Multiple Questions in French and Hungarian

A Lexical-Functional Analysis with Special Emphasis on the Syntax-Discourse Interface

Composition du jury :

Anne Abeillé
Mary Dalrymple
Jonathan Ginzburg
Ferenc Kiefer
Jean-Marie Marandin
Dávid Szabó

Université Paris Diderot - Paris 7; CNRS (co-directrice)
University of Oxford (pré-rapporteur)
Université Paris Diderot - Paris 7
Institut de Recherche en Linguistique
Académie des Sciences de Hongrie (pré-rapporteur)
CNRS-LLF (président)
Université Eötvös Loránd de Budapest (co-directeur)
Abstract

In this thesis, I aim at providing an analysis of multiple questions in Hungarian and French, in the framework of Lexical-Functional Grammar (LFG). The analysis concentrates on the syntactic and discursive aspects and on their interface. The study of multiple questions is embedded in a more comprehensive account of discourse functions and discourse structure, which both play a crucial role in the analysis.

It is argued that syntactic structure is discourse-neutral in the sense that it does not contain designated syntactic positions for discourse functions, such as topic and focus, even in discourse-configurational languages. Discourse functions are, in turn, defined semantically and pragmatically, systematically distinguished from formal highlighting.

The two main representational levels of LFG considered are the constituent structure and the information structure. In addition, a possible discourse structure representation is also explored. The proposed i-structure architecture does not contain the discourse functions as its primitives, but builds on two fundamental properties of those: prominence and D-linkedness. Although prominence is understood as a semantic notion, prominent elements are also formally highlighted, which is formalized with the help of correspondence functions between constituent structure and information structure on the one hand, and prosodic structure and information structure on the other. The former assumes that certain syntactic positions are associated with these information structure properties.

The interpretation of multiple questions (the type of answer they license) in Hungarian is directly derived from this constituent structure—information structure correspondence. Multiple questions that expect pair-list answers contain a D-linked question word, which appears in a position associated with this property. No such question word is present in multiple questions licensing single-pair answers. Syntax is not revelatory in French in this respect.

An important advantage of this approach is that it can also account for all-focus (“neutral”) sentences, which do not necessarily contain a (narrow-)focus or a topic, but a variety of other elements that share some properties with them. It also makes explicit the difference between D-linked and non-D-linked question words in multiple questions licensing a pair-list answer.

As far as discourse structure is concerned, it is proposed to contain the information structures of the individual sentences that constitute the particular discourse. Furthermore, it also specifies the discourse relation between the individual sentences.

The thesis thus not only offers a comprehensive analysis of all multiple question types in Hungarian and French, but also makes a contribution to research into the information and discourse structure in the LFG framework and beyond.
Acknowledgements

At the beginning of my doctoral studies, I could not imagine holding my dissertation in my hand. Clearly, it would not have been possible without the invaluable contribution of many people, to whom I would like to express my deepest gratitude. These two pages are not even enough to list them all, but I will still give it a try.

First and foremost, I would like to thank my advisors, Anne Abeillé at the Université Paris 7, and Dávid Szabó at the Eötvös Loránd University of Budapest, for consenting to "co-supervise" this thesis, for their expertise, help and comments on my work at various stages. I followed Anne Abeillé's classes during my Erasmus semester at Paris 7 where I got familiarized with the basics of Lexical-Functional Grammar, the theoretical framework of this dissertation. Two years later, she accepted me as her PhD student. She taught me a great deal about French linguistics, of which I can certainly make use in the future. I will never be able to thank her enough for trusting me from the beginning, for her support, guidance, encouragement, constructive criticism throughout these four years, and for her constant availability even when I was away from Paris.

I would like to thank the other members of my committee: Mary Dalrymple, Jonathan Ginzburg, Ferenc Kiefer and Jean-Marie Marandin for reading and commenting on the thesis. Mary Dalrymple’s comments on the formalization of my analysis were particularly useful and instrumental, and helped me improve that part significantly.

Besides them, three other people’s appearance on my linguistic stage played a particularly important role in shaping my ideas. In a chronological order, the first of them was András Komlósy, the supervisor of my master’s thesis at the Theoretical Linguistics Department in Budapest, and my teacher of LFG in Hungary. I am eternally grateful for his thorough criticism and comments on any piece of work I have ever shown him, for our lengthy discussions of any problem, his kindness and encouragements. I hope that at least some of the analyses presented in this work will meet the high standards he measures linguistic explanations against.

A key moment during my doctoral studies was the 13th Doctoral Seminar in Romance Studies (Warsaw, September 2008), where I met Zsuzsanna Gécseg. She got interested in my work and persuaded me to present it at other conferences as well, which proved to be a turning point in my research. Since then, I have had the possibility of working together with Zsuzsa in the framework of a project, and we realized that linguistics is far from being our only common interest. I thank her for her valuable comments on earlier versions of my dissertation and hope that our co-operation will continue fruitfully in the future.

I was lucky enough to be able to discuss various aspects of my research topic with Jean-Marie Marandin, who provided me with the suitable literature and helped me systematize my ideas. He has always been ready to sacrifice his time for our discussions, sometimes with the help of modern telecommunication when we were at a distance. A considerable part of my analysis owes to his explanations, and at the same time he always considered seriously and was listening attentively to all my half-baked ideas.
The completion of a thesis in linguistics also necessitates some technical expertise, which I lacked especially at the beginning. Had not it been for the help of Grégoire Winterstein, Frédéric Laurens, László Kálmán and László Fejes, I could have never succeeded in making progress in \LaTeX{}. I thank them for helping me any time when I could not cope with a problem. I hope I will need their help less and less often in the future.

The meetings of the French-Hungarian contrastive project *Balaton* have been a great opportunity for me to present my analyses at earlier stages and to receive valuable feedback on them from the other participants: Catherine Fuchs, Marie-Joséphe Gouesse, Ferenc Kiefer, Zsuzsanna Gécseg and Péter Balogh.

I could not emphasize enough the importance of the stimulating environment in my research community at the Laboratoire de Linguistique Formelle in Paris. I would like to thank all the members of the “labo” for their support and help.

Special thanks go to all my fellow PhD students, among them Gabriela Bílbíie, Margot Colinet, Xiaoliang Huang, Frédéric Laurens, Jana Strnadová, Delphine Tribout, Géraldine Walther and Grégoire Winterstein. They made my stay in Paris an unforgettable experience and I consider them more as my friends than simple colleagues.

I am to the same extent happy to be part of the linguistic community at the Research Institute for Linguistics of the Hungarian Academy of Sciences. When in Budapest, I always try to be ‘there’ on Thursday evenings, something one cannot miss.

At my doctoral school in Budapest, I have learnt a lot from Giampaolo Salvi. I thank him for his constant readiness to help and availability.

For their acceptability judgements, I am indebted to Gabriela Bílbíie, Jana Strnadová, Dimitrina Aleksandrova, Elena Borisova and Nicola Lampitelli.

My time in Paris would not have been the same without my friends outside the linguistic community: Eszter György, Júlia Wärmer, Dominika Rutkowska and Katarzyna Policzkiewicz. I am grateful to Dominika Rutkowska, Nicolas Falorni, Margot Colinet, Géraldine Walther and Benoît Sagot for putting me up in Paris on several occasions.

My family has always been there for me, even at the difficult moments. I hope once I will find the adequate way of expressing my gratitude to my mother, my father and my brother, without whom I would clearly not be who I am today. I appreciate all the values they transmitted to me and I hope I can assure the same support for them, any time in our lives.

Finally, a lifetime will not be enough to thank Gerhard for everything. He helped me in many, many ways, including proofreading in the last moments. More importantly, without his presence, love, attention and care I couldn’t have completed this thesis.

Merci Köszönöm Thanks Danke Dziękuję Děkuji Muştumesec Grazie Blagodarja
List of Abbreviations

The interlinear glosses in this thesis are mostly taken from the Leipzig Glossing Rules. Morpheme-by-morpheme glosses are not provided though, since they are not relevant with respect to the analysis.

#  pragmatically anomalous, but otherwise grammatical example
*  ungrammatical example, according to native speakers’ (or my own) intuitions
?, ?? doubtless example
(* item) the word would cause a structure to be ungrammatical
*(item) the word is obligatory in the structure, otherwise it is ungrammatical

NOM  Nominative
ACC  Accusative
DAT  Dative
INSTR  Instrumetal
CAUS-FIN  Causal-Final
TRANS-FACT  Translative-Factive
INESS  Inessive
SUPERESS  Superessive
ADESS  Adessive
SUBL  Sublative
DEL  Delative
ILL  Illative
ELAT  Elative
ALLAT  Allative
ABL  Ablative
TERM  Terminative
FORM  Formative
ESS-FORM  Essive-Formal
POSS  Possessive
TEMP  Temporal
F  Focus
T  Topic
CT  Contrastive Topic
C  Complementizer
INT.C  Interrogative Complementizer
TopP  Topic Phrase
FocP  Focus Phrase
CQ  Constituent Question
IS  Information Structure
QuD  Question under Discussion
ADV  Adverb
PRT  Particle
CL  Clitic
INT.CL  Interrogative Clitic
VM  Verbal Modifier
PART  Partitive
SBJ  Subjunctive
List of Symbols and Definitions

- **Symbols**
  The following symbols are used for a specific purpose, or differently from the mainstream LFG formalism in this thesis:
  - $Q$: a grammatical feature attribute, referring to interrogativity in the lexical entries of question words, its values are $+/-$
  - $\forall$: a grammatical feature attribute, identifying universal quantifiers in the lexical entries
  - INT: an i-structure attribute, used in order to indicate interrogativity in polar interrogatives (distinguishing them from declaratives) when it is present, its value is $+$
  - $\Rightarrow$: implication ("if . . . , then . . .")
  - $\rightarrow$: refers to the immediate sister constituent on the right of a node
  - $\leftarrow$: refers to the immediate sister constituent on the left of a node

The symbols are used in Chapters (7), (8) and (9).

- **Definitions**
  The following definitions of discourse functions are adopted in Chapter (4) and referred to in the remainder of the thesis:

  **Focus**: the semantically prominent and formally highlighted element of replies, answering or reacting to some previous discourse part

  **Information focus**: the narrow focus constituent that serves as answer to questions

  **Contrastive focus**: the subtype of focus whose alternatives, to which the focus is contrasted, are explicitly present in the discourse

  **Topic**: the constituent that links the sentence to the preceding discourse by introducing a subtopic of the discourse topic. It is usually, but not always, left peripheral, denotes specific entities and is given in a certain context, but none of these is obligatory

  **Thematic shifter**: a subtype of topic, present mostly but not exclusively in narrative contexts when the sentence does not continue the previous subtopic of the discourse topic

  **Contrastive topic**: a subtype of topic, always co-occurring with a focus (its correlate), introducing a complex discourse strategy (Büring, 2003)
### List of Figures

2.1 Monoclausal multiple question word structures ........................................ 17
2.2 Biclausal multiple question word structures ....................................... 17
2.3 The results of the questionnaire ......................................................... 45
3.1 Input OT-LFG .................................................................................. 119
3.2 Hierarchy of interrogative types, (Koenig, 2004, p. 141) .................... 121
3.3 Mycock’s architecture .................................................................. 128
3.4 Constituent and prosodic structure (Mycock, 2006, p. 106) .......... 129
3.5 I-structure (Butt and King, 1996) .................................................... 130
3.6 Scope of interrogativity in Japanese .................................................... 130
3.7 The prosody of the *wh*-sequence in Hungarian .............................. 131
4.1 Vallduvi’s Information Structure .................................................... 146
4.2 F-structure: *Ona pročitala knigu.* ................................................. 150
4.3 I-Structure units (Butt and King, 1996) ............................................. 150
4.4 Discourse-tree (Büring) .................................................................... 171
4.5 What did the popstars wear? .............................................................. 171
4.6 Would Fritz buy this suit? ................................................................. 171
4.7 Yes, OK, but what did JOHN eat? ......................................................... 172
4.8 Who ate what at the party? ............................................................... 184
5.1 The structure of the Hungarian sentence ............................................ 186
6.1 Syntactic structure of non-standard constituent coordination ........ 244
6.2 Final contours in French .................................................................. 246
7.1 I-structure 1: Who ate what? ............................................................... 274
7.2 I-structure 2: Who ate what? ............................................................... 274
7.3 F-structure: Who ate what? ................................................................. 274
7.4 C-structure 1: Who ate what? ............................................................... 275
7.5 C-structure 2: Who ate what? ............................................................... 275
7.6 Dalrymple and Nikolaeva (2011)’s architecture .............................. 276
7.7 Proposed i-structure ........................................................................ 284
7.8 I-structure: Who ordered what? .......................................................... 285
7.9 I-structure: Which student read which book? ................................... 286
7.10 I-structure: Ki és mikor ment moziba? .............................................. 287
7.11 I-structure: Who fell in love with whom? ......................................... 288
7.12 A Discourse-tree ............................................................................... 292
7.13 D-tree: John had a great evening last night... .................................. 295
7.15 D-tree: What did the pupils learn? Anna learnt French. ................ 297
7.16 Est-ce que Marie a réussi au bac? -Tous mes élèves ont réussi. .... 300
7.17 F-structure: Who remembers where Mary keeps which book? 302
7.18 I-structure 1 302
7.19 I-structure 2 302
7.20 D-tree: Anna tanult a gimnáziumban oroszul és Péter franciául. 304
7.21 D-tree: Anna nem oroszul, hanem franciául tanult. 305
7.22 D-tree: Nem! Anna tanult franciául! 305
7.23 D-tree: Narration-Hocus 306
7.24 I-structure roles and discourse relations 307

8.1 C-structure of Hungarian 312
8.2 The "focus" position 1 315
8.3 The "focus" position 2 315
8.4 C-structure: Ki mit hozott a bulira? 319
8.5 C-structure: János bort hozott... 319
8.6 F-structure: Ki mit hozott a bulira? 320
8.7 F-structure: János bort hozott ... 320
8.8 D-structure: Ki mit hozott a bulira? János bort, ... 320
8.9 C-structure: Mit ki hozott a bulira? 324
8.10 C-structure: A bort János hozta ... 324
8.11 D-structure: Mit ki hozott a bulira? A bort János, ... 325
8.12 C-structure: Bort hozott János... 328
8.13 C-structure: Honnan hová utazik János? 330
8.14 C-structure: Budapestről Szegedre utazik. 330
8.15 F-structure: Honnan hová utazik János? 331
8.16 F-structure: Budapestről Szegedre utazik. 331
8.17 D-structure: Honnan hova utazik? Budapestről Szegedre utazik. 332
8.18 C-structure: Ki hívott meg moziba kit? 335
8.19 C-structure: Mari hívta meg moziba Jánost. 335
8.20 F-structure: Ki hívott meg moziba kit? 336
8.21 F-structure: Mari hívta meg moziba Jánost. 336
8.22 D-structure: Ki hívott meg moziba kit? 337
8.23 D-structure: Ki hívott meg moziba kit? 339
8.24 C-structure: Hol...lehet és szabad munkát vállalni? 341
8.25 F-structure: Hol...lehet és szabad munkát vállalni? 342
8.26 D-structure: Hol...lehet és szabad diák munkát vállalni? 342
8.27 C-structure: Ki és miért bánalmazta a sérstettet? 345
8.28 F-structure: Ki és miért bánalmazta a sérstettet? 345
8.29 D-structure: Ki és miért bánalmazta a sérstettet? 346
8.30 C-structure: JÁNOS mit hozott? 347
8.31 F-structure: JÁNOS mit hozott? 347
8.32 D-structure: Jó, de JÁNOS mit hozott? 348
8.33 F-structure: Ki NEM A HAMLETET olvasta? 349
8.34 D-structure: Ki NEM A HAMLETET olvasta? 350

9.1 C-structure: Le directeur parlera demain à Marie. 354
9.2 C-structure: Cette femme, Marie n’a pas confiance en elle. 355
9.3 C-structure: Marie ne le connaît pas, cet homme. 355
9.4 C-structure: Quel livre Jean donne-t-il à Marie? 356
9.5 C-structure: Quel groupe a visité quel monument? 359
9.6 F-structure: Quel groupe a visité quel monument? 359
9.7 F-structure: Quels exercices as-tu donné à quel groupe? 360
9.8 F-structure: Tu as donné quels exercices à quel groupe? 360
9.9 D-structure: Quel groupe a visité quel monument? 361
9.10 C-structure: Quel groupe a visité quel monument? 362
9.11 D-structure: Quel groupe a visité quel monument? La Tour Eiffel, 363
9.12 C-structure: Qui chante quoi à la télévision? 364
9.15 C-structure: En famille qui dort où? 366
9.16 F-structure: En famille qui dort où? 366
9.17 D-structure: En famille qui dort où ? Les parents ..., les enfants ... 367
9.18 C-structure: Quand est-ce que tu as vu qui? 368
9.19 F-structure: Quand est-ce que tu as vu qui? 368
9.20 D-structure: Quand est-ce que tu as vu qui ? J’ai vu Jean ce matin. 369
9.21 C-structure: Tu rencontres qui à quelle heure? 369
9.22 F-structure: Tu rencontres qui à quelle heure? 369
9.23 D-structure: Tu rencontres qui à quelle heure? 370
9.24 C-structure: Où Jean a dormi quel jour? 372
9.25 F-structure: Où Jean a dormi quel jour? 372
9.26 Où Jean a dormi quel jour ? Lundi, Jean a dormi à Rome ... 373
9.27 C-structure: Quand est-ce qu’il est arrivé avec quelle intention? ... 375
9.28 F-structure: Quand est-ce qu’il est arrivé avec quelle intention? 375
9.29 D-structure: Quand est-ce qu’il est arrivé avec quelle intention? ... 376
9.30 C-structure: La conférence a eu lieu où et quand? 378
9.31 F-structure: La conférence a eu lieu où et quand? 378
9.32 D-structure: La conférence a eu lieu où et quand? 378
9.33 C-structure: Quand et pourquoi et voit-on circuler des trains ...? 379
9.34 F-structure: Quand et pourquoi voit-on circuler des trains ...? 379
9.35 Quand et pourquoi voit-on circuler des trains ...? 379
# Contents

## I Preliminaries

### 1 Introduction

1.1 Aim of the Thesis ....................................... 3
1.2 Problems Investigated .................................. 3
1.3 Methods and Framework of the Analysis .................. 5
1.4 Results of the Analysis ................................ 7
1.5 Scope of the Thesis .................................... 8
1.6 Organization of the Thesis .............................. 8

## 2 Multiple Questions: The Empirical Domain

2.1 Introduction .............................................. 13
2.2 Questions .................................................. 14
2.2.1 Types of questions .................................... 14
2.2.2 Formal highlighting .................................... 16
2.3 Multiple Questions: Generalities ....................... 16
2.3.1 Possible syntactic structures ........................ 16
2.3.2 Possible answers ...................................... 22
2.3.3 Asymmetry in pair-list questions .................... 24
2.4 Multiple Questions in Hungarian ....................... 29
2.4.1 All question words in a left-peripheral position ... 29
2.4.2 One left-peripheral question word, the other in situ .... 40
2.4.3 Coordinated left peripheral question words .......... 43
2.4.4 Clausal coordination with ellipsis in the second conjunct ... 62
2.5 Multiple Questions in French ............................ 64
2.5.1 One left-peripheral question word, the other in situ .... 65
2.5.2 All question words in situ .......................... 70
2.5.3 Coordinated question words on the left periphery .... 72
2.5.4 Coordinated question words in situ ................ 82
2.5.5 Clausal coordination (with ellipsis in the second clause) ... 82
2.5.6 Unexpected coordination ............................ 87
2.5.7 D-linkedness in French .............................. 94
2.6 Conclusion ............................................... 95

## 3 State of the Art

3.1 Introduction .............................................. 97
3.2 Patterns to be Accounted for ............................ 98
3.2.1 The order of the question words ..................... 98
3.2.2 Difference between two groups of “multiple fronting” languages 101
3.2.3 D-linkedness in pair-list questions .................. 105
3.2.4 Lack of multiple questions .......................... 110
3.3 Previous Approaches ................................. 112
## Analysis of Multiple Questions in Hungarian

8.1 Introduction ................................................................. 309
8.2 Syntactic Structure (C-structure) ........................................ 310
8.3 Multiple Questions ......................................................... 316
   8.3.1 All question words in the preverbal domain ....................... 317
   8.3.2 One left-peripheral and one right-peripheral question word .... 331
   8.3.3 Coordinated question words ........................................ 339
8.4 Beyond Multiple Questions ............................................... 346
8.5 Summary ................................................................. 351

## Analysis of Multiple Questions in French

9.1 Introduction ................................................................. 353
9.2 Syntactic Structure (C-structure) ........................................ 353
9.3 Multiple Questions ......................................................... 357
   9.3.1 One left peripheral/\textit{in situ} question word(s) ............... 357
   9.3.2 Multiple questions with coordination ................................ 375
9.4 Summary ................................................................. 381

## Conclusion

10.1 Main Results ............................................................. 383
10.2 Future Perspectives ..................................................... 386

Bibliography ................................................................. 387
Part I

Preliminaries
Chapter 1
Introduction

1.1 Aim of the Thesis
The aim of this thesis is to offer an analysis for multiple questions in Hungarian and French in the modular framework of Lexical-Functional Grammar (LFG). The two, typologically unrelated languages form the basis of this contrastive study, showing, at the same time, the power and flexibility of such a grammar that can be extended to other languages as well. The architecture of LFG itself reflects this perspective: one of its syntactic components, the c(onstituent)-structure is supposed to be language-specific, whereas other levels of representation (f(unctional)-structure and i(nformation)-structure) are conceived as universal. The interplay of these different levels of representation makes it possible to point out the common properties of the two languages, as well as the differences between them. In other words, the focus of the analysis is the syntax−discourse interface: it argues that despite the different syntactic properties, the information and discourse structure can reveal what is fundamentally the same in multiple questions in different languages, and, in turn, in what ways multiple questions can differ within one and the same language.

Although the present thesis concentrates on two languages, it aims at discussing and analyzing all the multiple question structures they exhibit, along with the possible answers they license, which, as we will see, are not always taken into consideration in their complexity by other analyses.

1.2 Problems Investigated
The problem investigated in this thesis is the syntax and interpretive import of multiple questions in Hungarian and French. The first issue to examine is the different readings multiple questions can license: some of them are answered by a pair-list (1), others by a single pair (2).
Chapter 1. Introduction

(1)  a. Q: -Who bought what?
    b. A: -John bought a suit, Mary a skirt, and Jane a pair of earrings.

(2)  a. Q: -When and where will we meet?
    b. A: -We will meet in front of the theatre at 7pm.

Although some general tendencies can be observed (the question words in the case of pair-list questions are usually clause-mates, and coordinated question words are mostly answered with a single pair), syntactic structure does not systematically determine the answer expected to a multiple question. In Hungarian, if all question words appear in the preverbal domain, the question is answered by a pair-list. The case of the other syntactic structures (which contain one preverbal and one right-peripheral question word, or coordinated question words) is less obvious. They are usually answered by a single pair, but in some cases they can also license pair-list answers. In French, multiple questions containing clause-mate question words are ambiguous between the two readings (especially if the question words are arguments), whereas multiple questions containing coordinated question words expect usually, but not exclusively single pair answers.

The interpretive difference is argued to come from the information structures associated with the questions. The prerequisite of a pair-list answer is that one of the question words denotes a contextually determined set, which is then mapped on the set denoted by the other question word. If the question lacks such a question word, the question is answered by a single pair.

This asymmetry of the question words is referred to as D-linkedness in the literature (Pesetsky, 1987; Comorovski, 1996), and the D-linked question word is often called the Sorting key of the question (Kuno and Takami, 1993). In Hungarian, D-linked question words precede non-D-linked ones in the wh-sequence, whereas in French, it is usually not identifiable by syntactic means. Again, what encodes the interpretive difference between the question words is the information structure (and the semantic information it contains, see Dalrymple (2010)).

Another problematic issue to account for is that of coordinated question words. Coordinated items are supposed to be identical to a certain extent, they share at least the same function (Sag et al., 1985; Peterson, 2004). However, in Hungarian, question words with different functions can be coordinated (3), which has to be accounted for by the analysis.

(3)  Ki és mikor ment moziba?
    who and when go.PST cinema.ILL
    Who went to the cinema and when?

It has been argued (Lipták, 2001; Skrabalova, 2006) that the common function of the question words is focus. However, this assumption cannot account for the Hungarian data, since the coordination of two (preverbal) foci is otherwise ungrammatical.

This leads us to the following problem. The syntactic structure of Hungarian is highly dependent on the information structure. This is why Hungarian is often referred to as a discourse-configurational language (É. Kiss, 1995). Based on these ob-
servations, it has been claimed that discourse functions are encoded in the syntactic structure of these languages, and this is represented by discourse-based functional projections in the structure (TopP, FocP, etc.). However, it will be demonstrated that discourse functions cannot be exclusively associated with particular syntactic positions, even in discourse-configurational languages.

The modular architecture of LFG makes it possible to account for the discourse—syntax interface without assuming discourse-functional projections in the syntactic structure. This necessitates two basic assumptions: a discourse-neutral syntactic structure (which contains, nevertheless, prominent positions), and the discourse-semantic definition of discourse functions, which can, depending on the discourse structure and on the syntactic structure of the particular language, appear in various syntactic positions. In the present thesis, I offer such an account for Hungarian and French, arguing that the decisive factors in the discourse—syntax mapping are prominence, the above mentioned D-linkedness and the way the constituent relates to the preceding discourse and to the discourse topic.

1.3 Methods and Framework of the Analysis

The majority of the data presented in this thesis have been collected from the Frantext corpus, from the Hungarian National Corpus (Váradi, 2002), from the internet (www.liberation.fr, www.index.hu, etc.) and from spontaneous speech. The analysis is thus based on attested examples. However, due to the complexity of most corpus examples, simpler ones are used in the argumentative sections, which make the argumentations easier to follow. Nevertheless, even these simplified examples have been tested with native speakers. In addition, 30 native speakers of Hungarian (half of them with linguistic background) filled out a questionnaire, in which they had to choose the appropriate multiple question for a given situation. The situations were planned to indicate if the answer expected was single pair or pair-list. Their judgements are also taken into consideration in the analysis. The French examples have also been discussed extensively with native speakers, mostly with linguistic background. On the whole, however, there is considerable variation in the acceptability judgments of native speakers of both languages concerning certain structures. Presumably, the explanation is that some multiple questions are somewhat more naturally used in embedded clauses (4)-(5) or as echo-questions (6) than out of the blue.

(4) So tell me who brought what!
(5) I still don’t know who wrote what.
(6) What did John put WHERE?

In addition, the context in which the question is uttered can also disambiguate an otherwise ambiguous syntactic structure. This is why, in some cases we can only talk about tendencies and not about general rules. Nevertheless, the analysis also attempts to account for these cases of hesitation in the acceptability judgments of native speakers.
As pointed out above, the analysis is conducted in the framework of Lexical-Functional Grammar (LFG). The modular architecture of LFG makes it possible to dissociate discourse functions from syntactic positions. The syntactic component consists of the c-structure, representing constituency and the linear order of elements, and the f-structure, encoding grammatical functions and predicate-argument relations. At an earlier stage of the LFG research, (syntactically) discourse functions were also included in the f-structure, associated with a grammatical function via functional uncertainty (for instance, topic and subject). Later on, more arguments were presented in favour of a separate i-structure (King, 1997; Choi, 1997), containing the following sets: TOPIC, FOCUS, COMPLETIVE INFORMATION, BACKGROUND INFORMATION. As a current development, the information structure also contains semantic information (Dalrymple and Nikolaeva, 2011).

Apart from these levels of representation, others have also been proposed (a(rgument)-structure, p(rosodic)-structure and m(orphosyntactic)-structure (Butt et al., 1996)), which all have their own specific architecture and primitives. The role of prosodic structure in the highlighting of question words and its interface with the information structure is described in detail in Mycock (2006), it is thus less treated in this thesis, but referred to, when necessary.

The levels of representation that play a crucial role in the present analysis are the c-structure and the i-structure, since it is their mapping that accounts for the syntactic structure of the question and for the type of answer expected. However, the conclusion is reached that the i-structure—c-structure mapping in itself cannot account for all the phenomena observed. The exact interpretation of discourse functions is highly dependent on the discourse and not only on the single sentence they are part of. For instance, (semantically defined) foci are the prominent elements of answers, corrections or parallel structures, all of them defined with respect to the preceding context. The same goes for questions: questions containing a focus, or asked for the second time differ from those asked out of the blue. Moreover, answers containing contrastive topics (Büring, 2003) are also best analyzed as indicators of an answering strategy, with respect to a preceding question. These considerations motivate an analysis beyond the sentence-level. The present analysis thus also contributes to research into discourse-structure in the LFG framework, which has already been discussed and formalized to varying degrees (King and Zaenen, 2004; O’Connor, 2005). The representations are therefore completed with a further level: d(iscourse)-structure. In addition, the f-structures are usually also provided.

As was mentioned above, semantic information can also be integrated into the i-structure in LFG. It is thus structured according to the discourse-function sets, but the sets contain the semantic descriptions of the elements. In this thesis, the Structured Meaning Approach for the semantics of questions (Krifka, 2001) is adopted and represented in basic lambda-calculus terms. However, the analysis is compatible with other frameworks (such as Ginzburg and Sag (2000)’s Propositional Abstract Theory) and other languages of logic (as the glue approach (Dalrymple, 1999, 2001)) as well.
1.4 Results of the Analysis

This section contains a brief summary of the main results of the present analysis.

First of all, a comprehensive analysis is given for all multiple question types in the two languages in the LFG framework. However, this necessitates a new LFG account of many other factors as well.

One of these is syntactic structure. A c-structure is proposed for both Hungarian and French. Following basic LFG assumptions, the c-structure is argued to be flat, non-configurational both in Hungarian, and in French, lacking a VP constituent.

Another important point is the architecture of the i-structure and its correspondence with syntax. It is demonstrated that the set of discourse functions to accommodate is larger than Butt and King (1996)’s architecture assumes. The most problematic example is that of Hungarian all-focus ("neutral") sentences, which do not contain any focused constituent, but elements that share some properties with focus. The content of the discourse function topic (thematic shifter or contrastive topic) also depends on the type of sentence and on the discourse. Therefore, first semantico-pragmatic definitions are offered for discourse functions, and then an alternative i-structure is proposed, which builds on the common properties of these discourse functions (D-linkedness, prominence) and can thus capture the phenomena observed. The fact that Hungarian (and possibly informal/spoken French) are discourse-oriented in their syntactic structure is formalized with the help of the c-structure−i-structure correspondence: these common properties are associated with certain syntactic positions. Due to the presence of D-linkedness as an i-structure set, the structure and interpretation of multiple questions is also derived in a simple way: in Hungarian, if a question word in a multiple question appears in a syntactic position associated with the i-structure category D-linked, the question will be answered with a pair-list, otherwise with a single pair. In French, syntax is usually not revelatory in this respect.

The first steps are also made in a comprehensive formalization of discourse-structure in LFG, in order to account for the discourse relations between the individual sentences. The two basic theories that are used is Büring (2003)’s account of contrastive topics in the answers, and Asher and Lascarides (2003)’s Segmented Discourse Representation Theory.

The coordination of question words with different functions is also touched upon in this thesis. It is demonstrated that in French, they have to share the same function at the level of f-structure, and both are + PROMINENT at the level of i-structure. F-structural identity is not always fulfilled in Hungarian, but in those cases semantic restrictions apply on the coordinated terms. However, this problem awaits further investigation.

---

1This is not always the case. For instance, Mycock (2006) includes more languages in her cross-linguistic LFG analysis of multiple questions, but treats only those that license pair-list answers. She also sets aside multiple questions containing coordinated question words, and those Hungarian multiple questions in which not all question words are preverbal.
1.5 Scope of the Thesis

Even a comprehensive study of multiple questions in two languages has its own limitations. Since multiple questions are defined in this thesis as more than one information gap in the domain of the same predicate, only monoclausal structures are considered, biclausal and elliptical ones are set aside. However, they are sometimes referred to in argumentations, in comparisons with monoclausal structures.

Echo questions are not addressed in this thesis either.

Importantly, some aspects of (multiple) question formation are beyond the scope of this thesis. Since the focus is on the syntax-discourse interface, the semantic and prosodic aspects have been accorded less attention than information-structure, discourse and syntactic phenomena. However, some prosodic and semantic issues are addressed when necessary. For a more detailed account of these aspects, see Mycock (2006).

1.6 Organization of the Thesis

The thesis is divided into three parts and ten chapters. Part I discusses the preliminaries of the analysis. In Chapter (2), the empirical domain of the study is presented. First, some basic issues about multiple questions are addressed (syntactic structures, possible answers, etc.), then the multiple question types of Hungarian and French are introduced and illustrated extensively with attested or corpus examples. The aim of this chapter is more the presentation of the data in a framework-neutral way, than arguing for a particular approach or analysis.

Chapter (3) presents some cross-linguistic problems that the analyses of multiple questions have had to face so far, such as the variety of syntactic structures, the subtle differences between two groups of languages that (apparently) exhibit the same syntactic structure, and the lack of multiple questions. This chapter is not specifically about Hungarian or French, its aim is rather to provide an overview of the phenomenon investigated.

The second part of the chapter presents some previous analyses offered and points out in what ways they could not account for all the phenomena adequately. The analyses examined include those proposed in the GB/Minimalist paradigm, in Optimality Theory, and in non-derivational frameworks, such as Head-driven Phrase Structure Grammar (HPSG) and LFG. Since Mycock (2006)’s analysis is also conducted in the LFG framework, it is examined in detail, in order to point out the assumptions that are kept in this thesis, and those that are rejected.

Part II offers the background for the analysis. Chapter (4) prepares the ground for the proposed information-structure representation of the present analysis. After considering previous approaches, the basic notions of information structure are defined (topic, focus, and their subtypes). It is argued that the formal (syntactic and prosodic) aspects are to be consistently distinguished from semantic and pragmatic factors. The former are referred to as highlighting, whereas the latter are exploited in the definitions offered.
1.6. Organization of the Thesis

The second part of the chapter deals with the semantic analysis of questions, since semantic information is also encoded in the i-structure in LFG. After introducing some relevant theories, the Structured Meaning Approach (Krifka, 2001) is adopted as the framework used in this thesis.

Finally, the semantic and pragmatic properties of answers are discussed, based on Marandin et al. (2009).

Chapter (5) deals with the relevant aspects of Hungarian syntax, prosody and information structure. Its aim is to provide the necessary background about the Hungarian language. First, the basic syntactic structure of the language is considered in a framework-neutral way. Then some of the most prominent analyses proposed in the past thirty years are presented, pointing out their shortcomings, especially with respect to the treatment of neutral ("all focus") sentences. Then the basic properties of Hungarian prosody are examined. Special emphasis is laid on the discussion of discourse functions in Hungarian, applying the definitions of Chapter (4) on the Hungarian data. This section also introduces new or so far less treated data containing a preverbal focus and a preverbal question word. Finally, the relevant issues on Hungarian question formation are presented.

Chapter (6) examines French from the same perspective. First, the syntactic structure of French is considered, with special emphasis on word order. It is shown that word order can considerably differ from the canonical SVO order in French: for instance the subject can be postverbal and intersperse with the complements in the case of subject NP inversion, or clitic objects can be incorporated in the verb forms. The concept of syntactic weight (Abeillé and Godard, 2000) is also introduced, along with its role in word order phenomena and in interrogatives. Furthermore, the prosodic manifestation of discourse functions (focus, thematic shifter) is also examined, both in assertions and in questions. Finally, the basic properties of question formation are discussed.

Part III introduces and illustrates the proposed analysis. Chapter (7) introduces the analysis. First, an account is proposed in the mainstream LFG framework, i.e. following Butt and King (1996)’s i-structure architecture. Adopting this aspect of Mycock (2006)’s analysis, it is assumed that some question words are prosodically highlighted (focused, in Mycock’s terminology), but the hypothesis that all question words are focused is rejected. It is shown that there are convincing arguments in favour of the assumption that D-linked question words are to be analyzed as topics at the level of i-structure.

However, an alternative i-structure architecture is also put forth, with the aim of capturing a wider range of data. It does not contain TOPIC or FOCUS as i-structure primitives, but considers them as the intersection of the two basic information structure properties (prominent, D-linked) on the one hand, and of discourse structure on the other. This is a formalized way of claiming, for instance, that ”focus is the prominent element of answers or corrections”.

Furthermore, a possible LFG representation of d(iscourse)-structure is also provided, although its exact formalization necessitates further research. It is represented as a tree-diagram, with the information structures of the individual sen-
tences as its nodes. It also specifies the discourse relations between the individual sentences, based on Büring (2003) and Asher and Lascarides (2003).

In Chapter (8), the analysis of the Hungarian data is presented. First a c-structure is proposed for Hungarian, with the necessary i-structure annotations, representing the mapping function between c-structure and i-structure. The proposed analysis is then applied to the different types of multiple questions. At each type, the c-, the f- and the d-structure of the question (and the answer) are provided, this latter containing the information structures of the individual sentences (which, in turn, contain semantic information). The LFG representation of sentences with a co-occurring focus and question word in the preverbal domain is also considered.

Chapter (9) presents the analysis of the French data in the same way. First the c-structure is presented and motivated, followed by the LFG representation of all multiple question types in French. The correspondences between c- and d-structures account for the relationships between structure and interpretation.

In Chapter (10), a general conclusion is provided.
2.1 Introduction

The aim of this chapter is to delimit the empirical domain of the present study. First I provide a general introduction of multiple questions (their possible syntactic structures and answerhood conditions) and explain why I examine only monoclausal structures. Then I go on to the presentation of the Hungarian and French data, illustrated extensively with a large set of attested examples. The aim of this part is to identify those particularities and difficulties which are then to be accounted for in the present analysis.
Chapter 2. Multiple Questions: The Empirical Domain

2.2 Questions

A question is a discourse act, with which the speaker indicates a gap of knowledge and awaits an answer that fills this gap. A question thus differs from other discourse acts (assertions, requests and exclamations) in a number of ways. Discourse acts are to be distinguished from sentence types, which constitute a formal classification of sentences into declarative, interrogative, imperative and exlamatory, based on various criteria (verbal mood, the presence of wh-words or certain complementizers, etc.). Sentence types are associated with different semantic contents (proposition, question/propositional abstract, outcome, fact, respectively, see Ginzburg and Sag (2000). There is no one-to-one correspondence between discourse acts and sentence types. For instance, a question can be expressed by an interrogative or a declarative sentence:

(1) Who did John see? (interrogative)
(2) John is coming, isn’t he? (questioning declarative)

2.2.1 Types of questions

Questions can be classified according to the answer they expect, which can correspond to some formal properties as well. The types of questions include (based on Kálmán (2001) and Beyssade et al. (2007)):

- Polar (yes/no) questions: polar questions are questions that can be answered by No. In this respect, the term yes/no is a misnomer, since in the case of some polar questions yes does not seem to be an adequate answer (at least in certain languages). Consider the following examples:

(3) a. Q: Could you tell me the time?
   b. A: # Yes.
   c. A: No.

(4) a. Q: Aren’t you coming to the party?
   b. A: # Yes.
   c. A: No.

(3-a) is an interrogative used as a polite request. Yes is inadequate as an answer, since the asker inquires about the time and not about the answerer’s capacity of answering the question. No, on the other hand, is a possible answer, in case the answerer is not in a position of telling the time (s/he does not have a watch). In the case of (4), the positive answer in some languages differs from the equivalent of yes: doch instead of ja in German, si instead of oui in French, and de igen instead of igen in Hungarian. Polar questions have two subtypes: partial and total.

