan Paul.
- Cecchetto, C. (1999). A Comparative Analysis of Left and Right Dislocation in Romance. *Studia Linguistica*, 53, 40-67.
- Cecchetto, C., & Chierchia, G. (1998). Reconstruction in dislocation constructions and the syntax/semantics interface. In S. Blake, E. Kim & K. Shanin (Eds.), *WCCFL XVII Proceedings*. Stanford: CSLI Publications.
- Chafe, W. (1976). Givenness, contrastiveness, definiteness, subjects, topics, and point of view. In C. Li (Ed.), *Subject and topic* (pp. 25-55). New York: Academic Press.
- Cheng, L. (1991). *On the typology of Wh questions*. Unpublished doctoral dissertation, MIT.
- Choi, H.-W. (1996). *Optimizing structure in context: scrambling and information structure*. Unpublished doctoral dissertation, Stanford University.
- Chomsky, N. (1957). *Syntactic Structures*. The Hague: Mouton.
- Chomsky, N. (1965). *Aspects of the Theory of Syntax*. Cambridge: MIT Press.
- Chomsky, N. (1971). Deep structure, surface structure, and semantic interpretation. In D. Steinberg & L. Jakobovits (Eds.), *Semantics: An Interdisciplinary Reader* (pp. 183-216). Cambridge: Cambridge University Press.
- Chomsky, N. (1976). Conditions on rules of Grammar. *Linguistic Analysis*, 2, 303-351.
- Chomsky, N. (1977a). *Essays on Form and Interpretation*. Amsterdam: Elsevier North-Holland.
- Chomsky, N. (1977b). On wh-movement. In P. W. Culicover, T. Wasow & A. Akmajian (Eds.), *Formal Syntax* (pp. 71-132). New York: Academic Press.
- Chomsky, N. (1981). *Lectures on Government and Binding: The Pisa Lectures*. Holland: Foris Publications.
- Chomsky, N. (1986). *Barriers*. Cambridge, MA and London: MIT Press.
- Chomsky, N. (1993). A minimalist program for linguistic theory. In K. Hale & S. J. Keyser (Eds.), *The view from building 20*. Cambridge, Mass.: MIT Press.
- Chomsky, N. (1994). Bare Phrase Structure. *MIT Occasional Papers in Linguistics*, 5.
- Chomsky, N. (1995). *The Minimalist Program*. Cambridge, Mass.: MIT Press.
- Chomsky, N. (1998). Minimalist inquiries: the framework. *MIT Working Papers in Linguistics*, 15.
- Chomsky, N. (1999). Derivation by phase. *MIT Working Papers in Linguistics*, 18.
- Chomsky, N. (2005). *On Phases*. Unpublished manuscript, MIT.

- Chomsky, N., & Halle, M. (1968). *The Sound Pattern of English*. New York: Harper and Row.
- Chomsky, N., & Lasnik, H. (1977). Filters and Control. *Linguistic Inquiry*, 8(3), 425-504.
- Cinque, G. (1977). The movement nature of left dislocation. *Linguistic Inquiry*, 8, 397-411.
- Cinque, G. (1983). 'Topic' constructions in some European languages and 'Connectedness'. In K. Ehlich & H. v. Riemsdijk (Eds.), *Connectedness in Sentence, Discourse and Text*. Tilburg: KBU.
- Cinque, G. (1990). *Types of A'-dependencies*. Cambridge, Mass.: MIT Press.
- Cinque, G. (1993). A null theory of phrase and compound stress. *Linguistic Inquiry*, 24, 239-267.
- Cinque, G. (Ed.). (2002). *Functional Structure in DP and IP: The Cartography of Syntactic Structures, Vol. 1*. Oxford: Oxford University Press.
- Conlin, F., Hagstrom, P., & Neidle, C. (2003). A Particle of Indefiniteness in ASL. *Linguistic Discovery*, 2(1).
- Contreras, J. M., & Roca, F. (in press). Un tipus especial d'oracions interrogatives: les interrogatives escindides. *Caplletra*.
- Cotton, E. G., & Sharp, J. M. (1988). *Spanish in the Americas*. Washington, D.C.: Georgetown University Press.
- Dahl, Ö. (1974). Topic-Comment structure revisited. In Ö. Dahl (Ed.), *Topic and comment, contextual boundedness and focus (Papers in Text Linguistics, 6)*. Hamburg: Helmut Buske.
- Dahl, Ö. (1976). What is new information? In N. E. Enkvist & V. Kohonen (Eds.), *Report on text linguistics: Approaches to word order*. Turku: Abo Akademi.