1Nevertheless, Yes. is also possible as a pragmatically marked answer.
- Partial questions: a part of the content is specifically questioned, the question contains a narrow focus

(5) Did you invite MARY? (Is it MARY that you invited?)

- Total questions: the whole content is questioned

(6) Is there a way to solve this problem?

- Alternative questions: they require a choice to be made from the alternatives present in the question (the alternatives are accented, indicated by the inverted commas)

(7) a. Q: Do you want ’tea or ’coffee?
   b. A: # Yes.
   c. A: # No.
   d. Tea./ Coffee.

Note that (7-a) can also be read as a polar interrogative (if the alternatives are not accented) which can be answered by yes or no. Here the question refers to the fact of drinking something and not to the choice between tea or coffee.

- Constituent questions: constituent questions contain a question word standing for the constituent that is questioned and conveying information about it (person, object, etc.)

(8) What did John eat?

- Multiple questions: In a multiple question, we find more than one information gap in a sentence. In syntactic terms, it means that a sentence contains more than one question word, like in the following examples:

(9) Who ate what?

(10) French:

Qui a dit quoi ?
who aux said what

Who said what?

The definition does not exclude the presence of a question word in a polar question either, which would also constitute at least two information gaps (polar question and constituent question). However, as will be shown in the following sections, such question are better analyzed as (the coordination of) two clauses and two questions.
• Echo questions: echo questions follow the structure of the sentence they reflect on, and express astonishment/anger, etc., or that the interlocutor could not hear all the first question

(11) John ate WHAT???

Echo questions are not treated in this thesis.

2.2.2 Formal highlighting

In constituent questions, question words are usually highlighted. This highlighting can happen, depending on the language, in the syntax and/or in the prosody. Syntactic highlighting means placing the question word in a salient position on the left periphery of the sentence (sentence-initial or preverbal position), which indicates, at the same time, that the sentence is interrogative. This happens in English, German, (formal) French, and Hungarian.

In polar questions, where there is no question word, questionhood is often indicated by subject-verb/auxiliary inversion (English, German), interrogative verb forms (French, see Chapter 6), interrogative particles (Japanese), or clitics (Hungarian and Russian in embedded contexts).

Prosody plays a crucial role in interrogatives in languages in which question words occupy the canonical position of the constituent they represent in the question, like in Chinese or Japanese. In this case, question words are prosodically highlighted (see Ishihara (2002); Mycock (2006)). Polar interrogatives can also be distinguished from their declarative counterparts only by prosodic means. This is the case in Hungarian or Russian root interrogatives, or in informal French (a more detailed description can be found in Chapters (5) and (6) about the individual languages).

In this thesis, I mainly consider the syntactic aspects of the formal highlighting in interrogatives. Nevertheless, I will also refer to prosodic highlighting at some points. After characterizing questions in general, in the next section I provide a brief description of multiple questions, before going on to the presentation of the Hungarian and French data.

2.3 Multiple Questions: Generalities

2.3.1 Possible syntactic structures

Sentences containing more than one question word can be mono-clausal or bi-clausal and the question words can be coordinated or not. Along these dichotomies and other factors, such as syntactic position, we can identify eight structures, represented in the following tree diagrams:2

---

2QW stands for Question word.
2.3. Multiple Questions: Generalities

Monoclausal structures

- No coordination
  - All QW extracted
    - (M-all-extr)
  - One QW extracted
    - (M-1-extr)
  - All QW *in situ*
    - (M-all-ins)
- Coordination
  - Extracted
    - (M-coord-extr)
  - *In situ*
    - (M-coord-ins)

Biclausal structures

- Coordination
  - (B-coord)
- Subordination
  - True QW
    - (B-sub-true)
  - Expletive QW
    - (B-sub-expl)

Figure 2.1: Monoclausal multiple question word structures

Figure 2.2: Biclausal multiple question word structures
Let us first illustrate all these structures one by one cross-linguistically with some examples.

- **M-all-extr**: Monoclausal structure without coordination, all question words in a left-peripheral position

  English and French do not exhibit this structure, but it is fully acceptable in Hungarian (12), Romanian (13), or in the Slavic languages (14)-(17):

  (12) **Hungarian:**
  
  Ki mit hozott a bulira?
  
  who what bring.PST the party.SUBL
  
  Who brought what to the party?

  (13) **Romanian:**

  Cine ce a scris?
  
  who what AUX written
  
  Who wrote what?

  (Laenzlinger and Soare, 2006, p. 2, 4b)

  (14) **Serbo-Croatian:**

  Ko koga/Koga ko voli?
  
  who whom/whom who loves
  
  Who loves whom?

  (Laenzlinger and Soare, 2006, p. 1, 3a)

  (15) **Polish:**

  Kto jak/Jak kto zareagował na nowosci z Kijawa?
  
  who how/how who react.PST on news from Kiev
  
  Who reacted how to the news from Kiev?

  (Laenzlinger and Soare, 2006, p. 2, 3b)

  (16) **Czech:**
2.3. Multiple Questions: Generalities

Kdo koho/Koho kdo vybral do dalšího kola?
who who.ACC/who.ACC who choose.PST to next round

Who elected whom for the next round?

(Laenzlinger and Soare, 2006, p. 2, 3c)

(17) Bulgarian:

Koj kakvo dade?
who what give.PST

Who gave what?

(Laenzlinger and Soare, 2006, p. 2, 4a)

• M-1-extr: Monoclausal structure without coordination, only one left-peripheral question word

This category includes, for instance, some Indo-European languages spoken in Western Europe, like English and French, and one type of multiple questions in Hungarian. In multiple questions, only one question word can appear on the left periphery, the others occupy the position of the corresponding argument (in situ). This is true, even if there is more than one question word in the sentence. Compare:

(18) Who said what?
(19) Who left when?
(20) Who said what when?
(21) *Who what when said?

The term left-peripheral is chosen, since question words are not always sentence-initial:

(22) French:

Jean, où l’as-tu vu ?
John, where CL.AUX you seen

Where did you see John?

(23) French:

Où, Jean, l’as-tu vu ?
where, John, CL.AUX you seen
Chapter 2. Multiple Questions: The Empirical Domain

Where did you see John?

• M-all-ins: Monoclusal structure without coordination, all question words in situ

In Chinese, Japanese and in the informal, spoken registers of French, all question words appear in situ, i.e. in the position of the corresponding constituent of the declarative sentence:

(24) Japanese:

Dare-ga nani-o tabeta no?
who-NOM what-ACC ate Q

Who ate what? (Laenzlinger and Soare, 2006, p. 1, 2a)

(25) Chinese:

John gei-le shei shenme?
John give-PERF who what

What did John give to who? (Laenzlinger and Soare, 2006, p. 1, 2b)

(26) French:

Il a donné quoi à qui ?
he AUX given what to who

What did he give to whom?

In French, the in situ position of the subject is, needless to say, sentence-initial or left-peripheral, thus it is difficult to decide (especially in a transformational analysis), whether the subject question word is fronted or it is in its base position. We will not be concerned with this question in this thesis.

• M-coord-extr: Monoclusal structure with extracted and coordinated question words

(27) When and where did John defend his thesis?

Coordinated question words usually have to share the same function, for instance adjunct (see Sag et al. (1985); Peterson (2004); Sag (2005)). However, there are remarkable counter-examples to this observation, to be illustrated in the next sections.

• M-coord-ins: Monoclusal structure with in situ coordinated question words

This structure is also typical in the spoken, informal register of French:
2.3. Multiple Questions: Generalities

(28) Le dernier concert a eu lieu où et quand?
The last concert AUX had place where and when
When and where did the last concert take place?

- B-coord: Biclausal structure with clausal coordination (and ellipsis in the second conjunct)

(29) When did Mary leave and why?
The arguments for the biclausal analysis of such structures will be presented in the next sections.

- B-sub-true: Biclausal structure with subordination, containing a question word in each clause

(30) What do you think John put where?
In (30), what bears the grammatical function object in the lower clause, but appears in the higher clause.

- B-sub-expl: Biclausal structure with subordination, containing an expletive question word in the main clause

When questions are embedded under non-interrogative predicates, in some languages, interrogativity has to be indicated at the beginning of the sentence as well, otherwise the sentence is ungrammatical. Consider the following example from Hungarian:

(31) *István azt gondolja, hogy János kinek mutatta be István that thinks that János who.DAT introduce.pst VM Marit?
    Mari.ACC
    *Steve thinks that to whom did John introduce Mary?

In embedded single questions, either the question word of the embedded question is extracted to the first clause (32), or the *wh*-expletive construction is used (33), in which an expletive *wh*-word is introduced into the main clause to indicate interrogativity (Kálmán, 2001; Mycock, 2006):

(32) Kinek gondolja Mari, hogy Anna bemutatta Jánost?
who.DAT thinks.DEF Mari that Anna introduce.pst János.Acc
To whom does Mary think that Anna introduced John?

(33) Mit gondol János, hogy István kinek mutatta be what thinks János that István who.DAT introduce.pst VM Marit?
    Mari.Acc
    What does John think, to whom did Steve introduce Mary?
In embedded multiple questions in Hungarian, however, only the *wh*-expletive construction can be used.³

(34) Mit gondol János, hogy István kit kinek mutatott what thinks János that István who ACC who DAT introduce PST be? VM

What does János think, whom did Steve introduce to whom?

The accusative case of the sentence-initial question word is arguably a default case and does not come from the matrix verb. See Horvath (1997) for a detailed analysis of the *wh*-expletive construction in a derivational (Minimalist), and Mycock (2004) in the LFG framework.

In this thesis, I deal only with monoclausal structures, in which the information gaps (represented by the question words) appear in the same clause and belong to the same predicate. In this respect, biclausal structures cannot be considered as true multiple questions and are thus not treated in this thesis, except for cases when they are used in an argumentation in order to distinguish mono- and biclausal structures.

In the next two sections, we examine two other general properties of multiple questions: the possible answers they expect, and the asymmetry between question words in certain structures.

### 2.3.2 Possible answers

It is often argued that a multiple question can be answered in two ways (Higginbotham and May, 1981). Some multiple questions license a pair-list, others a single-pair answer, some are ambiguous between the two. To decide which answer is appropriate, in certain cases only the context provides the clue, but some languages express this difference explicitly in the syntax. Let us examine such examples from French ((35) and (36)) and Romanian ((37) and (38)).

(35) French: pair-list

a. Q: -Qui est parti quand?
   who is left when
   Who left when?

³This can be explained by the fact that in multiple questions requiring a pair-list answer (see below, Section (2.3.2)), the question words (in most cases) must be clause-mates. However, if we consider questions requiring a single-pair answer, the sentence becomes more acceptable.

(i) ? Kit mondott János, hogy kibe szerett bele? who ACC said János that who ILL love PST into
   Who did John say that fell in love with whom?
b. A: -Jean est parti le matin, Marie l’après-midi et Pierre le soir.
John left in the morning, Mary in the afternoon, and Peter in the evening.

(36) French: single pair
a. Q: -Où et quand a lieu le dernier concert ?
   Where and when AUX place the last concert
   Where and when is the concert taking place?

b. A: -Le dernier concert aura lieu à Paris le 10 février.
   The last concert have.FUT place in Paris the 10 February
   The last concert is taking place in Paris on February 10th

(37) Romanian: pair-list
a. Q: -Ce ce a văzut ?
   Who what AUX seen
   Who saw what?

b. A: -Ion a văzut un elefant, Maria un tigru, iar Ioana un iepure.
   John AUX seen an elephant, Mary a tiger and Jane a rabbit.
   John saw an elephant, Mary a tiger, and Jane a rabbit.

(38) Romanian: single pair
a. Q: -Cine și ce a văzut?
   Who and what AUX seen
   Who saw what?

b. A: -Ion a văzut un elefant.
   John AUX seen an elephant
   John saw an elephant.

In French, the example in (35) asks for a pair-list answer, whereas (36) with coordination is more naturally answered with a single pair. In Romanian, it is possible to cumulate *wh*-words in sentence-initial-position (37), which licenses a pair-list answer, whereas the coordination of *wh*-words (38) expects a single-pair answer. As we will see later, the pair-list interpretation is due to a function applying between the two *wh*-words (Krifka, 2001), in which each element (or partition) of the set denoted by the first is paired up with one element denoted by the other *wh*-word. The

---

4I thank Gabriela Bilbii for these examples.
functional interpretation can also explain the observation that in a pair-list multiple question, the question words have to be clausemates (however, this generalization has to be somewhat refined with respect to the French data).

However, the above mentioned ambiguity is not assumed by all the proposed approaches. According to Higginbotham and May (1981), singular which-phrases carry a uniqueness presupposition and even in questions containing multiple singular which-phrases, the usual reading is single pair. Nevertheless, a pair-list (bijective) reading is not excluded either, in which the domains of quantification of the which-phrases are disjoint (39).

(39) Which student read which book?
(40) Who kills whom at the end of the film?

According to Higginbotham & May’s judgement, a pair-list answer to (40) would be anomalous, since the domains of quantification of both which-phrases is the set of the characters of the film.

Ginzburg and Sag (2000) argue, on the other hand, that multiple (which)-questions usually license pair-list answers and the existing single-pair answers are due to independent presuppositional factors: for instance, certain pronouns (like which) introduce a uniqueness presupposition, or verbs denoting unique events (to die, to kill). The authors usually refer to these questions as ‘quizmaster questions’:

(41) Which French emperor inspired which great Austrian composer to write which string quartet?
(42) Which criminal killed which American president?

However, the authors treat mainly which-phrases, which are special in the sense that singular ones carry a uniqueness presupposition, wheres plural ones also imply some familiarity in the context. Nevertheless, single pair readings are also possible when the question contains other wh-words, and in those cases it is difficult to relate the reading to the uniqueness presupposition introduced by the wh-word:

(43) Who is singing what on TV (at the moment)?

The single pair reading is rather related to the context and to the verb tense. Suppose that the asker is not in the room but can hear that someone is singing something on TV. The question will be answered with a single pair and not with a pair-list in this case. The idea of the pair-list/single pair answer ambiguity will thus be kept in this thesis, keeping in mind that some approaches assume only pair-list answers.

2.3.3 Asymmetry in pair-list questions

An important factor about pair-list questions is the observation that (at least) one of the question words denotes a contextually given set, which is salient in the discourse and which the locutor and the interlocutor can partition in the same way.
2.3. Multiple Questions: Generalities

(Comorovski, 1996). Comorovski (1996) calls this set D(iscourse)-linked, referring to the fact that the set has already been introduced into the discourse. The phenomenon of D-linkedness was first referred to by Bolinger (1978) and Pesetsky (1987), and later on by Hornstein (1995), when discussing Superiority violations. Superiority effects (Kuno and Robinson, 1972; Chomsky, 1973) refer to constraints on multiple constituent question formation, according to which (following the transformational terminology), wh-phrases cannot be fronted (moved) if there is another wh-phrase in the clause, which is closer to the left periphery. In other words, the potential candidate for a transformation is always the one closest to its landing site, obeying economical considerations preferring shorter movements to longer ones. This is supposed to account for the grammaticality and ungrammaticality of the following examples:

(44) Who saw what?
(45) *What did who see?

Since who is closer to the left periphery of the sentence than what, the fronting of what results in ungrammaticality. However, there are remarkable violations of this supposed superiority effect (of the subject over the object). For instance, object which-phrases can be fronted over subject which-phrases:

(46) Which book did which student read?

According to Pesetsky (1987) (also referred to by Ginzburg and Sag (2000, p. 247)) the use of certain wh-phrases requires a set of possible instantiators in the context for the argument role with which the D-linked expression is associated. Which-phrases are the prototypical D-linked interrogative expressions. In example (46), the set of books is supposed to be known and salient in the discourse, known by the discourse participants.

Other superiority violations include multiple questions with more than two wh-phrases (Bolinger, 1978):

(47) a. Who put WHAT WHERE?
b. What did WHO put WHERE?
c. Where did WHO put WHAT?

Although the questions in (47) are clear superiority violations, one should consider the fact that they are most probably uttered in echo-contexts, not out of the blue, i.e. in situations when the locutor could not hear all the necessary information about an event, or wants to express his amazement about what he has heard. After all, it seems quite improbable that someone asks a (non-echo) question about a putting event, without knowing any of the arguments of the 3-place predicate (the agent, the theme and the location).
Besides the above examples, Pesetsky (1987) also cites another one in which the set of possible instantiators of the fronted question word is contextually given, and it is not the subject:

(48) I know that we need to install transistor A, transistor B, and transistor C, and I know that these three holes are for transistors, but I’ll be damned if I can figure out from the instructions WHERE WHAT goes. (Pesetsky, 1987, p. 109)

*D-linkedness* is thus proposed to account for the order of question words in a multiple question in English: D-linked question words tend to appear on the left periphery of the sentence and to precede non-D-linked ones. In addition, as Comorovsky (1996) argues, in most contexts, the set determined by the subject *wh*-phrase is contextually given, thus subject *wh*-phrases are in most cases D-linked, and, as such, sentence-initial.

Ginzburg and Sag (2000) argue, on the other hand, that in some *which*-questions neither (none) of the sets has to be contextually determined, like in (49), thus the notion of *D-linkedness* is irrelevant in the description of multiple questions.

(49) Which recently published reports should be made required reading for which government departments? (Ginzburg and Sag, 2000, p. 248)

According to the authors, any new public official can ask (49), without knowing a defined set of reports. I believe, on the other hand, (see also Gazdik (2010b)) that without the context the issue is very difficult to judge. However, the modifiers *recently published* make it possible to establish a context and to restrict the possible reports to those published recently, which is, in a way, a contextually determined set, even if the new public official cannot list them by title. Note that without the above mentioned modifiers the sentence is less acceptable, if we suppose that the locutor does not know any of the reports (50):

(50) #Which reports should be made required reading for which government departments?

The authors also give another example in which D-linkedness does not seem to be relevant:

(51) I don’t know anything about cars. Do you have any suggestions about which car, if any, I should buy when I get a raise?

(52) I don’t know anything about cars. Do you have any suggestions about what car, if any, I should buy when I get a raise?

They argue that in neither of the above examples has the speaker in mind a range of felicitous answers and there is no possible interpretational difference between the *wh*-phrases in (51) and *whatcar*. They observe, at the same time, that the above mentioned *wh*-phrases have different properties with respect to multiple question formation: while it is completely acceptable with a *which*-phrase, it is less felicitous
in the case of a what-phrase:

(53)  Who did which president greet?
(54)  */#/ Who did what president greet?

However, this difference demonstrates that it is easier to make a partition of sets denoted by which-phrases, which presuppose the existence of this set in the discourse. What-phrases, on the other hand, do not carry such presuppositions. Needless to say, this does not mean that in (51) or in (53) the speaker has in mind the range of felicitous answers. What D-linkedness means here is that (as stated in Comorovski’s definition), but that the speaker and the hearer can partition the set in the same way. Since pair-list answers make a partition of one of the constituents in the question (most often one of the question phrases), the easiest candidate for this is which-phrases, which presuppose the existence of such a set.

Based on these facts, the notion of D-linkedness will be kept in the present analysis and formalized in Lexical-Functional Grammar, keeping in mind that it does not mean that all the elements of this set are known to the speaker and to the hearer. D-linkedness is especially relevant in the case of the Hungarian data.\(^7\)

Just like the type of answer expected to the question, D-linkedness can equally determine syntactic structure. In some languages, D-linked question words tend to precede non-D-linked ones, and similarly, the constituents corresponding to the D-linked question word precede their non-D-linked counterparts in the answer. Let us consider the following Hungarian examples (55)-(56):

(55)  a.  Q: -Ki mit hozott a bulira?
     who what bring.pst the party.subl
     Who brought what to the party?

     b.  A: -János bort, Mari pedig sütiket hozott.
     John wine.acc Mary and cookies.acc bring.pst
     John brought wine, and Mary cookies.

(56)  a.  Q: -Mit ki hozott a bulira?
     what who bring.pst the party.subl
     *What did who bring to the party?

     b.  A: -A bort János, a sütiket pedig Mari hozta.
     the wine.acc John the cookies.acc and Mary bring.pst
     The wine was brought by John, the cookies by Mary.

\(^7\)According to Jonathan Ginzburg, p.c., the relevance of the notion of D-linkedness can be questioned in the description of the English language. However, the notion of D-linkedness used here does not refer to the semantics of the question words, but to the fact that question words that function as the Sorting key of a multiple question, are considered to be D-linked. Thus, it refers to the usage of certain question words as Sorting keys, which are then thematized in the answer.

\(^8\)Both Hungarian és and pedig are translated as and into English, since English does not differentiate between the two and it is not relevant for the present purposes. In an enumeration, pedig introduces the last conjunct and expresses contrast between those conjuncts.
In example (55), the question asks for the enumeration of all the people present at the party and then about the thing each of them brought. In (56), on the other hand, all the things brought to the party are listed and then assigned to a person. This is also reflected by the different conjugation types of the verb in the answers in (55) and (56): in (55) hozott is in the indefinite, whereas hozta in (56) definite conjugation. In the English translation, apart from the passive, the definite article in the answer in (56) also expresses this change of perspective. This function of the D-linked wh-phrase is referred to by Kuno and Takami (1993), as the Sorting Key Hypothesis, where the linearly first question word determines the organization of the information in the answer. Note that É. Kiss (1993) accounts for these facts in terms of specificity, a notion subsumed by D-linkedness.

A more general observation is that pair-list answers are not always related to a D-linked question word in a multiple question, but to any element in the question that can be thematized (split into subgroups) in the answer. This is the case, for instance, when the subject is plural, and the question words are adjuncts:

(57) French
   a. Q: -Où et quand les enfants seront-ils rapatriés?
      where and when the children be.FUT.3SG-they sent home
      Where and when will the children be sent home?
   
      the boys in Paris tomorrow and the girls in Nantes the week next.
      The boys in Paris tomorrow, and the girls in Nantes next week.
   
   c. A2: -Demain à 8 heures.
      Tomorrow at 8 o’clock.
      Tomorrow at 8 am.

Nevertheless, as shown in (57-c), a single pair answer is also possible for this question. The choice of the answer depends on the treatment of the subject (the children): if it is treated is one group, the answer is single pair, if it is treated as a group consisting of subgroups, a pair-list answer is given to the question, providing an answer with respect to all the subgroups. We conclude, thus, that it is usually one of the question words that is thematized in the answer, but it is not necessarily the case. We will see other examples with a plural subject, in which the thematizing of this subject gives rise to a pair-list answer and not one of the question words.

In the next two sections, I present the Hungarian and the French data, respectively. In the examples, I do not distinguish between main and embedded clause

---

9 Transitive verbs in Hungarian agree with the definiteness of their object.
10 I thank Jean-Marie Marandin for calling my attention to this.
2.4. Multiple Questions in Hungarian

Among the above monoclausal structures, Hungarian exhibits M-all-extr, M-1-extr and M-coord-extr. We will now look at the Hungarian examples of these structures.\(^{11}\)

### 2.4.1 All question words in a left-peripheral position

Multiple questions of this structure usually license a pair-list answer, but this is not the only structure that can license this type of answer. In most cases, no element can intervene between the interrogative words, or between the rightmost interrogative word and the finite verb.

Main clause polar questions containing question words are ungrammatical in Hungarian. Main clause polar interrogatives differ from declaratives only in their intonation:

\[(58) \text{*Jön mikor Mari a bulira?/Jön Mari a bulira come.3sg when Mari the party.subl/come.3sg Mari the party.subl mikor?} \]

\*Is Mary coming to the party when?

One of the reasons for the ungrammaticality of such examples is pragmatic: one cannot ask a question about an event itself (Mary’s coming to the party) and about one aspect of it in the same sentence: in order to answer the constituent question, the answer to the polar question has to be provided. This is the reason why the co-occurrence of a polar question and a constituent question is only grammatical if a given structure is presumably biclausal.

The same goes for embedded polar questions, which are indicated by the interrogative clitic -e:

\[(59) \text{*Nem tudom, jön-e Mari a bulira mikor.} \]

\*I don’t know if Mary is coming to the party when.

Generally, all interrogative words can appear with another preverbally; however, the linearly first one is subject to some restrictions. These question words have to denote sets of individuals, which are then paired up with an element of the set denoted by the immediately preverbal question word (see \((55)\)). The denotation of such questions words is thus a set of referential and (sometimes) specific individuals, or other individualizable sets, like places or points in time.

\(^{11}\)The examples come from the Hungarian National Corpus (\textit{Magyar Nemzeti Szövegtár Váradi (2002))}, from the Internet, and from spontaneous speech.
Nevertheless, this does not exclude that both question words denote a contextually determined set. In this case, what is asked for is the matching of all the members of the first set with the members of the second:\footnote{Krifka (2001) and Comorovski (1996) refer to this type as “matching questions”. For the sake of precision, I call matching questions only that subset of pair-list questions, in which both question words are D-linked.}

(60) a. Q: \textit{Melyik csoport melyik helyi láttnivalót nézte meg?}  
Which group which local landmark.\textit{ACC see.\textit{PST VM}}  
Which group visited which local landmark?  
{the linguists, the psychologists, the physicists}, {the Eiffel-tower, the Arc de Triomphe, the Louvre}

b. A: -A nyelvészek a Louvre-t, a pszichológusok az  
the linguists the Louvre.\textit{ACC}, the psychologists the  
Eiffel-tornyot, a fizikusok pedig a Diadalivet  
Eiffel-tower.\textit{ACC}, the physicists as for the Arc de Triomphe.\textit{ACC}  
nézték meg.  
see.\textit{PST VM}  
The linguists visited the Louvre, the psychologists the Eiffel-tower  
and the physicists the Arc de Triomphe.

Nevertheless, in other examples it is enough if only the first question word denotes a contextually given set:

(61) (Nem tudom, hogy) \textit{kinek mi a baja.}  
not \textit{know.1SG, that who.DAT what the problem.Poss.3SG}  
I don’t know who has got what kind of a problem.

Interrogative words not denoting a set of individuals, such as \textit{miért} (why), or \textit{hogyan} (how) can rarely appear as linearly first question words:

(62) *\textit{Miért kit hívtál meg a bulira?}  
why who.\textit{ACC invite.\textit{PST.2SG VM} the party.SUBL}  
*Why did you invite who to the party?

(63) *\textit{Hogyan kinek mondtad el a hírt?}  
how who.\textit{DAT tell.\textit{PST.2SG VM the news.ACC}}  
*How did you tell who the news?

In the reverse order of the question words, however, the question is grammatical:

(64) \textit{Kit miért hívtál meg a bulira?}  
who.\textit{ACC why invite.\textit{PST.2SG VM the party.SUBL}}  
Who did you invite why to the party?
2.4. Multiple Questions in Hungarian

(65) **Kinek hogyan mondtad el a hírt?**
who.DAT how tell.PST.2SG VM the news.ACC
Who did you tell the news how?

This also supports the claim that the immediately preverbal question word can be D-linked as well, but this is not obligatorily so.

The above examples ((62) and (63)) are more acceptable if the first interrogative words are replaced by a periphrastic, inherently specific interrogative expression:

(66) **Milyen/melyik okból kit hívtál meg a**
which reason.ELAT who.ACC invite.PST.2SG VM the
bulira?
party.SUBL
For which reason did you invite who to the party?

(67) ?? **Melyik/milyen módon kinek mondtad el a hírt?**
which way who.DAT tell.PST.2SG VM the news.ACC
Which way did you tell the news who?

I conclude that in Hungarian multiple questions containing left-peripheral question words, the structure of the question reveals that a pair-list answer is expected and it also indicates the structure of the expected answer: D-linked question words tend to precede non-D-linked ones in the question, like the corresponding constituents in the answer.

Let us see now some other examples with the variation of the various question words starting with questions containing two arguments:

- *ki-kit* (who-who.ACC)

(68) Majd hetven nap-ja arról szól tehát a múisor, hogy almost seventeen days-for about talks thus the programme, that ki kit képes kijteni az ötmillióért who who.ACC can eliminate the five million.CAUS-FIN zajló versenyből. taking place contest.ELAT
Thus for seventeen days the programme has been about who can eliminate who from the contest for five million.

(Tamás Ungár: Egy jó családból való vállalkozó meséi, Magyar tükör, Népszabadság, 2000/10/16)

(69) A Magyarok Világszövetségében jelenleg nem the Hungarians World Alliance.INESS presently not állapítható meg, hogy ki kit képvisel a acknowledgeable VM, that who who.ACC represent the
Presently it cannot be determined who represents who at the assembly of the Alliance of the World’s Hungarians, and the exact number of the members of the alliance either.

(Újraszámolnák a világmagyarokat, Népszava, 2000/07/13)

I sat down and chocked up the newspaper in order to read who killed who - recently everyone has been killing everyone, the kidnapper the policeman, the policeman the maffioso, the black the white, the white the indonesian.

(Tibor Gyurkovics, Gyerekregények, 2001)

- **ki-mit** (who-what.ACC)

Egyebek közt az is szóba került, hogy ki mit others among that also words.ILL come.PST, that who what olvasson fel a miskolci estélyen.

Among others we talked about the question of who should read what at the evening party at Miskolc.

(Áron Tamási, Gondolat és árvaság, 2000)

- **mi-mit** (what.NOM-what.ACC)

III. Richard rimánkodása bizonyítja, helyzete válogatja,

III. Richard begging.POSS proves, situation.POSS select,

what what.ACC is worth
III. Richard’s beggings prove, it depends on the situation what is worth what.

(Anna Jókai, Perc-emberkék dáridója, 1996)

• *mi-mivel* (what-what.INSTR)

(73) **Mi miivel áll szemben?**

what what.INSTR stands opposite

What stands opposite what?

(http://www.c3.hu/szf/Szofi97/Sz97-04/Sz97-04-Area-2.htm)

• *honnan-hová* (from where-to where)

(74) Akkor hogyan tudja elsorolni fejből dátum szerint, hogy honnan hová irányítothatták?

from where to where direct.pst.3pl

Then how can you list by heart, from where to where you were directed?

(Magyar Elektronikus Könyvtár [Hungarian Electronic Library])

It is even possible to have three arguments in the preverbal field:

• *ki-honnan-hová* (who-from where-to where)

(75) Ezen a gépen lehet figyelni programokkal, hogy ki honnan hová ment.

You can observe on this machine with programmes who went where from where.

(Zsuzsa Ladányi, Csongrádi internet körkép, Magyar Rádió Hírműsorok, 2000/08/14)

In the next examples, *ki* (who) and *mi* (what) are combined with adjuncts. Observe that the order of question words clearly depends on the structure of the answer expected:

• *mi-hol* (what.nom-where)
(76) A rovatvezetők a többi napilapot olvassák, the columnists the other daily papers.ACC read, összehasonlítják, hol mi a téma. compare, where what the topic
The columnists read the other daily papers and compare which of them deals with what topic.

(Zoltán Csorba, A slapajtól a főszerkesztőig - " Csak olvasd az újságot! ", 2005/03/11)

• ki-hol (who-where)

(77) Ma méginkább a protekción dolgozza meg a jól fizető today even more the contacts decides VM the well-paying munkahelyek sorsát, mint az, hogy ki hol végzett. workplaces fate.Poss.ACC, than that, that who where finished
Today, the distribution of well-paid jobs depends even more on the contacts one has than on where one graduated from.

(google, 2009/03/18)

(78) Ha így hiszel nekem, akkor menjünk el estére if so believe.2sg me, then let's go VM in the evening együtt szórakozni Lakkival, akkor legalább lesz tanú, together have fun Lakki.Instr, then at least will be witness, hogy ki hol volt. that who where was
If you believe me so, then let's go out this evening together with Lakki, then at least there will be a witness, who can justify who was where.

(Freddy Krüger, Index Fórum, Rovat: Törzsasztal: Vérgőzfürdő, 1998/10/14)

• ki-mikor (who-when)

(79) A szerző pontosan megfogalmazta, ki mikor megy át a the author precisely said, who when go across the színen (és mit csinál). scene.Superess (and what does)
The author gave precise instructions about who should cross the scene when (and what they should do).
2.4. Multiple Questions in Hungarian


(80) Az esztergomi Suzuki személyzeti osztályvezetője arról the of Esztergom Suzuki staff department head about that számolt be, hogy a gyárban májusban készítenek speak.PST VM, that the factory.INESS May.INESS make tervet, ki mikor pihenhet.

The head of the staff department of the Suzuki factory in Esztergom reported that they would make plans in May about who can have a rest when.

(A szabadságidő szabályozása, Népszava, 1999/04/28)

In the last group of examples, *ki* (who) and *mi* (what) stand with non individualizable adjunct interrogative words. Notice that these latter can never precede the former:

• *ki-hogyan* (who-how)

(81) **Ki hogyan** oldotta meg az előző nap who how solve.PST VM the preceding day eseményeinek tálalását?

events.Poss.1sg.dat presentation.Poss.3sg.acc

Who could find which way of presenting the events of the preceding day?

(http://www.hhrf.org/magyarszo/arbiva/foszerk/csorta.htm)

(82) Nincs nagydobra verve, hogy ki hogyan próbált pénzt keresni, not public, that who how try.PST money earn, mert ez mellékfoglalkozás, így vagy úgy legális. because this side job, this way or that way legal

It is not common knowledge who tried to earn money how, because this is a side job, this way or that way legal.

(google, 2009/03/18)

(83) Ami viszont különség, az az, hogy ki hogyan távozott what however difference, that that, that who how leave.PST a pártból.

the party.ELAT

What makes a difference is that who left the party how.
Szerintem megérne egy beszélgetést, hogy ki miért érzi a fizikát érthetetlennek, amit mellesleg én why feels the physics ununderstandable, which besides I rendkívül sajnálok.

extremely regret

In my opinion, it should be discussed, who considers physics ununderstandable why, which, besides, I extremely regret.

It is unnecessary to start a debate about who is here why.

Indeed it is hard to know who doesn’t answer one’s questions why.

Of course I don’t blame them for this, but in reality I don’t understand what they do why.
Nevertheless, there is a type of multiple questions in Hungarian in which all question words precede the finite verb, but the answer expected is not a list of pairs. Consider the following examples:

(88) a. Q: -Honnan hova utazik János?
   from where to where travels János
   From where to where is János travelling?

b. A: -BUDAPESTRŐL SZEGEDRE utazik.
   Budapest.DEL Szeged.SUBL travels
   He is travelling from Budapest to Szeged.

(89) a. Q: -Mettől meddig tartott a buli?
   when.ABL when.TERM last.PST.3SG the party
   From when to when did the party last?

b. A: -ESTE 8-TŐL HAJNALI 5-IG tartott.
   evening 8-ABL dawn 5-TERM last.PST.3SG
   It lasted from 8 pm till 5 am.

(90) a. Q: -Miből mivé változott át a varázsló?
   what.ELAT what.TRANSL-FACT transform.PST VM the magician
   From what into what did the magician transform?

b. A: -BÉKÁBÓL KIRÁLYFIVÁ változott.
   frog.ELAT prince.TRANSL-FACT transformed
   He transformed into a prince from a frog.

What we see in the above examples is, however, a special type of multiple questions in that the question words refer to the starting and final point of one and the same process or event. This structure will be analyzed as a single question containing a complex question word (see Chapter 8).

Cross-linguistically the same phenomenon can be observed:

(91) French:

D’où à où voyage-t-il ?
from where to where travels-he
From where to where is he travelling?

(92) French:

De quand à quand sont les vacances de Pâques ?
from when to when are the holidays of Easter
From when to when are the Easter holidays?
In French, like in English, it is impossible to extract two question words apart from the above example:

(93) French:

*Qui où est allé?
who where is gone

*Who where went?

2.4.1.1 The problem of the preverbal position

Mycoc (2006) observes that the two question words in pair-list questions in Hungarian constitute an inseparable prosodic and syntactic unit. She points out that only the immediately preverbal interrogative word receives the characteristic pitch accent of focused constituents, the other bears a high monotone contour. Referring to Varga (2002), Mycock proposes that the question words cannot be separated by intervening material because the first one is the prosodic preparation (forward pointing) for the prominence of the second. It points forward to the preverbal pitch accent.

In some cases, however, the two question words are separated by certain elements, and these are, arguably, auxiliaries. Kálmán C. et al. (1985) examine Verb + infinitive constructions in Hungarian, and argue that only in those constructions can the finite verb be considered as an auxiliary, in which the infinitive can appear in the prominent preverbal position and the verb, deaccented, cliticizes on it even in level-prosody, neutral sentences. Compare:

(94) A vevők utálnak várákozni.
the customers hate to wait
The customers hate waiting.

(Kálmán C. et al., 1985, p. 50, 10)

(95) A vevők perelné akarnak.
the customers to sue want
The customers want to sue.

(Kálmán C. et al., 1985, p. 50, 11): 50, 11

According to this approach, akar (to want) is an auxiliary, whereas utál (to hate) is a main verb.

This phenomenon works slightly differently in the case of interrogatives. In a single question, the interrogative word occupies the preverbal position, followed by the auxiliary and the infinitive:

(96) Ki akar perelné?
who wants to sue
Who wants to sue?
In multiple questions, both preverbal question words can precede the auxiliary, but there are also cases in which the auxiliary follows the linearly first question word, but precedes the second:

(97) Konkrétumokról, például, hogy (több évtizedes késéssel) concrete things.DEL, for instance, that (more decades delay.INSTR), mennyi pénzt kéne mire fordítni how much money.ACC should what.SUBL spend (emphasis mine, (G. A.)), és hogy milyen miniszteriális and that what kind of ministerial rendeleteket kéne azonnali hatályal meghozni, nem orders.ACC should immediate validity.INSTR accept, not ejtettek szót, de hát miért is legyenek mention.PST.3SG word.ACC, but well why on earth be.IMP.3PL pápábbak a pápánál. more pope the pope.ADESS

They didn’t mention anything concrete, for instance about how much money should be spent on what (with several decades of delay), and what kind of ministerial orders should be accepted with immediate validity, but well, why should they be more like the pope than the pope himself.

(Élet és irodalom [weekly newspaper], Vol. LIII., Number 6., February 6th 2009)

According to László Kálmán (p.c.), a possible approach would be that the auxiliary always takes the second position in the verbal modifier/focus/interrogative word + Verb sequence, no matter how many elements the structure consists of. Further support comes from other interrogative data, where the prominent preverbal position is occupied by two coordinated question words:

(98) Kér dés, hogy mit akarunk és hol vacsorázni. question, that what want.1PL and where to dine The question is what we want to have for dinner and where.

(spontaneous speech)

(99) Meg 1-2 gyakorlati kérdés, hogy szerinted milyen bérletet plus 1-2 practical question, that according to you what kind of pass vegyek, vagy az ilyen fényképeset hol kell meg hogyan buy.IMP.1SG, or the such with photo.ACC where must plus how intézni. arrange And 1-2 practical questions, in your opinion, what kind of pass should I buy, or where and how can these passes with a photo be arranged?

(google, 2009/03/19)

Nevertheless, the above order is not obligatory, the auxiliary can also follow the question words and precede the infinitive:
Mit milyen szempontból lehet tanítani?
What can be taught from which point of view?

We will discuss tomorrow when and where we should meet, and then I’ll call you to discuss this.

A megbeszélést nem én tartom, úgyhogy nem tudom, hogy mit és mennyi ideig fogunk megbeszélni. It is not me heading the discussion, so I don’t know what we will discuss and for how long.

I do not pursue this issue further here (see also Chapter (5) for more discussion), the main aim having been to call the reader’s attention to this phenomenon and to show that some elements can intervene between the question words in the preverbal sequence. I now turn to another possible multiple question structure in Hungarian.