- de Cuba, C. (2007). *On (Non)Factivity, Clausal Complementation and the CP-Field*. Unpublished doctoral dissertation, Stony Brook University.
- Diesing, M. (1997). Yiddish DP order and the typology of object movement in Germanic. *Natural Language and Linguistic Theory*, 15, 369-427.
- Dolci, R. (1986). *Algunas construcciones con anteposición de constituyentes oracionales en español: su determinación y análisis sintáctico*. Unpublished doctoral dissertation, Università degli Studi di Venezia.
- Domínguez, L. (2004). *Mapping Focus: The Syntax and Prosody of Focus in Spanish*. Unpublished doctoral dissertation, Boston University.
- Donati, C. (2000). A note on negation in comparison. *Quaderni del Dipartimento di Linguistica*, 10, 55-68.
- Donati, C., & Nespor, M. (2003). From focus to syntax. *Lingua*, 113(11), 1119-1142.
- Dretske, F. I. (1972). Contrastive statements. *Philosophical Review*, 81, 411-437.
- Dretske, F. I. (1981). *Knowledge and the flow of information*. Cambridge: MIT Press.
- É. Kiss, K. (1998). Identificational focus versus information focus. *Language*, 74, 245-273.

- Engdahl, E. (1985). Parasitic gaps, resumptive pronouns, and subject extractions. *Linguistics*, 23, 3-44.
- Erdmann, P. (1990). *Discourse and Grammar. Focussing and defocussing in English*. Max Niemeyer Verlag.
- Erteschik-Shir, N. (1973). *On the nature of island constraints*. Unpublished doctoral dissertation, MIT.
- Erteschik-Shir, N. (1986). Wh-questions and focus. *Linguistics and Philosophy*, 9, 117-149.
- Erteschik-Shir, N. (1997). *The dynamics of focus structure*. Cambridge: Cambridge University Press.
- Erteschik-Shir, N. (to appear). On the architecture of topic and focus. In V. Molnar & S. Winkler (Eds.), *The architecture of topic and focus*. Berlin: Mouton.
- Escobar, M. Á. (1995). *Lefthand satellites in Spanish*. Unpublished doctoral dissertation, Universiteit Utrecht.
- Fabra, P. (1956). *Gramàtica Catalana. Prefaci de Joan Coromines*. Barcelona: Teide.
- Face, T. L. (2001). Focus and early peak alignment in Spanish intonation. *Probus*, 13, 223-246.
- Fortin, C. (in press). Some (Not All) Nonsententials Are Only a Phase. *Lingua*.
- Frascarelli, M. (2000). *The Syntax-Phonology Interface in Focus and Topic Constructions in Italian*. Dordrecht: Kluwer.
- Frascarelli, M., & Puglielli, A. (in press). The focus system in Cushitic Languages. A comparative-typological analysis. In P. Fronzaroli (Ed.), *Proceedings of the 10th Meeting of Hamito-Semitic Linguistics*: Università di Firenze.
- Gardent, C. (1997). *Focus: Syntax and Semantics* (course notes). Aix en Provence (France): 9th European Summer School in Logic, Language and Information.
- Geilfuss-Wolfgang, J. (1995). *Über gewisse Fälle von Assoziation mit Fokus*. Unpublished doctoral dissertation, Universität Tübingen.
- Giannakidou, A. (2000). Negative...concord? *Natural Language and Linguistic Theory*, 18, 457-523.
- Grice, H. P. (1975). Logic and conversation. In P. Cole & J. Morgan (Eds.), *Syntax and Semantics, 3: Speech Acts* (pp. 41-58). New York: Academic Press.
- Grice, H. P. (1989). *Studies in the Way of Words*: Harvard University Press.
- Grimshaw, J. (1991). *Extended Projection*. Unpublished manuscript, Brandeis University.
- Grimshaw, J. (2005). *Words and structure*. Stanford, California: CLSI Publications.
- Groenendijk, J., & Stokhof, M. (1982). Semantic analysis of wh-complements. *Linguistics and Philosophy*, 5(2), 175-233.
- Grohmann, K. K. (2006). Top issues in questions. In L. L.-S. Cheng & N. Corver (Eds.), *Wh-Movement. Moving On* (pp. 249-288). Cambridge, Mass.: MIT Press.

- Grosz, B., & Ziv, Y. (1998). Centering, Global Focus, and Right Dislocation. In M. Walker, A. Joshi & E. Prince (Eds.), *Centering in Discourse* (pp. 39-51). Oxford: Oxford University Press.
- Haegeman, L. (1995). *The syntax of negation*. Cambridge: Cambridge University Press.
- Hagstrom, P. (1998). *Decomposing questions*. Unpublished doctoral dissertation, MIT.
- Hagstrom, P. (2003). What questions mean. *Glott International*, 7(7/8), 188-201.
- Hajicová, E. (1976). Question and Answer in Linguistics and in Man-Machine Communication. *Statistical Methods in Linguistics*, 30–46.
- Hajicová, E. (1983). On Some Aspects of Presuppositions of Questions. In F. Kiefer (Ed.), *Questions and Answers* (pp. 85-96). Dordrecht: Reidel.
- Hajicová, E. (1984). Topic and Focus. In P. E. Sgall (Ed.), *Contributions to Functional Syntax. Semantics and Language Comprehension (LLSEE, 16)* (pp. 189-202). Amsterdam / Philadelphia: John Benjamins.
- Halle, M., & Vergnaud, J.-R. (1987). *An essay on stress*. Cambridge, Mass.: MIT Press.
- Halliday, M. A. K. (1967). Notes on transitivity and theme in English, part II. *Journal of Linguistics*, 3, 199-244.
- Hamblin, C. L. (1958). Questions. *Australasian Journal of Philosophy*, 36, 159-168.
- Hamblin, C. L. (1973). Questions in Montague English. *Foundations of Language*, 10, 41-53.
- Hankamer, J. (1979). *Deletion in coordinate structures*. New York: Garland.
- Hayes, B. (1989). The prosodic hierarchy in meter. In P. Kiparsky & G. Youmans (Eds.), *Rhythm and meter* (pp. 201-260). Orlando: Academic Press.
- Hayes, B. (1995). *Metrical Stress Theory*. Chicago: The University of Chicago Press.
- Hedberg, N. (2002, Dec. 6-8). *Topic-Focus Controversies*. Paper presented at the Informationstruktur - kontrastivt, Lund, Sweden.
- Hedberg, N., & Sosa, J. M. (2001). *The Prosodic Structure of Topic and Focus in Spontaneous English Dialogue*. Paper presented at the Topic and Focus: A Workshop on Intonation and Meaning, University of California, Santa Barbara.
- Heim, I. (1983). File Change Semantics and the familiarity theory of definiteness. In R. Bäuerle, C. Schwarze & A. v. Stechow (Eds.), *Meaning, use, and interpretation of language* (pp. 164-189). Berlin/New York: de Gruyter.
- Hernanz, M. L., & Brucart, J. M. (1987). *La Sintaxis*. Barcelona: Crítica.
- Hiraiwa, K. *Dimensions of Agreement*. Unpublished manuscript.
- Hockett, C. F. (1958). *A course in modern linguistics*. New York: Macmillan.
- Householder, F. (1957). Accent, juncture, intonation, and my grandfather's reader. *Word*, 13, 234-245.
- Huang, C.-T. J. (1982). *Logical relations in Chinese and the theory of grammar*. Unpublished doctoral dissertation, MIT.
- Inkelas, S., & Zec, D. (1995). Syntax-phonology interface. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 535-549). Cambridge, Mass. and Oxford: Blackwell.

- Inkelas, S., & Zec, D. (Eds.). (1990). *The phonology-syntax connection*. Chicago: University of Chicago Press.
- Jackendoff, R. S. (1972). *Semantic interpretation in generative grammar*. Cambridge, Mass.: MIT Press.
- Jacobs, J. (1986). The syntax of focus and adverbials in German. In W. Abraham & S. de Meij (Eds.), *Topic, focus, and configurationality* (pp. 103-127). Amsterdam: John Benjamins.
- Jacobs, J. (1988). Fokus-Hintergrund-Gliederung und Grammatik. In H. Altmann (Ed.), *Intonationsforschungen* (pp. 89-134). Tübingen: Niemeyer.
- Jaeggli, O. (1982). *Topics on Romance Syntax*. Dordrecht: Foris.
- Jayaseelan, K. A. (2001). IP-internal Topic and Focus phrases. *Studia Linguistica*, 55(1), 39-75.
- Jespersen, O. (1924). *The philosophy of grammar*. New York: H. Holt and company.
- Jiménez, M. L. (1997). *Semantic and pragmatic conditions on word order in Spanish*. Unpublished doctoral dissertation, Georgetown University, Washington, DC.