2.4.2 One left-peripheral question word, the other in situ

This multiple question type is often called ‘the English type’ in the literature. In this kind of questions in Hungarian, one interrogative word appears in the preverbal position, and the other in situ, or, arguably, in a sentence-final position. In Hungarian, this question type is subject to various constraints. Consider the following example:

Ki ment moziba kivel?
Who went to the cinema with whom?

According to the literature (see Kálmán (2001), for example), these questions can only be answered by a single pair. For instance, in (103), the locutor knows that someone went to the cinema with someone else, but does not know who exactly they were. A pair-list answer would seem to be less acceptable. Lipták (2001) observes that in this type of interrogatives, the question words must be of the same lexeme, i.e. they can differ only in their cases, otherwise the question is ungrammatical.
In other words, this means that the question should contain two (or more) forms of the interrogative word ki (who), mi (what), or hol (where).

According to a plausible explanation (Kálmán, 2001), the expected answer to such a question can be a direction, or, more precisely, the direction of the relation between the denotations of the two question words and not the denotations themselves:

(105)  **Kí szeretett bele kibe?** (János Mariba, vagy Mari Jánosba?)  
who love.PST.3SG VM who.ILL (János Mari.ILL or Mari János.ILL)  
Who fell in love with whom? (John with Mary, or Mary with John?)

(106)  **Ilyenkor az a szöveg, amikor kiderülne, ki on such occasions that the text, when would turn out, who lopott meg ki gyilkolt meg ki.**  
steal.PST.3SG VM who.ACC, who kill.PST.3SG VM who.ACC  
On such occasions what they say is like when it would turn out who stole something from whom, who killed whom.  

(from google, 2009/03/19)

From this it follows that the "lexematic" identity of the question words is important only in such "direction" questions, since in these the sets denoted by the question words must be projected onto each other in both directions.

However, the correct direction is not the only possible answer to this question type. Again, we can only talk about tendencies. First of all, questions containing two melyik (which) phrases are grammatical and can license even pair-list answers, just like the previous structure, in which all question phrases are extracted (108):

(107)  **Melyik diák választotta melyik könyvet?**  
which student choose.PST which book.ACC  
Which student chose which book?

(108)  **Melyik diák melyik könyvet választotta?**  
which student which book.ACC choose.PST  
Which student chose which book?

For some speakers, questions with two different interrogative words are completely acceptable, even if the relation is not potentially reversible. In such cases, however, both question words are arguments, including the subject. Questions containing an argument and an adjunct are clearly ungrammatical (104).

(109)  ?? **Kí mondott mit?**  
who say.PST.3SG what
Who said what?

(110) ?? **Ki** tört **össze mit?**
who break.PST.3SG VM what.ACC
Who broke what?

(111) **Mi** esett **rá kire?**
what fall.PST.3SG VM who.SUBL
What fell on who?

The data suggest that the least acceptable case is the one in which one question word is the subject, and the other corresponds to the object. The reason might be pragmatic. Situations when the locutor asks a non-echo question about an event in connection with which he does not know any of the participants (the agent or theme/patient) are rare.

Until now we have adopted the view that multiple questions of this type can license only single pair answers. However, Surányi (2006) claims that in some cases they are compatible with pair-list answers as well:

(112) a. **Q:** -**Ki** nézett **rá kire?**
who glance.PST.3SG at who.SUBL
Who glanced at who?

b. **A:** -János nézett **rá Maríra, Péter Julíra, és Pista János** glance.PST.3SG at Mari.SUBL, Péter Juli.SUBL and Pista Annára.
John glanced at Mary, Peter at Julia and Steve at Anna.

(based on Surányi (2006))

This assumption is confirmed by the following example from the Hungarian translation of a 20th century short story:

(113) **Mi** sem természetesebb, mint hogy mind a három lány what not more natural, than that all the three maids szívét előntötte a bámulat meg a szerelem heart.POSS.1SG.ACC overwhelm.PST the fascination and the love Lejbel íránt. Hiszen ŏ úzte el Doboszovát, Lapitutot Lejbel toxards since he chase.PST VM Dobosova.ACC, Lapitut.ACC meg a többi örögfajzatot! De persze tudták, hogy and the other diabolic creatures.ACC but of course know.PST.3PL that Lejbel csak egyiküket veheti feleségül, s végére Hersel is, Lejbel only one of them.ACC can marry, and all in all Hersel too, Velvel is helyre legény. No de **ki** vegyen el **kit?** Velvel too, handsome lad so but who marry.IMP VM who.ACC
It goes without saying that all the three maids were overwhelmed with fascination and love with Lejbel, since it was him who had chased Dobosova, Lapitut and the other Diabolic creatures. But of course, they knew that
Lejbej could marry only one of them, and, all in all, Hersel and Velvel were also nice lads. But who should marry who?

(Isaac Bashevis Singer: A félelmetes fogadó. Történetek gyermekeknek [The Fearsome Inn. Stories for children])

In sum, in the second type of questions, the main type of answer expected is a single pair, but pair-list answers are not excluded either. In some of them the direction of the relation between the denotations of the question words is expected. In such cases, the interrogative words must refer to the same type of set (animate individuals, inanimate individuals, locatives, specific individuals).

### 2.4.3 Coordinated left peripheral question words

In the third type of multiple questions in Hungarian, we find coordinated question words in front of the finite verb:

(114) **Ki és mikor ment moziba?**
  who and when went cinema.I.I.L.  
  Who went to the cinema and when?

(115) **Hol és mikor találkozunk?**
  where and when meet.1.I.PL  
  When and where will we meet?

This question type licenses only single pair answers according to Kálmán (2001). It complements the previous type, in that it is less acceptable with two question words denoting the same type of set. Such questions usually contain either an adjunct and a complement, or, even more frequently, two adjuncts.

(116) ?? **Ki és kivel ment moziba?**
  who and who.INSTR went cinema.I.I.L.  
  *Who and with whom went to the cinema?

As we have seen above, there is some hesitation in the use of *ki* (who) and *mi* (what) and their inflected forms, especially when both of them are arguments with respect to the 'English-type'. Since this third structure complements the 'English type', in that it licenses a single pair answer, but permits an argument and an adjunct question word as well, it is interesting to examine if the combinations of these two question words are grammatical or not. Intuitively, as pointed out above, it is difficult to construct the grammatical (non-echo) question, since questions asking about the subject and the object at the same time are somewhat rare. Such examples can indeed be attested, although not each of them is unanimously accepted by all speakers:

(117) Én tudós vagyok – mondja Sas Béla -, engem nem befolyásol-hat,  
  I scholar am - says Sas Béla - me.ACC not influence-can,  
  **ki és mit érez az igazság hallatára.**  
  who and what fells the truth hearing.SUBL
I am a scholar, says Sas Béla, I cannot be influenced by who feels what when hearing the truth.

(118) \[ \text{Ki és mit énekelte a tévében tegnap esté?} \]
\[ \text{who and what sing.pst the TV.iness yesterday evening} \]
Who sang what on TV yesterday evening?

These question words do not seem to be more grammatical in different structures either:

(119) \[ \text{Ki énekelte mit a tévében tegnap esté?} \]
\[ \text{who sing.pst what the TV.iness yesterday evening} \]

(120) \[ \text{Ki énekelte és mit a tévében tegnap esté?} \]
\[ \text{who sing.pst and what the TV.iness yesterday evening} \]

In a questionnaire, native speakers of Hungarian had to choose the most adequate multiple question for a given situation. Here, the situation was as follows: ”You have heard about an invention, but you don’t know who invented something and what was it, so you ask.”. They had to choose between the four questions:

1. (121) \[ \text{ki mit talált fel.} \]
\[ \text{who what invent.pst vm who invented what} \]

2. (122) \[ \text{ki talált fel mit.} \]

3. (123) \[ \text{ki és mit talált fel.} \]

4. (124) \[ \text{mit és ki talált fel.} \]

The answers reflected the above suggested hesitation: Figure (2.3).

According to the answers, most people preferred preverbal coordination, although a significant number of people chose the second answer as well.

The questions are more acceptable if the interrogative words do not correspond to the subject and the object (only to one of them):

(125) \[ \text{Kinek és mit mondta?} \]
\[ \text{who.dat and what say.pst.2sg} \]
To who did you say something and what was it?

The other two structures are marked or ungrammatical (in non-echo contexts) in this case as well:

(126) \[ \text{?? Kinek mondta mit?} \]
\[ \text{who.dat say.pst.2sg what} \]

(127) \[ \text{?? Kinek mondta és mit?} \]
\[ \text{who.dat say.pst.2sg and what} \]
Concerning the status of the coordinated preverbal question words in (114), it has been argued (Bánréti, 2007) that they have a similar structure to the following question:

(128) **Ki ment moziba és mikor?**

who go.PST cinema.ILL and when

Who went to the cinema and when?

Bánréti claims that both (114) and (128) can be analyzed as elliptical structures, the ellipsis being backward in (114) and forward in (128). Lipták (2001) and Skrabalova (2006) argue, on the other hand, that only the structure in (128) can be considered as elliptical; in (114) what we see is the coordination of two constituents without ellipsis (*i.e.* a true multiple question). Let us list the arguments for and against the elliptical analysis of the third type of questions:

- Both question words are arguments

  When both interrogative words are arguments, only the preverbal coordination is fully acceptable (*i.e.* in a non-echo context).

(129) **Ki és mit talált fel?**

who and what invent.PST VM

Who invented something and what was it?

(130) **Ki talált fel és mit?**

who invent.PST VM and what

*Who invented and what?*

This can be explained if we assume that the question words are in the same clause in (129), but in different ones in (130). If we suppose that (130) is a bi-clausal, elliptical structure, the original sentence would be the following:
(131) **Ki talált fel és mit talált fel?**  
who invent.PST VM and what invented VM

The ungrammaticality of the bi-clausal version is obvious: the object is missing from the first clause and the subject from the second, the predicates are not saturated. Since the preverbal coordination of the question words is grammatical, we can assume that they must be in the same clause in that structure.

However, some transitive verbs have intransitive uses without an explicit object in the sentence. Such verbs are *to sing* or *to eat*. Unlike the above examples, the object of such optionally transitive verbs can be coordinated also sentence-finally:

(132) **Ki és mit énekelt?**  
who and what sing.PST  
Who sang something and what was it?

(133) **Ki énekelt és mit?**  
who sing.PST and what  
?? Who sang and what did s/he sing?

In this case, however, the verb is used intransitively in the first conjunct, and transitively in the second.

- **Argument and adjunct**

On similar grounds, when one of the question words is an argument and the other an adjunct, the coordination of the argument at the end of the sentence (forward ellipsis) is ungrammatical:

(134) **Mikor és ki járt itt?**  
when and who come.PST here  
Who came here and when?

(135) **Ki és mikor járt itt?**  
who and when come.PST here  
Who came here and when?

(136) **Ki járt itt és mikor?**  
who come.PST here and when  
Who came here and when?

(137) *Mikor járt itt és ki?*  
when come.PST here and who

From this it follows that example (137) is indeed the coordination of two clauses, since the subject is missing from the first conjunct and thus the whole sentence is ungrammatical. If it is an adjunct that is coordinated sentence-finally, or both are in the preverbal field, the sentence is grammatical. Notice,
that if (135) was the coordination of two clauses, the subject would be missing from the second clause, leading to ungrammaticality. This supports the view that in (134) and (135) constituents and not clauses are coordinated.

- *it*-reading and *at all* reading (Gračanin-Yuksek, 2007)

If we consider again optionally transitive verbs, if the question words (one argument and one adjunct) are preverbal, only the transitive reading is possible:

(138) Mit és miért olvasott?
     what and why read.PST
     What did s/he read and why did s/he read it?

(139) Miért és mit olvasott?
     why and what read.PST
     What did s/he read and why did s/he read it?

On the other hand, if one question word is coordinated clause-finally, two readings are possible:

- if the adjunct is clause-final, the verb is interpreted transitively in both clauses; this is referred to as the *it*-reading:

(140) Mit olvasott és miért?
     what read.PST and why
     What did s/he read and why did s/he read it?

- if the argument (the object) is clause-final, two readings are possible: the above-mentioned *it*-reading, and the *at all*-reading (where the verb is interpreted as intransitive in the first clause and as transitive in the second)

(141) Miért olvasott és mit?
     why read.PST and what
     Why did s/he read (what s/he was reading) and what did s/he read?/Why did s/he read at all and what did s/he read?

- The difference between the definite and the indefinite conjugation

Transitive verbs agree with the definiteness of their object in Hungarian. Lipták (2001) points out that when the verb follows the two coordinated *wh*-phrases, its conjugation (indefinite) is different from what we would expect (definite) if the sentence was biclausal and elliptical:

(142) Nem érdekel, hogy mit készítesz és hogyan
     not interests, that what prepare.2SG.INDEF and how
     készítédek készítesz.
     prepare.2SG.DEF
     I am not interested in what you do and how you do it.
(143) Nem érdekel, hogy [mit és hogyan] készítesz.
    not interests, that what and how prepare.2SG.INDEF
    I am not interested in what you do and how.

The definite conjugation in (142) implies that the question in the first clause is already solved and adds another question to it, whereas the indefinite conjugation in (143) suggests that neither of the questions is solved. It cannot be elliptical, since in this case we would expect a definite verb form.

Nevertheless, Bánréti (2007) argues that the rule of ellipsis in Hungarian is less strict with respect to definiteness agreement than in the case of mood or tense endings. This means that even if the pronounced verb and the one undergoing ellipsis exhibit different conjugations, the sentence can be grammatical.

- First question solved

Another argument supporting the biclausal analysis of interrogatives with sentence-final coordination is polar questions containing a coordinated question word. As opposed to monoclausal questions without coordination, the interrogative clitic can co-occur with a coordinated question word. In this case, the question in the second clause presupposes that the first is already solved:

(144) Léci, léci, jelezzen, aki még nem tette, hogy és hányan!!!
    please please sign.IMP.3SG, who yet not did, that comes-CL.INTERR and how many
    Please please, tell me if you come and if so, how many of you!

    (invitation to a party; email)

(145) *Léci, léci, jelezzen, aki még nem tette, hogy és hányan jön-e!!!

This is impossible in the case of coordinated preverbal question words.

Taking these arguments into consideration, I will consider (128) as an elliptical structure and (114) as the coordination of constituents. However, several problems emerge if we take a closer look at the coordination of two interrogative words.

The first is that the function of the coordinated items is supposed to be identical (see Sag et al. (1985); Sag (2005), etc.). This accounts for the ungrammaticality of the following example, in which the subject and an adjunct are coordinated:

(146) *János és tegnap ment moziba.
    János and yesterday go.PST cinema.ILL
    *John and yesterday went to the cinema.
When both question words are adjuncts (115), this condition on the functional identity of the conjuncts is met, but this is not the case if one question word is an argument and the other an adjunct.

Peterson (2004) and Skrabalova (2006) suppose that even question words with unlike functions can be coordinated, because they are stressed and they share a common (discourse) function: focus. According to Lipták (2001), stressed elements can be coordinated in Hungarian, independently of their functions. She provides an example with universal quantifiers:

    here everyone and always everyone and always enter-can
    EVERYONE can ALWAYS enter here.

The second problem is related to the first. Although the above illustrated universal quantifiers and interrogative words are all stressed and prominent, all stressed and prominent constituents cannot be coordinated in Hungarian. When there is more than one focused constituent in the sentence, strictly only one is preverbal, and the other comes at the end of the sentence:

(148) */?? JÁNOS és KÉT FILMET nézett meg.
    JOHN and two films,ACC watch,PST VM
    *JOHN and TWO FILMS watched.
(149) JÁNOS nézett meg KÉT FILMET.
    JOHN watch,PST VM two films,ACC
    It was John who watched TWO FILMS.

Furthermore, as we have seen, not all question words can be coordinated, for instance those that denote the same type of set:

(150) ?? Ki és kibe szeretett bele?
    who and who,ILL love,PST VM
    Who fell in love with whom?

From this, we can conclude that the condition on coordination of prominent elements in Hungarian builds on some semantic identity of the conjuncts (universal quantifiers or interrogative words), specified in the lexical entries, although this condition has to be even more specific in the case of these latter. Focus, on the other hand, is not a lexical property of interrogative words.

Another approach is to claim that the common function of the question words is not focus, but the function ”extracted”. Extracted constituents are supposed to bear a grammatical function, referred to as filler in the HPSG, and Op(erator) (Alsina, 2008) or UDF (Unbounded Dependency Function) (Asudeh, 2010) in the LFG framework. As Skrabalova (2006) argues, question words with different functions can be coordinated in Czech, provided that they are extracted, but they cannot be coordinated in situ. However, even if we assume that the common function of these elements is a version of ”extracted”, we have to account for the fact that (150) is not acceptable for a significant number of locutors in Hungarian (although
the question words are extracted), and that certain semantic factors also come into play.

A typological remark is due here. Apart from Hungarian and Czech (Skrabalova, 2004, 2006), the possibility of coordinating question words with different functions has been reported about Russian (Kazenin, 2010), and Romanian (Comorovski, 1989). Kazenin shows, based on similar observations and arguments, that multiple questions containing coordinated preverbal question words are monoclausal, as opposed to the elliptical biclausal alternative with a sentence-final coordinated question word (usually an adjunct). He proposes, in a movement-based framework, that the indices of both question words percolate onto the head of the coordinate phrase and thus it can govern the traces of both question words. In the LFG framework, these facts have to be accounted for without such assumptions.

Let us now consider other corpus data. First we examine preverbal coordination of question words. We start with questions in which one question word is an argument and the other an adjunct.

- **ki-hol** (who-where)

(151) Na erre **ki és hol** tud válaszolni?
well this.SUBJ who and where can answer
And who could answer this question, and where?

(google, 2009/03/19)

- **ki-mikor** (who-when)

(152) Az ellátást szervező szolgáltató elszámolását
the attendance.ACC organizing service accounting
valamennyi, az érintett lakosság ellátásával
all, the concerned inhabitants attendance.INSTR
kapcsolatban keletkezett természetbeni kiadás terheli,
in connection with created natural expense loads,
függetlenül attól, hogy az ellátást **ki és**
independently that.ABL that the attendance.ACC who and
mikor kezdeményezte a modellkísérlet időtartama alatt.
when initiate.PST the model experiment period.PSS during
All natural expenses coming from the attendance of the concerned inhabitants are loaded onto the accounting of the service organizing the attendance, independently of who and when initiated the attendance during the period of the model experiment.

(http://www.min.hu/43-1999korm.htm)

(153) Korábban csak a tejesembertől vagy a postástól lehetett
before only the milkman.ABL or the postman.ABL could
információkat megtudni arról, hogy éppen **ki és**
information.ACC get about, that at the moment who and
mikor nyaral, vagy általában melyik napszakban
when is on holiday, or usually which part of the day.

nem tartózkodik otthon, manapság viszont a
not is at home, nowadays on the other hand the

betörők internetes kutatómunkával készülnek fel a
burglars internet research work. INSTR prepare VM the

kiszemelt házak, lakások kifosztására.
chosen houses, flats burglary.Poss.Subl

Before one could get information only from the milkman or the post-

man about who went on holiday and when, or usually who is not

at home in which part of the day. Nowadays, on the other hand,
burglars prepare for the burglary of the chosen houses and flats with

some research on the internet.

(http://hetivalasz.hu/kronika/megdobbento-tenyek-derultek-ki-a-

facebookrol-30869)

• ki-hogyan (who-how)

(154) Ki és hőgyan döní el azt, hogy például egy
who and how decides VM that, that for example a

cosmetikai cég reprezentatív célból, vagy csak a

cosmetic firm representative aim. ELAT, or only the

reklám kedvéért ajándékozta meg termékeivel
advertisement for the sake made a gift VM products. INSTR

üzleti partnereit?
business partners. Poss. 1sg. Acc

Who decides and how, if, for example, a cosmetic firm made a gift

of its products to its business partners with a representative aim, or

for the sake of advertising?

(http://www.origo.hu/itthon/20010105reprezentacio.html)

• ki-miért (who-why)

(155) Máig nem derült ki, hogy ki és miért fuvarozta
until today not turn. PST VM, that who and why transport. PST

Szabó László, akit tavaly a Gyulai Városi Bíróság 3
Szabó László. ACC, who. Acc last year the Gyula Town Court 3
év szabadságvesztésre ítélte.
sence imprisonment. Subl sentence. PST. many plaintiff. DAT

természetesen fogalmazott sincs a rendőrségi kihallgatás során,
naturally idea. Poss not the police questioning during,
hogy ki és miért bántalmazta.
that who and why hurt. PST

Until today it has not turned out who transported Szabó László

and why. He was sentenced by the Gyula Town Court to 3 years
of imprisonment. Naturally, many plaintiffs don’t have the slightest idea during the questioning at the police, who hurt them and why.

(google, 2009/03/19)

* mit-miért (what-why)

(156) Mindig akadnak kifogások, hogy mit és miért nem lehet. always emerge excuses, that what and why not is possible There are always excuses about what is not possible and why.

(György Varga, a Magyar Köztársaság kassai főkonzula, Interjú, Forrás: Új Szó, 2000/12/07)

(157) Mert ha valami megjelenik a tévében vagy az because if something appears the TV or the indexen Morvai Krisztináról, akkor itt az ő index.SUPERESS Morvai Krisztina.DEL, then here the her véleményét is el lehet olvasni. (ma pl. opinion.ACC too VM is possible to read (today for example teljesen értelmesen összefoglalta, hogy miért és mit completely sensibly sum.PST UP, that why and what dolgozott az MSZP-nék). work.PST the MSZP.DAT)

Because if something appears on TV or on Index (internet news site) about Morvai Krisztina, then her opinion can also be read here (today, for instance, she gave a completely sensible summary about why she had worked with the MSZP (Hungarian Socialist Party) and what she had done).

(Internet forum, 2010/01/20)

* hol-mit (where-what)

(158) Ne tévesszen meg senkit, hogy Hornról két év senki not mistake VM no one.ACC, that Horn.DEL two years no one nem tudja, hogy hol és mit csinál. not knows, that where and what does No one should be mistaken by the fact that no one knows about Horn, where he is and what he does.

(http://matula.hu/index.php?section=article&rel=39&id=464)

(159) Egy idő után sajnos tényleg nem számít, hogy hol és a time after unfortunately really not counts, that where and mit végeztél, még itthon sem. what finish.PST.2SG, even at home either
After a certain time, unfortunately, it does not matter what you studied and where, not even in your home country.

(Internet forum, 2010/05/14)

- **hol-milyen feltételek mellett** (where-on which conditions)

(160) **Hol és milyen feltételekkel lehet és szabad diák munkát vállalni?**

Where and on what conditions can one undertake a student job?

(http://www.stop.hu/forum/entries.php?lstresults=1&forum_topics_id=232082)

- **mit-hogyan** (what-how)

(161) **De gyakorlati info kellene arról, mit és hogyan lehet, kell but practical info needed about, what and how can, must könyvelni. Nem mindegy, hogy kifizetek havi 200 Eurot book, not equal, that pay.1sg monthly 200 Euros.acc egy könyvelőnek ezért vagy megesnálatom magam, ha tudom. an accountant.dat for this or vm.do myself, if can Some practical information would be needed about what can be and must be booked and how. It really matters if I pay 200 euros an accountant every month for this, or I do it myself if I can.

(Internet forum, 2009/05/16)

(162) **Szélviharosan indult a tegnapi plenáris, a képviselők windy start.pst the yesterday plenary, the deputies igyekeztek nem közvetlenül a Kossuth téri fák alá try.pst not directly the Kossuth square trees under parkolni – ki tudja, mit és hogy dönt az időjárás. park - who knows, what and how decides the weather The yesterday plenary had a windy start, the deputies trying not to park their cars under the trees on Kossuth square - who knows on what the weather decides and how.

(google, 2010/03/08)

- **mennyi-mire** (how much-what for)

(163) **Az ÁPV rt-nél tegnap még nem tudták megmondani, the APV Ltd-at yesterday yet not know.pst.3pl vm.to tell, hogy mennyi pénzt és mire utalnak át a that how much money.acc and what.subl transfer vm the
At the APV LTD yesterday they could not tell yet how much money would be transferred to the company and for what.

(http://www.origo.hu/itthon/20001221kedden.html)

In the following examples, both question words are oblique arguments, and thus they share the same function.

- **hol-mikor** (where-when)

(164) Jegyek rendben, igyekszem levelet írni, hogy **hol és mikor** találkozunk.

The tickets are OK, I'll try to write a letter to tell you where and when we meet.

(email, 2010/02/18)

(165) Arisztotelész elvárja, hogy a nézők azonnal értesüljenek róla, hogy a cselekmény **hol és mikor** játékának rólát, ahol a trójai háború kezdetei.

Aristotle expects the spectators to be immediately informed about where and when the plot takes place: for example in the city of Troy at the beginning of the war.

(Géza Hegedüs, Világirodalmi arcképcsarnok, Magyar Elektronikus Könyvtár [Hungarian Electronic Library], 1998/12)

(166) It has not yet been decided when and where the cancelled contest will be organized.

(http://www.hhrf.org/ujeszto/2001/47/sport.htm)

In the following examples, both question words are adjuncts, sharing the same function.

- **hol-miért** (where-why)
(167) Megvilágította annak az okát, hogy hol és miért clarified that.Poss the reason.Acc, that where and why alakult ki az a többinél fejlettebb develop.Pst Vm that the than the others more developed társadalom, ami megszülte az ipari forradalmat és society, that create.Pst the industrial revolution.Acc and annak alapján a klasszikus kapitalizmust. that.Poss base.Superess the classical capitalism.Acc He clarified the reason for where and why could develop that society that was more developed than the others and that created the industrial revolution, and, on the basis of that, the classical capitalism. ([http://mek.niif.hu/01400/01430/html/01.htm](http://mek.niif.hu/01400/01430/html/01.htm))

• hol-hogyan (where-how)

(168) Ez egy kicsit felvidította, igyekezett elképzelni, hogy this a bit make.Pst him happy, try.Pst to imagine, that hol és hogyan bünteti meg Pötyikét. where and how punish Vm Pötyike.Acc This made him a bit happier, he tried to imagine where and how he would punish Pötyike. ([http://dia.jadox.pim.hu](http://dia.jadox.pim.hu))

(169) Az, hogy hol és hogyan alakul egy nép sorsa that, that where and how develops a people destiny.Poss ezekben a századokban, vagyis korunkban, az these.Iness the centuries.Iness, i.e. age.Poss.1Pl.Iness, that minden történelmi előzménynél nagyobb mértékben fiúgg all historical antecedents bigger measure.Iness depends attól, milyen erősek és milyen természetiek az érintett that.Abl, how strong and what nature the concerned területeken az ott élő népek polgári tradíciói. area.Superess the there living peoples civil traditions.Poss The question of where and how the destiny of a people develops in these centuries, i.e. in our age, depends fundamentally on how strong and of what type are the civil traditions of the people living in a given area, much more than on historic antecedents. ([http://mek.niif.hu/01400/01423/01423.htm](http://mek.niif.hu/01400/01423/01423.htm))

(170) Igenis a szülő nő joga eldönteni, hogyan és yes the giving birth woman right.Poss to decide, how and hol szül. where gives birth
It is unquestionably the right of the woman giving birth, to decide how and when she gives birth to her child.

(MHO, Lockerbie-ügy: egy bűnös és egy ártatlan, Magyar Hírlap, 2001/01/31)

• mikor-miért (when-why)

Kérdés, hogy mikor és miért jön el ez a pillanat, és question, that when and why comes VM this the moment, and kérdés az is, hogy mit tehet a magyar kormány question that too, that what can do the Hungarian government és a magyar országgyűlés e pillanat eljövetelének and the Hungarian parliament this moment coming,POSS megakadályozásáért. prevention.CAUS-FIN

The question is when this moment comes and why, and it is also a question what the Hungarian government and parliament can do to prevent this moment from coming.

(http://dia.pool.pim.hu)

Töbnyire nem lehet azt később megmondani, hogy mostly not possible that.ACC later to VM.say, that miért s mikor szerettünk meg egy költöt. why and when liked.PST.1PL VM a poet.ACC

Mostly it is not possible to say later why and when we started to like a poet.

(http://dia.pool.pim.hu)
The question why and when the husband became an alcoholic has not been clear to me yet from Macs’s writings, although I don’t consider it unimportant.

(http://forum.index.hu/Article/showArticle?t=1001722)

- **hogyan-miért** (how-why)

(175) Arról is szeretnék még beszélgetni píct, hogy a tévé és a könyv **hogyan és miért** egészen másfajta kommunikáció, és hogy ennek társsalómi okai are I would like to talk a bit about the question how and why TV and reading books constitute completely different types of communication and that this has social reasons.

(176) Azt feszegettem ebben, hogy **miért, hogyan** vált az európai mítosz immár kitöröhetetlen become.PST the European myth now indelible részévé Auschwitz, a mindaz, ami e part.TRANSL-FACT Auschwitz, the everything, that this fogalomkörhöz tartozik. notion.ALL belongs In this, I analyze why, how Auschwitz and everything associated with this notion became a now indelible part of the European myth.

(google, 2009/06/18)

(177) Tanulmányom központi kérdése az, **hogyan és miért** study.POSS.1SG central question that, how and why minősít a közösség egy személyt bolondnak, azaz hatalom qualifies the community a person.ACC fool, i.e. power nélküli, a társsalomból kizárt individuumnak. without, the society.ELAT excluded individual The central question of my study is how and why the community qualifies a person as a fool, i.e. an individual without power, excluded from society.

(http://epa.oszk.hu/00400/00458/00012/12k05.htm)
more than two question words

(178) Pontosan máig sem tudjuk – talán a történetírók exactly till today either know.1pl - perhaps the historians egyszer földerítik -, hogy hol, mikor és miért következett once reveal - that where, when and why happen.PST be fordulat ebben az ügyben.

Even today we don’t know exactly - perhaps one day the historians will reveal it - where a turning point occurred in this issue, when and why.

(179) Hogy az ingatlant ki, mikor és miért hasította négyfelé, that the estate.ACC who, when and why cut.PST into four, nem szól semmiféle okmány. not tells any kind of document

There is no document about who cut the estate into four, when and why.

(180) Az ellenzéknek igazából nincs beleszólása abba, hogy mi the opposition in reality no influence into, that what miért és hogyan történik. why and how happens

The opposition cannot influence what happens, why and how.

Contrary to what we have seen so far, it has gone unnoticed that the coordination of two preverbal question words can license pair-list answers as well. As we shall see now, this reading possibility does not necessarily come from an inherent property of the syntactic structure or the question words. We have already seen that in structures with multiple preverbal question words, the linearly first interrogative word is usually a complement denoting a set of individuals. That structure is inherently associated with a pair-list reading (apart from complex question words). In the present case, the expected reading is usually a single pair, since the question is about the circumstances of a single event. However, there are also questions referring to a sequence of events or multiple subjects, in connection with which pairs of constituents are expected as answers. Still in these cases, miért (why) and hogyan (how) are rare as linearly first question words, but they are not excluded. Crucially, whenever the pair-list reading is available, the question words are coordinated preverbally. This also supports the analysis of coordinated preverbal question words as a monoclausal structure, since in order for a pair-list reading to be available, the question words have to be clausemates. Let us consider the following corpus data:
• *hol-mikor* (where-when)

(181) Minden relatív, így az is, hogy *hol* és *mikor* van szükség all relative, so that also, that where and when is need forgalomirányító rendőrökre, s ha kellenek, akkor traffic policemen.SUBJ, and if needed, then mennyi idő alatt érnek a helyszínre. how much time under get the scene.SUBJ

Everything is relative, also the question, when and where are traffic wardens needed, and when they are needed how much time it takes them to get to the scene.

(( h. l. l. ) Káosz karkelzésre, Magyar Nemzet 1999/07/07)

In this example, place and time pairs are expected as answers. The places are enumerated on the first place and distribute over different times. Compare this with the following:

(182) Bevezetőben leírta, hogy *hol* és *mikor* született, introduction describe.PST, that where and when be born.PST, hogy hívták anyját, apját és how call.PST mother.Poss.ACC, father.Poss.3sg.ACC and testvéreit, és hogy azok *hol* s *mikor* siblings.Poss.3sg.ACC, and that those where and when születtek.

be born.PST.

He described in the introduction where and when he was born, how his mother, his father and siblings were called and where and when they were born.

((google, 2009/04/16)

In the above example, the pair-list reading is due to the multiple subjects (his mother, his father and siblings), in connection with which pairs of answers are expected. In what follows we will see more examples of both types of pair-list answers.

• *mikor-hogyan* (when-how)

(183) Az viszont biztosan nem visz előre, ha két that on the other hand surely not brings forward, if two párt – mellesleg azonos oldalon állva – parties - besides identical side.SUPERESS standing - azon vitázik, újságokon át üzengetve, hogy that.SUPERESS quarrel, newspapers across sending messages, that *mikor* és *hogyan* kellene együttműködni. when and how should cooperate
On the other hand, it is no step forward if two parties - standing on the same side, by the way - quarrel, sending messages across newspapers, about when and how they should cooperate.

(Székfoglaló, Népszava, 2001/09/17)

(184) A növényeknél a fejlődési ritmus és az évszakok válta kozása is befolyásolja, hogy mikor és hogyan kell alternation also influences, that when and how must öntözni.

Concerning plants, their rhythm of development, as well as the changing of the seasons influence when and how one must water them.

(Női tükör, Szomjazó szobanövények, Dunántúli Napló 2001/05/30)

• mikor-miért (when-how)

(185) Az engedély birtokosainak naplót kell vezetniük arról, the permission owners.DAT diary.ACC must lead that.DEL, hogy mikor és miért használtak megkülönböztető jelzést, that when and why use.PST distinguishing sign.ACC, és ezt a rendőrség ellenőrzi.

The owners of the permission have to note in a diary when and why they used the distinguishing sign and the police control this.

(hírösszetevők, Be akarja vonni a BM a kék lámpát a miniszterektől, Népszabadság, 2000/11/4)

• hol-hogyan

(186) Küldött a színésznek egy gyönyörű hangos zenei levelet, send.PST the actor.DAT a wonderful loud musical letter.ACC, amiben megszólaltatja nemsak a közismert which.INESS makes speak not only the commonly known változatot, hanem a különböző variációkat, és elmondja, version.ACC, but the different variations.ACC, and tells, hogy azokat hol és hogyan éneklik.

She sent the actor a wonderful loud-speaking musical letter, in which not only the commonly known version is presented, but the different variations as well and tells where and how those are sung.

(A megfelelés kényszerében, Magyar Hírlap, 1998/07/04)
2.4. Multiple Questions in Hungarian

(187) A régészek szinte a kincs előkerülése óta the archeologists almost the treasure discovery.POSS since probálják kideríteni, kié lehettek a nemesfémértéküket try to reveal, whose could be the precious metal value.ACC messze fölülmúlóan beces - különösen szépen far overriding valuable - extraordinarily beautifully kidolgozott - dísz- és használati tárgyak, hol és worked - decoration and usable objects, where and hogyan készülhettek, illetve miként és miért kerülhettek how could be made, and as what and why got feltalálási helyükre.

discovery place.POSS.3PL.SUBL

Almost since the discovery of the treasure, the archeologists have been trying to find out to whom belonged the extraordinarily beautiful decorative and useable objects that well override their value in precious metal, where and how they were made and as what and why they got to the place where they were discovered.

(Zsófia Pethő, A szilágyosmolyói kincselet, Heti Világgazdaság, 1999/05/22)

(188) A kémkedés minden titkába beavattak, hogyan the spying all secret.POSS.3SG.ILL introduce.PST, how és hol kell üzeneteket hagyni, hogyan lehet elmaszkórozní and where must messages leave, how can mask magunkat; ezerszer intettek éberségre és ourselves, a thousand times warn.PST awareness.SUBL, and különösen az ügynőköt fizető legnagyobb veszélyre: a especially the agents.ACC threatening biggest danger: the nőkre!

women!

They introduced us all the secrets of spying, how and where one should leave messages, how to mask ourselves, they warned us a thousand times to be aware and also against the biggest danger threatening agents: women!

(http://mek.niif.hu/02200/02241/html/01.htm)

• hol-miért (where-why)

(189) Sorolhatnánk még, hogy hol és miért nincs jegyző, enumerate.COND.1PL still, that where and why not notary, illetve milyen módon próbálják hiányát legalább and which manner try lack.POSS.1SG.ACC at least átmenetileg pótolni.

temporarily substitute

We could keep on listing, where and why there is no notary, and which way they try to substitute for the lack of it.
(M. A., Baranyai tükör, Sok jegyző hiányzik a megyében, Dunántúli Napló, 2001/01/29)

(190) Most értettem meg, hogy miért és hol maradt now understand.PST.1SG VM, that why and where remain.PST meg műemléknek az a néhány középkori várom és VM as monuments that the some medieval castle ruin and templomomladék, ami úgy-ahogy megmaradt. church ruin, partly remain.PST It has become clear to me now why and where those few medieval castle and church ruins that got preserved, got preserved to become monuments.


2.4.4 Clausal coordination with ellipsis in the second conjunct

After the examples with preverbal question words, let us consider some others in which one interrogative word is coordinated sentence-finally (as noted above, this biclausal type does not qualify as a true multiple question):\footnote{Note that only the single pair reading is available in these questions.}

- ki-mikor (who-why)

(191) **Ki** döntött és mikor, amikor a kormány 1995. március who decide.PST and when, when the government 1995 May 12-éi határozata azt mondja ki a 11. oldalon, 12 decision that.ACC declares VM the 11th page.SUPERESS hogy elő kell készíteni a felsőoktatás that VM must to prepare the higher education fejlesztéséről szóló törvényt? development.POSS.DEL about law.ACC Who decided and when, if the 12 March 1995 decision of the government declares on page 11 that the law about the development of higher education is to be prepared?

(Dr. Krisztina Dobos (MDF), Országgyűlési Napló, 1995/04/11)

(192) Nemtommár **kivel** vitatkoztam az don’t know any longer who.INSTR quarrell.PST.1SG the olívabogyókról és mikor, de baromira igazam volt olives.DEL and when, but absolutely right.POSS was I don’t know who I quarrelled with about olives and when, but I was absolutely right.

(email, 2009/11/12)
2.4. Multiple Questions in Hungarian

- mit-miért (what-why)

(193) Ti persze biztosan nagy tudorok vagytok, de pontosan mit you of course surely big scholars are, but exactly what hazudott volna és miért? lied would and why
You are surely big scholars, but what would he have lied to me and why?

(http://forum.index.hu/Article/showArticle?t=1002303&go=107905&p=1)

(194) Holnap délben megyek megint vizsgálatra, és majd tomorrow at noon go.1SG again check-up.SUBL, and then beszámolok, hogy mit döntött a dokim és tell.1SG, that what.ACC decide.PST the doctor.POSS.1SG and miért. why
I am going to a medical check-up tomorrow at noon and then I’ll tell how my doctor decided and why.

(Törzsasztal: Szülés, terhesség, Index Fórum, 1999/06/01)

- hol-miért (where-why)

(195) Valahol el kéne kezdenem, de nem tudom, hol somewhere VM should to start.1SG, but not know.1SG, where vagyok és miért. am and why
I should start somewhere, but I don’t know where I am and why.

(http://dia.jadox.pim.hu/jetspeed/displayXhtml?docId=0000001183 &secId=0000116540&mainContent=true&mode=html)

- hogyan-miért (how-why)

(196) És meddig tartott ez a gyöngyélet, hogyan and until when last.PST this the life in clover, how ért véget és miért. come.PST to an end and why
And for how long lasted this life in clover, how did it come to an end and why.