- Johnson, K. (1997). A Review of The Antisymmetry of Syntax. *Lingua: International Review of General Linguistics (Lingua)*, 102(1)(May), 21-53.
- Kadmon, N. (2001). *Formal Pragmatics*. Malden, Mass. / Oxford: Blackwell.
- Kaga, H. (1997). Suuryoushi to bubun-hitei. In *Shiji to Shouou to Hitei, Kenkyuu-sha* (pp. 92-178). Tokyo.
- Karttunen, L. (1977). Syntax and semantics of questions. *Linguistics and Philosophy*, 1, 3-44.
- Kato, Y. (1985). Negative sentences in Japanese. *Sophia Linguistica*, 19.
- Kayne, R. S. (1975). *French Syntax*. Cambridge: MIT Press.
- Kayne, R. S. (1994). *The Antisymmetry of Syntax*. Cambridge, Mass.: MIT Press.
- Kenesei, I. (1986). On the logic of Hungarian word order. In W. Abraham & S. de Meij (Eds.), *Topic, focus, and configurationality* (pp. 143-159). Amsterdam: John Benjamins.
- Kiparsky, P., & Kiparsky, C. (1971). Fact. In D. D. Steinberg & L. A. Jakobovits (Eds.), *Semantics. An interdisciplinary reader in philosophy, linguistics and psychology*. Cambridge: Cambridge University Press.
- Koopman, H. (1982). Control from COMP and comparative syntax. *Linguistic Review*, 2, 365-391.
- Koopman, H. (1984). *The syntax of verbs. From verb movement rules in the Kru languages to Universal Grammar*. Dordrecht: Foris.
- Krifka, M. (1992). A framework for focus-sensitive quantification. In C. Barker & D. Dowty (Eds.), *Proceedings of the Second Conference on Semantics and Linguistic Theory* (pp. 215-336). Columbus: Ohio State University.
- Krifka, M. (2001-2002). *Topik und Fokus* (seminar handouts). Humboldt-Universität Berlin.

- Laca, B. (1986). Notes per a un estudi del pleonasme pronominal en català. In *Estudis de llengua i literatura catalanes XIII (Miscel·lània Antoni Badia i Margarit)* (Vol. 5, pp. 65-88). Barcelona: Publicacions de l'Abadia de Montserrat.
- Ladd, D. R. (1980). *The structure of intonational meaning: Evidence from English*. Indiana University Press.
- Laka, I. (1990). *Negation in syntax: On the nature of functional categories and projections*. Unpublished doctoral dissertation, MIT, Cambridge, Mass.
- Lambrecht, K. (1994). *Information structure and sentence form. Topic, focus, and the mental representations of discourse referents*. Cambridge, UK: Cambridge University Press.
- Landau, I. (2004). *Chain Resolution in Hebrew V(P)-fronting*. Unpublished manuscript, Ben-Gurion.
- Larson, R. (1988). On the Double Object Construction. *Linguistic Inquiry*, 19(3), 335-391.
- Larsson, E. (1979). *La dislocation en français: étude de syntaxe générative* (Vol. 28). Lund: CWK Gleerup.
- Lasnik, H., & Saito, M. (1992). *Move alpha: Conditions on Its Application and Output*. Cambridge, Mass.: MIT Press.
- Lasnik, H., & Stowell, T. (1991). Weakest cross-over. *Linguistic Inquiry*, 22, 687-720.
- Lee, F. (2001). Wh- and Focus are not the same projection. *Snippets*, 3, 10-12.
- Lenerz, J. (1977). *Zur Abfolge nominaler Satzglieder im Deutschen* (Vol. 5). Tübingen: Narr.
- Liberman, M. (1975). *The intonational system of English*. Unpublished doctoral dissertation, MIT.
- Liberman, M., & Prince, A. (1977). On stress and linguistic rhythm. *Linguistic Inquiry*, 8(2), 249-336.
- Lope Blanch, J. M. (1972). *Estudios sobre el español de México*. México: Universidad Nacional Autónoma de México.
- López-Cortina, J. (1997). *Acerca de algunas construcciones interrogativas del español*. Paper presented at the 47th MIFLC, East Carolina University.
- López-Cortina, J. (2003). The Structure of Split Interrogatives. In P. Kempchinsky & C.-E. Piñeros (Eds.), *Theory, Practice, and Acquisition* (pp. 140-155). Somerville, Mass.: Cascadilla Press.
- López-Cortina, J. (2006). *Question tags, split questions, and remnant movement*. Paper presented at the 2006 GRAPHSY, Georgetown University, Washington DC.
- López-Cortina, J. (2007). *Split questions, extended projections, and dialect variation*. Paper presented at the XVII Conference on Generative Grammar, Girona.
- López, L. (1999). The Syntax of Contrastive Focus: Evidence from VP-Ellipsis. In J. Gutiérrez-Rexach & F. Martínez-Gil (Eds.), *Advances in Hispanic Linguistics: Papers from the 2nd Hispanic Linguistics Symposium* (Vol. 2, pp. 412-427). Somerville, Mass.: Cascadilla Press.

- Lorenzo González, G. (1994). Qué expletivo en preguntas dislocadas. *Archivum Ovetensis*, XLIV-XLV(1), 423-446.
- Lorenzo González, G. (1995). La sintaxis de los clíticos del asturiano en el contexto de las lenguas romances. Desarrollo y evaluación de algunas hipótesis. *Verba*, 22, 191-227.
- Lyons, J. (1977). *Semantics*. Cambridge: Cambridge University Press.
- Manzini, R. M., & Savoia, L. M. (1999). *Clitics: Cooccurrence and mutual exclusion patterns*. Paper presented at the Siena 'Cartography' Workshop.
- Martín, J. (2003). Against a uniform wh-landing site in Spanish. In P. Kempchinsky & C.-E. Piñeros (Eds.), *Theory, Practice, and Acquisition* (pp. 156-174). Somerville, Mass.: Cascadilla Press.
- Martorana, C. (2004). *The Structure of the Left Periphery in Sicilian*. Unpublished manuscript, Georgetown University.
- Mathesius, V. (1915). O passivu v moderní anglictine. *Sborník filologický*, 5, 198-220.
- McCloskey, J. (2006). Resumption. In M. Everaert, H. v. Riemsdijk, R. Goedemans & B. Hollebrandse (Eds.), *The Blackwell Companion to Syntax* (Vol. IV). Oxford: Blackwell.
- Merchant, J. (2001). Why no(t)? : University of Chicago.
- Merchant, J. (2004). Fragments and Ellipsis. *Linguistics and Philosophy*, 27, 661-738.
- Millán, J. A. (2005). *Perdón, imposible*. Barcelona: RBA.
- Müller, G. (1998). *Incomplete category fronting : a derivational approach to remnant movement in German*. Dordrecht, Boston: Kluwer.
- Müller, G. (2000). *Shape conservation and remnant movement*. Paper presented at the NELS 30.
- Munaro, N., Poletto, C., & Pollock, J.-Y. (2001). *Eppur si muove! On comparing French and Bellunese wh-movement*. Unpublished manuscript, Padua & Amiens.
- Muysken, P. (1995). Focus in Quechua. In K. É. Kiss (Ed.), *Discourse configurational languages* (pp. 375-393). New York, Oxford: Oxford University Press.
- Ndayiragije, J. (1999). Checking economy. *Linguistic Inquiry*, 30, 399-444.
- Neeleman, A., & Reinhart, T. (1998). Scrambling and the PF Interface. In M. Butt & W. Geuder (Eds.), *The Projection of Arguments. Lexical and Compositional Factors* (pp. 309-353). Stanford, CA: CSLI Publications.
- Neeleman, A., & Szendrői, K. (2004). Superman sentences. *Linguistic Inquiry*, 35(1), 149-159.
- Neidle, C. (2002). Language across modalities. ASL focus and question constructions. *Linguistic Variation Yearbook*, 2, 69-96.
- Neidle, C., Kegl, J., Bahan, B., Aarons, D., & MacLaughlin, D. (1997). Rightward Wh-Movement in American Sign Language. In D. Beerman (sic), D. LeBlanc & H. v. Riemsdijk (Eds.), *Rightward Movement* (pp. 247-278). Amsterdam & Philadelphia: John Benjamins.

- Neidle, C., Kegl, J., MacLaughlin, D., Bahan, B., & Lee, R. G. (2000). *The Syntax of American Sign Language: Functional Categories and Hierarchical Structure*. Cambridge, MA: MIT Press.
- Neidle, C., & MacLaughlin, D. (in press). The Distribution of Functional Projections in ASL: Evidence from Overt Expressions of Syntactic Features. In G. Cinque (Ed.), *Functional Structure in the DP and IP: The Cartography of Syntactic Structures*. Oxford: Oxford University Press.
- Neidle, C., MacLaughlin, D., Lee, R. G., Bahan, B., & Kegl, J. (1998a). The Rightward Analysis of Wh-Movement in ASL: A Reply to Petronio and Lillo-Martin 1997. *Language*, 74(4), 819–831.