(http://dia.jadox.pim.hu/jetspeed/displayXhtml?docId=0000000821 &secId=0000076702&mainContent=true&mode=html)

- more than two question words
After this we expect our readers’ answers in the comments: on which belongings of which stars would they spend money and how much?

(http://www.stop.hu/articles/article.php?id=679247)

We have seen that there are three types of multiple questions in Hungarian. The interpretation of these questions cannot always be associated with a syntactic structure. In the case of multiple preverbal question words, the pair-list answer is required (the single pair answers is also possible in a specific case). In the second and third cases, the single pair answer is preferred, but the pair-list is not excluded either. What provides the clue for the interlocutor to give the right type of answer to a multiple question if the syntax is not a reliable source? The obvious answer seems to be the context, prosody, and (not independently of these two), the information structure of the question (which cannot be directly experienced). In the next section we go on to the French data.

2.5 Multiple Questions in French

In this section, I examine multiple questions in French, and attempt to describe both formal and informal French data (arguing against views according to which they constitute two different grammars). In French, we can find the structures M-1-extr, M-all-ins, M-coord-extr, M-coord-ins of the above presented multiple question types. In two of them, the question words are coordinated (M-coord-extr, M-coord-ins), whereas in the other two they appear without coordination (M-1-extr, M-all-ins). In both of these groups, it is possible to extract question words (M-1-extr, M-coord-extr), or to leave all of them in situ (M-all-ins, M-coord-ins), depending on the register. Note that the extracted-in situ difference does not influence the possible readings of the question. As we will see, types B and C cannot be exclusively associated with a type of reading (pair-list or single-pair), but the argument vs. adjunct status of the question words influences the preferred interpretation.

A fifth type can also be identified, containing clause-final coordination. As expected, this variant can mostly be answered by a single pair (there are some well-defined exceptions to this, as we will see). Like in Hungarian, these will be analyzed as the coordination of two single questions, and not as a true multiple question.

Let us now have a look at these types one by one, using, like in the previous section, a lot of attested examples.14

---

14The examples come from the Frantext corpus, from the French press (www.liberation.fr), from the internet, and from spontaneous speech.
2.5. Multiple Questions in French

2.5.1 One left-peripheral question word, the other *in situ*

- Polar questions containing a question word
  
  Like in Hungarian, polar questions containing a question word are ungrammatical in French, both in main and embedded clauses (in non-echo contexts):

  (198) *Est-ce que Paul a été où ?
      int.c  Paul aux been where
      *Has Paul been where?

  (199) *Est-ce que Marie a mangé quoi ?
      int.c  Marie aux eaten what
      *Did Mary eat what?

  (200) *Paul demande si Marie a été où.
      Paul ask.3sg if Marie aux been where
      *Paul asks if Mary was where.

- Arguments
  
  This type is ambiguous between a pair-list and a single-pair reading, especially when the question words are both arguments. Consider the following examples:15

  (201) a. Q1: Qu’as-tu donné à qui ?
       what have you given to whom
       What did you give to whom?
  b. Q2: Qu’est-ce que tu as donné à qui ?
       what is it that you have given to whom
       What did you give to whom?

  Both questions are compatible with both answers:

  (202) a. Single pair:
       A1: J’ai donné un livre à Marie.
       I have given a book to Mary.
       I gave a book to Mary.

  b. Pair-list:
       A2: J’ai donné un livre à Marie, et un stylo à Paul...
       I have given a book to Mary, and a pen to Paul...
       I gave a book to Mary, and a pen to Paul.

15See Chapter (6) about the variety of French interrogatives.
The answer expected is highly context-dependent. Consider the following examples:

- *qui*.NOM-*qui*.ACC (who-whom)

  Pair-list:

  (203) La multitude de dessins réunis ici est le fruit d’un
  the multitude of drawings collected here is the fruit of a
  principe simple: les artistes de la galerie invite un artiste
  principle simple: the artists of the gallery invite an artist
  de leur choix à exposer. L’enchevêtrement des œuvres
  de leur choix à exposer. The entanglement of the works
  invite le visiteur à chercher les relations qui les
  invite the visitor to look for the relations which them
  unissent: *Qui* a invité *qui* et pourquoi?
  unite: who has invited whom and why
  The multitude of drawings collected here is the fruit of a simple
  principle: the artists of the gallery invite an artist of their choice
  to expose. The entanglement of the works initiate the visitor
  to seek the relations which unite them: who invited whom and
  why?


  Pair-list:

  (204) Assez déçu par la fin du film... je ne vois pas
  quite disappointed by the end of the film... I see not
  une annonce de l’Allemagne nazie, ou une fresque
  an advertisement of the Germany Nazi, or a portrait
  d’enfants pervers, tout simplement un film gore où le
  of children perverse, all simply a gore film where the
  réalisateur ne savait pas lui même la réponse à la
  director knew not himself the answer to the
  question *qui* tue *qui*.
  question who kills whom
  I was quite disappointed by the end of the film... I can’t see
  any advertisement of Nazi Germany, or a portrait of perverse
  children, but very simply a gore film, where the director himself
  didn’t know the answer to the question: who kills who.

  (http://www.zerodeconduite.net/blog/?itemid=18705)

- *qui*-quoi (who-what)

  (205)
Qui fait quoi sur la une dans cette station?
who does what on the one in this station

Who does what on line one in this station?

(RATP, métro Paris)

– qui-quel (who-which)

(206) Pair-list:

On verra/ on sait pas qui aura quel cadeau.
one see/FUT one knows not who get/FUT which present

We’ll see/ we don’t know yet who will get which present.

(about a ‘Secret Santa’ Christmas party, email, 2009/12/15)

– qui-à qui (who-to who)

(207) Single pair:

Qui parle à qui? Lear à sa fille? L’art à who talks to who Lear to his daughter the art to
l’innocence? Un vieil homme à lui-même? La jeune the innocence an old man to himself the young
actrice (Julie-Marie Parmentier) qui lui répond, actress (Julie-Marie Parmentier) who him.DAT answers,
à peine, est miraculeuse. hardly, is miraculous

Who is talking to who? Lear to his daughter? Art to innocence? An old man to himself? The young actress (Julie-Marie Parmentier), who is hardly answering him, is miraculous.

(http://www.estroiscoups.com/article-27647114.html)

• Argument and adjunct

When one or both question words are adjuncts, the pair-list reading is preferred, although the single-pair reading is also possible in some limited cases.

– qui-où (who-where)

(208) Pair-list:

a. Q: En famille, qui dort où ?
in family, who sleeps where
In the family (home) who sleeps where?

The parents sleep on the first floor and the children on the second.

– *qui-quand* (who-when)

(209) Pair-list:

**Qui** est parti *quand* et qui est encore au Havre?
who left *when* and who is still *in* Le Havre

Who left when and who is still in Le Havre?

In some examples which would (pragmatically) expect a single-pair answer, the question is ungrammatical. This shows that this question type is mostly answered by a pair-list:

(210) *Qui a tué Henri IV quand?*
who *aux* killed Henri IV when
*Who killed Henri IV when?*

Since asking about unique events requires a single-pair answer, the coordination of two single questions is preferred in French, when one question word is an argument:

(211) **Qui a tué Henri IV et quand?**
who *aux* killed Henri IV and *when*
Who killed Henri IV and *when?*

(212) **Qui avait tué Manu et pourquoi?**
who *had killed* Manu and *why*
Who *had killed* Manu and *why?*

(J-C. Izzo, Total Khéops, 1995, p. 327, 15 Où la haine du monde est l’unique scénario)

• **Adjuncts**

When both question words are adjuncts, again, both readings are possible depending on the context, but the pair-list reading is preferred. A plausible
2.5. Multiple Questions in French

explanation is that if at least one adjunct question word is present in the question, an alternative structure (two single questions with ellipsis, with the sentence-final coordination of the adjunct question word) is also possible that requires exclusively single-pair answers.

(213) Pair-list:

a. Où a dormi Jean quel jour ?

Where did John sleep on which day?

• Lundi, Jean a dormi à Rome, mardi à Nice, mercredi à Cannes.
  Monday John slept in Rome, Tuesday in Nice, Wednesday in Cannes.
  On Monday, John slept in Rome, on Tuesday in Nice, on Wednesday in Cannes.

(214) Single pair:

a. Quand est-ce qu’il est arrivé avec quelle intention?

When did he arrive and what was his intention with it?

• Il est arrivé lundi pour rencontrer Jean.
  He arrived on Monday to meet John.

Another piece of evidence for the assumption that when one question word is an adjunct, the pair-list reading is preferred, is illustrated by the following example, which is less natural than the ones above:

(215) ?? Où a dormi Jean pourquoi ?

Where did John sleep why?

A possible explanation is (just like in the case of Hungarian) that contrary to other question words pourquoi (why) does not denote a set of individuals, but to a set of possible reasons, which are more difficult to list than individuals. This is why the pair-list reading is available only in those contexts in which the set of possible reasons pourquoi (why) refers to is explicit: for instance, if the question word is pour quelle raison (for which reason). Multiple questions with pourquoi contain, in most cases, clausal coordination, since, in order to answer the question introduced by pourquoi, the answer to the first part of the question has to be accommodated. We shall see this later.

In French, it is also possible to have a complex question word in the preverbal domain:
The sequences of elements d’où à où and de quand à quand are analyzed as complex question words (see also Bonami (1999)), since the multiple extraction of question words is clearly ungrammatical in French.

### 2.5.2 All question words *in situ*

The interpretation and answerhood conditions of this type of multiple questions is exactly the same as those of the previous type. Let us consider the corresponding examples:

- **Arguments**

  \[(218)\] Tu as donné quoi à qui ?

  you have given what to whom

  What did you give to whom?

  This question is also compatible with both answers: single pair (219-a) and pair-list (219-b).


      I have given a book to Mary.

      I gave a book to Mary.

  b. A2: J’ai donné un livre à Marie, et un stylo à Paul...

      I have given a book to Mary, and a pen to Paul...

      I gave a book to Mary, and a pen to Paul.

- **Argument and adjunct**

  Again, both readings are possible, although the pair-list reading (220) is preferred.

  \[(220)\] a. Q: Tu vas chercher qui à quelle heure ?

      you go pick up whom at which hour

      Whom are you going to pick up when?

  b. A: Je vais chercher Max à 17 heures et Léa à 18 heures.

      I go pick up Max at 17 hours and Léa at 18 hours

      I’m going to pick up Max at 5 pm and Léa at 6 pm.
Multiple Questions in French

(221)  a. Q: Tu es allé où avec quelle intention ?
        you are gone where with what kind of intention
        Where did you go (and) what was your intention with it?

b. A: Je suis allé chez Jean pour lui dire la vérité.
        I am gone to John to him tell the truth
        I went to John to tell him the truth.

• Adjuncts

If both question words are adjuncts, again both readings are possible:

(222)  a. Jean a dormi où quel jour ?
        John AUX slept where which day
        Where did John sleep on which day?

• Lundi, Jean a dormi à Rome, mardi à Nice, mercredi à Cannes.
        Monday John AUX slept in Rome, Tuesday in Nice, Wednesday in Cannes
        On Monday, John slept in Rome, on Tuesday in Nice, on Wednesday in Cannes.

(223)  a. Il est arrivé quand avec quelle intention?
        he is arrived when with what kind of intention
        When did he arrive and what was his intention?

b. Il est arrivé lundi pour rencontrer Jean.
        he is arrived Monday to meet John
        He arrived on Monday to tell John the truth.

In situ question words can also form complex question words and qualify as a single question:

(224)  Y aurait-il une alternative pair à pair, et si oui, de qui à
        CL would have.CL an alternative pair to pair, and if yes, from who to
        qui et d’où à où ?
        who and from where to where
        Would there be a peer-to-peer alternative, and if so, from who to who,
        and from where to where?
        (http://www.poptronics.fr/La-vie-peer-to-peer-mode-d-emploi)

(225)  C’est un vol qui va d’où à où ?
        this is a flight which goes from where to where
        From where to where is this flight going?
        (google, 2009/10/15)
Ces voitures vous transportent d’où à où, si on est sur une petite île?

(226) From where to where do these cars transport you if we are on a small island?

(227) Two different arguments:

*Qui et quoi fait ?

who and what do

*Who and what is doing?

(228) Argument and adjunct:

*Qui et quand est parti ?

who and when is left

*Who and when left?

The coordinated question words can appear on the left periphery, or, like in the first type, in situ.

• Arguments

(229) Extracted:

Quand et où a eu lieu le dernier concert ?

when and where AUX had place the last concert

When and where did the concert take place?

2.5.3 Coordinated question words on the left periphery

Now we consider multiple question in which question words are coordinated on the left periphery. Unlike in Hungarian, in French the coordinated question words must have identical functions, otherwise the question is ungrammatical (see, however, Section 2.5.6):
2.5. Multiple Questions in French

(230) Extracted:

Quel âge et quel grade a M. Martin ?
which age and which rank AUX Mr Martin

How old is Mr Martin and what is his rank?

The coordination of two arguments with the same function, on the other hand, cannot always be considered as a multiple question, since it does not contain two information gaps in the sentence. Its function is rather the enlargement of the domain set of the question, for instance, to cancel the restriction imposed by qui (who) that the answer has to be [human +].

(231) Qui ou quoi a favorisé l’allaitement maternel ?
who or what AUX approved the breastfeeding maternal
Who or what did approve of maternal breastfeeding?

Note that although quoi is a weak form that cannot be extracted unless it is the complement of a preposition (see Chapter 6), it can be extracted if it is coordinated. Since neither que nor quoi can be extracted in French, the form qu’est-ce qui is used in these contexts. It can also be coordinated with qui as a subject:

(232) Qui ou qu’est-ce qui vous fait rire ?
who or what you.Acc makes laugh
Who or what makes you laugh?

(233) Mais sur qui et sur quoi, au juste, s’exercerait son autorité ?
but on who and on what, to just, CL.applied his authority

But to who and to what did his authority apply?

(234) Mais à qui et à quoi a-t-on jamais été capable
d’assigner un rôle si précis ?
of assigning a role so precise
But to who and to what have we never been capable of assigning a so precise role?
Tu es jaloux de qui ou de quoi?
Who or what are you jealous of?

Elle se sentait parfaitement épuisée et elle contemplait les petites rides naissantes au coin de ses paupières, de sa bouche, en se demandant ce qu'elles signifiaient, de qui ou de quoi elles pouvaient venir. She felt extremely exhausted and was contemplating the little wrinkles on the corner of her eyelids and mouth, asking herself what they meant and from who or from what they could come.

As the complement of être (to be, the copula), où (where) and qui (who) can be coordinated:

Où et qui serions nous?
where and who would be we
Who would we be and where?

The other possibility is the coordination of two adjunct question words. Let us consider other examples with various combinations of adjunct question words (like in Hungarian, the pair-list reading is sometimes also available):

- quand-pourquoi (when-why)

Pourquoi et quand avez-vous décidé d'arrêter vos études universitaires?
Why and when have you decided to stop your studies?

– J’ai décidé de les arrêter en 2001, parce que je n’en voyais plus aucune perspective.

I decided to stop them in 2001, because I couldn’t see any perspective thereof anymore.

(239) Pair-list:

a. Q: **Quand et pourquoi** voir-on circuler des trains sans voyageurs?

When and why can we see trains without passengers?

(SNCF ; http://www.infolignes.com/article.php3?id_article=3505)

– A: Par mesure de sécurité, chaque matin avant les premières circulations commerciales, un TGV-balai effectue un aller-retour sur toutes les lignes à grande vitesse; après un incident, une rame peut repartir à vide quand les voyageurs ont été orientés vers une rame de substitution; en période de vacances, la SNCF fait venir un grand nombre de trains de province dans la capitale pour assurer tous les départs, etc.

For security reasons, every morning before the first commercial trips, a TGV-broom carries out a return trip on all the lines with high speed; after an accident, a train can leave empty, if the passengers have been oriented towards a substitution train; during the holidays, the SNCF (French State Railways) bring a big number of trains from the country into the capital in order to assure all the departures, etc.

(238) is answered by a single pair and (239) by a pair-list. The possibility of the pair-list reading suggests that the question words are in the same clause, and that the structure contains constituent and not clausal coordination with ellipsis in the first conjunct. As we will see, the pair list reading is not available in the fifth type, which, in turn, will be considered as clausal coordination.

– *ou-pourquoi* (where-why)

(240) Pair-list:
Où et pourquoi peut-on utiliser les emotes ?

Where and why can one use the emoticons?

(_google, 2009/10/15)

(241) Single pair:

Il ne s’est même pas soucié de la disparition des deux cyclistes quand ils ont filé dieu sait où et pourquoi.

He didn’t even worry about the disappearance of the two cyclists when they have left God knows where and why.

(Claude Simon, L’Acacia, 1989, p. 294, X 1940)

(242) Pourquoi et où se produisent les séismes ?

Where do earthquakes happen and why?

(http://www.irsn.fr)

– pourquoi-comment (why-how)

(243) Comment et pourquoi donc ont fini chez nous la première, la deuxième et la troisième républiques ?

So how and why did the first, the second and the third republics come to an end in France?

(Charles de Gaulle, Mémoires de guerre : t. 3 : Le Salut (1944-1946), 1959, p. 650)

In the above example, the pair-list reading is due to the plurality of the subject and not to the question words i.e. an answer is needed with respect to all the three republics in France.

(244) Comment et pourquoi nous vient cette espèce d’illumination soudaine du passé, je n’en sais rien, mais elle est profondément troublante.

How and why this sudden illumination of the past comes, I have
no idea, but it is highly troublesome.

(Julien Green, Journal : t. 5 : 1946-1950, 1950, p. 142)

(245) **Pourquoi et comment** il a choisi et délimité son
why and how he chosen and shaped his
sujet ?
topic
Why and how did he choose and shape his topic?

(http://www.gbeu.ch/spec/gbeu/download/Dialogue/De la
connaissance historique Marrou notes et commentaires.pdf)

– *quand-comment* (when-how)

(246) **Quand et comment** les différents groupes humains se
when and why the different groups human CL
sont détachés du tronc commun, la question reste
are detached from the trunk common, the question stays
controversée.
controversial
The question, when and why the different groups of humans
got detached from the common trunk, is still controversial.

(http://unesdoc.unesco.org/images/0015/001577/157730fb.pdf)

(247) **Quand et comment** je quittai le véranda pour me
when and how I left the veranda in order to me
mettre en marche, je ne sais.
put into march, I know
When and how I left the veranda to start marching, I don’t
know.

(Joseph Kessel, Le Lion, 1958, p. 16)

(248) **Quand et comment** avez-vous découvert la situation
when and how have you discovered the situation
dépendante de cette presque moitié de la France ?
dependent of this almost half of the France
When and how did you discover the dependent situation of this
almost half of France?

(Christiane Rochefort, Ma vie revue et corrigée par l’auteur à partir
d’entretiens avec Maurice Chavardès, 1978, p. 325, Questions)

(249) Si le vieux continent devait tomber tout entier à la
if the old continent had to fall all entire to the
discrétion des Soviets, je ne vois pas comment et
discretion of the Soviets, I see not how and
**quand**, dans notre pauvre monde, l’emporterait la liberté ?
when, in our poor world, CL. bring the liberty

If the entire old continent was to fall into the discretion of the Soviets, I can’t see how and when, in our poor world, it would gain liberty?


(250) **Comment et quand** était-il venu, lui, mourant la how and when was he come, him, dying the veille ?
day before
How and when did he come, him who died the day before?

(Henri Vincenot, Le Pape des escargots, 1972, p. 211)

(251) Je ne sais **ni comment ni quand** est morte ma I PRT know neither how nor when is died my propre grand-mère qui m’éleva jusqu’à presque own grandmother who me.ACC brought up until almost neuf ans. nine years
I know neither how nor when my own grandmother died, who brought me up until I was nine.

(Annie Duperey, Le voile noir, 1992, p. 62, La lionne)

– **où-comment** (where-how)

(252) **Où et comment** le repos trouve-t-il des situations where and how the rest finds he situations privilégiées ? privileged
Where and how does rest find privileged situations?

(Gaston Bachelard, La Poétique de l’espace, 1957, p. 18)

(253) **Où et comment** Patricia aurait-elle eu where and how Patricia would have she had l’occasion et le temps d’apprendre le sens de la the occasion and the time of to learn the sense of the mort ? death
Where and how would Patricia have had the occasion to lean the sense of death?
2.5. Multiple Questions in French

(Joseph Kessel, Le Lion, 1958, p. 239)

(254) Et, d’ailleurs, comment et où
and, as a matter of fact, how and where
voudriez-vous vous exprimer?
would like you CL express
And, as a matter of fact, how and where would you like to express yourself?

(Philippe Sollers, Le Secret, 1993, p. 50)

(255) Où et comment a-t-il passé les vacances de telle ou
where and how AUX he spent the holidays of this and
telle année?
that year
Where and ow did he spend the holidays in this or that year?

(R. Martin du Gard, Souvenirs autobiographiques et littéraires, 1955, p. CX11, 1942)

(256) Il ne suffit pas de savoir où et comment les
It PRF suffice not to know where and how them
trouver, il faut aussi, et surtout, savoir, quels
to find, it must also, and above all, to know, which
documents chercher.
documents to look for
It is not enough to know where and how to find them, one must
also, and, above all, know, which documents to look for.

(H-I. Marrou De la connaissance historique, 1954, p. 76)

(257) S’il voulait savoir où et comment elle passait son
if he wanted to know where and how she spent her
temps, il n’avait qu’à se rendre boulevard
time, he PRF had only to CL go boulevard
Saint-Germain n 1, quatrième étage, porte en face.
Saint-Germain number 1, fourth floor, door opposite
If he wanted to know where and how she spent her time, the
only thing he had to do was to go to boulevard Saint-Germain,
number 1, fourth floor, opposite door.

(Alphonse Daudet, Soutien de famille, 1897, p. 272)

– comment- preposition + qui (how- preposition + who)

(258) Dis-moi comment et par qui tu as su ce que
tell me how and by who you have learned what
tu sais, et je te dirai ce que tu ne sais pas et
you know, and I CL tell.FUT what you FUT know not and
brûles de savoir.
burn  to know
Tell me how and from who you learned what you know, and I’ll
tell you what you don’t know and burns to know.

(Philippe Labro, Des bateaux dans la nuit, 1982, p. 375)

– quand-où (when-where)

(259) Quand on lui demandait où et quand elle était née,
when we her asked where and when she was born,
Dida répondait que c’était à Bagnani, sur l’Anio,
Dida answered that it was in Bagnani, on the Anio,
il y avait bien longtemps, avant même que le roi ne
there was well long time, before even than the king
fut entré dans Rome.
was entered in Rome.
When we asked her, where and when she was born, Dida an-
swered that it was in Bagnani on the Anio, a long time ago,
even before the king has entered in Rome.

(Marguerite Yourcenar, Denier du rêve, 1959, p. 252)

(260) Quand et où l’avait on vu se fixer, s’arrêter ?
when and where did we see it fixing, stopping?

(260) Quand et où l’avait on vu se fixer, s’arrêter ?
when and where CL had we seen CL to fix, CL. to stop
When and where did we see it fixing, stopping?

(261) Mais je ne sais ni quand ni où je pourrai
but I know neither when nor where I can.

manger de nouveau.
to eat again
But I don’t know neither when, nor where I will be able to eat
again.

(Pierre Mendès-France, OEuvres complètes. 1. S’engager. 1922-1943.,
1984, p. 504)

(262) Je ne sais pas quand et où elle est morte.
I don’t know not when and where she is died
I don’t know when and where she died.

(Georges Perec, Entretiens et conférences II [1979-1981], 2003, p. 161)

(263) Où et quand comptez-vous le trouver, maintenant ?
where and when do you expect to find him now?

Where and when do you expect to find him now?

(Alain Robbe-Grillet, Les Gommes, 1952, p. 104)
2.5. Multiple Questions in French

(264) Où et quand a été fabriquée la première china en France ?

Where and when was the first china produced in France?

(Georges Fontaine, La Céramique française, 1965, p. 91)

– More question words

(265) Où et quand et comment, il lui avait raconté sa vie, Dunkerque ?

Where, when and how did he tell him about his life, Dunkerque?

(Jean-Pierre Chabrol, La Folie des miens, 1977, p. 133)

(266) Et qu’ils disent où, quand et comment ils ont pu me voir.

And they should tell me where, when and how they could see me.

(Zoé Oldenbourg, Les Cités charnelles ou l’histoire de Roger de Montbrun, 1961, p. 373)

(267) Les invitations de tel(le) à tel(le) dont les pères-et-mères ont souvent demandé en douce aux instructions bien informées qui est qui et quand et où on fera quoi, la valse des carnets de santé, des bulletin mensuels, des cahiers de reportage, mensuels, ART.INDF exercise books of texte et autres paperasseries pour mineurs. The invitations from this to that, whose fathers and mothers often asked softly from the well-informed instructors, who is who and when and where we will do what, the waltz of the health booklets, monthly reports, textbooks and other paper chores for minors.

2.5.4 Coordinated question words in situ

- **Arguments**

  The question words can also appear in situ:

  (268) Le concert a eu lieu quand et où ?
        the concert AUX had place when and where
        When and where did the concert take place?

  (269) M. Martin a quel âge et quel grade ?
        Mr Martin AUX which age and which rank
        How old is Mr Martin and what is his rank?

- **Adjuncts**

  Two adjuncts can also be coordinated sentence-finally. In this latter case, however, the structure is difficult to distinguish from the biclausal coordination of two single questions with ellipsis (270):

  (270) Il a dormi où et quand?
        he AUX slept where and when
        Where did he sleep and when?

  In other examples, the fact that an (optional) argument of the verb can follow the coordinated question words suggests that they are in the same clause:

  (271) Tu pars quand et pour combien de temps dans ton
désert ?
        you leave when and for how much time in your
desert
        When do you leave and for how much time in your desert?

2.5.5 Clausal coordination (with ellipsis in the second clause)

In this fifth type, the structure consists of two coordinated single questions with ellipsis in the second clause. That is why this type is not considered as a true multiple question. In this case, only the single pair reading is usually available, which we can consider as an argument for the biclausal analysis (see however the observations about sentential modifiers below). Like in the previous types, it is possible to extract the question word of the first clause, or to leave it in situ:

(272) Extracted:

   A qui as-tu parlé et pourquoi ?
   to whom have you spoken and why

   Who did you talk to and why?
In situ:

Tu as parlé à qui et pourquoi ?

Who did you talk to and why?

In these examples, the content of the first clause is substituted into the second without the question word:

(274)  A qui as-tu parlé et pourquoi (lui) as-tu parlé ?

to who AUX you spoken and why (him.DAT) have you spoken

Who did you talk to and why did you talk to him?

(275) *A qui as-tu parlé et pourquoi as-tu parlé à qui ?

to who AUX you spoken and why have you spoken to who

Let us see other arguments supporting the biclausal analysis.

• first question resolved

First of all, like in Hungarian, the interrogative complementizer of polar questions can co-occur with a coordinated question word (both in main and embedded clause questions in French). In this case, the second part of the sentence can presuppose that the first question is already answered and can thus refer only to one of the possible answers to the question.

(276)  Est-ce que tu viens et à quelle heure ?

Are you coming, and at what time?

(277)  Est-ce que quelqu’un sera là et qui ?

Will there be anyone, and if so, who?

(278)  Dites-nous si vous venez et à combien, pour qu’on puisse say us if you come and how many, in order that we can s’organiser en fonction.

Tell us if you come, and if so, how many of you, so that we could take it into consideration during the organization (of the party).

(invitation to a party; email, 2009/11/18)

A corresponding structure with right node raising (RNR) is also possible (see Section (2.5.6)): 

Tell us if you come, and if so, how many of you, so that we could take it into consideration during the organization (of the party).

In all these cases, in order to answer the constituent question à quelle heure, qui, à combien, it has to be accommodated (Ginzburg, 1997) that the answer to the polar question is yes.

* Arguments and adjuncts

Another argument for the biclausal analysis is the fact that when both question words are arguments, the second one must be optional. This is possible only in the case of verbs that have both a transitive and an intransitive use:

(280) Qui va parler et de quoi ?
     who FUT talk and about what
     Who is going to give a talk and about what?

The original question can be reconstructed as:

(281) Qui va parler et de quoi va-t-il parler ?
     who FUT talk and about what FUT he talk
     Who is going to give a talk and what is he going to talk about?

The supposed reconstruction of the content of the first question including the question word is ungrammatical:

(282) *Qui va parler et de quoi qui va parler ?
     who FUT talk and about what who FUT talk

The example is ungrammatical with obligatory arguments:

(283) *Qui va faire et quoi ?
     who FUT do and what
     *Who is going to do and what?

When the verb has an optional argument, like partir (to leave), and this argument is coordinated sentence-finally, we suppose that in the first clause the verb is used without an argument, whereas it is used with an argument in the second:

(284) Il est parti quand et où ?
     he is left when and why
     When and why did he leave?
The structure is perfectly possible with an argument and an adjunct (285) or with two adjuncts (286) if the adjunct is sentence-final:

(285) Qui a encore décroché le récepteur et pourquoi ?
who AUX again picked up the receiver and why
Who has again picked up the receiver and why?

(Genet, Jean (1959) Les bonnes. p. 80)

(286) Pourquoi voulait-il l’aider et comment ?
why wanted he him help and how
Why did he want to help him and how?

(François-Régis Bastide, Les Adieux, 1956, p. 52)

According to what we have seen so far, the pair list reading is only possible if the question words are in the same clause. I assume that, in the first four types, the distinction between the single-pair and pair-list readings is enormously context-dependent. In the first two, both readings are possible, but the pair-list reading is preferred if one of the question words is an adjunct. In the second and the third, the single pair reading is more frequent, but the other is equally possible. In order to have a pair-list reading, (at least) one question word has to denote a contextually determined set (D-linked) and it has to distribute over the other. When this is impossible, what we obtain is the single pair reading. As we have already seen, in questions with pourquoi (why) the preferred structure is still the biclausal one, possibly because of the already mentioned accommodation of the first question before pourquoi.

However, in French we can find some exceptions to the observation that biclausal structures license only single-pair answers. Consider the following examples:16

(287) a. Q: Quelles catégories socio-professionnelles ont voté FN
which categories socio-professional have voted FN
dimanche dernier et pourquoi ?
Sunday last and why
Which socio-professional categories voted FN (Front National) last Sunday and why?

b. A: Les paysans à cause de la crise, les retraités à cause de
the.PL peasants because of the crisis, the pensioners because of
l’insécurité . . .
the uncertainty
The peasants because of the crisis, the pensioners because of the uncertainty . . .

16I thank Jean-Marie Marandin for pointing this out to me.
a. Q: En matière économique, qui est responsable et à quel niveau ?
In the economic domain, who is responsible and at which level?

b. A: Les gouvernements sont responsables au niveau national, le Parlement de Strasbourg est responsable au niveau européen...
The governments are responsible at the national level, the parliament in Strasbourg on the European level...

The additional observation is the following: adverbs that are sentential modifiers (that take the whole event/the whole proposition as an argument) cannot appear in the same clause as V or VP level modifiers (for a typology of adverbial modifiers see Geuder (2000, 2004)); they have to be coordinated in a biclausal structure, but can still trigger a pair-list reading. This can also explain why pourquoi cannot appear in the same clause as another interrogative word. The observation about pair-list readings has to be modified as follows: in French, in multiple questions triggering a pair-list reading, the interrogative words have to be clausemates in case they are V or VP level modifiers (if they are adjuncts). Sentential modifiers have to be coordinated in a separate clause.

After this short detour, in what follows, the biclausal structure is illustrated with other examples:

(289) Antiagrégants : Faut-il arrêter, quand et pourquoi ?
Antiagregants: must one stop, when and why?
(http://formathon.fr/fr/spip.php?article58)

(290) Comment l’aurais-je trompé et avec qui ?
How would I have cheated on him and with who?
(Marguerite Duras, Cahiers de la guerre et autres textes, 2006, p. 65)

(291) Comment avons-nous fait notre Rassemblement et pourquoi ?
How did we do our "Rassemblement" (RPF, political party) and why?
Parce que tu ne sais même pas ce que tu voudrais être, où, because you know even not what you would like to be, where, avec qui, comment, tu ne sais rien !

Because you don’t even know what you would like to be, where, with who, how, you don’t know anything!

(François-Régis Bastide, Les Adieux, 1956, p. 193)

Impossible d’imaginer ce qu’elles font, comment elles vivent, avec qui, pourquoi ou en dépit de quoi.

It is impossible to imagine what they do, how they live, with whom, why or in spite of what.

(Philippe Sollers, Le Secret, 1993, p. 228, III)

This type is also possible in infinitival structures:

Contre qui se battre, à présent, et pourquoi ?

Against who fight, presently and why

Against who should one fight and why?

(Zoé Oldenbourg, Les Cités charnelles ou l’histoire de Roger de Montbrun, 1961, p. 560)

Où intervenir et pourquoi ?

Where intervene and why

Where should one intervene and why?

(http://www.monde-diplomatique.fr/2009/03/LEMOINE/16938)

2.5.6 Unexpected coordination

Generally, unexpected coordination means that the apparently coordinated question words do not share the same function in the clause. However, as I will show, all of these ”unexpected” cases can be analyzed as biclausal, elliptical structures, containing right node raising or sluicing.

• Right-node raising (RNR)
  – Infinitival structures

The coordination of question words is also possible in infinitival structures. In the following examples they share the function adjunct:

Quand et pourquoi s’adresser à la Défenseur des enfants ?

When and why turn to the defender of the children

When and why should one turn to the Children Protection?
Quand et pourquoi crypter/chiffrer ses connexions ?
When and why should one encrypt his/her connections?

Pourquoi et où utiliser des logiciels libres en Guadeloupe ?
Why and when should one use free softwares in Guadeloupe?

Où et pourquoi acheter du vin bio ?
Where and why should one buy organic wine?

Qui et comment suivre sur twitter?
Who should we follow on twitter and why?

Qui suivre et comment suivre sur Twitter ?
Who follow and why should we follow them on twitter?

Comment et où habiter s’il n’existe aucun lieu enchanté au milieu des destructions ?
How and where should one live, if there exist no enchanted place in the middle of the destructions?

The above example is the elliptical equivalent of:

(300) Qui suivre et comment suivre sur twitter?
Who should we follow on twitter and why?

(301) Qui suivre et comment suivre sur Twitter?
Who should we follow and why should we follow them on twitter?

(302) Comment et où habiter s’il n’existe aucun lieu enchanté au milieu des destructions ?
How and where should one live, if there exist no enchanted place in the middle of the destructions?
Comment should one live and where should one live, if there exist no enchanted place in the middle of the destructions?

Qui et quand should we consult and when, in case if a couple has problems having a child?

Who should we consult and when should we consult them, in case if a couple has problems having a child?

Who should we consult and when?

Unlike the ungrammatical example in (228), in which question words with different functions were coordinated within the same clause, the sentences in (300) and (304) are elliptical, and the object argument of the verb is optional. In all the examples in (300)-(304), the verb can be used in one of the clauses without its argument because it is used in a specific context: in a philosophical sense in (302), in a medical sense in (304), and with respect to twitter in (300). Abeillé and Mouret (2010) observe that right node raising of the finite verb is ungrammatical or degraded in French, whereas it is more acceptable in the case of participles or the infinitive. This can be explained by the assumption that the omission of a complement is more felicitous with non-finite verb forms. Compare the above examples with the following ones:
We must know who answered and who hasn’t answered.

(308) ?? Il faut savoir quand Jean et quand Marie viendra. must know when Jean and when Marie come.FUT We must know when John is coming and when Mary is coming.

(309) ?? Qui et quand on consulte quand on a un problem quand on a un problem Who does one consult and when, when one has a problem?

(310) ? Qui et quand consulte-t-on quand on a un problem quand on a un problem Who does one consult and when, when one has a problem?

However, RNR is not completely excluded with finite verb forms either.

– Polar question and constituent question

There are examples which lend themselves to an analysis in terms of RNR, and they contain a finite verb form:

(311) Pourquoi et à qui va-t-elle téléphoner ? why and to who FUT she telephone To who will she telephone and why?

(René-Victor Pilhes, La Rhubarbe, 1965, p. 25)

The non-elliptical version of the example would be the following:

(312) Pourquoi va-t-elle téléphoner et à qui va-t-elle why FUT she telephone and to who FUT she telephone ?

Robert Kubica : “Impossible de prédire si et quand impossible to predict if and when

Interestingly, the interrogative complementizer can also be coordinated with a question word preverbally. This is unexpected as well, since they do not share the same function. The following structures are also analyzed as structures containing the RNR of the verb. Consider the following examples:

* si et quand (if-when)

(313) Robert Kubica : “Impossible de prédire si et quand
2.5. Multiple Questions in French

il reviendra”
he come.FUT BACK
Robert Kubica: impossible to predict if he can come back,
and if so, when.


(314) Le secrétaire général des Nations Unies, Ban
the secretary genaral of the Nations Unites, Ban
Ki-moon, s’est déclaré disponible à se rendre à
Ki-moon, cl.is declared disposed to cl go to
Pyongyang ”si et quand cela sera nécessaire”
Pyongyang if and when this be.FUT necessary
The general secretary of the United Nations, Ban Ki-moon
declared that he would go to Pyongyang if and when this
is necessary.


(315) Est-il possible de savoir si et quand le destinataire
is it possible to know if and when the recipient
a lu mon mail?
AUX read my email
Is it possible to know if the recipient has read my email
and if so, when?

(http://fr.answers.yahoo.com/question/index?qid=20100908024731AAoUYAS)

Compare with the following English example:

(316) If and when such a grouping is born, March 15th 2011
will likely be seen as the date of its conception.

(http://www.economist.com/blogs/easternapproaches/2011/03/protest_hungary)

* si et où (if-where)

(317) Savez vous si et où on peut en trouver
know you if and when one can CL.PART fond
d’occasion pour un budget de 400 EUR?
used for a budget of 400 euros?
Do you now if we can find a used one for a budget of 400
euros, and if so, where? (pianist looking for a piano)

(http://www.partoch.com)

(318) ):accepterais-tu de te renseigner pour savoir si et
accept.COND YOU to CL inform to know if and
où ces jeans pourraient être vendus en Belgique?
where these jeans could be sold in Belgium
Merci!
Thanks
Would you accept to find out if these jeans could be sold in Belgium, and if so, where? Thanks!

(http://forum.aufeminin.com/forum/beaute2/-f23440-beaute2-
Jeans-taille-ultra-basse-ou.html)

(319) Un certain nombre d’examens sont prescrits
a certain number of examinations are prescribed
afin de déterminer si et où le cancer
in order to determine if and where the cancer
bronchique s’est propagé (c’est-à-dire son stade).
bronchial cancer is developed (i.e. its stage)
A certain number of examinations are prescribed in order
to determine if the bronchial cancer has developed and if
so, where (i.e. its stage).

(http://www.e-cancer.fr/les-cancers/cancers-du-poumon/le-
diagnostic)

The interrogative complementizer of embedded polar questions can also
appear in a sequence of question words:

(320) Dans ces conditions personne dans le monde, en
in these conditions nobody in the world, in
particulier personne en Amérique, ne peut dire si, où,
particular nobody in America, can say, if, where,
quand, comment, dans quelle mesure, les armements
when, how, in which measure, the weapons
nucléaires américains seraient employés à défendre
nuclear weapons would be applied to defend
the Europe.
In these circumstances, nobody in the world, in particular, no
one in America can say, if, where, when, how, to what extent,
the American nuclear weapons would be applied to defend Eu-
rope.