- Neidle, C., MacLaughlin, D., Lee, R. G., Bahan, B., & Kegl, J. (1998b). Wh-Questions in ASL: A Case for Rightward Movement. *American Sign Language Linguistic Research Project*, 6.
- Nespor, M., & Vogel, I. (1982). Prosodic domains of external sandhi rules. In H. v. d. Hulst & N. Smith (Eds.), *The Structure of Phonological Representations* (Vol. 1, pp. 222-255). Dordrecht: Foris.
- Nespor, M., & Vogel, I. (1986). *Prosodic Phonology*. Dordrecht: Foris.
- Partee, B. (1999). Focus, quantification, and semantic-pragmatics issues. In P. Bosch & R. v. d. Sandt (Eds.), *Focus. Linguistic, cognitive, and computational perspectives*. (pp. 213-231). Cambridge: Cambridge University Press.
- Pesetsky, D. (1987). Wh-in-situ: Movement and Unselective Binding. In E. Reuland & A. t. Meulen (Eds.), *The representation of (in)definiteness*. Cambridge, MA: MIT Press.
- Petronio, K., & Lillo-Martin, D. (1997). Wh-Movement and the Position of SpecCP: Evidence from American Sign Language. *Language*, 73(1), 18-57.
- Pierrehumbert, J. (1980). *The phonology and phonetics of English intonation*: MIT doctoral dissertation, published by the Indiana University Linguistics Club.
- Plann, S. (1982). Indirect questions in Spanish. *Linguistic Inquiry*, 13, 297-312.
- Poletto, C. (2000). *The Higher Functional Field: Evidence from Northern Italian Dialects*. New York & Oxford: Oxford University Press.
- Poletto, C. (2004). Internal and external Topic and Focus: the relation between the CP and the vP in Old Italian. Talk at Georgetown University.
- Poletto, C., & Pollock, J.-Y. (2004). On the left periphery of some Romance wh-questions. In L. Rizzi (Ed.), *The Structure of CP and IP* (Vol. 2). Oxford: Oxford University Press.
- Poletto, C., & Zanuttini, R. (2000). *Marking new information in the left periphery: the case of pa in Central Rhaeto-Romance*. Paper presented at the VI Giornata di Dialettologia, University of Padua.
- Pollock, J.-Y. (2002). Three Arguments for Remnant IP Movement in Romance. In A.-M. Di Sciullo (Ed.), *Asymmetry in Grammar, Volume 1: Syntax and Semantics* (pp. 251-277). Amsterdam/Philadelphia: John Benjamins.

- Postal, P. (1991). *An apparent French extraction anomaly*. Unpublished manuscript, IBM Research Center.
- Potsdam, E. (1997). NegP and Subjunctive Complements in English. *Linguistic Inquiry*, 28, 533-541.
- Prince, E. F. (1981). Toward a taxonomy of given-new information. In P. Cole (Ed.), *Radical Pragmatics* (pp. 223-256). New York: Academic Press.
- Prince, E. F. (1986). On the syntactic marking of presupposed open propositions. In A. Farley, P. Farley & K.-E. McCullough (Eds.), *Papers from the Parasession on Pragmatics and Grammatical Theory, 22nd Regional Meeting, Chicago Linguistic Society* (pp. 208-222).
- Puglielli, A. (1997). Focus system in Cushitic Languages. *Afroasiatica Neapolitana, 'Studi africanistici. Serie etiopica'*(6), 189-203.
- Pullum, G., & Zwicky, A. (1988). The Syntax-Phonology Interface. In F. J. Newmeyer (Ed.), *Linguistics: The Cambridge Survey* (Vol. 1, pp. 255-280). Cambridge: Cambridge University Press.
- Py, B. (1971). *La interrogación en el español hablado de Madrid*. Bruxelles: Aimav.
- Reichenbach, H. (1947). *Elements of symbolic logic*: Macmillan Co.
- Reinhart, T. (1981). Pragmatics and linguistics: an analysis of sentence topics. *Philosophica*, 27, 53-94.
- Reinhart, T. (1995). *Interface strategies*: OTS Working Papers.
- Rivero, M. L. (1978). Topicalization and Wh-Movement in Spanish. *Linguistic Inquiry*, 9, 513-517.
- Rivero, M. L. (1980). On Left-Dislocation and Topicalization in Spanish. *Linguistic Inquiry*, 11, 363-393.
- Rizzi, L. (1990). *Relativized Minimality*. Cambridge, Mass.: MIT Press.
- Rizzi, L. (1997). The fine structure of the left periphery. In L. Haegeman (Ed.), *Elements of Grammar* (pp. 281-337). Dordrecht: Kluwer.
- Rizzi, L. (2001). On the position "Int(errogative)" in the left periphery of the clause. In G. Cinque & G. Salvi (Eds.), *Current Studies in Italian Syntax offered to Lorenzo Renzi* (pp. 287-296). Amsterdam; New York: Elsevier.
- Rizzi, L. (2004a). Locality and Left Periphery. In A. Belletti (Ed.), *Structures and Beyond. The Cartography of Syntactic Structures* (Vol. 3). Oxford: Oxford University Press.
- Rizzi, L. (Ed.). (2004b). *The Structure of CP and IP. The Cartography of Syntactic Structures, Vol. 2*. New York: Oxford University Press.
- Roberts, C. (1998). *Information Structure in Discourse: Towards an Integrated Formal Theory of Pragmatics*. Unpublished manuscript, Ohio State University.
- Rochemont, M. S. (1986). *Focus in generative grammar*. Amsterdam; Philadelphia: John Benjamins.
- Rodman, R. (1977). On left dislocation. *Papers in Linguistics*, 7, 437-466.
- Rohlf, G. (1970). *Le Gascon. Études de philologie pyrénéenne*. Pau: Marrimpouey.

- Rooth, M. (1985). *Association with Focus*. Unpublished doctoral dissertation, University of Massachusetts, Amherst.
- Rooth, M. (1992). A theory of focus interpretation. *Natural Language Semantics*, 1, 75-116.
- Rooth, M. (1996). Focus. In S. Lappin (Ed.), *The handbook of contemporary semantic theory* (pp. 271-297). Oxford: Blackwell.
- Ross, J. (1967). *Conditions on variables in syntax*. Unpublished doctoral dissertation, MIT, Cambridge, Mass.
- Ross, J. (1983). *Inner islands*. Paper presented at the Tenth Meeting of the Berkeley Linguistics Society.
- Searle, J. (1989). How performatives work. *Linguistics and Philosophy*, 12, 535-558.
- Selkirk, E. (1984). *Phonology and syntax*. Cambridge: MIT Press.
- Selkirk, E. (1995). Sentence Prosody: Intonation, Stress, and Phrasing. In J. A. Goldsmith (Ed.), *The Handbook of Phonological Theory* (pp. 550-569). Oxford: Blackwell.
- Sgall, P. E., Hajicová, E., & Benesova, E. (1973). *Focus and Generative Semantics*. Democratic Republic of Germany: Scriptor Verlag.
- Sgall, P. E., Hajicová, E., & Panevová, J. (1986). *The meaning of the sentence in its semantic and pragmatic aspects*. Dordrecht: Reidel.
- Solà, J., & Gavarró, A. (2006). Subextraction in Romance interrogatives. In J. Doetjes & P. González (Eds.), *Romance Languages and Linguistic Theory 2004* (pp. 297-315). Amsterdam: John Benjamins.
- Sosa, J. M. (1999). *La entonación del español*. Madrid: Cátedra.
- Stechow, A. v. (1981). Topic, Focus and Local Relevance. In W. Klein & W. Levelt (Eds.), *Crossing the boundaries in Linguistics* (pp. 95-130): Reidel.
- Stechow, A. v. (1996). Partial Wh-Movement and Logical Form. An Introduction. In U. Lutz & G. Müller (Eds.), *Papers on Wh-scope Marking* (pp. 1-36). Stuttgart: Universität Stuttgart.
- Stechow, A. v., & Uhlmann, S. (1986). Some remarks on focus projection. In W. Abraham & S. de Meij (Eds.), *Topic, focus, and configurationality* (pp. 295-320). Amsterdam: John Benjamins.
- Steedman, M. (2000). Information structure and the syntax-phonology interface. *Linguistic Inquiry*, 31(4), 649-689.
- Stepanov, A. (2005). *The hows and whys of how and why: Unselective binding and merger of wh-adjuncts*. Paper presented at the Mayfest 2005, University of Maryland, College Park.
- Suñer, M. (1982). *Syntax and semantics of Spanish presentational sentence types*. Washington, D.C.: Georgetown University Press.
- Suñer, M. (1991). Indirect Questions and the Structure of CP: Some Consequences. In H. Campos & F. Martínez-Gil (Eds.), *Current Studies in Spanish Linguistics* (pp. 238-312). Washington, DC: Georgetown University Press.