(Charles de Gaulle, Discours et messages. 4. Pour l’effort. 1962-1965,
1970, p. 73)

These examples can also be analyzed as clausal coordination with ellipsis
in the first conjunct, all the more so, because, in this case clearly, the
first (polar) question has to be resolved in order to answer the second.
This can be seen in the English translations and if so, where/when?.

(319) Un certain nombre d’examens sont prescrits
a certain number of examinations are prescribed
afin de déterminer si et où le cancer
in order to determine if and where the cancer
bronchique s’est propagé (c’est-à-dire son stade).
bronchial cancer is developed (i.e. its stage)
A certain number of examinations are prescribed in order
to determine if the bronchial cancer has developed and if
so, where (i.e. its stage).

(http://www.e-cancer.fr/les-cancers/cancers-du-poumon/le-
diagnostic)
2.5. Multiple Questions in French

A more detailed analysis of this phenomenon is, however, left for future research.

- Sluicing

Sluicing refers to such structures in which an interrogative phrase stands for a whole elliptical clause (see Merchant (2001)), like in the following English example:

(321) Someone came, but I don’t know who (= who came).

(322) J'étais content de ce monde et d'avoir été convié, je ne sais pas trop pourquoi ni par qui, à y passer quelque temps.

I was happy with this world and to have been invited, I don’t know exactly why, nor by whom, to spend there some time.

(Jean d'Ormesson, Tous les hommes sont fous, 1986, p. 273)

We concluded above that in French only question words with identical functions can be coordinated. However, we do find marginal examples in which adjuncts (pourquoi, comment (why, how)) are coordinated preverbally with an argument, mostly qui (who), and the adjunct precedes the other question word. Consider the following examples, generally not considered as grammatical by the locutors:

(323) (?/?*) En effet, dans son sac à dos, Ariane transportait une petite somme d’argent en liquide, ses papiers ainsi qu’un téléphone portable. Mais ces biens ont été retrouvés sur les lieux du crime. « But these properties have been found on the places of the crime. »

En effet, dans son sac à dos, Ariane transportait une petite somme d’argent en liquide, ses papiers ainsi qu’un téléphone portable. But these properties have been found on the places of the crime. So, who killed the decent young teacher
and why, without stealing her money? In spite of the investigations lead
in all directions, the mystery is getting thicker and thicker every day.

(http://www.leparisien.fr/faits-divers/les-crimes-inexpliques-l-enigme-du-chateau-de-nice-12-08-2002-2003316830.php)

(324) (??/*) Les intoxications : comment et qui est concerné ?
the.PL intoxications how and who is concerned
Intoxications: who is concerned and why?

(http://www.guide-assurance.com/eviterlesintoxications.htm)

These examples can also be analyzed as the coordination of a verbless ques-
tion (Pourquoi, Comment? (Why, How?)) and a second question. Notice that
the question introduced by qui (who) can by no means assumed to be solved and
possibly this is the reason why the question containing the adjunct can precede the
one containing the argument.

In the next section, we examine the expression of D-linkedness in French.

2.5.7 D-linkedness in French

In this section, we examine how the D-linkedness is indicated on French question
words. As we have already seen in the case of Hungarian, it is possible to express
this in the syntax. As opposed to Hungarian, French has a more rigid SVO word
order that does not allow the placement of any question word in sentence-initial
position, in which it can precede all the others. In French, what we observe is that
a pair-list question, just like the answer to it, is ambiguous between interpretations
corresponding to the D-linked nature of the question words:

(325) Quel groupe est allé voir quel monument?
which group is gone see which monument
Which group went to see which monument?

(326) a. A1: Les linguistes sont allés voir la Tour Eiffel, les psychologues
the linguists are gone see the Tower Eiffel, the psychologists
l’Arc de Triomphe...
the Arc de Triomphe
The linguists went to see the Eiffel Tower, the psychologists the Arc
de Triomphe...

b. A2: La Tour Eiffel, c’est les linguistes, l’Arc de Triomphe,
The Eiffel Tower, it is the linguists, the Arc de Triomphe,
c’est les psychologues...
it is the psychologists...
The Eiffel Tower was visited by the linguists, the Arc the Triomphe
by the psychologists...
2.6 Conclusion

In these examples, (as reflected by the answers) either the set of monuments, or the set of groups, or both of them can be contextually determined, the syntax provides no clue to decide this. Still, in the answer, it is the set of the D-linked question word that gets exhaustively paired up with one element of the set denoted by the other question word (this latter set, on the other hand, can contain elements left out from the pairing). A1 is also ambiguous between the two readings, whereas the left dislocation in A2 implies that the monuments are D-linked.

I propose that the information about the D-linkedness of question words in French comes from the information structure and not from syntax. I formalize this assumption in the framework of LFG in Chapter (9).

2.6 Conclusion

In this chapter, I provided a general introduction of multiple questions, considering the syntactic structures occurring cross-linguistically and the possible answers they license. I argued that biclausal structures do not qualify as true multiple questions; thus they are not extensively studied in this thesis, only in comparison with monoclausal structures. In the second part of this chapter, I presented the Hungarian and French data to be analyzed in this work. Hungarian exhibits three possible structures, which are typically but not exclusively associated with a certain type of reading (pair-list or single pair). The particularities of Hungarian multiple questions include that D-linkedness is explicitly indicated in syntax in the case of multiple preverbal questions, that in multiple questions with one left- and one right-peripheral question word the question words must denote the same type of set, and that question words of unlike functions can be coordinated preverbally.

In French, I identified four structures, differing in the extracted-in situ position of the question words and coordination. The answer expected depended also on the argument/adjunct status of the question words. If one of the question words is an adjunct, the pair-list reading is more frequent, since the biclausal multiple question with final coordination is unanimously answered with a single pair. In French, I also attested examples with apparent coordination of question words with unlike functions (in a coordination, the conjuncts must share the same function in French). However, I argued that they could be analyzed as right-node raising (in infinitival structures) or as biclausal elliptical structures. The type of answer expected by a multiple question (and the organization of the answer) is crucially dependent on the presence of a D-linked question word in the question. It can easily be identified in the first type of Hungarian multiple questions, but not in the other types (where it is not excluded either), and syntactic structure is not revelatory in this respect in French either. Taking into consideration also the variety of syntactic forms of multiple questions, and the fundamentally identical answer possibilities, I conclude that the type and exact structure of the answer, although it has syntactic correlates, cannot be exhaustively identified in the syntax, but the semantic description of the question words, the information structure of the question, and the discourse in which the question is uttered all play a crucial role in the interpretation of multiple questions. The analysis will be conducted in the modular framework
of Lexical-Functional Grammar, since it permits the dissociation of these levels (syntax, prosody, information structure, discourse, etc.). The analysis is presented in Chapter (7). In the next chapter, I examine some typological perspectives and point out what kind of difficulties the study of multiple questions has had to face in different frameworks.
3.1 Introduction

In the previous chapter, we examined different types of structures containing more than one question word. I argued that out of these, only the five monoclausal ones count as true multiple questions. Multiple questions have been extensively analyzed both in transformational and non-transformational frameworks. Interestingly, most analyses consider only the structures M-all-extr, M-1-extr, and M-all-ins, i.e. the ones that do not contain coordination. However, as I showed in the previous chapter, the coordination of question words cannot always be analyzed as clausal coordination. In certain cases it clearly contains coordination on the level of the question words, which means that the structure is a true multiple question and has to be analyzed along with the structures M-all-extr, M-1-extr, and M-all-ins.

The multiple question types are most often named after the terminology of transformational grammars: structure M-all-extr is referred to as multiple fronting, M-1-extr as single fronting, or the ”English type”, and M-all-ins as in situ. However, since the transformational framework (Government and Binding Theory (GB) (Chomsky, 1981), The Minimalist Program (MP) (Chomsky, 1995) is not the approach adopted

\footnote{Works on the coordination of question words include Comorovski (1989); Kazenin (2010); Skrabalova (2006); Merchant (2009); Haida and Repp (2011)}
in this thesis, other category names are chosen, such as *left-peripheral question words*, etc.

We will consider four approaches: the transformational analyses, Optimality Theory, Head-Driven Phrase-Structure Grammar (HPSG) and Lexical-Functional Grammar (LFG). Let us first examine some problems related to multiple questions that these analyses had to account for.

### 3.2 Patterns to be Accounted for

#### 3.2.1 The order of the question words

It has been observed that the order of the question words in a multiple question is not free. In transformational grammars, this phenomenon is described as a result of *Superiority effects*, discussed in the previous chapter. However, the *Superiority effects* of the transformational analyses cannot account for the fact that subject question words tend to precede object ones even in languages in which subjects do not necessarily precede other arguments in declarative sentences. In such languages, like Hungarian or Bulgarian, discourse structure and discourse functions play a more important role in sentence structure than syntactic functions (subject, object, etc.). This is why other hierarchies have been proposed, according to which in multiple questions, subject *wh*-words tend to precede objects and adjuncts. The order of these latter varies from language to language, and from speaker to speaker: Laenzlinger and Soare (2006) argue that in Romanian, adjunct *wh*-words (*when*, *where*, *how*) follow all [human+] interrogative elements, but precede the [human−] ones (*what*) for a group of speakers, whereas for others adjunct *wh*-phrases follow only [human+] subjects, but precede all other [human+] and [human−] *wh*-words. What we observe here is thus most probably a feature of human perception of the world and information structuring, with respect to which speakers tend to start with [human+] and subject entities (subjects being in most cases [human+]) and then go on to objects/adjuncts and [human−] ones. However, speakers do not always follow this hierarchy. In some cases, the object is more prominent than the subject for the communicative needs, for example when the subject is not identified, or its identity is evident. Some languages allow to start the sentence with the object, like Russian, Hungarian or German (needless to say, these are languages in which the rich case system makes it possible to identify the object and the subject independently of syntactic position):

1. **German:**

   Diesen Autor kenne ich nicht.
   
   this.ACC author know I not
   
   I don’t know this author.

2. **Russian:**
3.2. Patterns to be Accounted for

Etot pulover ja nikogda ne pokupala by.
this pulover I never not buy.PST PRT

I would never buy/have never bought this pullover.

(3) Hungarian:

Ezt a színdarabot sokan lát-ták.
this.ACC the play.ACC many see.PST

Many people have seen this play.

In other languages, when the object constituent (theme or patient) is more prominent than the subject, the sentence can be passivized and the supposed agent is often left out. Thus many passive sentences do not have an active counterpart from which they have been "transformed". Such languages are English and French, and passivization is also possible in German. Another way of deviating from the canonical SVO word-order is the preposing of the object (and other complements), which is also possible in English and (spoken) French:

(4) Bagels, John likes.
(5) Sandy, they named their dog.
(6) (spoken) French:

300 euros, ça m’a coûté.
300 euros that cl.aux cost

It cost me 300 euros.

(7) (spoken) French:

Alexandre, il s’appelle.
Alexandre he cl.refl.calls

His name is Alexandre.

NP preposing is a case of syntactic extraction (and not the OSV word order of the base clause) and often analyzed as topicalization or focalization. It is clearly not associated with only one of these discourse functions.\(^2\)

In multiple questions in which the object is more prominent, the versions in which the object is sentence-initial, preceding the subject, is ungrammatical or less natural than the subject-initial one. Passivization is often applied in this case in order to indicate the prominence of the object:

\(^2\)See Abeillé et al. (2008) for arguments against an analysis of NP preposing in terms of narrow focalization in French.
(8) *What did who buy?

(9) What was bought by whom?

In German multiple questions (where word order is more flexible than in English), both passivization and word order can express the prominence of the object, although the former is preferred to the latter:

(10) Was hat wer gesagt?
    what AUX who said
    Who said what?

(11) Was wurde von wem mitgebracht?
    what AUX from who.DAT brought
    What was brought by whom?

In this respect, French seems to be similar to English:

(12) Qui a dit quoi?
    who AUX said what
    Who said what?

(13) *Qu’a dit qui?
    what AUX said who
    Who said what?

(14) Qu’est-ce qui a été ajouté par qui?
    what AUX been added by whom
    What was added by whom?

However, we can also find examples in French, in which object question phrases precede subject question phrases (note that these are *which-phrases*, which have a less fixed word order even in English):

(15) Quelle information a apporté quel témoin?
    which information AUX brought which witness
    Which piece of information was contributed by which witness?

The structure is even more acceptable if the subject is inanimate:

(16) Quelle information a apporté quel témoignage?
    which information AUX brought which testimony
    Which piece of information has been contributed by which testimony?
3.2. Patterns to be Accounted for

However, according to native speakers’ judgements, the active, subject (agent)-initial sentences can also be used in contexts where the object (theme, patient) is more prominent in the question than the subject (for example, when the asker, looking at the table full of wine, beer, other drinks and food, wonders who brought each of those):

(17) Qui a apporté quoi/quelle boisson à la fête ?
    who AUX brought what/which drink to the party
    Who has brought what/which drink to the party?

Thus, the given situation is a determining factor in the interpretation of multiple questions, and these subtle differences are not always expressed by linguistic means (word order, passivization, etc.).

3.2.2 Difference between two groups of “multiple fronting” languages

According to the seminal paper of Rudin (1988), two groups of languages can be identified within the category A (all question words on the left periphery), Polish, Serbo-Croatian and Czech belonging to the first, Romanian and Bulgarian to the second. Consider again the corresponding examples from Laenzlinger and Soare (2006) (repeated from the previous chapter):

(18) Serbo-Croatian:

    Ko koga/Koga ko voli?
    who whom/whom who loves

    Who loves whom?

(19) Polish:

    Kto jak/Jak kto zareagował na nowosci z Kijawa?
    who how/how who react.PST on news from Kiev

    Who reacted how to the news from Kiev?

(20) Czech:

    Kdo koho/Koho kdo vybral do dalšího kola?
    who who.ACC/WHO.ACC who choose.PST to next round

    Who elected whom for the next round?

(21) Bulgarian:
Koj kakvo dade?
who what give.PST

Who gave what?

(22) Romanian:

Cine ce a scris?
who what AUX written

Who wrote what?

As can be seen in the above examples, the order of the question words is relatively free in the first group, whereas Romanian and, to a certain extent Bulgarian, follow the above-mentioned hierarchy, in that [human+] and subject question words precede [human−] ones, objects and adjuncts. Interestingly, objects tend to be the linearly last one in the wh-sequence.

According to Rudin (1988), the underlying difference between the two groups of languages is that in Polish, Czech and Serbo-Croatian, the question words do not constitute an inseparable sequence, whereas in Bulgarian and in Romanian they do. In transformational terms, this means that in the syntactic structure, only one question word occupies the Spec,CP position and the others are IP-adjoined in the former group, whereas all of them appears in Spec,CP in the latter group. I do not adopt this syntactic structure in this thesis. The other differences between the language groups, supporting this view, are the following:

- multiple wh-extraction from a clause

Rudin (1988) argues that the extraction of multiple wh-phrases from a subordinated clause is grammatical in Romanian and Bulgarian, but ungrammatical in the other group of languages:

(23) Czech:

*Kde kdy* si myslíš, že budeme spát?
where when REFL think.2SG that will.1PL sleep

Where do you think we will sleep when?

(Rudin, 1988, p. 455, 15a)

(24) Bulgarian:

Boris *na kogo* kakvo kaza [ce ste dade]?
Boris to whom what said that will give.3SG
3.2. Patterns to be Accounted for

What did Boris say that (he) would give to whom?

(Rudin, 1988, 451, 7a)

However, as Skrabalova (2010) shows, the long distance extraction of two *wh*-words is indeed possible in Czech:

(25) Czech:

Kdo komu bys myslel, že se nakonec omluvil?

who who.DAT COND.2SG thought that REFL in the end apologized

Finally, what do you think, who apologized to whom?

(Skrabalova, 2010, p. 7, 9)

She argues that in case the clitic (*bys*) follow both/all interrogative words, the expected answer is a single pair and not a pair-list in Czech. The extractability of question words might also depend on their grammatical function. In Skrabalova’s example both question words are obligatory arguments (subject and indirect object), denoting entities, whereas in Rudin’s examples one of the question words is an adjunct, which is more difficult to extract than a complement. I will not pursue the issue further here, only note that the supposed difference between these two groups of languages is no longer so clear-cut as it was claimed by Rudin (1988). Let us have a look at the other arguments.

• evidence for constituent structure:
  
  – clitic position
  
  The position of clitics can often be revelatory with respect to constituent structure and the syntactic position of elements. Rudin (1988) observes that in Bulgarian, clitics cannot split up the *wh*-sequence, but have to come right after it. In Romanian, clitics are attached to the verb, and therefore their position does not tell anything about the status of the *wh*-sequence. In the other group of languages, clitics usually come after the first *wh*-word, which is supposed to prove that the *wh*-cluster does not form a constituent (in Serbo-Croatian and Czech for example, clitics are always clause second, in Polish they can also immediately precede the verb). However, according to Lambova (2003), certain elements can intervene between the first and the second *wh*-phrases in Bulgarian, including different types of (often dialectal) particles, such as the vocative feminine *ma* (non-standard) and masculine *be* (non-standard), *mari* or *mori, bre* or *ba*.

  – parentheticals and related phenomena
  
  Similarly to clitics, adverbs and other parentheticals cannot interrupt the *wh*-sequence in Bulgarian and in Romanian, according to Rudin (1988),
whereas they can in the other group. Interestingly, according to native speakers' judgements, adverbials such as možebi (perhaps), verojatno (probably) and sigurno (surely) can be inserted between the question words in Bulgarian, especially after the first one (Dimitrina Aleksandrova, Elena Borisova, p.c.):

(26) **Koj možebi kakvo e napravil?** who perhaps what AUX.PRS.3SG do.PST Who has perhaps done what?

(27) **Koj verojatno kakvo koga e kazal?** who probably what when AUX.PRS.3SG say.PST Who has probably said what when?

(28) **Koj sigurno kakvo e kupil?** who surely what AUX.PRS.3SG buy.PST Who has surely painted what?

(Lambova, 2003, p. 130-131)

In Lambova (2003), we find other examples supporting this view (also cited by Laenzlinger and Soare (2006)) with intervening parentheticals, like null subject clauses such as kazvam (say), iskam (want). Some other examples are given below:

(29) **Koj, čunkim, kakvo mi e dal?** who PRT what me.CL AUX.PRS.3SG give.PST Who, for God's sake, has given me what?

(Lambova, 2003, p. 336)

(30) **Koj, spored tebe, kakvo pie?** who according to you what drink.PRS.3SG Who, according to you, is drinking what?

(31) **Koj, iskaš da znaeš, kakvo šte kaže?** who want.PRS.2SG PRT know.PRS.2SG what will say Who, do you want to know, will say what?

(Lambova, 2003, p. 342)

- **wh-word order**

Rudin (1988) observes that the order of the question words is more fixed in Romanian and in Bulgarian, than in Czech, Polish or Serbo-Croatian. She argues that a complicated hierarchy can be observed in Bulgarian, according to which question words in the accusative always follow those in the nominative, etc. Such orderings are not present in the other group of languages.
We can conclude again that the difference between the two groups of languages is not so clear-cut as in Rudin (1988). Bulgarian shares some features of the other group of Slavic languages: namely that the order of the wh-words is no longer strictly fixed (at least in spoken Bulgarian) and that the wh-sequence can be interrupted. On the other hand, multiple wh-extraction seems to be possible in Czech, and the separate syntactic position of the first wh-word is not always motivated either.

3.2.3 D-linkedness in pair-list questions

An issue to examine, with respect to multiple question types, is the answer expected to the question, and in the case of a pair-list answer, the way the D-linked/non-D-linked difference is manifested in a given language. For instance, Rudin (1988) does not take into consideration, when classifying languages, the type of answer (single pair or pair-list) a certain question licenses. However, even languages with a relatively fixed word order (like English and French) can express the type of the answer expected in the question (questions licensing a single pair as an answer usually contain coordination):

(32) Pair-list:
    Who left when?

(33) Single pair:
    **When** and **where** shall we meet?

Another important factor, with respect to pair-list questions, is the way languages indicate the asymmetry (or D-linkedness) between the question words, i.e. the fact that at least one of the question words denote a salient, contextually given set, the elements of which are then to be paired up with an element of the other set. It seems that in Hungarian, and in one group of Slavic languages (Polish, Czech, Serbo-Croatian), the order of the question words reflect this difference (the question word denoting the contextually given set precedes the other). This means that the order of the question words is far from being free in these languages. The orders follow semantico-pragmatic rules. In the previous chapter, we presented examples from Hungarian ((55) and (56)).

In Romanian, the usual order in a wh-sequence follows the order: subject > complement > adjunct. However, depending on the salience of the interrogative phrases in the discourse, the inverse word order is also possible in the case of some interrogative phrases (Guţu-Romalo, 2005):

(34) **Cine cui** i-a spus?
    who who.DAT CL.DAT-AUX said
    Who told it to whom?

(35) **Cui cine** i-a spus?
    who.DAT who CL.DAT-AUX said
    Who told it to whom?
However, in other cases, only one wording of the questions is possible (Gabriela Bilbiie, p.c.):

(36) a.  Q: -**Cine ce** a adus la petrecere?
          who  what AUX brought to party
        Who brought what to the party?

        b.  A: -Ion a adus merele, Maria perele și Paul
            John AUX brought apples.ART Mary pears.ART and Paul
            bananele.
            bananas.ART
            John brought the apples, Mary the pears and Paul the bananas.

(37) *Ce cine** a adus la petrecere?
       what who AUX brought to party

If the set of food and drinks is contextually given (the asker can see them on the table, but does not know/cannot see all the people present), it is again question (36-b) that would be asked, and contextually interpreted. However, in this case, the answer could be different:

(38) Merele le-a adus Ion, perele Maria, iar
    apples.ART cl.ACC-AUX brought John, pears.ART Mary, and
    bananele Paul.
    bananas.ART Paul
    As for the apples, John brought them, the pears Mary, and the bananas
    Paul.

Other examples include:

(39) a.  **Cine ce** a zis?
       who  what AUX said
       who said what

       b.  *Ce cine a zis?

(40) a.  **Cine cu cine** a vorbit?
       who with who AUX talked
       Who talked with whom?

       b.  *Cu cine cine a vorbit?

(41) a.  **Cine pe cine** a lovit?
       who MRQ.ACC who AUX beaten
       Who beat whom?
b. *Pe cine cine a lovít?

Consider now the following examples from Czech:

(42) a. Q: **Kdo koho** vybral?
   
   who whom choose.PST
   
   Who chose/elected whom?

   b. A: Marie vybrala Petra, Coline (vybrala)
   
   Marie chose/elected Petr.ACC, Coline (choose/elect.PST)
   
   Pavla.
   
   Pavel.ACC.
   
   Marie chose/elected Peter, Coline (chose/elected) Pavel.

(43) a. Q: **Koho kdo** vybral?
   
   whom who choose/elect.PST
   
   Who was elected by whom?

   b. A: Marii vybral Petr, Pavla vybrala
   
   Marie.ACC choose/elect.PST Petr, Pavel.ACC (choose/elect.PST)
   
   Coline.
   
   Coline.
   
   Marie was chosen/elected by Peter, Pavel was chosen/elected by Coline.

The question in (42-a) is more appropriate in a context in which a group of friends are discussing the results of the elections, asking who voted for which candidate. On the other hand, example (43-a) would be more appropriate to ask when talking about the candidates elected, estimating who (which social class, which region) elected them (Jana Strnadová, p.c.).

The situation is more complicated in Bulgarian. The order of question words obey a certain hierarchy, they thus do not always show the above presented difference between the question words. According to native speakers, multiple questions can be ambiguous between two readings, in one of them the first, in the other the second question word being contextually determined:

(44) **Koj kogo** vizda?
   
   who whom see.PST
   
   Who saw whom?/ Who was seen by whom?

However, in spoken Bulgarian there are ways of changing the otherwise fixed word order. According to Jaeger and Gerassimova (2002); Jaeger (2003), topicalized objects are clitic-doubled in spoken Bulgarian, and in object-initial multiple questions the object question word is obligatorily clitic doubled, otherwise the question is ungrammatical:
Chapter 3. State of the Art

(45) **Koj** kakvo donese?
    who what bring.PST
    Who brought what?

(46) **Kakvo/Koe** koj go donese?
    what/which who CL bring.PST
    What/Which was brought by whom?

(47) *Kakvo koj* donese?
    what who bring.PST

Neverthelesss, when the inherently D-linked question word koe (which) is used, it
does not have to be clitic doubled:

(48) **Koe** ot kogo e?
    which from whom is
    Which is from whom?

Clitics play a role in the interpretation of Czech multiple questions as well. In
Czech, clitics come in the second position of the clause. In subordinated clauses,
they follow the complementizer (˘ze):

As Skrabalova (2010) shows, in questions, clitics follow the question words, in
multiple questions they follow the first question word:

(49) **Komu** jsi dal tu knihu?
    who.DAT CL.AUX.2SG give.PST this book.ACC
    To whom did you give this book?

(50) **Komu** jsi co dal?
    who.DAT CL.AUX.2SG what.ACC give.PST
    What did you give to whom?

A homonym form of the complementizer ˘ze can appear on the left periphery of
the sentence:

(51) **Kdo** ˘ze mu dal tu knihu?
    what.NOM that CL.DAT.3SG give.PST this book.ACC
    Who gave him/her this book?

Skrabalova analyzes this form as a particle, which is homonymous with the com-
plementizer. It plays a pragmatic role in spoken language, expressing astonishment,
or the fact that the asker has already heard the answer, but does not remember it.
Skrabalova proposes to analyze ˘ze in this use as a focus particle (which is supported
by the assumption that question words are often associated with foci). There are
some cases, however, where the focus particle and the clitics can follow more than one question word in multiple questions:

(52)  \( \text{Kdo (že) si koho váží víc?} \)

\( \text{what.NOM (PRT) CL.REFL who.ACC appreciates more} \)

Who appreciates whom more?

(Skrabalova, 2010, p. 17, 24a)

(given x and y, is it x who appreciates more y or is it y who appreciates more x?)

(53)  \( \text{Kdo koho (že) si váží víc?} \)

\( \text{who.NOM who.ACC PRT CL.REFL appreciates more} \)

Who appreciates whom more?

(Skrabalova, 2010, p. 17, 24b)

(for each \( x \), \( x = \text{person} \) and for each \( y \), which \( x \) appreciates which \( y \)?)

In example (53), as explained by Skrabalova, the existence of a pair of individuals is presupposed, and the question does not refer to their identity, but rather to the (direction of the) relationship between them. On the other hand, in example (52), the answer expected is a pair-list, in which each member of a set of people is to be paired up with a person whom they appreciate. According to Skrabalova, thus, only in those cases can two interrogative words precede the clitics and the particle, when they denote the same type of set, since otherwise the relationship between them would not be reversible. This would suggest, following from the clause-second position of the particle and the clitics, that the two question words form one constituent and occupy the same position together. This analysis has been proposed also for Romanian (Rudin, 1988; Laenzlinger and Soare, 2006).

However, in spoken Czech, the particle and the clitics can even interrupt a \( \text{wh} \)-sequence of three question words (follow two but precede one of them), even if the question words do not denote the same type of set (one is an argument, the others are adjuncts):

(54)  \( \text{Koho kam že jsi kdy zavezl?} \)

\( \text{whom where PRT CL when take.PST} \)

Whom did you take where and when?

The status of the particle \( Že \) in Czech is thus not easy to determine. As the above data show, it clearly does not constitute any evidence for the syntactic structure proposed by Rudin (1988), since it can also follow the second \( \text{wh} \)-word in a sequence of three \( \text{wh} \)-phrases, nor is it true that whenever it follows two \( \text{wh} \)-phrases, the question words form a constituent and the structure triggers a single-pair reading referring to the direction of the relation of the entities denoted by the question words. As can be seen in (54), it can follow two question words even if they do not
denote entities and there is no reason to consider that they form a single constituent without the third question word. Although a complete account of the particle is beyond the scope of this thesis, we can follow native speakers’ judgement (Jana Strnadová, p.c.) in concluding that it expresses that the locutor was already aware of the answer to the question, but has forgotten it, and thus asks a reminding question. This is not the same as an echo question, since it can be uttered much later than the previous question, or there has to be no previous question at all, a reminding question just indicates that the asker has forgotten something he knew before. Similar particles exist in French, German and in Hungarian. Interestingly, they are both homonymous with another frequently used adverb or particle, like in Czech. Consider the following examples (the meaning is particularly transparent in German):

(55) Hungarian:

Mikor *is* volt a konferencia?
when PRT(= too) was the conference?

When was the conference again?

(56) German:

Wann war *noch mal* die Konferenz?
when was PRT(= once more) the conference?

When was the conference again?

(57) French:

C’était quand *déjà* la conférence?
it was when ADV(= already) the conference

When was the conference again?

However, if this line of analysis is essentially correct, Skrabalova (2010)’s approach claiming that *že* is a focus particle has to be rejected. Even though it follows question words, which are usually associated with focus, based on its semantico-pragmatic role, it refers rather to already known, discourse-old information, which is not the usual property of focused constituents.

### 3.2.4 Lack of multiple questions

Some languages seem not to allow multiple question formation at all. For instance, in Italian (Calabrese, 1984), neither multiple foci, nor multiple questions are al-
3.2. Patterns to be Accounted for

Calabrese (1984) explains this with the fact that narrow focus is situated on the right periphery of the sentence, and thus the clause-initial position is incompatible with this discourse function. More generally, Italian could constitute indirect evidence for the focal nature of question words. On the other hand, Frascarelli (2000) quotes marginal examples with multiple foci in Italian:

(60) ?? Chi ha dato che cosa a MARIA?
    who AUX given what to Maria
    Who gave what to Mary?

(Frascarelli, 2000, p. 92)

(61) ?? GIOVANNI ha dato un bacio a MARIA (non Antonio a Luisa)
    Giovanni AUX given a kiss to Maria (not Antonio to Luisa)
    JOHN gave a kiss to MARY (and not ANTONY to LOUISA).

(Frascarelli, 2000, p. 91)

Note that both of the above examples belong to a somewhat special case. In (61), two prosodically (and most probably also syntactically) prominent/highlighted constituents are contrasted to a pair of parallel constituents in a second clause. Although it consists of two parts, they form a complex focus (Krifka, 1992), not two single independently focused elements. As for (60), sentences in which both interrogative and non-interrogative foci appear can differ from those in which we find more than one focused constituent or more than one question word (usually it is already the second occurrence of the focused constituent), see Chapter (5) for more discussion.

Other counterexamples are listed by Renzi et al. (2001), in which interrogative phrases can be coordinated, forming multiple questions:

(62) Che cosa e dove cercare?
    what and where to look

However, according to other native speakers, these are acceptable in informal, spoken language (Nicola Lampitelli, p.c.).
It seems from the above examples that multiple questions requiring a single-pair as an answer are permitted in Italian (especially if the interrogative phrases are coordinated). A complete account of multiple questions in Italian would be beyond the scope of this thesis.

3.3 Previous Approaches

3.3.1 Transformational analyses

Concerning the transformational analysis of multiple questions in general, it is supposed that the derivation of the sentence continues even after reaching its surface structure, so that at the level of LF (logical form) the linear order of elements reflect their relative scopes. In this approach, all types of linguistic information are accumulated in the syntax, like semantics and scope relations. Therefore, in *in situ* or "single fronting" multiple questions, non-initial question words are supposed to move covertly to the left edge of the sentence, where they take their scope. Instead of the counter-intuitive and in many cases *ad hoc* movement analyses, I will adopt Mycock (2006)'s analysis, who shows that scope information can come either from syntax, or from prosody. Thus even *in situ* question words can take scope over the rest of the sentence.\(^4\) We will examine her analysis in a later section.

In the derivational analyses (GB/MP), movements (of elements/features towards the left periphery of the sentence) are supposed to account for the problem of surface sentence structure and predicate-argument locality, *i.e.* the fact that some elements do not appear in the proximity of the predicate with which they are in a syntactic relationship (*ex. case*), but on the periphery of the sentence. One example is a question:

(65) German:

Welchen Film schaut Peter sich an?

which.\textit{ACC} film see Peter him.\textit{DAT} VM

Which film is Peter watching?

\(^4\)Not all analyses assume that question words take part in scope relations, see for instance Ginzburg and Sag (2000). In this case, the aim of covert movement would be the indication of interrogativity at the beginning of the sentence.
The constituent *welchen Film* is the object of *sich anschauen* and is in the accusative case, but as an interrogative phrase, it appears at the beginning of the sentence, informally, in order to indicate that the sentence is a (non-echo) question. This opposition (that an element belongs to two places at the same time, but is actually pronounced only once) is resolved in the transformational analyses via movement. In some languages, all question words obligatorily "move" to the left periphery, in others none of them do, and in the third type one and only one can move. However, in the transformational analyses it is supposed that all interrogative phrases end up on the left periphery of the sentence (in non-echo questions), to take their scope. The ones that are not pronounced there, undergo covert movement. Covert movement is the mechanism proposed to account for scope-phenomena in *in situ* languages like Japanese, or for *in situ* question words in the English-type languages. It is not my aim here to question the intuitiveness of such models; however, we should notice the sometimes ad hoc explanations and mechanisms introduced in these frameworks in order to capture the complicated facts. This is reflected by the multiplicity of analyses proposed, which sometimes even modify the assumptions of the theory itself in order to accommodate certain data.

As Laenzlinger and Soare (2006) show, there have been three main approaches to multiple wh-movement in the transformational literature.

The first group is that of the adjunction analyses, in which the wh-words right-(or left-) adjoin to CP or IP nodes. An analysis in these terms was proposed by Rudin (1988). The main lines of Rudin’s analysis of multiple questions has been presented above. She was the first to acknowledge that not all "multiple fronting" languages share the same structure and presented mostly syntactic reasons to show that these languages have two different syntactic structures. Unfortunately, she does not discuss the interpretations associated with these structures, *i.e.* the types of answers expected to these questions, which could contribute to a better clarification of the issue. The structures she proposes are the following. In the first, all wh-words are in Spec,CP, whereas in the other type, one occupies the Spec,CP position, while the others are in an IP-initial position, adjoined to IP. In the first case, (proposed for Romanian and Bulgarian), the wh-words form an inseparable cluster, whereas in the other languages the linearly first wh-word can be separated from the rest of the sequence, by intervening clitics, adverbs, or auxiliaries (see Hungarian). However, as shown above, Rudin’s arguments can be falsified, mostly by referring to contemporary spoken language in the case of Bulgarian and Czech. In Bulgarian, the wh-sequence can be interrupted by adverbs, whereas in Czech, clitics and particles can even follow the linearly second interrogative word. From this it follows that the proposed split between these languages is no longer as clear-cut as it is suggested by Rudin’s approach. The difference between them is better analyzed as gradual.

The second is the multiple specifier approach, either at the CP, or at the IP

---

5 The previous analyses proposed that the wh-words are in some S-external position. In the case of multiple questions, the question words were supposed to occupy the same position or to appear in neighbouring positions (S’, S, or Spec,CP, and C respectively).
level, and characterizes mostly the Minimalist Program, where the main motivation behind the movement of elements is the so-called feature-checking, i.e. heads (such as I and C) attract elements carrying certain features (like WH and Foc(us)) that check their features against those heads locally. As for the multiple IP specifier approach is concerned, some authors, Dobrovie-Sorin (1994), for instance, account for the strict order of wh-words (following animacy constraints, for instance in Romanian) by proposing that first the linearly first wh-phrase (in most cases the subject) moves to the Spec,IP position, and then the other wh-phrases (objects or adjuncts) tuck in underneath the first into separate IP-specifier positions, by forming iterable I-bars. This “tucking in” process accounts for superiority effects, and for the (almost always) first position of the subject.

The multiple CP specifier approach is advocated by Bošković (1998, 2002), who argues that in multiple wh-fronting languages, the first moved wh-element is attracted to C to check its wh-feature, while the other wh-elements are attracted to C to check their focus feature. He argues that C in Bulgarian (and Romanian) is a sort of syncretic head, having both an ”Attract one feature” wh-feature and an ”Attract all-F” focus feature. Thus in those languages, all wh-phrases move to Spec,CP positions, the highest wh-phrase first, followed by movement of the remaining wh-elements in order to check their strong focus-feature. After the first wh-phrase has moved to the external specifier of C, the others ”tuck in” as inner specifiers of C. C in the other group of languages attracts only one wh-phrase, whereas the others are supposed to check their focus features at the IP domain.

A slightly different approach is taken by Surányi (2006, 2007), who argues that in Hungarian ”multiple fronting” questions, only the linearly last question word is focused (raising to Spec,FocP), checking both its F and WH feature against the Foc head, but the others raise into multiple Spec,FocP positions above the linearly last one, to check their WH features. He remarks that although the linearly non-last wh-phrases cannot be raised by quantifier- or focus-movement, nor by topicalization, at interfaces (discourse-semantic), they are interpreted as topics. Concerning the ”English-type” questions, he proposes that the sentence-final question word is only optionally focused. When it is focused, it moves covertly to the left periphery of the sentence deriving pair-list answers, similarly to ”multiple fronting” questions, and when it is not focused, a single-pair answer is licensed. Two remarks are due concerning Surányi’s analysis. First, it does not make explicit, how an element can be interpreted as topic at certain interfaces, if it has not been moved by topicalization, a commonly assumed syntactic process in the literature on Hungarian. Furthermore, their syntactic status is not clear either. They are in iterable Spec,FocP positions, without carrying Foc features, checking their WH features inside the FocP projection. Secondly, the author also calls attention to the fact that in Hungarian only one preverbal focus is allowed. However, he claims that in English-type questions, the sentence-final question words can optionally be focused and if so, they move covertly to a preverbal position, even if in multiple focus (declarative) constructions, the second focus is obligatorily sentence-final, just like the second question word in English-type multiple questions.

Although these analyses can account for many problems concerning the struc-
3.3. Previous Approaches

ture and the interpretation of multiple questions, as pointed out above, they stay inconclusive in some other respects. In general, it is not clear why FocP, a discourse-semantically defined syntactic-functional projection can at the same time ”check” the Foc and WH features of elements. The motivation for raising to the multiple specifiers is thus different in the case of the different elements. In addition, we should note that no independent evidence is available to support the claim that some heads in some languages attract all features, whereas others only one. These claims are made upon considering the data in order to account for the same set of data, which leads to circularity in the argumentation.

Bošković (2002) also proposes that multiple wh-fronting languages constitute no separate type on their own, they are dispersed among the other types (French being treated as a separate type, in which wh-movement is optional.) He shows that Bulgarian is parallel to English in that it is always subject to superiority effects (the order of wh-words is fixed), whereas movement is always obligatory in English. The same way, Serbo-Croatian is the ”multiple fronting” counterpart of French, which does not always present superiority effects, just like movement is optional in French. Finally, Russian is supposed to be like Chinese, the former never showing superiority effects, the latter never having overt wh-movement in its surface syntactic structure. The movement of the maximally one question word is motivated by the checking of the wh-feature. The fact that in the above mentioned Slavic languages all the question words appear on the left periphery requires additional explanations, and in Bošković (2002) this is focus-movement, more precisely the checking of the focus feature. The problem with this approach is firstly that superiority effects in multiple fronting languages and the obligatoryness vs. optionality of movement in English, French, Chinese, and in the languages alike are neither logically, nor intuitively related. Secondly, as we have already seen in the preceding section, word order is not always so fixed in Bulgarian as Bošković claims, nor is it freer in Serbo-Croatian, or so free in Russian (as the author himself notes it in a footnote).

The third main axis in the literature is the split-CP approach, which divides the CP domain into a set of different projections and claim that all wh-phrases move to a different one. Such approaches are based on Rizzi (1997). This is followed by Grohman (2006), who claims that in German wh-phrases move to FocP and FP, and are separable by topicalizable elements, as well as by Skrabalova (2010) on Czech, Krapova and Cinque (2005) on Bulgarian and É. Kiss (1992a) on Hungarian, who argues that linearly first/D-linked wh-phrases are universal quantifiers moving to QP/DistP. The problem with this approach is that in some languages (like Romanian and unlike Bulgarian) the wh-phrases cannot be separated by any intervening material. This is accounted for by analyses claiming that wh-phrases form a cluster (Grewendorf, 2001), or a chunk (Laenzlinger and Soare, 2006), occupying the same position.

As can be seen from the above enumerated approaches, movement-based analyses attempt to give a uniform account of question formation in different languages, within a theory which claims that movement as such is an inherent property of hu-
man language. In order to achieve that, they have to assume the covert movement of elements, null elements, cluster formation with invisible wh-phrases, and diverse feature-checking mechanisms changing from language to language and from analysis to analysis, which, after all, questions the intuitiveness, elegance and simplicity of such approaches. In what follows, other (not movement-based) accounts will be presented, which might shed a different light on these facts.

3.3.2 Optimality Theory

Optimality Theory (OT) deviates from traditional transformational grammar in that it introduces a relative conception of grammaticality, stating that expressions are not ungrammatical in themselves, but because they are suboptimal compared to the optimal (grammatical) candidate.

OT was first introduced by Prince and Smolensky (1993), who assume that Universal Grammar consists of a set of soft (violable) constraints on well-formedness, and individual grammars are constructed by the reranking of these constraints. The expression that satisfies the highest or higher ranked constraints is the optimal, thus, the grammatical one, the others are then considered suboptimal, thus ungrammatical.

It should be noticed that OT, as such, is not a linguistic theory in itself, but rather an approach that does not interpret elements of a system in an all-or-nothing fashion, but rather ranks them relative to each other. When applied in linguistics, it functions as a metatheory, since several syntactic theories can lend themselves to it as its grammatical formalism.

The architecture of an OT grammar is the following. It consists of an Input that contains, partly based on the lexicon, predicate-argument structures, the necessary lexical items, information, illocutionary (Legendre et al., 2001) and functional features, like tense, aspect, or WH+/− (Bakovic and Keer, 2001). Out of this input, Gen (a universal structure generator) generates several candidates that correspond, more or less, to the input. These candidates then undergo the Evalutation process, which, taking into consideration the interaction of the language-specifically ranked constraints, selects the optimal candidate (Kager, 1999).

A general problem concerning OT and potentially a number of other theories is the question of universality and independence in the architecture of the grammar. Constraints, although they are supposed to be universal and capture facts about all the languages possible in the world via the different rankings, are still defined based on one (or a few) particular language, in order to account for the facts of that same language. This, in turn, leads to circularity in the argument and has little explanatory adequacy.

Concerning question formation and the analysis of multiple questions, most OT analyses are based on a derivational framework, Government and Binding Theory or The Minimalist Program. One of the first analyses was proposed by Grimshaw (1997), who accounts for English interrogative phenomena (subject-auxiliary inversion, the fronting of wh-words and some auxiliaries and do-insertion) with the help of, among others, the following constraints:
3.3. Previous Approaches

- Op-Spec: operators are in specifier positions
- Ob-Hd: obligatory heads
- Stay: economy of movement
- No-Lex-Mvt: a lexical head cannot move

Based on these constraints, Grimshaw (1997) accounts for the fact that in English lexical verbs do not "move", but since the *wh*-operator appears in the specifier of CP, its head position must also be filled, this is why dummy *do* is inserted in case there is no other auxiliary (modal or aspectual) that could front to fill that position. As we can see, the constraints, which are supposed to be universal in that they can be used in the OT grammar of other languages, clearly follow the properties of English interrogatives, and are thus less applicable to languages in which, from a syntactic point of view, there is no difference between lexical verbs and auxiliaries, or in which operators are not clause-initial.

In later works, attempts were made to exploit the different constraint rankings in order to show that they can account for a greater variety of languages. Müller (1997) proposes an OT account for cross-linguistic question formation and partial *wh*-movement in German and Hungarian, Legendre et al. (1998) concentrate more on English and Chinese. Ackema and Neeleman (1998a, b) investigate the structure of multiple questions cross-linguistically from an OT perspective. Let us examine their proposal in more details. The representative languages of the different types are English, Bulgarian, Czech, and Chinese/Japanese. As we have already discussed, in English, the highest *wh*-element and an auxiliary is fronted; in Bulgarian, all *wh*-phrases (according to one approach) move as a cluster to the specifier of the highest functional projection; in Czech, the verb stays *in situ* and the *wh*-phrases are adjoined to the highest VP projection; finally, in Chinese and Japanese nothing moves. Other important assumptions are the following:

- functional heads mark their complements (for example, a *wh* operator in the specifier of a functional projection can mark the verb in the head position, and through the verb, the whole clause as interrogative); a proposition must therefore be question-marked, and since its syntactic representation is the VP, the VP is question-marked by the *wh*-operator (the most straightforward case is that of the complementizer *if*)

- there is no rigid skeleton in sentence structure, the functional architecture varies across languages and constructions (although the basic X-bar principles are part of, thus respected by Gen)

- functional projections are only present if their head position is filled

The constraints used are the following:

- Q-Marking: in a question, [Q+] feature is to be assigned to the constituent corresponding to the proposition
• Q-Scope: \( [Q+] \) elements must c-command the constituent corresponding to the proposition

• Stay: Do not move, economy of movement

The proposed hierarchy for English is: Q-Marking > Stay > QScope. Since Stay is higher ranked than QScope, only one wh-expression moves, to question mark the clause, but the others stay in situ and do not front in order to indicate their scope. In Bulgarian, QScope ranks higher than Stay: Q-Marking > QScope > Stay. Therefore, all wh-words front, as an indivisible constituent cluster to the specifier of some functional projection. (We have already seen above that the cluster approach is not always adequate.) In Czech, the order of the constraints is the following: QScope > Stay > Q-Marking, which results in structures in which all wh-elements are VP-adjoined and take scope over the whole proposition, but do not form a cluster. In Chinese and Japanese, Stay outranks both other constraints, which can come in any order, the result will be the same: all wh-elements will appear in situ.

There are, however, languages that do not correspond to any of the above described patterns. One is Irish (and perhaps Italian) that do not permit multiple questions at all, and French, in which a certain optionality can be observed between the English and the in situ type (see Chapter (6)):

\[(66)\]

\begin{align*}
\text{a. Où est-ce que Jean est allé en été ?} \\
\text{where COMP.INT Jean AUX gone in summer} \\
\text{Where did John go in the summer?}
\end{align*}

\begin{align*}
\text{b. Jean est allé où en été ?} \\
\text{Jean AUX gone where in summer} \\
\text{Where did John go in the summer?}
\end{align*}

In order to account for the problem of Irish, an additional constraint, Parse is introduced, which says that elements in the numeration must be realized. If Q-Marking is ranked higher than Parse, the null parse will satisfy all the constraints except for Parse, whereas all the other multiple question types would commit more and higher violations, thus the null parse will be the optimal candidate.

In the case of French, there are several paths to take. First, it can be assumed that written/standard and colloquial French constitute different grammars. In addition, the equal ranking of two constraints could also account for the optionality: Stay >> Q-Marking > QScope. A third possible modification would be to claim that variants with and without a complementizer que\(^6\) or the grammaticalized interrogative expression also in the C position est-ce que do not compete with each other.

\(^6\)In colloquial and substandard French, it is possible to extract the question word without subject-verb inversion (interrogative verb form), or even to insert a complementizer after it:

\[(i)\]

\begin{align*}
\text{Où tu vas ?} \\
\text{where you go} \\
\text{Where are you going?}
\end{align*}
since they belong to different numerations. Hence, again, there is no optionality at all.

It is obvious from these movement-based OT analyses that all take for granted that movement is a necessary syntactic operation in human language. This is reflected by the wording of the constraints (Stay, No Lexical Movement, etc.), and is assumed by the way the constraints are supposed to function. However, as was discussed above, movement-based analyses have some serious shortcomings, which prevent them from serving as an adequate framework for a cross-linguistic study.

Nevertheless, this does not mean that OT, as a metatheory, cannot be used in hand with other linguistic theories. A different approach in the OT-literature is alignment syntax, which eliminates X-bar type and any other hierarchical structure, and uses alignment constraints, requiring that constituents be aligned with respect to each other (left or right) or with respect to domains. See Newson (2004); Newson and Maunula (2006); Newson (2008, 2010) for more detailed descriptions of alignment syntax, and analyses conducted in this framework. Alignment syntax is also exploited by Payne and Chisarik (2000), who account for the relative order of the interrogative elements, focus, and the negative element, and their complementary distribution in the preverbal position in Hungarian. Although their analysis is compelling, the same problem emerges in connection with it, as with some minimalist-based analyses: the constraints are designed based on some phenomena of a particular language, in order to account for the same phenomena.

Another non-derivational direction in the literature is that of OT-LFG, which amalgamates elements from the architecture of both theories. The obvious difficulty when constructing an OT-LFG Grammar is the combination of the different dimensions they are conceived in. An OT Grammar consists of linear, consecutive steps from the input till the selection of the optimal candidate, whereas LFG assumes structures parallel to each other. A possible OT-LFG Grammar is presented briefly based on Bresnan (2000); Sells (2001); Kuhn (2001). Within this framework the input corresponds to an underspecified f-structure, which represents the main information a sentence expresses (predicate-argument relations, lexical items and the basic functional features). As the correspondences with c-structure nodes are not indicated at this stage, the input is considered language-independent. A typical input is illustrated in Figure (3.1):

```
Figure 3.1: Input OT-LFG
```

(ii) Ou que tu vas?
where C you go
Where are you going?
The grammar (Ginviol) is based on an LFG-type set of rewrite rules that yield various c-structures. This component ensures that all candidates respect basic linguistic assumptions and are possible expressions in a human language. Compared to traditional X-bar rewrite rules, here the key factor is the optionality of practically all the nodes in the tree and the assumption that any daughter node can become the head of a phrase. The structure generator (Gen) can be defined as a function between the set of f-structures and that of the candidates (the power set of the analyses in Ginviol) that correspond to possible structures generated according to Ginviol. In this version, thus, there is an abstract universal component that allows a wide range of analyses and instead of generating the candidates, Gen has a mapping role between an input f-structure and the candidates, which always constitute a subset of Ginviol analyses. The candidates in this framework are LFG-type annotated c- and f-structure pairs. The evaluation part of the grammar happens in an OT manner. Depending on the constraint-ranking of the language concerned, more and more structures are gradually ruled out until one is qualified as the optimal one. LFG-OT is used as a framework for the analysis of French interrogatives in Gazdik (2008), modeling the optionality between the use of the interrogative complementizer est-ce que and interrogative verb forms in French, as well as question formation with in situ question words in spoken/informal French. The main shortcoming of that analysis is that it makes use of the constraints adopted in movement-based frameworks. Even though these constraints are adjusted to the non-derivational framework, the constraints lose their simplicity, elegance and generality, which would otherwise be needed. This, however, does not rule out a possible OT-LFG approach in which the constraints assumed do not contain the reminiscent features of movement-based accounts, or builds on the above mentioned alignment syntax. In the next subsections we turn to two non-transformational, constraint-based theories, HPSG and LFG.

### 3.3.3 Head-driven Phrase Structure Grammar (HPSG)

Ginzburg and Sag (2000) provide a detailed analysis of English interrogatives, from a semantic, pragmatic and formal point of view. They propose a constructional analysis with a hierarchy of interrogative types and account for a wide range of English examples: Figure (3.2).

The five basic subtypes of interrogative phrases are: polar, subject (su-wh-int-cl) and non-subject (ns-wh-int-cl) wh interrogatives, in situ (is-int-cl) and sluiced (slu-int-cl) interrogatives. Polar interrogatives are built through subject-auxiliary inversion (sai-ph) constructions, whereas subject and non-subject wh interrogatives through head-filler constructions (hd-fill-ph). In situ interrogatives come in two types: direct (dir-is-int-cl) or reprise (repr-int-cl). The former are clarification questions about the identity of an implicit argument or adjunct in the previous context, showing that not only echo-questions can contain in situ question phrases:

(67) (Policeman to suspect:)
And you saw what?

A constructional, HPSG analysis of interrogatives has been adapted to or proposed
3.3. Previous Approaches

Figure 3.2: Hierarchy of interrogative types, (Koenig, 2004, p. 141)
in connection with other languages as well: French (Abeillé and Godard, 1999c, 2002, 2011), Spanish (Metcalfe, 2003), and German (Holler, 2009).

Another important observation of Ginzburg and Sag (2000) concerning wh-words, the authors convincingly argue against the long-standing view that wh-expressions, as syntactic operators, are quantificational, which undermines assumptions claiming that wh-words undergo movement to take their scope. As we have already mentioned above, É. Kiss proposes a quantificational treatment of Hungarian pair-list multiple questions claiming that non-sequence-final wh-phrases are universal quantifiers, since not only do they occupy the same syntactic position as universal quantifiers, but have the same semantic properties as well. On the other hand, Surányi (2006) demonstrates that the universal quantifier analysis is inadequate both from a a semantic (interpretational) and a distributional point of view. Ginzburg and Sag (2000) also present semantic evidence against the quantificational treatment of wh-phrases. First of all, multiple wh-interrogatives are not synonymous with wh/generalized quantifier (GQ) interrogatives, as demonstrated by the following examples:

(68) (It is known that Robin phoned Dale or Dale phoned Robin.)
   a. Q1: -Who phoned whom?
   b. Q2: -Who did each person phone?

   (Ginzburg and Sag, 2000, p. 141, 18-19)

In that particular context, the natural response to Q1 is A1, and to Q2 is A2. The quantifier, contrary to the wh-phrase, cannot be used in disjunctive contexts, i.e. in those in which not the denotation/referent of the entities in the relation is questioned, but the direction of their relation.

Another interesting difference between the two types of questions is the possibility of short, functional answers:

(69) a. Q: -Who does each Japanese woman admire most?
   b. A: -Her karate instructor.

   (Ginzburg and Sag, 2000, p. 14, 22)

(70) a. Q: -In your experience, which person does which Japanese woman typically admire most?
   b. A1: # -Her karate instructor.
   c. A2: -The academic woman admires Banana Yoshimoto, whereas the businesswoman admires Hiday Nakata.

   (Ginzburg and Sag, 2000, p. 141, 23)

The semantic anomaly with A1 in (70-b) is, according to the authors, that since both wh-phrases introduce a parameter, the question is a two-place proposition abstract, thus both “holes” have to be filled in the congruent answers. It has been argued that in the functional reading, the quantifier outscopes the wh-phrase.
3.3. Previous Approaches

However, Ginzburg and Sag (2000) (quoting Engdahl 1980, 1986), offer a non-scopal account to this problem claiming that "the queried property in such a case is not a property of individuals, but rather a property of functions between individuals" and conclude that *wh*-phrases scope wider than generalized quantifiers, since the former scope over the whole proposition(al abstract), whereas the latter scope internal to a proposition. We do not pursue the issue any further here, but note that we consider it now supported that *wh*-phrases are not quantificational, they do not scope over one another, thus they cannot be collapsed with generalized quantifiers, as É. Kiss (1992a) proposed.

Let us now see the characterization of multiple questions Ginzburg and Sag (2000) propose in the HPSG framework. The authors assume that a feature WH distinguishes between the interrogative and the exclamative uses of *wh*-words, in that the WH value of exclamative *wh*-phrases is a quantifier, for example unusual in the following example:

(71) How tall they are!

The quantifier analysis adds to the interpretation of this fact the nuance, that they are unusually or unexpectedly tall. Considering now interrogative *wh*-phrases, the authors propose that every interrogative *wh*-expression constitute two lexical entries, one which is *wh*-specified, and another which is not. The first type can appear in an extracted position, clause-initially in English, and can be modified by expressions like the hell. In situ *wh*-phrases are not *wh*-specified, but can contribute a restriction to the whole question. The authors also point out that whereas clause-initial *wh*-phrases are optionally accented, in situ ones obligatorily bear a pitch accent, this difference being also encoded in their different lexical entries.

Considering now the analysis proposed for multiple questions, there can only be one *wh*-phrase per clause that is WH-specified and appears extracted in a clause-initial position. The other is obligatorily [WH { }] and appears in an unfronted position. However, even though these *wh*-words are not WH-specified, semantically they contribute a parameter (a variable) to the propositional abstract (see Chapter (4)).

The HPSG approach offers an interesting account of the ambiguity in the case of multiple questions embedded under a question. Consider the following example:

(72) a. Who remembers where Mary keeps which book?

b. Bill remembers, where Mary keeps which book.

c. Joe remembers where Mary keeps Aspects, Max remembers where Mary keeps Syntactic Structures.

According to Ginzburg and Sag (2000)'s analysis, the parameter the *wh*-phrase *which book* imports, can be retrieved either in the subclause, or in the main clause, the first resulting in the (72-b), the second in the (72-c) reading. We will come back to the issue in Chapter (7).
However compelling the above sketched analysis is, it is not the HPSG framework that I adopt in this thesis. Nevertheless, I do not exclude the possibility of an HPSG analysis of the data presented in Chapter (2).

First of all, positing two different lexical entries, as Mycock (2006) also remarks, is an ad hoc way of accounting for the accentual and syntactic differences between the *wh*-words in different positions in English multiple questions. It seems that the WH-specified and non-WH-specified lexical dichotomy is based on the observation that in English interrogatives one and only one *wh*-phrase fronts, but it fronts obligatorily, and then it is used to formalize and explain the very same facts. Moreover, it is also not clear why postlexical prosodic information should be encoded in the lexicon. In this respect, Mycock (2006)’s approach is more plausible, claiming that prosody is a way of indicating the focusing of interrogative words, which is obligatory in English, if the *wh*-phrase is not fronted (this is, in turn, an instance of syntactic focusing), and concluding that all *wh*-words must be focused at some level (syntax, prosody, morphology, etc.), and that even within one language different ways of focusing can coexist.

Furthermore, Ginzburg and Sag (2000) represent the single pair answers by adding a uniqueness presupposition to the regular pair-list representation. This means that the default interpretation of multiple questions is pair-list and the ones expecting single pair answers are specific sub-cases of those (the typical examples being *which*-phrases). In this thesis, however, the different readings are related to the information structure of the question and neither the pair-list, nor the single pair reading is analyzed as the default case. Moreover, it is not clear in Ginzburg and Sag (2000)’s analysis, how the asymmetry of the question words is indicated in pair-list questions. Even if the term D-linkedness is not applied, usually one of the question words has to be exhaustively answered and this has to be clear from the representation.

Another argument for the LFG framework is that in spite of Engdahl and Vallduví (1994, 1996), and works by De Kuthy and Meurers, there is still little consensus on the exact integration of the information structure in the HPSG framework and on its interface with syntax, which plays a crucial role in the present analysis.

Based on all these considerations, the present analysis is conducted in the framework of Lexical-Functional Grammar, building on previous work in that framework (King, 1997; Choi, 1997; Mycock, 2006; Gazdik, 2010a; Dalrymple and Nikolaeva, 2011).

3.3.4 Lexical-Functional Grammar

The first in-depth analysis of (multiple) interrogatives in the LFG framework is presented in Mycock (2006). Since this is the framework in which the present analysis is conducted, and in some aspects it builds on Mycock’s analysis, a short summary of her analysis is provided, along with the elements that are kept in the present study and those that are, in my view, problematic.

In her thesis, Mycock aims at providing a new typology of constituent ques-
tion formation cross-linguistically. The name *Constituent Question (CQ)*, used throughout the thesis, already indicates that the analysis proposed, on purpose, deviates from the usual term *wh-question*, usually used in English-centred approaches. Most analyses, conducted in the framework of Government and Binding Theory/Minimalism, take English as their starting point, and assume that movement is an inherent property of human language. Some interrogative words move overtly to the left periphery of the clause (precisely one in English, and all of them in Slavic languages, Hungarian and Romanian), whereas others are supposed to do so covertly, since in the structure that is pronounced they appear in situ, or unfronted position.

The first claim Mycock posits is that without exception all question words are focused. Focusing is related, cognitively, to highlighting an information gap in a question, and as such, it should apply to all human languages. Focusing can be indicated in at least 3 ways: syntactically, morphologically, or prosodically, and more than one means can be used within one single language. *In situ* languages focus question words by means of prosody (and morphology via question particles in written Japanese, which is not obligatory in the spoken language though), and multiple ”fronting” languages focus all question words syntactically, which, nevertheless, does not mean that prosody does not play any role in the focusing of question words in those languages. In this sense, English-type languages constitute a hybrid of the two focusing strategies. One question word per clause is syntactically focused, whereas all the others appear *in situ* and undergo prosodic focusing.

One motivation behind the hypothesis that *in situ* question words move covertly to the left periphery was the question of scope. Question words, as syntactic operators take scope over a whole clause (matrix or embedded) and mark it as interrogative. In the following English example, *what* bears a grammatical function (object) in the lower clause, but appears in the higher one to mark that interrogativity extends to the matrix clause:

(73) **What** do you think Mary put on the table?

Consider now the next example, in which the question word takes scope only over the embedded clause:

(74) John asked [*what* Mary put on the table].

(75) *John asked [Mary put what on the table].

In Japanese, on the other hand, the position of the question word(s) does not indicate if interrogative scope extends only to the embedded or also to the higher clause.

Thus, a study that claims that the notion of movement is irrelevant in the formation of constituent questions, has to provide an alternative theory of scope as well, explaining how only prosodically focused question words can take scope. Mycock’s second hypothesis is that interrogative scope is unequivocally marked by a question phrase, in terms of the way it is focused (be it prosodic or syntactic focusing), and in the case of multiple question words, the scope-marking one is
distinguishable from the others. Another factor to account for is the fact that some languages (apparently) do not allow multiple question formation at all. Such languages are, as we have seen, Italian (Calabrese, 1984) and Irish (Mycock cites McCloskey 1979).

Since, according to the thesis, both focusing and scope-marking can happen at different levels (mostly syntax or prosody), derivational frameworks are not adequate for capturing the basic points of this analysis. This is why the Lexical-Functional Grammar, with its parallel architecture, where prosody can play the same role as syntax in the assignment of discourse functions (such as topic or focus) and scope, is adopted as theoretical framework of the analysis.

The hypotheses are tested in 4 case-study languages. The focusing of question words and scope-marking are tested in three out of these: Japanese (in which all question words appear in situ), Hungarian (in which all of them are supposed to be syntactically focused), English (where there is one clause-initial question word); and an additional one only for scope-marking: Malay.

After motivating the choice of the theoretical framework (LFG), the author offers a brief presentation of the treatment of interrogatives in different semantic frameworks and motivates her choice, that of propositional abstract theory (Ginzburg and Sag, 2000), which she then integrates into the LFG framework. The choice of a unified semantic framework is highly important, since the meaning of (multiple) constituent questions, despite the formal differences, is supposed to be fundamentally the same cross-linguistically. The parameter (the variable) the question word introduces is part of its lexical entry and part of the f- and i-structures in the LFG analysis. According to King (1997), the bare predicate values of the lexical words appear at the level of i-structure instead of the PRED value (containing all the arguments of a predicate) in order to avoid circularity in the assignment of discourse functions. Since the PRED value of interrogative pronouns is PRO, it is their PARAM value that will be focused. Apart from a parameter (variable), another piece of information necessary in the lexical entry of a question word is its restriction that it introduces about the "role-filler": who (human), what (non-human), etc.

The other important proposal in the domain of semantics concerns the scope of interrogative elements. Since the approach to describe meaning adopted in the LFG literature is the glue approach (Dalrymple, 1999), the propositional abstract theory has to be expressed in glue-theoretic terms. The name glue approach refers to the way in which in linear logic "the meanings of the parts of an utterance can be "glued together" to form the meaning of the whole utterance" (Dalrymple, 2001, p. 230). In the glue approach, all lexical entries contain semantic information, an expression called meaning constructor indicating the meaning of the given item and the way it combines with other items in order to get complete structures (like sentences). In the case of entities, the meaning constructor associates a meaning with a semantic structure, whereas the meaning constructor of intransitive verbs associates the meaning (a one-place predicate, possibly a lambda expression) with the logic formula, indicating that if the resource of the meaning of a subject is found, then the meaning of a sentence (a complete semantic unit) can be constructed. I will
not go into further details of the glue approach here. This short introduction was necessary to show how Mycock (2006) treats the problem of scope in her analysis. She introduces two new meaning constructors that can be introduced both at the level of c- and p-structure (syntax or prosody) by a configuration that indicates the scope of interrogativity or question particle. The meaning constructor \([Qs]\) characterizes the question words which delimits the extent of interrogative scope, whereas the meaning constructor \([\text{interrog-scope}]\) belongs to the proposition in the scope of the question word.

Mycock (2006) fully exploits the modular system of LFG, which consists of parallel levels of representation. The structures she makes use of are constituent and functional structure (representing the syntactic component), prosodic structure, represented as a tree diagram, and information structure (as discussed in Chapters (4) and (7)). She also points out the relevance of morphological structure (for instance, in the case of question particles), and provides semantic descriptions as well. Since her analysis is based on the interrelation of these levels and even on new correspondences between them, she proposes new correspondence functions that are crucial in the analysis. These new mapping functions include \(\epsilon\), which maps p-structure to s-structure (since in some cases prosody can influence meaning without the intermediary of other structures), \(\xi\) that maps p-structure to i-structure, a function \(\beta\) that relates the prosodic form of the string to its prosodic structure, and the function \(\gamma\) mapping m-structure to s-structure (reflecting the way morphology can influence meaning.) (Figure (3.3), from Mycock (2006, p. 71))

Generally, Mycock (2006)’s prosodic analysis uses, as primitives, 3 prosodic units, the Prosodic Word, the Accentual Phrase, and the Intonational Phrase and represents the prosodic structure of a sentence as a mirror image of its syntactic structure, as a tree diagram. This representation makes the indication of scope-taking mechanisms and the extent of interrogativity particularly clear. Consider the c- and p-structures of the question What did Charlie put where?: Figure (3.4).
Figure 3.3: Mycock’s architecture
Figure 3.4: Constituent and prosodic structure (Mycock, 2006, p. 106)
Concerning the architecture of the information structure, Mycock adopts that of Butt and King (1996), consisting of four sets, defined by the combinations of two features (new and prominent): Figure 3.5.

\[
\begin{align*}
\text{TOP} & \quad \text{PROM} + / \text{NEW} - \\
\text{FOC} & \quad \text{PROM} + / \text{NEW} - \\
\text{COMPL INF} & \quad \text{PROM} - / \text{NEW} + \\
\text{BACKGR INF} & \quad \text{PROM} - / \text{NEW} - 
\end{align*}
\]

Figure 3.5: I-structure (Butt and King, 1996)

However, as it is shown in Chapters (4) and (7), these sets defined this way seem to be too simplified to account for all the data attested. For instance, the focused constituent is not always new information, its prominence is rather related to the role the sentence containing it plays in the discourse. Moreover, as we will see, it seems unjustified to assign all prominent constituents (for instance, question words) into the focus set at the level of information structure. In this case, an alternative solution has to be provided. We will come back to this problem in Chapter (7).

The first case-study language is Japanese. After giving a detailed introduction of its basic syntactic and prosodic properties, the author concludes that no syntactic focusing takes place in Japanese, but both the highlighting of question words and the delimitation of interrogative scope take place in the prosody. Since pitch accent is a lexical property of certain words, prosodic focusing must happen by different means, by means of pitch range manipulation, an initial pitch range expansion, followed by pitch range compression. The prosody of question words corresponds to the focus prosody in declarative sentences. In addition, the total period of focus prosody in questions (single or multiple) correlates with the scope of interrogativity (i.e. it can extend until the end of the embedded clause or to the matrix clause, in case this latter continues after the end of the embedded elements).

\[(27)\] Naoya-wa [Mari-ga nani-o nomiya-de noz Desu] Yumi-ni
Naoya-top Mari-NOM what-ACC bar-LOC drink.PAST that Yumi-DAT
morishita no.
divulge.PAST Q-PART

‘What did Naoya divulge to Yumi that Mari drank at the bar?’

Figure 3.6: Scope of interrogativity in Japanese


The second language examined is Hungarian. In the data considered in this
3.3. Previous Approaches

analysis, all question words appear in the preverbal “focus” position, thus they are syntactically focused. As we will see in Chapter (5) about Hungarian, focus is indicated by a sharp falling pitch accent on the element in the preverbal position and the deaccenting of all the elements following it (at least, no other pitch accent can follow the focused constituent), constituting a pitch range expansion followed by a low plateau. Interestingly, in a multiple question in Hungarian, only the rightmost, immediately preverbal question word gets prosodically focused, characterized by heavy stress, a pitch accent, and a characteristic sharp fall. The non-final question words are thus not part of the focus prosody, and carry high monotone contour. Mycock (2006) referring to Varga (2002) assumes that this intonation pattern indicates the pointing towards something, the preparation for the pitch accent of the focused constituent. This assumption could, in Mycock’s view, also explain the fact that the question words form an indivisible cluster, since nothing can intervene between the question words.\(^7\) The prosody of the *wh*-sequence is illustrated by the following example and Figure (3.7) (adopted from Mycock (2006)):

(76) **Ki kinek** mutatta be Marit?
who who.DAT introduce.PST VM Mari.ACC
Who introduced Mary to who?

Figure 3.7: The prosody of the *wh*-sequence in Hungarian

(Mycock, 2006, p. 222)

Another assumption in connection with what has been presented so far is that the function of delimiting scope (contributing [Qs] to the meaning of the question) subsumes the function of Sorting key (Kuno and Takami, 1993). According to Kuno and Takami (1993) the Sorting key is usually the leftmost interrogative word, which corresponds to the question word indicating interrogative scope in Japanese. In Hungarian, however, the question word prosodically distinct from the others is the rightmost one, and Mycock argues that it is this question word that delimits interrogative scope and serves as the Sorting key at the same time. Upon these considerations she rejects Kuno and Takami (1993)’s Sorting key hypothesis and adopts instead the Qs Sorting key hypothesis:

\(^7\)In a multiple CQ [Constituent Question], the question phrase which acts as the

\(^7\)We have seen in Chapter (2) that some auxiliaries can interrupt the sequence of *wh*-words.
We will come back to the evaluation of this hypothesis at the end of this section, after a brief overview of the key assumptions of this thesis.

The third language considered in this typology is English, as the representative of single syntactic focusing languages. Focus prosody in English is also characterized by pitch range expansion on the focused constituent followed by a pitch range compression on the following prosodic units. Interestingly, interrogative and non-interrogative foci behave differently in English with respect to focusing, in that non-interrogative foci carry focus prosody, whereas fronted question words are not accented. Lambrecht and Michaelis (1998) stipulate that they are in different syntactic positions, which would account for the fact that the former are accentable and the latter not. However, in situ question words are, on the other hand, accented, which clearly refutes the view that interrogative words are unaccentable. According to Mycock’s analysis, thus, the clause-initial question word is unaccented, but syntactically focused, delimiting, at the same time, interrogative scope, whereas the others are only prosodically prominent (focused).

The rest of the thesis offers an LFG account of other ways of scope-taking, like the wh-expletive construction (for instance, in Hungarian, see Chapter (2), and Mycock (2004)) and bare scope-marking (in Malay). Since this part is not our main concern here, we refer the reader to consult these chapters and the references therein.

Let us now go back to some claims and analyses proposed in Mycock’s thesis, which are, in my view, problematic and have to be reviewed. Since the present dissertation is concerned with French and Hungarian, and Hungarian is also a language treated in detail by Mycock, most of the following comments will be about some problematic issues in the proposed analysis of Hungarian. The main problems can be summarized as follows: the analysis of all question words as foci is not universally tenable; all multiple question types (especially those containing coordination) are not considered; the context and the answer expected to the questions are not included; and questions licensing single pair answers are not considered. Now we turn to the details.

The first problem concerns the inconclusive treatment of interrogative and non-interrogative focus. A basic assumption of Mycock’s thesis is that all question words are focused. In Chapter (7) we will speculate about the focus status of question words and conclude that the matter is not easy to decide and the focus status of question words is far from being obvious. When talking about the general properties of constituent question formation cross-linguistically, Mycock concludes that

"[c]ross-linguistically, differences between interrogative and non-interrogative fo-
cus are attested. For instance, multiple syntactic foci are only grammatical in Hungarian when each one is a question phrase; only a single non-interrogative phrase can be syntactically focused in Hungarian. This motivates a distinction between the two types of foci at the level of i-structure. I propose that all focused constituents are members of the set FOCUS at i-structure, but that this set is divided into interrogative and non-interrogative foci.” (p. 92)

Then, when she presents her analysis of Hungarian, she refers back to this chapter saying

“[g]iven their syntactic position, it is reasonable to assume that the question phrases in a Hungarian multiple CQ are multiple foci because they all appear preverbally. Multiple non-interrogative preverbal foci are not possible in Hungarian. This is not problematic for my proposed analysis because a distinction is made between question focus and other types of focus.” (p. 210)

This argumentation thus seems to be circular, without providing a real explanation to this contradictory phenomenon: how is it possible that (supposing that all question words are focused) two (or more) question words can appear preverbally in Hungarian, whereas this is strictly ungrammatical in the case of non-interrogative focused constituents? Mycock distinguishes between interrogative and non-interrogative foci in the information structure, based on the Hungarian data, and at the same time she accounts for the Hungarian data by claiming that interrogative and non-interrogative foci are anyway different in nature. The argumentation would necessitate some independent evidence in order to posit different categories within the information structure.

Nevertheless, we have seen that in Hungarian, only the rightmost, immediately preverbal question word has the characteristic focus prosody. The prosody of the others are closer to that of topics or other elements in the preverbal domain (see Chapter (5)). In addition, the claim that the question words form an undividable cluster, in which the first question words constitute the preparatory part to the focused one, is not supported by the facts. The question words, as we have seen, can indeed be divided in Hungarian by intervening material, although the phenomenon is not widespread. One typical example is auxiliary verbs:

(77) **Ki akar kit kinek bemutatni?**
    who wants whom who.DAT introduce
    **Who wants to introduce who to whom?**

In addition, embedded multiple questions are not considered in the thesis either. Arguably, the question words in embedded questions (especially those embedded in factive predicates, see Chapter (4)) cannot be analyzed as foci, even if their main
clause counterparts share some properties with foci.

Moreover, the above mentioned facts about the prosody of English interrogative and non-interrogative foci would also support the view that question words are not simply a subclass of focus. Mycock also refers to the example of Italian Calabrese (1984), in which no multiple foci or multiple questions are allowed Mycock (2006, p. 384). However, we can find counterexamples, in which a sentence can contain both a focus and a question word. For a more detailed description of this phenomenon see Chapter (2) and (5) on similar Hungarian examples, but note that this also supports the claim that focus and interrogative phrases do not have the same status and this difference is more than being two subclasses within the focus set at the level of information structure.

Another claim that I criticize here is the Qs Sorting Key Hypothesis. This assumption subsumes two fundamentally different properties: a formal one, i.e. the syntactic or prosodic highlighting of a question phrase, which delimits interrogative scope, and a semantic one, i.e. being the Sorting key in a multiple question, that is, the question word according to which information is organized in the answer. Nevertheless, these two can overlap in languages in which the formally highlighted question word, being the leftmost at the same time, does not only mark interrogative scope, but serves as the Sorting key as well (denotes the salient set which has to be exhaustively treated in the answer). Note that it is a long-standing view that given/already known parts in a sentence usually come before new information and thus integrates the upcoming new information in the common ground. It is also a common assumption about topics that usually they occupy a position on the left periphery of the sentence.

However, based on formal criteria (syntactic position and pitch accent), in Hungarian, the rightmost, immediately preverbal question word has the special, scope delimiting function ([Qs]), and Mycock assumes that it is the sorting key of the multiple question at the same time. She bases this claim on native speakers’ judgements:

"My informants do not regard (46) and (47), nor (48) and (49a-b) as being equivalent in meaning. The difference, they explained, is that the single prosodically focused question word in the preverbal group is what the entire multiple CQ is 'about'. The question phrase which is prosodically and syntactically focused therefore has sorting key status." (p. 225)

However, it is not entirely clear what the informants mean by saying what a question is about. A multiple question can be about the Sorting key, since this question word determines the structure of the answer, but it can also be about the other question word (which is not salient in the discourse), since this is the question word the question really requests new information about.

In Hungarian, the syntactic structure of the (multiple) question clearly determines the structure of the answer. Consider the following examples:
As can be seen from the above examples, in both cases, the answer exhaustively enumerates the elements of the set denoted by the leftmost question word and pairs it up with an element from the set denoted by the other question word. It is thus the leftmost, not the immediately preverbal one that functions as the Sorting key of the question. The rightmost, preverbal one is highlighted in the same way as the focused constituents, since this is the part of the question in connection to which the answer reveals information new to the speaker. Consider another example in which the set denoted by the preverbal question word is clearly not part of the common ground or of the speaker’s knowledge:

(80) Hungarian:

Kinek mi a baja?
who.DAT what the problem.POSS

What is the problem of who? (Tell me about everyone what their problem is.)

Taking these facts into consideration, the Sorting key hypothesis can be kept and Mycock’s Qs Sorting Key Hypothesis must be rejected. It can also be concluded that the prosodic and semantic differences between the preverbal question words in Hungarian indicate that the leftmost (non-preverbal) question words do not have the same focus status at the level of information structure as the preverbal one. Their different status has to be kept even in an analysis claiming that question words are not a subclass of foci at all. We will come back to this problem in
An additional problem with Mycock’s account is that it does not take different types of answers into consideration. She makes it clear from the beginning that she deals only with multiple questions licensing pair-list answers, thus it is not surprising that interrogative structures answered rather by a single pair are ignored by the thesis. However, by looking at the pair-list answers of multiple questions in Hungarian in which the order of the same question words is different, the sorting key status of the leftmost question word would have been clearer. Questions are therefore not to be examined in isolation, only in a context where at least the immediate answer is provided, in order to draw consequences from it.

Another important factor is subsumed by the previous one. A complete treatment of constituent questions and within it, multiple questions, cannot rely on only one type of answer these questions license. Even if polar and echo questions are not treated by a study, multiple questions licensing single pair answers are a necessary part of any study of multiple questions, all the more so because Hungarian multiple questions have two more possible syntactic structures. In one of them, the question words are coordinated in a clause-initial position, in the other, one of the question words is preverbal and the other appears on the right periphery. In order to obtain a complete picture about (multiple) question formation mechanisms in a language, all these structures have to be considered. A possible analysis of all multiple question types in Hungarian is provided in Chapter (8).

In the present analysis, therefore, some of Mycock’s key assumptions are kept, but some of them are rejected. I also support the claim that the formal highlighting of question words (and focused constituents) cannot be reduced only to syntactic mechanisms and that other modules, like prosody or morphology can also contribute to this effect in the same way as syntax. Since focusing/highlighting, after all, is a semantic phenomenon, it does not have to take place only at one formal level. As Mycock notes, ”[i]mpressionistically, a question word is focused in order to highlight the information gap which it represents. Such a lack of a speaker knowledge must ultimately be regarded as semantic rather than syntactic or functional in nature.” (p. 90.) Thus, I adopt Mycock’s analyses on prosodic focusing/highlighting in the different languages.