- Suñer, M. (1993). About Indirect Questions and Semi-Questions. *Linguistics and Philosophy*, 16(1), 45-77.
- Suñer, M. (1994). V-Movement and the Licensing of Argumental Wh-Phrases in Spanish. *Natural Language and Linguistic Theory*, 12(2), 335-372.
- Suñer, M., & Yépez, M. (1988). Null definite objects in Quiteño. *Linguistic Inquiry*, 19, 511-519.
- Swingle, K. (1993). *The role of prosody in Right Node Raising*. Paper presented at the LSA meeting, Los Angeles.
- Szabolcsi, A. (1981). The semantics of topic and focus articulation. In J. Groenendijk & M. Stokhof (Eds.), *Formal methods in the study of language* (pp. 513-540). Amsterdam.
- Szabolcsi, A. (1994). All quantifiers are not equal: The case of focus. *Acta Linguistica Hungarica*, 42, 171-187.
- Szendrői, K. (2001). *Focus and the Syntax-Phonology Interface*. Unpublished doctoral dissertation, University College London.
- Szendrői, K. (2004). Focus and the interaction between syntax and pragmatics. *Lingua*, 114(3), 229-254.
- Thiers, C. (1985). *VP and Scrambling in the German Mittelfeld*. Unpublished manuscript, University of Tilburg.
- Torrego, E. (1984). On inversion in Spanish and some of its effects. *Linguistic Inquiry*, 15, 103-129.
- Tsimpli, I. M. (1994). Focusing in Modern Greek. In K. É. Kiss (Ed.), *Discourse configurational languages* (pp. 176-206). New York and Oxford: Oxford University Press.
- Umbach, C. (2001). *Contrast and contrastive topic*. Paper presented at the ESSLLI 2001 Workshop on Information Structure, Discourse Structure and Discourse Semantics.
- Uriagereka, J. (1988). *On government*. Unpublished doctoral dissertation, University of Connecticut.
- Uriagereka, J. (1995). Aspects of the Syntax of Clitic Placement in Western Romance. *Linguistic Inquiry*, 26(1), 79-123.
- Valioulis, M. (1994). Anaphora, agreement, and right dislocations in Greek. *Journal of Semantics*, 11, 55-82.
- Vallduví, E. (1990). *The informational component*. Unpublished doctoral dissertation, University of Pennsylvania. Published in 1992 with the same title. New York; Garland.
- Vallduví, E. (1995). Structural properties of information packaging in Catalan. In K. É. Kiss (Ed.), *Discourse Configurational Languages* (pp. 122-152). New York: Oxford University Press.
- Vigara Tauste, A. M. (1992). *Morfosintaxis del español coloquial*. Madrid: Gredos.

- Vilkuna, M. (1994). Discourse configurationality in Finnish. In K. É. Kiss (Ed.), *Discourse configurational languages* (pp. 244-268). New York and Oxford: Oxford University Press.
- Villalba, X. (1996). Sobre la dislocació a la dreta. *Llengua & Literatura*, 7, 209-234.
- Villalba, X. (2000). *The Syntax of Sentence Periphery*. Unpublished doctoral dissertation, Universitat Autònoma de Barcelona, Barcelona.
- Wall, R. (1972). *Introduction to mathematical linguistics*. Englewood Cliffs, New Jersey: Prentice-Hall.
- Ward, G. (1985). *The semantics and pragmatics of preposing*. Unpublished doctoral dissertation, University of Pennsylvania.
- Watanabe, A. (2003). Wh and operator constructions in Japanese. *Lingua*, 113(4-6), 519-558.
- Zampini, M. (1988). *La focalización en español*. Unpublished manuscript, Georgetown University.
- Zanuttini, R. (1990). Two types of negative markers. In *Proceedings of NELS XX* (Vol. 2, pp. 517-530). University of Massachusetts, Amherst: GLSA.
- Zanuttini, R. (1991). *Syntactic properties of sentential negation: a comparative study of Romance languages*. University of Pennsylvania, Philadelphia.
- Zanuttini, R. (1997). *Negation and clausal structure: A comparative study of Romance languages*. Oxford: Oxford University Press.
- Zec, D., & Inkelas, S. (1990). Prosodically constrained syntax. In S. Inkelas & D. Zec (Eds.), *The phonology-syntax connection* (pp. 365-378). Chicago: University of Chicago Press.
- Zubizarreta, M. L. (1998). *Prosody, Focus, and Word Order*. Cambridge, Mass.: MIT Press.