At the same time, my analysis will differ from Mycock’s, first of all in its scope: it concentrates only on two languages, but aims to describe all the possible multiple question structures in those, along with the answers they license. Secondly, I reject the claim that all question words are focused. One kind of formal highlighting, which characterizes both focused constituents and some (but clearly not all) question phrases, does not mean that these expressions have the same status at the level of semantics or information structure. I will attempt to provide an analysis that accounts for both the common and different properties of foci and the question words. Thirdly, not only are answers important in the analysis of questions, but the whole context in which the question is uttered as well. This is then also a factor to be considered by the present analysis.
3.4 Summary

In this chapter, I presented some controversial issues in the analysis of multiple questions, such as the constraints on the order of the question words, the differences between languages exhibiting multiple questions with more than one left-peripheral question word, D-linkedness in pair-list questions, and the lack of multiple questions in some languages. I examined transformational, OT and non-transformational accounts of multiple questions and pointed out some problems in these approaches. Concerning derivational analyses, I concluded that the mechanisms they introduce are sometimes inconclusive (feature checking, covert movement, etc.) and the hypothesis that movement is an inherent property of human language is erroneous. With respect to OT, I showed that the constraints proposed are often based on the English language, and are not readily applicable to other languages, even if the aim of the whole theory is a cross-linguistic analysis. Moreover, some of the constraints also assume movement, which is problematic in itself. I also mentioned the HPSG analysis of (multiple) questions, which is also built on English data, does not explain how the formal properties of the question can (in some languages), account for the different readings of multiple questions. Finally, I considered Mycock (2006)’s analysis in the LFG framework, and concluded that the hypothesis of syntactic and/or prosodic focusing (highlighting) of questions words is clearly a path to follow. However, I also pointed out some problems in her analysis, for instance the “focusing” of question words in Hungarian, which I attempt to solve differently in this thesis.
Part II

Background
Chapter 4

The Meaning Side: Information Structure and Semantics

Contents

4.1 Introduction .................................................. 142

4.2 Information Structure and History ...................... 143
  4.2.1 Sámuel Brassai and his contemporaries .......... 143
  4.2.2 The Prague School and its followers .......... 144
  4.2.3 Recent developments .............................. 145

4.3 The Basic Architectures ................................. 145
  4.3.1 Monodimensional architectures .................. 146
  4.3.2 Multidimensional architectures ................. 146

4.4 Information Structure in Different Frameworks ....... 148
  4.4.1 Derivational frameworks .......................... 148
  4.4.2 Lexical-Functional Grammar (LFG) ............. 149
  4.4.3 Head-driven Phrase Structure Grammar (HPSG) .. 150

4.5 Information Structure Units ............................ 152
  4.5.1 The Focus ........................................ 152
  4.5.2 The Topic ....................................... 161

4.6 Semantic Approaches .................................... 174
  4.6.1 Proposition set theories .......................... 174
  4.6.2 Alternative semantics ............................. 175
  4.6.3 The structured meaning approach .............. 177
  4.6.4 Questions as propositional abstracts .......... 178

4.7 Possible Answers .......................................... 181
  4.7.1 Congruent and non-congruent answers .......... 181
  4.7.2 Implicative and non-implicative answers ....... 182
  4.7.3 Direct and indirect answers ..................... 183
  4.7.4 Answers to multiple questions .................. 183

4.8 Summary .................................................... 184
4.1 Introduction

Approaches claiming that information structure is an independent level of linguistic analysis base their claim on the fact that the semantic content a sentence describes can be expressed by different formal means in prosody and/or syntax. For instance, the sentence *John is going to Vienna by car tomorrow* has different versions in Hungarian, which are not compatible with the same contexts (Lambrecht (1994) refers to these as *allo-sentences*). In Hungarian, the difference between them concerns mostly word order and prosody, but in other languages only one of the two, and other factors can play a role as well.¹

(1) a. JÁNOS utazik holnap autóval Bécsbe.
   János travels tomorrow car.instr Vienna.ill
   It is JOHN who is travelling to Vienna by car tomorrow.

   b. BÉCSBE utazik János holnap autóval.
      Vienna.ill travels János tomorrow car.instr
      It is to VIENNA that John is travelling tomorrow by car.

   c. AUTÓVAL utazik János holnap Bécsbe.
      car.instr travels János tomorrow Vienna.ill
      It is BY CAR that John is travelling tomorrow to Vienna.

   d. HOLNAP utazik János autóval Bécsbe.
      tomorrow travels János car.instr Vienna.ill
      It is TOMORROW that John is travelling by car to Vienna.

(based on Marandin (2006a))

We can observe that each of these sentences is a possible answer to a different question:

(2) a. Ki utazik holnap autóval Bécsbe?
   who travels tomorrow car.instr Vienna.ill
   Who is travelling tomorrow by car to Vienna?

   b. Hova utazik János holnap autóval?
      where travels János tomorrow car.instr
      Where is John travelling tomorrow by car?

   c. Mivel utazik János holnap Bécsbe?
      what instr travels János tomorrow Vienna.ill
      How (by which means of transport) is John travelling to Vienna tomorrow?

¹The capitals refer to prosodic prominence (pitch accent).
4.2. Information Structure and History

The four sentences have the same semantic content, but formally they differ: the information structure is to account for the relationship between the semantic identity and the different forms. In most theories, the above difference can be captured by the focus-background structure of the sentences, the focus being the prominent constituent answering the question, and the background the part that is common in the question and in the answer. I will define what is meant by these terms in this thesis later in this chapter.

4.2 Information Structure and History

4.2.1 Sámuel Brassai and his contemporaries

This section is based on É. Kiss (1981). One of the first linguistic works that dealt with questions that would belong to the information structure today was a lecture delivered by the Hungarian linguist, Sámuel Brassai, at the Hungarian Academy of Sciences in 1853: Tapogatódások a magyar nyelv körüli [Exploring the Hungarian language], which was further elaborated in a study that appeared between 1860 and 1865 entitled A magyar mondat [The Hungarian Sentence]. An important innovation in Brassai’s linguistic work was that he considered the sentence as the basic unit of language and not the word, this latter being in the center of attention in historical linguistics that time. He argued that the subject-predicate/verb dichotomy of the sentence cannot be considered as a linguistic universal, since the subject is a logical term that gained importance in the description of Western European languages. He observed that the role of the subject cannot be central, as in a number of languages it does not even appear obligatorily in all sentences, and because in some cases it is difficult to define (the constituent in the nominative case, or the sentence-initial constituent...?). From this it follows that the central part of the sentence is supposed to be the verb, the subject being only one of its ”complements” (he used the term adverbs).

However, there is another dichotomy that, according to Brassai, can be discovered in all sentences, which he called the preparatory section (inchoatívum) and main section (zöm) (topic and comment in contemporary linguistics). He showed that this structure characterizes even languages with relatively free word order, such as Hungarian, and the difference between languages can be captured in determining which complements (and how many of them) can appear in the preparatory section of a sentence (in Western European languages, it is mostly the subject, and in German, there can be only one preparatory constituent). He further elaborated his theory in A mondat dualismusa [The Dualism of the Sentence] (1885) and Szörend és accentus [Word Order and Accent]. About Brassai’s work and its impact on

---

²Most probably after medieval Arab grammar, talking about mubdada (beginning) and habar (news) (see Krifka (2008)).
today’s linguistics see É. Kiss (2008b). Unfortunately, since it was published in Hungarian, Brassai’s work remained unnoticed by his contemporaries and in the first half of the 20th century.

Another important figure of the early research on information structure was the German linguist, Georg von der Gabelentz, who, in his articles in 1875 and 1879, argued that sentences are made up of a psychological subject and a predicate. Brassai was aware of Gabelentz’s work (É. Kiss, 1981) and criticized it at important points. First of all, the terms psychological subject and predicate are remainders of the logical subject-predicate distinction, and if the former is defined as part about which the rest of the sentence predicates something, it would be difficult to accommodate in the theory such sentence-initial constituents that do not denote individuals (like certain adverbs). Secondly, Gabelentz does not treat the question whether both the psychological subject and the predicate are both obligatory in the sentence, or one of them can be left out. On the other hand, Brassai explains that only the main part is obligatory, while the preparatory section can be left out, for instance, if it is already known by the hearer, or if the speaker cannot or does not want to include it. Furthermore, contrary to Gabelentz, Brassai also examines how the two parts are divided in a sentence. He argues that the boundary between them is indicated by word order in German and by prosody in Hungarian (the first pitch accent falls on the beginning of the main part). Finally, according to Brassai, Gabelentz does not clearly define how the psychological subject and the predicate can be identified. As for Brassai’s analysis, he claims that the introductory part can be recognized by the fact that it is in the sentence-initial position, that it conveys already known facts to the hearer, and by its preparatory function. In turn, the main part is informative in that it conveys new or partially new information to the hearer.

The later Prague school was also greatly influenced by the work of another XIXth century French linguist, Henri Weil (De l’ordre des mots dans les langues anciennes comparées aux langues modernes, Paris, Joubert, 1844), from whose work basic notions, such as the theme and rheme originate. In addition, he was the first to use the ‘interrogative test’ in the identification of the focus, the ‘most important’ constituent of the information structure.3

Since Brassai’s ideas remained unnoticed outside Hungary for a long time, just like Weil’s work in the early XXth century French linguistics, Gabelentz’s claims are more widely known in the linguistic literature and he is often recognized as the precursor of today’s information structure theories.

4.2.2 The Prague School and its followers

The already mentioned topic-comment dichotomy originates from Mathesius (1915 and later on), who divided the sentence into a part indicating what the speaker wants to speak about and what is said about it. Later on, the linear order of constituents in a sentence was accounted for by the concept of Communicative Dynamism (Firbas), referring to the fact that elements that push the communication

3I thank Jean-Marie Marandin and Ferenc Kiefer for calling my attention to Henri Weil’s work.
forward (for example by conveying new information) are supposed to follow less communicative elements. However, it has to be noted that although such an approach is compatible with relatively free-word order languages, they might work less adequately in the case of languages like English and French. This corresponds to the analysis of Sgall et al. (1968), who adopt the Topic-Focus articulation of the sentence and argue that the Topic, which is generally contextually bound, is the least dynamic, and the focus, which is contextually not bound, the most dynamic part of the sentence.

The topic-comment dichotomy appears in more recent concepts of the information structure as well. According to Vallduvi, a sentence can be divided into what the sentence is about and what is said about it (we will examine the Vallduvian framework in the next section). Gundel (1988) argues that any constituent in a sentence-initial slot (which she calls the syntactic topic position) must be interpreted as the topic of the sentence, whereas all topics (in the pragmatic sense of the term) are not necessarily in the sentence-initial position.

4.2.3 Recent developments

The term Information Structure was coined by Halliday (1967), who argued that the units theme and rheme have to do "with the relation of what is being said to what has gone before in the discourse, and its internal organization into an act of communication" (p. 199). As for its architecture, Halliday proposed another dichotomy: focus - presupposition or focus - open proposition. In these frameworks, the focus is supposed to be a subset of the rheme (the informative part of the sentence) and it is complemented by the information already established in the discourse (the presupposition or background).

The focus-background architecture is promoted in Jackendoff (1972), who excludes thus the dimension aboutness (the predicative relation between what the sentence is about and what is said about it) from the architecture. Since then, many information structure architectures have been proposed, which differ mainly in the place of the information structure in the model, and in the complexity/number of dimensions it includes. This is what we turn to in the next section.

4.3 The Basic Architectures

It is a long-standing view, as we have seen in the previous section, that a sentence, as Gyuris (2002) puts it, usually contains two basic parts: one that links it to the discourse in which it is uttered, and another that conveys new information. There is cross-linguistic variation between ways in which languages express this difference (prosody, word order, morphological markers, etc). Some authors argue that these distinctions simultaneously exist at different levels (logic, semantic, pragmatic, etc.) (Östman and Virtanen, 1999; Gécség and Kiefer, 2009). Marandin (2006a) mentions two such possible divisions. In the first, the distinction is made between what is given in the context and what is added to it, the latter being thus more informative (focus-background frameworks). In the second, a predicative
(aboutness) relation is defined between a predicate and the unit about which it predicates something (logical subject, topic) (topic-comment frameworks). A third dimension, according to Marandin (2006a), is contrast, more precisely, the question if all focused constituents are inherently contrastive, or the information structure units have a contrastive and a non-contrastive use. We will deal with contrast in a later section.

4.3.1 Monodimensional architectures

In Jackendoff (1972) the only dimension present is the presupposition-new information, the aboutness and contrastive distinctions are neglected. Lambrecht (1994) already talks about presupposition/background - focus distinction and proposes that it has three types: predicate focus, argument focus and all focus. In the first, the predicate itself (most often the verb) is focused, whereas in the second, some argument of it. Both characterize categorical judgements, in which a property is predicated about a given entity, as opposed to all focus constructions, in which the whole sentence constitutes new information. This is related to thetic judgements, describing a state-of-affairs without predicating anything about a given entity.

Other approaches emphasize the aboutness relation as the main factor when determining the information structure architecture. As we have seen in the previous section, such accounts include Mathesius, Gundel, Firbas, and other followers of the Prague school in the second half of the 20th century. However, it has also become clear that a monodimensional architecture cannot capture all the phenomena related to the information structure. Therefore, there have been attempts to integrate more dimensions and have a better understanding of these facts.

4.3.2 Multidimensional architectures

In the case of Vallduví (1992), the aboutness dimension is integrated into the focus-presupposition one. In his system, which is built on the Catalan language, but is supposed to adequately characterize other languages as well, the focus - background distinction is enriched with a further distinction in the background part: the link and the tail, the former corresponding more or less to topics, and the latter to that part of the background that does not play the role of a ”pointer”, determining what the sentence is about (see Figure (4.1)).

![Figure 4.1: Vallduví's Information Structure](image-url)
In Catalan, this order correctly depicts the surface order of constituents in that topics (links) are always on the left periphery, and focused constituents on the right periphery of the sentence. Topics are present in a sentence only if they introduce a new subtopic, i.e. the sentence does not continue the subtopic of the previous sentence. Nevertheless, since topics are to a certain degree given, or discourse-linked, all possible topics relate to the discourse topic or Question under Discussion (QuD). Consider the following example from (Erteschik-Shir, 2007):

(3) Catalan:

Quant al Joan i la Isidora no t’ho sé dir, doncs **el Joan**
as for the Joan and the Isidora no cl know.1sg to say since the Joan
el veiem ben POC.
cl.acc see.1pl quite little

As for Joan and Isidora I can’t say, since Joan we see very little of.

(Erteschik-Shir, 2007, p. 10)

In the embedded clause, *el Joan* appears as the link (topic), since the clause does not continue the previous subtopic (Joan and Isidora). The QuD might be *What about Joan and Isidora?*. The tail (POC) appears on the right periphery of the sentence.

In the architecture of Halliday (1967), formalized by Steedman (2000), the focus-background distinction is integrated into the theme-rheme logico-informational distinction. According to this structure, both the theme and the rheme part contain a highlighted part (the focus) and a given part (the background). The highlighted parts correspond to the pitch accents in English. It would be interesting to investigate if they correspond systematically to certain highlighted/focused constituents in other languages as well. In this thesis, it will be proposed that the highlighted parts can be associated with prominent syntactic positions in Hungarian (although Steedman (2000)’s architecture will be modified).

In Vallduví and Vilkuna (1998) the focus - presupposition level (focus - link - tail) is completed by a contrastive dimension, meaning that each of the above units can be contrastive or non-contrastive. This means that, in this approach, all the three above mentioned dimensions are included, aboutness integrated into the focus-background division, and contrast as the second dimension.

In the next section, we examine where the place of the information structure is in the grammar according to some contemporary frameworks.

---

4 The notion subtopic is understood in relation to the discourse topic, which is assumed to be split into subtopics. The "topic" constituents are supposed to introduce a new subtopic.
4.4 Information Structure in Different Frameworks

4.4.1 Derivational frameworks

The so-called derivational (movement-based) frameworks can mostly be characterized by *syntactocentrism* in that the information structure is integrated into the syntax. The X-bar tree structures encoding syntactic functions (subject, object, adjunct) in lexical projections (NP, VP, PP, etc.) are enriched with functional projections, encoding, on the one hand, information under or abstracted away from the word-level. For instance, the IP hosts inflections and indicates if the clause is finite or non-finite, and the CP contains complementizers and *wh*-words indicating illocutionary force. On the other hand, information structure units, like topic and focus, also appear in the specifier positions of functional projections, such as TopP (Topic Phrase) and FocP (Focus Phrase) (see Brody (1990); Rizzi (1997); E. Kiss (1994, 2002)). In this sense, this means that both the topic-comment and the focus-background/presupposition dimensions are present at the same level of analysis, their prominent elements having a phrase of their own. This means that all what is not topic belongs to the comment, and all what is not focus, belongs to the presupposition. In some approaches, even the contrastive dimension is included via CTopP (Contrastive Topic Phrase). The most detailed description of all functional projections can be found in Rizzi (1997). The concept of presenting information structure categories as functional projections at the level of syntax faces some serious problems, however. Firstly, the semantic content of such projections has never been formalized and exactly defined. It is thus difficult to determine what certain elements occupying the SpecTopP or Spec,FocP position have in common. Discourse functions, such as topic and focus, as opposed to syntactic functions, have to do with the semantic content of the sentence, whereas syntactic functions (subject, object) express relations between two syntactic objects (the ”subject” of the ”sentence”), which can also be indicated by syntactic agreement. It seems, thus, that the integration of discourse functions into the syntactic structure is not well-founded from a semantic viewpoint, even if these discourse functions can directly be associated with certain syntactic positions in languages like Hungarian. Secondly, discourse functions or information structure units do not always correspond to syntactic constituents:

(4) a. Q: -What happened to the dishes?
   b. A: ~JOHN WASHED them.

(Erteschik-Shir, 2007, p. 1, 2b)

(5) It was the RED shirt that John wore at the party.

(6) StalagTIT and not StalagMIT.

In (4), the subject and the verb together constitute the answer to the question, at the level of information structure both of them are to be considered as focused expressions. In (5), only the colour red is focused (it is contrasted to some other colour), but syntactically the whole constituent is clefted (which is supposed to be
4.4 Information Structure in Different Frameworks

4.4.2 Lexical-Functional Grammar (LFG)

LFG is a non-transformational framework that (according to most analyses) contains no traces or empty categories (however, see Bresnan (1995, 1998, 2001) for an alternative view). It consists of parallel levels of representation that are interrelated via correspondence functions. A detailed description of the LFG framework can be found in Bresnan (2001); Dalrymple (2001); Falk (2001) and Komlósy (2001) (in Hungarian). The level of syntax is represented in two structures: c(onstituent)-structure, which is a tree diagram, based on flexible X-bar principles (no binary-branching constraint, constituents can be exocentric), representing dominance and linear precedence relations; and f(unctional)-structure, a feature matrix encoding grammatical functions and predicate-argument relations. Since the beginning of the research in the LFG framework, many other levels of representation have been proposed that encode other aspects: argument structure, prosodic structure, semantic structure and information structure. Since in the present analysis, functional and information structure will play an important role, I give a short introduction of the LFG treatment information structure.

In earlier versions of the LFG framework, discourse functions were integrated in the functional structure, linked via functional uncertainty (one syntactic unit was associated with two functions at the same time, for instance topic and subject). The projection of the information structure as a separate level of representation was motivated by the following problems.

First of all, King (1997) argues that encoding discourse functions in the f-structure leads to circularity, in case if it is only the verb, without its arguments, that is focused. Let us look at the following Russian example:

(7) Ona PROˇCITALA knigu.
she read.pst book
She READ the book.

(King, 1997, p. 5, 9)

The f-structure corresponding to (7) is illustrated in Figure (4.2).

As can be seen in this structure, it is impossible to focus the predicate without its arguments. This is why King (1997) proposes an independent level of representation, where discourse functions are encoded, with their bare predicate value (without their arguments).

Another reason why a separate level of information structure is necessary is that
syntactic constituents do not correspond systematically to constituents of information structure (as demonstrated above in (4) and (5)). In (5), the semantic-syntactic difference can be captured if clefting and focusing (RED) are represented at different levels.

Butt and King (1996) propose that the information structure consists of 4 sets, which are defined by the combination of two features: new +/- and prominent +/-.

The TOPIC set contains elements that are prominent, but not new, the (Information) FOCUS set contains new and prominent elements, whereas old and not prominent elements belong to BACKGROUND and new but not prominent ones to COMPLETIVE INFORMATION (see Figure 4.3).

<table>
<thead>
<tr>
<th></th>
<th>Topic</th>
<th>Focus</th>
<th>Background Information</th>
<th>Completive Information</th>
</tr>
</thead>
<tbody>
<tr>
<td>New</td>
<td>−</td>
<td>+</td>
<td>−</td>
<td>+</td>
</tr>
<tr>
<td>Prominent</td>
<td>+</td>
<td>+</td>
<td>−</td>
<td>−</td>
</tr>
</tbody>
</table>

Figure 4.3: I-Structure units (Butt and King, 1996)

Although this classification simplifies the definition of discourse functions, for instance, foci are not always new, and prominence also has to be defined, I will propose an analysis which is compatible with this architecture, but I will also put forth an alternative architecture which could capture the facts presented in this thesis more adequately. As a recent development, Dalrymple and Nikolaeva (2011) propose to encode semantic information in the information structure. In this representation, which I also follow in this thesis, the values of the discourse function attributes are the meaning constructors of the elements that appear in a given set.

4.4.3 Head-driven Phrase Structure Grammar (HPSG)

Contrary to LFG, which has a modular architecture separating syntactic, prosodic and semantic information and representing them at different levels, another well-known non-transformational framework is HPSG (see Pollard and Sag (1994)), which represents all kinds of linguistic information in one unit: the sign. Engdahl and Vallduví (1996) integrate the Vallduvian information structure architecture into the HPSG framework. Any substructure of a sign contains information about the subcategorizational, semantic and contextual properties of the particular item or constituent. The authors consider that information structure is not purely semantic, but also contextual information, this is why it appears at the CONTEXT.
4.4. Information Structure in Different Frameworks

attribute. At the same time, their semantic content is identical to the value of the CONTENT attribute of the node in question. The information structure attribute has three values: FOCUS, LINK and TAIL, which means that all items in a sentence qualify as one of these. Moreover, the authors define three prosodic values that can appear at the PHON attribute: A-accent, B-accent and unaccented.\(^5\) The accent, or the fact that the constituent in question is unaccented is revelatory in that it indicates its information structure status. Constituents with a B-accent are supposed to be links, the A-accent is related to focused constituents, and unaccentedness means that the constituent belongs to the tail.

However, De Kuthy and Meurers call the attention to some inconsistencies that emerge if this approach is applied. In HPSG, features and the corresponding values are inherited from basic syntactic units (the level of words) until the level of the sentence. This means that, similarly to the LFG problem described in King (1997), whenever some part of the VP is focused (for example the verb itself), at the VP level, it will share its CONTENT with the whole VP, indicating incorrectly that the whole VP is focused. The focusing of the complement also leads to contradiction in that although its focus value is co-indexed with its content value, at the VP level this content value will already correspond to the value of the VP CONTENT attribute, which, in turn, is co-indexed with the head (the verb), and not with its complement.

In more recent HPSG analyses, these problems are treated differently (for instance, in Ginzburg and Sag (2000), the verb, the VP and the sentence already have different semantic types), see also Sag et al. (2003) and Copestake et al. (2005) (for an introduction into Minimal Recursion Semantics, which has been integrated into the HPSG framework). Beyssade and Marandin (2007) describe the prosody of French and its implications on the information structure, and a HPSG-compatible formalism was presented in Marandin et al. (2010).

All in all, the main problem with the HPSG framework, as far as the information structure is concerned, is that it integrates all types of linguistic information into one unit (the sign). In this thesis, the analysis will be conducted in the framework of Lexical-Functional Grammar, which, with its modular architecture, can illustrate the interface of the different levels of linguistic analysis better than a compact linguistic sign.

\(^5\)Bolinger (1958) distinguishes between a fall-rise accent, which he calls B-accent, and a fall accent, which he calls, in turn, A-accent. Jackendoff (1972) identifies the B accent with topics, and A-accents with foci. Informally speaking, in question-answer pairs, B-accented constituents thematize the answers in saying in what respect they answer the question, whereas constituents with an A-accent provide the non-presupposed part of the answer to the question, see also Büring (2003).
4.5 Information Structure Units

4.5.1 The Focus

4.5.1.1 Semantic and pragmatic considerations

We have already seen above that the focus can have different interpretations according to the concept of information structure chosen. In the given/anchored/discourse-old vs. new/informative dichotomy, the focus represents the new, unexpected information in the sentence. This is often related to the assumption that focused constituents are the ones that answer constituent questions, i.e. the part in the answer that corresponds to the interrogative word in the question. Supposedly, answers always contain new information to fulfill communicative needs, otherwise one would not ask the question (apart from rhetorical or echo questions). However, this is not always the case. Foci do not necessarily represent new information in the sense of introducing a new discourse referent. Consider the following example:

(8) Hungarian:
   a. Q: -Kit hívtál meg, Jánost vagy Marit?
      who invite.PST.2SG VM, János.ACC or Mari.ACC?
      Who did you invite, John or Mary?
   
   b. A: -JÁNOST hívtam meg.
      JÁNOS.ACC invite.PST.1SG VM
      I invited JOHN.

The correspondence of focus and new information is not complete from the other perspective either. There are also examples, in which not all the new information is focused. Butt and King (1996) propose an account of information structure in Hindi/Urdu, a discourse-configurational language. Discourse functions are encoded syntactically, topics appear sentence-initially, foci immediately before the verb, and background material is postverbal. However, there are preverbal elements that are not sentence-initial and do not precede the verb immediately. These are referred to as completive information in Butt and King (1996)’s model, and characterized by the fact that they convey non-presupposed information, but they are less prominent in a particular discourse than focused material, since they do not constitute answers to a preceding question. Consider the following example:

(9) Urdu:
   a. Q: -Naadyaa kahaa-se aa rah-ii hai
      Nadya.F.NOM where-from come Stat-F.SG be.PRS.3SG
      Where is Nadya coming from?
   
   b. A: -Naadya to abhii tofii bazaar-me xarid
      Nadya.F.NOM indeed just now toffee.F.NOM market.M-in buy
In example (9), *at the market* is the focus, since it is immediately preverbal and conveys non-presupposed information. *Toffee*, however, is not immediately preverbal, although it also conveys non-presupposed information. On the other hand, it is preverbal, thus not part of the background material either. The authors’ explanation is that it is less prominent in this context than *at the market*, since it does not constitute the answer to the constituent question.

Hence, instead of the *focus as new information* approach, another possible definition would be that the focus is the prominent constituent in the answer that corresponds to the particular interrogative word. This second definition even captures the fact that the focus is more adequately defined in referring to the discourse or to certain contexts. *Beaver and Clark* (2008, p. 7-8.) approach focus phenomena in a similar way:

"[w]e use the term focus to describe an information structural property that an expression may have which affects truth-conditional and non-truth-conditional meaning. Focus phenomena include the fact that prosodic prominence in an answer corresponds to the *wh*-phrase in a question. In many languages, word order and morphology can serve a similar function to focus in English”.

However, the prominent constituent of answers approach still does not cover all cases in which we can find focused constituents. In Hungarian, and many other languages, the focus is used to express corrections, contrast, parallelism, or exclusivity (as the associate of focus-sensitive operators ex. *only*). What is common in all these uses is that a sentence containing a focus cannot be uttered *out of the blue*, but always as a reaction to the preceding context: this can be a question, or simply a statement that the interlocutor considers untrue or does not agree with some parts of it:

(10) **Hungarian:**

a. Q: -Ki ment el?
   Who go.PST VM
   Who left?

b. A: -JÁNOS ment el.
   János go.PST VM
   JOHN left.
   Correction:
As can be seen in the above example, the focus answering a question (JÁNOS) and the one expressing contrast in a corrective sentence occupy the same syntactic position in Hungarian: the verbal modifier usually precedes the verb and forms a prosodic and lexical unit with it; if a focus (or other semantically/prosodically prominent element) is present, it will occupy the immediately preverbal position and the verbal modifier must follow the verb. Thus, the postverbal position of the verbal modifier indicates that the focus is immediately preverbal in both cases. Moreover, the focused constituents bear a pitch accent and all the postfocal sentence-part is deaccented (or exhibits reduced stress throughout).

It is thus preferable to define the focus in the following way:

**Definition 1.** The focus is the semantically prominent and formally highlighted element of replies, answering or reacting to some previous discourse part.

Two remarks are due here. First, semantic prominence is understood in the sense of Jacobs (1984) as the constituent *most affected by the illocutionary operator* associated with the utterance. Second, contrastive topics (carrying a B-accent), are also prominent elements of answers. However, foci and contrastive topics can be distinguished not only in terms of prosody, but also considering the fact that the latter relates to the discourse topic by splitting it up into subquestions (Büring, 2003), whereas the former does not play such a role in the discourse (see Section (4.5.2.5)).

### 4.5.1.2 Formal aspects: syntax and prosody

When I claim that the focus is the prominent part of replies, I do not mean only that it is the constituent that corresponds to the interrogative word of the question or to the constituent contrasted or corrected of the previous sentence, but also that it is the constituent that is formally highlighted. From the syntactic point of view, highlighting of the focus means placing the constituent in salient syntactic positions (see Choi (1999)). The notion of salient syntactic position is language-dependent, but it is commonly observed that syntactically highlighted constituents appear on the left or right periphery of the sentence (German, French), or in the preverbal position (Hungarian, Hindi/Urdu). The syntactic and semantic/pragmatic applications of the term *focus* often lead to confusion, thus in this thesis the term will be used in its semantico-pragmatic sense. "Syntactic focusing" will be referred to as syntactic *highlighting*.

Féry (2008) also argues that the traditional notions of information structure have no designated or invariant correlates in the grammar (syntactic position, prosodic
Nevertheless, this does not mean that certain tendencies cannot be observed (like the sentence-initial position of topics), but discourse functions and other information-structure related notions cannot be exhaustively defined based only on formal criteria.

As far as the prosody of focus is concerned, focus is also prosodically highlighted, and its exact realization is also language-specific. It is indicated by a sharp falling pitch accent in Hungarian, and the deaccenting of (or reduced stress on) postfocal material, whereas in French, the illocutionary boundary tone is anchored at the left edge of the focused constituent (see a detailed description in Chapters (5) and (6)). The same way, "prosodic focusing" will be referred to as prosodic highlighting.

4.5.1.3 Contrastive/Identificational focus vs. Information Focus

In her seminal paper, É. Kiss (1998a) calls for a distinction between contrastive/identificational and information focus. She bases her claim on Hungarian data, but extends it to other languages as well. É. Kiss observes that focus (in the sense of non-presupposed material) can appear in two, structurally different positions in the Hungarian sentence and have different semantic/interpretive properties accordingly. Compare the following examples from Hungarian:

(11) Tegnap este MARINAK mutattam be Péter.
yesterday evening Mari.dat introduce.pst.1sg vm Péter.acc
It was to Mary that I introduced Peter yesterday.

(É. Kiss, 1998a, 247, 5a)

(12) Tegnap este bemutattam Péter Marinak.
yesterday evening vm.introduce.pst.1sg Péter.acc Mari.dat
Yesterday evening, I introduced Peter to Mary.

(É. Kiss, 1998a, 247, 5b)

É. Kiss argues that in (11) the immediately preverbal constituent (MARI-
NAK) exhaustively identifies the person to whom I introduced Peter. This means
that I introduced Peter only to Mary and to no one else. In (12), on the other hand,
where Marinak is in the postverbal domain, what the sentence says is that an event
happened yesterday (I introduced Peter to Mary), but it says nothing about other
people I might have introduced Mary to. É. Kiss calls the first type of focus iden-
tificational focus, and the second, which purely conveys new or non-presupposed
information, information focus. We will deal with the issue in more details in the
chapter on Hungarian, however we note here that this example is not the best illus-
tration of the contrast between the two types of foci, since the sentences cannot be
used in the same context. Example (11) contains narrow focus (of one constituent),
and it can answer the following question:

(13) Kinek mutattad be tegnap este Péter?
who.dat introduce.pst.2sg vm yesterday evenig Péter.acc
To whom did you introduce Peter yesterday evening?

(Marinak) would even be enough as an answer, the rest of the sentence in (11) could be left out. This is not the case with (12), which would answer the following question:

(14) Mi történt?
    what happen.PST
    What happened?

In (12), not only Marinak constitutes non-presupposed information, but the whole sentence (all-focus/wide-focus sentence). However, we can construct cases in which non-presupposed material is postverbal. Consider the following example, based on Kratzer and Selkirk (2009):

(15) a. (Gary is an art dealer. Lately he's been very picky about which museum he deals with; he doesn't do business with the Metropolitan or the Guggenheim.)

b. Így (csak) a LOUVRE-NAK adná ezt az új Modiglianit.
    thus only the Louvre.DAT would give this.ACC the new Modigliani.ACC
    So he would only offer this Modigliani to the Louvre.

Note that (in the very same context) the constituent ezt az új Modiglianit could appear on the left periphery as topic as well, and in that case it would shift the subtopic of the previous sentence (Gary, the art dealer) to a painting (also included in the discourse topic), but I ignore this possibility here. What is important is that (a LOUVRE-NAK), the museum, which is contrasted to the previously mentioned ones, appears in the preverbal position, whereas the non-presupposed painting can be postverbal. This confirms the claim that certain types of non-presupposed constituent occupy the preverbal position in Hungarian (and are syntactically and prosodically highlighted), whereas others can be postverbal.

Generally, looking at the contexts of any sentence, elements that are explicitly contrasted to another, or that provide the answer to a question, are formally highlighted and qualify as identificational focus, according to É. Kiss’s terminology. Before establishing a classification of different types of focus, let us have a look at the notion of contrast in this respect.

4.5.1.4 Contrast

According to the definition of Rooth (1992), phrases have an ordinary and a focus semantic value. The focus semantic value of a phrase is "the set of propositions obtainable from the ordinary semantic value by making a substitution in the position
corresponding to the focused phrase” (p. 76). Substitution is made by using the other members of the set of alternatives the focus introduces. In this respect, all uses of focus are contrastive. Even all-focus sentences can be contrasted to any alternative sentence:

(16) a. Q: - What happened?
    b. A: - John went to the cinema with Mary. / Mary had a picnic with her friends. / John spent the whole week-end at home writing his thesis.

The presence of alternatives and contrast is, nevertheless more evident in the case of corrective sentences, where one of the alternatives is present in the preceding discourse and the focused constituent is contrasted to it (like JÁNOS and PÉTER in example (10), or the museums in (15)). Alternatives can be present in the question-answer context as well, but this is not necessarily the case. Compare:

(17) a. Q: - Who came?
    b. A: - MARY.

(18) a. Q: - Who came, John or Mary?
    b. A: - MARY.

According to Erteschik-Shir (2007), focus is contrastive only if the alternatives are explicit and present in the context:

(19) a. Q: - Would you like ’tea or ’coffee?
    b. A: - I’d like COFFEE please.

Now, the question is if the presence of alternatives (and thus contrast) is only part of certain semantic formalisms (which is an abstract way of accounting for the meaning of focused constituents), or we should assume that every highlighted constituent induces a set of alternatives and is contrasted to them. The answer to this latter question is, I believe, no. Contrast is not necessarily an inherent property of focus. Let us examine this question in more details.

It has been proposed in the literature about English that all prosodic peaks (pitch accent) receive a contrastive meaning (Marandin (2006c), citing Bolinger 1961 and Jackendoff 1972). However, as Kratzer and Selkirk (2009) demonstrate, there are prosodic differences between (as they put it) contrastive focus and new information focus. They show that the former show greater pitch prominence than the latter within the same sentence. In between-sentence comparisons, the former show greater duration than the latter in corresponding syntactic positions.

However, in Hungarian, both contrastive and new information foci can occupy the same preverbal position, thus it seems that a further distinction has to be made in order to accommodate all focus-related phenomena. I will distinguish between two subtypes of focus: contrastive focus and information focus, and I define them in the following way:

- Contrastive focus

I follow Erteschik-Shir (2007) in claiming that
Definition 2. Contrastive focus is the subtype of focus whose alternatives, to which the focus is contrasted, are explicitly present in the discourse.

This is the type that appears in corrections, parallel structures, or in answers to questions that contain possible alternatives for the answer, and other contrastive contexts, (10). We will see in Chapter (7) how these relations between sentences can be formalized.

• Information focus

Definition 3. Information focus is the narrow focus constituent that serves as answer to questions.

This means that contrastive and information focus can fall together in certain contexts (when alternatives are explicitly present in the question), but this is not necessarily the case. Both information focus and contrastive focus are formally highlighted in Hungarian, they bear a pitch accent and reside in the preverbal position.

In Italian, however, contrastive and information focus are realized differently: if contrast is explicit, focus is preverbal. Focus answering a question, without the presence of alternatives, is postverbal:

(20)  a. Q: -Who has broken the vase?
    b. A: -Il vaso ha rotto MARIA.
        the vase aux broken Maria
        Maria has broken the vase.

(21)  a. Q: -Was it Peter who broke the vase?
    b. A: -MARIA ha rotto il vaso.
        Maria aux broken the vase
        (No,) Maria has broken the vase.

Choi (1997) shows that contrastive focus and information focus behave differently in German as well, with respect to scrambling:

(22)  a. Q: -Was hast du dem Kassierer gegeben?
      what aux you. dat cashier given
      What did you give the cashier?
        I aux the.dat cashier [the money] given
        I gave the cashier the money.
   I AUX the money the cashier given
   I gave the money to the cashier.

   (Choi, 1999, p. 4, 9)

(23) weil Hans [das BUCH] dem Mann gegeben hat (nicht die Zeitung)
    because Hans the book the.DAT man given AUX not the newspaper
    because Hans gave the book to the man, not the newspaper

   (Choi, 1999, p. 4, 10)

As the examples show, contrastive focus (23) can scramble, whereas a non-contrastive focal element (22) (information focus answering a question) cannot. This also motivates a distinction between the two.

Note that the presence of alternatives is not always easy to determine. Erteschik-Shir (2007) refers to the focus in the answer of the following example as restrictive focus:

(24) a. Q: -Which movie did John watch yesterday evening?
    b. A: -He watched Star Wars.

Although the alternatives are not explicitly present in the discourse, which-expressions denote contextually determined sets (like here, the set of films that were on the day before, or the set of DVDs John has at home, or borrowed recently, etc.). As opposed to contrastive foci, in this case, the set of alternatives is not clearly defined, and therefore the focused element is not explicitly contrasted to the other elements. I thus consider this as a case of information focus.

As É. Kiss argued, information focus (her identificational focus) is both exclusive and contrastive in that it excludes all the other potential answers to the question (the rest of the alternative set) and exclusively identifies the element about which the predicate holds. I argue, on the other hand, that the exclusive and exhaustive interpretation of information focus comes from pragmatic factors, i.e. from the fact that answers are expected to be exhaustive (see Chapter (5) on Hungarian and Brody and Szendrői (2011)). Information focus is thus not contrastive in the sense that it is explicitly contrasted to alternatives in the context, but since it is an answer, it is implied that other possible alternatives do not come into play. Consider the following dialogue in French:

(25) a. Q1: -Qu’est-ce qu’il a pris au déjeuner ?
    what he has taken at lunch
What did he eat for lunch?

b. A: -Il a pris une pomme.
   he AUX taken an apple
   He ate an apple.

c. Q2: -Ah bon ? C’est tout ?
   Oh really it is all
   Oh really? Was that all?

Since answers are supposed to be exhaustive, the locutor is surprised, since he expects to hear a main course and not just a piece of fruit. The locutor takes the answer as exhaustive, i.e. he expects the interlocutor to name the main course that constituted the lunch of the third person under discussion that day. Of course, this does not mean that it is impossible to complete this answer and add:

(26) A2: -Oui, mais après voyant les autres manger, il a pris une grosse salade.
   yes, but then seeing the others eat, he AUX taken a big salad.
   Yes, but then, seeing the others eat, he took a big salad.

However, according to Jean-Marie Marandin (p.c.), there can be prosodic or syntactic differences between exhaustive and non-exhaustive answers in French. (This is impossible in Hungarian, since the constituent in the preverbal position is always interpreted as an exhaustive answer.) Compare the following examples:

(27) a. Q: -Qui est venu ?
   who AUX come
   Who came?

   It’s Marie.
   It was Mary.
   # et d’autres encore (and others as well)

c. A2: -Marie (est venue). ]L%
   Marie (AUX come).
   Mary (came).
   # et d’autres encore (and others as well)

d. A3: -Marie (est venue). ]H%
   Marie (AUX come).
   Mary (came).
   et d’autres encore (and others as well)

The first answer is exhaustive and this is indicated by syntactic means (clefting). In the second answer, exhaustivity, is indicated by the illocutionary boundary tone
4.5. Information Structure Units

(161). The third answer, with a different illocutionary boundary tone is non-exhaustive. It can be continued by adding other people who came. However, since such answers as A3 are not exhaustive, askers tend to add another question of the type: And who else?, which shows that pragmatically, exhaustive answers are expected to questions.

A related notion is the already mentioned Complettive information. I follow Butt and King (1996) in claiming that it refers to information “new to the hearer, [...] but not of primary importance to the information structure of the discourse at hand” (p. 2). The lack of prominence with respect to focus can be explained by the fact that Complettive information does not constitute the answer to a constituent question and is not contrasted to another element. Accordingly, complettive information is not syntactically and prosodically highlighted in the same way as contrastive and information foci.

When using the term focus, I will thus refer to contrastive and information focus as defined here, in the semantico-pragmatic sense of the term. If it is necessary, a distinction will be made between the two. The formal ”focusing” of the two will be referred to as highlighting.

4.5.2 The Topic

4.5.2.1 Semantic and pragmatic issues

The definition of the topic is as controversial as that of the focus. The topic is also a pragmatically defined information structure category that has certain formal properties in a number of languages (sentence-initial position is Hungarian, left dislocation in French, etc.). What makes the definition especially difficult is that the phenomena to be captured are diverse, within one language and cross-linguistically as well.

The problem of the topic goes back to the observation that in a sentence, word order can differ from the canonical word order in certain contexts. This divergence can be optional in languages with relatively fixed word order, but it is far from being arbitrary or context-independent. For instance, the grammatical subject (the NP in the nominative case with which the finite verb agrees) is not always the constituent about which the rest of the sentence conveys new information, and it is not even always in a sentence-initial position. Consider the following examples from English and French:

(28) This book I don’t like.

(29) French:

Cet homme, je ne le connais pas.

this man I PRT CL.ACC know NEG

This man I don’t know.

In a variety of relatively free word-order languages (Slavic languages, German, Hun-
the sentence-initial constituent does not necessarily correspond to the grammatical subject, but to a constituent that can be described in several ways. The phenomenon can be approached from the syntactic and the pragmatic side, although both play a role. As we have already seen in the section about the focus, from an informational point of view, there are different dimensions according to which a sentence can be divided into two parts. One of these distinctions is the old vs. new information, and another is the aboutness relation. The topic has been defined with respect to both. According to the former, the topic is supposed to be the constituent denoting an already known discourse referent, i.e. the part that links the sentence to the preceding discourse and provides the connection with the upcoming new information. In the latter, the topic constituent is the one the rest of the sentence predicates something about. The two overlap in most cases, but according to the aboutness approach, the topic can denote new information, whereas according to the old-new distinction, the topic is not necessarily the (only) constituent the sentence is about.

Deciding what a sentence is about is in itself problematic. Consider the following example:

(30) a. Q: -Who brought the flowers?
   b. A: -John (L %) brought the tulips (H %L %).

It is difficult to explain in what way the answer in (30) is more about John (A-accent, focus) than about the tulips (B-accent, contrastive topic).

Concerning the old-new distinction, it is also controversial how ”old” should be defined. According to Marandin (2007b), the term old is ambiguous at least in two respects. First, it can refer to a discourse referent already introduced into the discourse, or to some piece of old information. This latter can be understood in two ways as well: old is what is part of the common ground, also called hearer-old, or what is discourse-old (it has been activated in the immediately preceding discourse).

There are some clear cases in which the topic of the sentence is not discourse-old in the sense that it has been mentioned before. However, it is identifiable to the hearer. Consider the following examples:

(31) a. Q: -How are you?
   b. A: -I am fine.
(32) The president is elected directly by the people.
(33) That sofa has a nice colour.

Erteschik-Shir (2007) calls these Permanently Available Topics, that, even if not mentioned in the discourse, are identifiable, either because of their unicity in a given context (the president, my mother, the Moon, etc.), or because they are present at the current scene at the moment of the conversation (this chair, that sofa). It is a well-known observation, related to the old-information status of topics, that the referent of the topicalized constituent is specific (definite expressions, pronouns, proper names), or at least identifiable to both speaker and hearer. This is why
4.5. Information Structure Units

non-specific indefinites are supposed to be ungrammatical if they are the topic of the sentence. However, they can be rendered specific, like in the following example:

(34) Danish:

En pige som jeg modte i går gav jeg en god bog.  
a girl that I meet.pst yesterday gave I a good book

I gave a good book to a girl that I met yesterday.

(Erteschik-Shir, 2007, p. 8, 2e)

The topicalization in the above example can be accounted for by the fact that the relative clause makes the topic referent specific, identifiable for the hearer. There are, however, several exceptions, or, at least problematic points that emerge with respect to this observation about non-specific indefinites. Gécseg and Kiefer (2009) identify examples in Hungarian, in which the sentence-initial constituent is clearly non-specific, and discourse-new/non-given and it still can appear sentence-initially:

(35) Egy gyerek leesett a villamosról.  
a child VM.fall.pst the tram.DEL

A child fell from the tram.

(Gécseg and Kiefer, 2009, p. 585, 2a)

(36) Valaki bekopogott az ajtómon tegnap.  
someone VM.knock.pst the door.POSS.1SG.SUPERESS yesterday

Someone knocked on my door yesterday.

(Gécseg and Kiefer, 2009, p. 597, 26)

The authors consider the sentence-initial constituent of the Hungarian sentence as a logical subject, i.e. a constituent that does not necessarily correspond to the grammatical subject of the sentence, but which is subject to less restrictions as the topic, since it is context independent (it can be discourse-new). We discuss Gécseg and Kiefer (2009)’s theory in Chapter (5) about Hungarian.

Referring to Gundel (1974), Erteschik-Shir (2007) (see also Krifka (2008) citing Marty 1884) even presents the view that all-focus sentences (answering questions like What happened?), that seemingly have no topics, do have so called Stage Topics, the particular spatio-temporal parameters of the sentence, the situation about which it is asserted.

Based on these, Erteschik-Shir (2007) prefers the definition given to discourse-old, since the former captures even contextually present, but not yet mentioned items.

A related notion is accommodation. Topics that are discourse-new, have to be acknowledged by the hearer. This process is called accommodation. Accommodation is facilitated by the above presented relative clauses that allow the hearer to
identify, or at least categorize the topic referent. The following are similar examples:

(37) My sister lives in London.
(38) My car broke down.
(39) John was murdered yesterday. The knife lay nearby.

(37)–(39) (Erteschik-Shir, 2007, p. 19, 22)

The initial constituents in these examples can be, in all cases, discourse-new. However, the first two are easily accommodated by the hearer, because of the possessive structure, even if he/she was not aware of the fact that the locutor had a sister or a car. In the last example, the knife has not yet been mentioned. Nevertheless, it is completely acceptable as a topic, since the context (the murder) entails the existence of all possible weapons, thus the hearer accommodates that the murder was committed with a knife.

4.5.2.2 Formal aspects

From a formal perspective, the topic is usually, but not always, the sentence-initial constituent. In languages that are claimed to reflect information structure in syntax (E. Kiss, 1995), the canonical word order contains the topic as the linearly first constituent. In other, more fixed word order languages, topics also often appear on the left periphery, detached from the rest of the sentence. However, topicalization is a syntactic operation akin to left dislocation, which lacks a resumptive pronoun within the rest of the sentence (see Lambrecht (1994) and Delais-Roussarie et al. (2004) on French) and is by no means to be identified with topics in the pragmatic sense of the term. Syntactic topicalization does not always involve the information structure role topic (see Prince (1981) and the discussion on French in Chapter (6)). The universal observation that topics tend to appear on the left periphery, is related to the fact that the linear order of constituents has, in most cases to do with what is known and what is unknown in a particular context. Before conveying new information, speakers tend to mention a link that connects it to the preceding discourse (this was referred to as Communicative Dynamism in the Prague school, describing Czech data). Nevertheless, this is rather a tendency, not a general rule, especially in configurational languages. Gécseg and Kiefer (2009) argue that the constituent which is obligatorily initial in Hungarian, is the logical subject, not the topic (in their terminology).6 This does not mean, however, that, according to the authors, there can be no topic in the Hungarian sentence. The topicalized constituent is, however, context-dependent, and not necessarily sentence-initial. It can fall together with the logical subject, but this is not necessarily so. Consider the examples with topics and logical subjects in Hungarian:

6The authors define the logical subject as the constituent denoting the individual(s) of which the logical predicate holds, independently of a particular context. The logical subject-logical predicate division reflects the propositional articulation of the clause, based on the concept of aboutness. The logical subject is assumed to be exclusively sentence-initial, unlike topics that can appear in different positions as well.
4.5. Information Structure Units

(40) A gyerekeim nagyon szeretik a palacsintát.  
the children.POSS.1SG very like the pancake.ACC  
My children love pancakes.  

(Gécseg and Kiefer, 2009, p. 609, 53)

(41) Valaki kiabál odakint.  
somebody screams outside  
Somebody is screaming outside.  

(Gécseg and Kiefer, 2009, p. 609, 54)

(42) Nagyon szeretik a gyerekeim a palacsintát.  
very like the children.POSS.1SG the pancake.ACC  
My children love pancakes.  

(Gécseg and Kiefer, 2009, p. 610, 55a)

(43) Kiabál valaki odakint.  
screams somebody outside  
Somebody is screaming outside.  

(Gécseg and Kiefer, 2009, p. 610, 56b)

The initial constituent in (40) is at the same time the topic and the logical subject of the sentence. (41) does not contain a topic in the pragmatic sense of the term, the sentence initial constituent is the logical subject. In (43), a thetic sentence (not predicating a property about any subject), valaki is neither topic, nor logical subject (logical subjects are, according to this framework, always sentence-initial). However, in (42), as the authors argue, the individual-level predicate enforces a categorical reading, which leads to the assumption that a gyerekeim must be the topic of the sentence. These examples illustrate that the left-peripheral position of the topic is also a tendency, but not a general rule, if we consider that the topic is a semantico-pragmatic notion that cannot be exclusively identified with one particular syntactic position. (Delais-Roussarie et al., 2004) show that right dislocation is also possible in French, which is used for foregrounding non-presupposed material, thus in French as well, it is not only the left periphery which is associated with topics in the sentence.

4.5.2.3 Entity or question

This last observation leads us to the following problem concerning the definition of topichood. According to both the old-new and the aboutness definitions, topics are entities, denoted by the topicalized constituent. However, in some cases, the topic of the sentence is not an individual, but some point in space or time, which can still be a good candidate for the part of the sentence about which the rest predicates
something or which connects it to the preceding discourse.

(44) Hungarian:

Ma reggel nem mentem iskolába.

Today morning not go.pst.1sg school.ill

This morning, I didn’t go to school.

(45) French:

Dans les montagnes, il neige souvent.

In the mountains, it snows often

It snows often in the mountains.

Another approach is to claim that the topic of the sentence cannot be captured without referring to the discourse in which the sentence is uttered. The discourse itself has a topic (the discourse topic). According to Reinhart (1981) ”local entries corresponding to sentence topics can be further organized under more global entries, thus constructing the discourse topics” (p. 24), which, according to this definition, are the discourse-level equivalents of sentence-topics, but they are not fundamentally different from them, i.e. they are the same type of entities. As opposed to this, the discourse topic has also been modeled as a hierarchy of questions (Marandin, 2007a). In a successful exchange of information between the participants, the main discourse question gets answered. However, in most cases it is decomposed into subquestions that are easier to deal with or to answer in one question-answer turn-taking. The discourse topic, which is considered the Question under Discussion, can determine the eventual further questions to be asked and answered in the particular dialogue. Büring (1997) proposes a hierarchical, question-based discourse model. His analysis mainly deals with contrastive topics in question-answer pairs. His approach, however, can be enlarged to sentences containing ”ordinary” topics as well. We will examine Büring’s theory in details in a later subsection. Let us compare the two discourse-topic conceptions with the help of the following examples:

(46) a. Q: -What did the family do during the weekend?
   b. A: -The father repaired the car, the mother went to the swimming pool, and the children played computer games.

(47) (What about Mary?)

Mary was an only child. She got spoilt by her parents. At school, she was very smart and won a lot of competitions. She had some close friends. Ann was her best friend, with whom they studied on together at university. At the faculty of law, she met her future husband...

In (46-b), the original question is split into implicit subquestions in the answer (What did the father do?, What did the mother do?, What did the children do?),
and it would also be possible to answer only some (and not all) of these. In (47), the discourse topic is represented as a question, not as an entity (Mary), since strictly speaking, not all sentences are about Mary, but all of them are related to a global question about her life. In this interpretation, the underlined constituents are all subtopics of the discourse topic and indicate a shift with respect to the previous subtopic.

To sum up what we have dealt with so far in this section, we can conclude that none of the well-known topic definitions is without exceptions. Although they overlap and cover a great part of topicalized constituents, it is never difficult to find counterexamples. Topics are not always discourse-old. A better term is that they are identifiable for the hearer, from the preceding discourse, from the physical context, from world knowledge, or it is easily accommodated. Semantically, the topic is not always an entity the rest of the sentence is about, since expressions other than entities can also appear sentence-initially (nevertheless, other definitions would not consider them as topics). From the formal point of view, according to some analyses, topics are always sentence initial, while others claim that only logical subjects appear strictly in that position. After all, if topics are semantically-pragmatically-contextually defined, they are not necessarily associated with one syntactic position, and, as such, they do not necessarily correspond to syntactic constituents either.

What makes the definition of topics especially difficult is the complexity and multidimensionality of the question. Topics have been defined syntactically, semantically and pragmatically, and every language has its own topic definition that fits its information structure the better.

Based on these observations, the definition of topic adopted in this thesis is the following:

**Definition 4.** The topic is the constituent that links the sentence to the preceding discourse by introducing a subtopic of the discourse topic. It is usually, but not always, left peripheral, denotes specific entities and is given in a certain context, but none of these is obligatory.

I distinguish between two types of topics: thematic shifters and contrastive topics. The former are illustrated in (47), the latter in (46-b).

4.5.2.4 Thematic shifters

Thematic shifters, also referred to as sentence topics, as opposed to the discourse topic, can appear both in narrative contexts and dialogues.

Erteschik-Shir (2007), citing Daneš (1974) claims that there are three ways in which a sentence topic can be connected to the discourse topic. She refers to this as Thematic progression. In the first, which she calls topic chaining the theme of the previous sentence is kept constant (48). In the second, the theme is derived from the rheme (new information) of the previous sentence (49). Erteschik-Shir calls this focus chaining. In Daneš’s third third case, the sentence topic is derived
from a hypertheme (a discoursally restricted set of elements) (50). Let us illustrate these with the following examples:

(48) John likes to read. He is intelligent and industrious and will go far.

(Erteschik-Shir, 2007, p. 3, 3)

(49) There’s a girl in the class who the teacher likes. She answered all the questions the teacher asked.

(Erteschik-Shir, 2007, p. 3, 4a)

(50) I’ll tell you about my friends, John, Paul, and Mary. John is an old friend from school, Paul, I met at college, and Mary is a colleague at work.

(Erteschik-Shir, 2007, p. 3, 5a)

In the above sense, (48) and (49) do not introduce a new subtopic, which is indicated by the fact that the sentence-initial constituents are pronouns. In (50), however, the three friends constitute the thematic shifter in the clauses of the second sentence. According to the above definition, unaccented pronominal subjects are not considered as topics, since they do not introduce a new subtopic of the discourse topic. Note that the Hungarian equivalents of examples (48) and (49) do not contain a pronominal subject, since it is a pro-drop language:

(51) János szeret olvasni. (#Ò) Intelligens, szorgalmas és sokra fogja János likes to read (he) intelligent, industrious and much.SUBL will reach

(52) Van egy új lány az osztályban, akit nagyon szeret a tanár. (#Ò) is a new girl the class.INESS, whom very likes the teacher (she) Mindenre tudott válaszolni, amit a tanár kérdézett. all.SUBL could answer, that the teacher ask.PST

The presence of the pronoun, if not interpreted contrastively, would render the sentences pragmatically anomalous. On the other hand, the subjects have to be present in the Hungarian equivalent of (50):

(53) Mesélekn neked a barátaimról, Jánosról, Paliról és tell.PRS.1SG you.DAT the friends.POSS1SG.DEL, János.DEL, Pali.DEL and Mariról. János egy régi iskolai barátom, Palit a Mari.DEL. János an old school friend.POSS.1SG, Pali.ACC the főiskoláról ismerem, Marival pedig együtt dolgozom. college.DEL know.1SG, Mari instr and together work.PRS.1SG

Erteschik-Shir (2007) presents topicalization in two languages, Danish and Catalan. She argues (referring to Vallduví (1992)) that in Catalan, only constituents that were not the topic of the previous sentence can be topicalized (the link in Vallduví (1992)’s terms). She calls these elements thematic shifters and argues that
they can represent a clear case of focus chaining (however, they are not obligato-
riely the focus/rheme of the previous sentence either). About Danish, she observes
that the topicalized element can continue the topic of the previous sentence, as
far as it is specific, or clearly identifiable for the hearer. She argues that although
continued topics are definitely possible in Germanic languages, they also tend to
topicalize shifted topics more. Two remarks are due here concerning the Catalan
topicalization.

The first is that shifted topics are, paradoxically, old and new at the same time.
They are new, since they cannot have appeared as the topic of the previous sentence.
On the other hand, if they are topicalized, the interlocutor is expected to be able
to identify or be aware of them. The explanation is related to the discourse topic.
Thematic shifters are new in a particular sentence, but at the same time they are
supposed to be related to the discourse topic, or the Question under Discussion.
From this respect, they are never brand-new and are not introduced out of the blue.

As we will see in the chapter on Hungarian, topicalization in Hungarian follows
the Catalan pattern, i.e. only shifted topics are possible, which do not correspond
to the topic of the previous sentence.

Based on these examples, Danes’s categorization seems to be somewhat too re-
stricted. In Catalan and in Hungarian, shifted topics are not necessarily derived
from a hypertheme that is contextually given. They are rather related to the dis-
course topic (they represent a subtopic), whose subtopics need not be explicitly
present in the discourse.

4.5.2.5 Contrastive topics

The other type of topic I distinguish is contrastive topics. The problem of contrast
has emerged concerning topics and foci as well, as much as it was even integrated
into some information structure architectures. Vallduví and Vilkuna define a two-
dimensional architecture, one being the already presented Link-Focus-Tail struc-
ture. At the second level, each of these, as a constituent, can be contrastive or
non-contrastive. Contrastiveness is thus a property of individual constituents and
not of whole sentences.

In this subsection, we take a look at contrastive topics. I rely on Büring’s work
(Büring, 1997, 2003) concerning contrastive topics (Büring’s model on discourse
structure goes back to Bonnie Webber’s, Livia Polanyi’s and Remko Scha’s works
for inspiration). According to Büring, the role of contrastive topic is the indication
of a certain strategy the interlocutor chooses when they answer a question. This
means splitting up the original question into implicit subquestions and answer (some
of) them. The choice of a certain strategy can be motivated by the intention
of indicating the structure of the answer, for instance, if the question cannot be
answered at once, but some intermediary steps are needed. Sometimes even an
implicit superquestion has to be assumed (see Figure (4.5)). In Büring’s model
thus, the contrastive topic indicates that the interlocutor divides the Question under
discussion into (implicit) subquestions, either because they do not know all the
details of the expected answer, or because they do not want to give an exhaustive
answer and they imply this to the locutor. Büring calls this kind of topic *S-topic* (Sentence topic), as opposed to the *Discourse topic* (this latter is equivalent to the Question under discussion, and does not necessarily correspond to a syntactic constituent). In his approach, the contrastive topic is one of the three types of S-topics. These are the following (Büring, 1997, p. 56):

1. **Partial topic**

   (54)  
   a. Q: -What did the pop stars wear?  
   b. A: -The [FEMALE] pop stars wore [CAFTANS].

   (55)  
   a. Q: -How was the play?  
   b. A: -[THE MUSIC] was good.

2. **Contrastive topic**

   (56)  
   a. Q: -Do you think that Fritz would buy this suit?  
   b. A: -Well, [I] certainly [WOULDN’T].

3. **Implicational topic**

   (57)  
   a. Q: -Did your wife kiss other men?  

Note that the topic types in all these examples are pronounced with a specific intonation pattern (B-accent, see Jackendoff (1972)) in English. These examples would all be pronounced with a rising tone (specific to contrastive topics) in Hungarian, and with a C-accent (see Chapter (6)) in French (indicating discourse moves). In what follows I will not distinguish between partial, contrastive and implicational topics, but use the term *contrastive topic* for all of them. As can be seen from the examples, the interlocutor does not give a complete answer to the questions. In (54-b) he decomposes the set of pop stars to two partitions and the answer concerns only one of them. At the same time, the partial answer suggests that either the same answer is not true for the other partition (*i.e.* the male pop stars wore something else, but not caftans) or that concerning that partition, the interlocutor does not have the necessary information, or does not want to / cannot convey it. In (56-b) and (57-b) the answers clearly imply that the other partition of the set introduced by the question is not covered by the answer itself. Büring represents his analysis with so-called *Discourse-trees*, which indicate that the topic of the discourse, the Question under discussion is divided into subquestions:

The discourse trees of (54-b) and (56-b) are represented in Figures (4.5) and (4.6), respectively.

In Büring’s model, every node in a discourse tree is called a *Move*. Every move corresponds to a sentence. The contrastive topic indicates a strategy that can be defined as a substructure of the discourse tree that is dominated by a question. In other words, the contrastive topic indicates that a question is answered through its
discourse: dressing habits of popstars

<table>
<thead>
<tr>
<th>question: What did the popstars wear?</th>
</tr>
</thead>
<tbody>
<tr>
<td>(subq: What did the female popstars wear?)</td>
</tr>
<tr>
<td>answer: caftans</td>
</tr>
<tr>
<td>(subq: What did the male popstars wear?)</td>
</tr>
<tr>
<td>(answer: not caftans)</td>
</tr>
</tbody>
</table>

Figure 4.5: What did the popstars wear?

(superguestion: Who would buy this suit?)

question: Would Fritz buy this suit?  (question: Would I...?)  (question: Would X...?)

<table>
<thead>
<tr>
<th>answer: (probably not)</th>
</tr>
</thead>
<tbody>
<tr>
<td>answer: no</td>
</tr>
</tbody>
</table>

Figure 4.6: Would Fritz buy this suit?
subquestions, even if those are implicit in the discourse. It is thus contrastive with respect to the other potential answers to the question. The value of the contrastive topic can be defined with the help of the following algorithm:

1. We replace the focus in the answer with the appropriate interrogative word
2. We create a set of questions by replacing the contrastive topic in the first answer by its alternatives

1. John ate chocolate cake. → WHAT did John eat?

The subquestion representing a strategy is, however, not always implicit. We can often hear questions that are asked because the locutor did not get a complete and satisfactory answer to his original question. Consider the next dialogue:

(58) a. Q1: -Who ate what at the party? / What did the people eat at the party?
   b. A: -Julia ate COOKIES, George CHOCOLATE CAKE, Mary PIZZA...
   c. Q2: -Yes, OK, but what did JOHN eat?

In this case, the asker recognizes the strategy the answerer followed and asks a new question, which is at the same time part of the strategy and a new question node, see Figure (4.7):

- discourse: the party
- question: Who ate what at the party?
- (subq: Julia?)
- (subq: George?)
- (subq: Mary?)
- q & subq: What did John eat?
- A: COOKIES  A: CHOCOLATE CAKE  A: PIZZA

Figure 4.7: Yes, OK, but what did JOHN eat?

In the Hungarian equivalent of the above example, the prominent constituent of the repeated question combines the properties of contrastive topics and foci. It bears a pitch accent and appears in the preverbal field (like the focused constituent). However, unlike foci, it is not in the immediately preverbal position, which is, in turn, occupied by the interrogative word:

(59) a. Q1: -Ki mit ettek a bulin? / Mit ettek az who what eat.PST the party.SUPERESS / what eat.PST the emberek a bulin? people the party.SUPERESS
b. A: Juli SÜTIT evett, Robi CSOKITORTÁT, Mari Cookies. Juli eat.pst, Robi chocolate cake. Mari PIZZÁT... pizza. Juli ate COOKIES, Robi CHOCOLATE CAKE and Mary PIZZA.

c. Q2: -Jó, de JÁNOS mit evett?
   OK, but János what eat.pst?
   Yes, OK, but what did JOHN eat?

Such questions are typically introduced by Jó, de... Na, de... (Yes, OK, but...), which justifies the claim that such questions cannot start a dialogue, but are uttered as part of a strategy. We can thus conclude that the prominent constituent in such repeated questions shares properties of topics and foci. It identifies the single individual in connection with which the answer is expected, contrasting it to the other elements of the relevant set (the people at the party). Its intonation follows that of focused material. On the other hand, this is the part of the sentence that indicates its role in the strategy. It is in parallel with the contrastive topics of the previous answers. In addition, it does not occupy the canonical focus position because it is filled by the question word. This latter, although in the focus position, is not accented. Following Onea (2009a), this can be considered as a case of second occurrence focus. If the same focus appears for the second time in a discourse, it will not follow the same intonation pattern (pitch accent), as at its first occurrence. Sometimes the second question does not narrow down, but repeats the first, since the answers provided by the interlocutor were not relevant. According to Büring, a proposition is relevant in a given discourse if it answers the QuD. A question, in turn, can be considered as relevant if at least one of the answers given to it answers the QuD as well.

We have again observed that information structure categories such as (contrastive) topic and focus cannot be unambiguously related to syntactic positions. Interestingly, we find that the above mentioned contrastive topic/focus constituent can appear sentence-finally as well:

(60) -Jó, de mit evett JÁNOS?
   OK, but JÁNOS what ate?
   Yes, OK, but what did JOHN eat?

The repeated question in (60) can be used in exactly the same contexts as the one in (59-c). Gécseg (2001) even derives contrastive topics from sentence-final prominent constituents.

Another phenomenon closely related to contrastive topics is frame setting (Jacobs, 2001; Krifka, 2006). Frame setters restrict the predication to some per-
spective that has not yet been identified by the context, and they can even delimit
the predication that can be made. Consider the following examples:

(61) a. Q: -How is John?
   b. A1: -As for his personal life, he is fine.
   c. A2: -# As for his personal life, he has spent two weeks in hospital
      recently.

Frame setters are marked in the same way (B-accent) prosodically as contrastive
topics in English, which also shows that they both delimit the validity of a predicate
in a certain way, thus I consider frame setters as a type of contrastive topics.

How could we thus capture the connections between the prominent elements of
the information structure and the prominent syntactic positions? I believe that this
is not possible without examining the discourse in which the particular sentence is
uttered. Sometimes a sentence contains two foci, because it answers a multiple
question with a single pair (and thus the answer does not contain a strategy à la
Büring, decomposing the main question into subquestions), but it is also possible
that it is contrasted to a previous sentence containing multiple foci. As we have
seen, in some cases a question repeats a previously asked question, focusing the con-
stituent in connection with which the answer is expected, but leaving the repeated
question word in the immediately preverbal position. However, when a multiple
question is answered with the help of a strategy, one of the prominent elements
is a contrastive topic (which, in turn, cannot always easily be distinguished from
focussed constituents). The prominent elements appear in different syntactic posi-
tions following these discourse patterns in Hungarian. The information structure
assumed in this thesis is presented in Chapter (7).

Since the present analysis is conducted in the framework of Lexical-Functional
Grammar, in which semantic information is also present in the information struc-
ture (Dalrymple and Nikolaeva, 2011), the next section provides an overview of the
different semantic approaches treating questions.

### 4.6 Semantic Approaches

Semantic theories of questions often refer to the **focus** of the answer, which is
supposed to be tightly related to the question word. (Information) focus, as we
have seen, is the constituent that answers the question (the rest of the answer can
even be left out). Let us examine some semantic theories dealing with questions:

#### 4.6.1 Proposition set theories

These analyses go back to Hamblin (1973) and Karttunen (1977) and assume that
the meaning of a question is the set of its possible answers, which means a set of
propositions.

(62) a. -What did John eat?
   b. -{John ate CHICKEN, John ate TURKEY, John ate FISH... }
Karttunen (1977) refines this, arguing that the meaning of indirect (embedded) questions is the set of propositions that constitute a true and complete answer to the question. He relates indirect interrogatives to declarative sentences and argues that the meaning of direct questions can be treated as equivalent to the corresponding indirect interrogative, embedded under certain verbs:

\[(\text{I ask you to tell me}) \text{ what John ate.}\]

As for polar interrogatives, theoretically they have two possible answers in this framework, a positive and a negative, no matter if the interrogative itself is positive or negative:

\[(64) \begin{align*}
\text{a. Did John go to the party?} & \{\text{Yes, he did.}; \text{No, he didn’t}\} \\
\text{b. Didn’t John go to the party?} & \{\text{Yes, he did.}; \text{No, he didn’t}\}
\end{align*}\]

Thus, according to this approach, (64-a) and (64-b) are synonymous. As we have seen above, the positive answer to negative polar interrogatives is marked in a couple of languages, in that it differs from the usual positive answer given to polar questions. Nevertheless, the lexical difference does not exclude the existence of positive answers in those cases.

This approach is criticized by Krifka (2001), to be discussed in section (4.6.3).

### 4.6.2 Alternative semantics

Rooth proposes an analysis of foci in terms of the alternatives focused constituents are supposed to introduce, but also treats questions (question-answer pairs) in his account. According to Rooth (1992), a sentence containing a focused constituent has both an ordinary and a focused semantic value. This latter means a set of alternative propositions, which differ only in the alternatives introduced by the focus:

\[(65) \begin{align*}
\text{a. John bought APPLES on the market.} \\
\text{b. \{John bought APPLES on the market, John bought PEARs on the market, John bought CHERRY on the market, John bought PLUMS on the market... \}}
\end{align*}\]

Question words are also supposed to introduce alternatives, which are, in turn, the set of propositions potentially answering the question. Contrary to the case of assertions, the set of propositions is considered to be the ordinary semantic value of questions. Concerning questions, Rooth follows Hamblin’s analysis, and assumes that in the case of a congruent question-answer pair (see below), the ordinary semantic value of the question is a subset of the focus semantic value of the answer. Beaver and Clark (2008) also adopt the Rooth-Hamblin approach to formalize the semantics of questions and foci.

\[(66) \begin{align*}
\text{a. Who travelled to Berlin?}
\end{align*}\]
b. \{ JOHN travelled to Berlin, MARY travelled to Berlin, ROSE travelled to Berlin, FRED travelled to Berlin... \}

The analysis assumes that the question in (66-a) presupposes that someone travelled to Berlin and asks for the identification of this person. However, the question can clearly be answered in the following way too:

(67) a. Q: -Who travelled to Berlin?
   b. A: -No one.

Now, we can suppose that the alternative set of the question word contains the empty set as well, or another way of approaching the issue is to claim that the answer in (67-b) cancels the presupposition that someone came, and the cancellation of this presupposition has to be accommodated by the asker. Haida (2007) argues for this latter option, showing that an answer containing only the presupposition of the question is inadequate:

(68) a. Q: -Who travelled to Berlin?
   b. A: # -Someone (travelled to Berlin).

The question is, now, if the alternative set of the focus can also contain the empty set or not:

(69) John bought APPLES/ PEARS/ PLUMS/ NOTHING on the market.

I believe that the alternative set of the focus does not contain the empty set. Nevertheless, it can be used in the place of the focus, canceling the original presupposition, and it has to be accommodated by the interlocutor. Consider the following corrective examples from Hungarian:

(70) Hungarian:
   a. János ALMÁT vett a piacon.
      János apple.ACC buy.PST the market.SUPERESS
      John bought APPLES on the market.

   b. Nem! (János) KÖRTÉT vett (a piacon).
      No (János) pears.ACC buy.PST (the market.SUPERESS)
      No! He bought PEARs.

   c. Nem! Nem vett SEMMIT.
      No not buy.PST nothing.ACC
      No! He didn’t buy anything.

   d. Nem! SEMMIT sem vett.
      No nothing.ACC not buy.PST
      No! He didn’t buy anything.
4.6. Semantic Approaches

In the first correction (70-b), the presupposition is not cancelled, but the focus is replaced with one of its alternatives. Consider now the second correction (70-c), in which we find two negative elements in the sentence (Nem and SEMMIT). The first negates the verb (the truth of the proposition), whereas the second is contrasted to the focus of the previous sentence. This illustrates that (in Hungarian), when a sentence is a correction of another one containing a focus in a way that it introduces the empty set, it has to cancel the presupposition of the sentence (verbal negation) and contrast the focused constituent with the constituent indicating the empty set. Sentence (70-d) shows that when SEMMIT precedes the negative particle and the verb, the former undergoes agreement (nem-sem). Nevertheless, the matter necessitates further research.

However, question words contain certain restrictions concerning the semantics of the answer, which is not the case of focused material. For instance, questions containing who license only [human +] entities as answers, whereas questions containing what expect inanimate/[human −] entities in the answer. It is not obvious how the Proposition set theories account for these differences and further specifications would be expected in alternative semantics as well. Let us now have a look at an approach that captures this side of the problem.

4.6.3 The structured meaning approach

Krifka (2001) argues that the Structured Meaning Approach is, from many points of view, more adequate in describing the phenomena related to interrogatives, than theories considering the semantics of questions as the set of possible answers. According to Krifka (2001) (and the precursors⁷), questions are functions, that, when applied to the answer, give a proposition as a result. A question consists of two parts, which constitute an ordered pair: the restriction and the background. The background is a lambda-expression replacing the question word with a variable. The role of the restriction is to delimit the domain of the function: for example, in the case of who, to the set of human entities, in the case of what to the set of inanimate entities. Let us have a look at the following example:

(71) Who did Mary see? \((\lambda x \ [\text{saw(M)(x)}], \text{humans})\)

According to the structured meaning approach, if John is substituted for the variable, we get a proposition. The answers are similar ordered pairs: they contain a focused constituent and the background.

(72) (She saw) JOHN. \((\lambda x \ [\text{saw(M)(x)}], J), \text{where John } \in \text{humans}\)

The question-answer pair is congruent if the background in the question and in the answer is the same, and the focus is an element of the restriction of the interrogative word.

In fact, with the lambda expression, the Structured Meaning Approach divides the question itself into two parts: the interrogative word (represented by the vari-

able) and the rest (the presupposition). In the proposition set approaches these two are not separated; all the propositions are full answers, differing only in the element that substitutes for the question word in the question. The background is the same in all of them. It is thus common in the two types of approaches that they treat the meaning of questions as the set of possible congruent answers. However, in the proposition set approaches, this set contains the set of full answers to the question, whereas the *Structured Meaning Approach* leaves only the focus part of the question/answer open. According to Krifka (2001), this has several advantages. For instance, it makes it possible to interpret pair-list multiple questions as functions (which would account for many other characteristics of these), or to exclude answers containing over- or underfocused material with respect to the question (in this case, the background of the question and the answer would not be the same).

As far as polar interrogatives are concerned, they are also represented as a lambda expression and a restriction. *Yes* and *no* are assumed to be propositional operators that retains and reverse the truth value, respectively.

\[(73) \quad \text{Did John go to the party?} \quad \langle \lambda f[\text{GO}(P)(J)], \{\lambda p[\neg p], \lambda p[p]\} \rangle\]

As for alternative interrogatives, their meaning is constructed differently from that of polar interrogatives:

\[(74) \quad \text{Does John want 'tea or 'coffee?} \quad \langle \lambda x [\text{want}(J,x)], \{\text{tea, coffee}\} \rangle\]

As can be seen in the example, the restriction of the interrogative is restricted to the alternatives appearing in the question. The answer has to be one of those.

Another advantage of this approach is that it adequately distinguishes between term answers and full answers. Term answers are *John* in (72). If the question meaning is applied to the term answer, what we get is the full answer *She saw John*.

### 4.6.4 Questions as propositional abstracts

Ginzburg and Sag (2000) propose to analyze questions as *propositional abstracts*. In their semantic ontology, the authors shed light on subtle differences between declaratives and interrogatives, emphasizing that not all declarative sentences denote propositions, and that interrogatives have also different interpretational properties in main clauses and in independent contexts. Let us examine the proposed ontology in more details with the help of a few examples. Most declarative clauses denote *propositions*, *i.e.* semantic entities, which can be either true or false. However, we can also find examples in which a fact gets semantically contributed. This is the case of some cognitive predicates:

\[(75) \quad \text{Jack knows that Mary doesn’t like fish and chips.}\]
\[(76) \quad \text{Peter forgot that Suzy’s favourite colour was red.}\]
\[(77) \quad \text{Sam discovered that Kim was hanging out with Bill.}\]

The same is true for some emotive predicates:
4.6. Semantic Approaches

(78) John regrets to have insulted Bob.

(79) It surprised Robert that Kate could cook so well.

The complement clauses of factive verbs are presuppositional, *i.e.* the negation of the sentence do not affect the presupposition (Kiefer, 2007):

(80) Jack doesn’t know that Mary doesn’t like fish and chips.

(81) Peter didn’t forget that Suzy’s favourite colour was red.

(82) Sam didn’t discover that Kim was hanging out with Bill.

(83) John doesn’t regret to have insulted Bob.

(84) It didn’t surprise Robert that Kate could cook so well.

The truth of these facts is thus already established in the discourse universe, prior to the above illustrated assertions. These are called factive or resolutive predicates. Compare the above examples with the so-called TF (true-false) predicates, which can co-occur with complements about which truth can be predicated:

(85) Jack believes that Mary doesn’t like fish and chips.

(86) Peter asserted that Suzy’s favourite colour was red.

(87) Sam denies that Kim was hanging out with Bill.

Interestingly, factive verbs can also appear with interrogative complements:

(88) Jack knows whether Mary likes fish and chips.

(89) Peter forgot what Suzy’s favourite colour was.

(90) Sam discovered who Kim is hanging out with.

But not TF predicates:

(91) *Jack believes if Mary doesn’t like fish and chips.

(92) *Peter asserted what Suzy’s favourite colour was.

(93) *Sam denies who Kim is hanging out with.

These examples clearly show that interrogatives cannot be analyzed as denoting propositions, like declaratives do. On the other hand, there are also so-called Question Embedding verbs, which, in turn, take only interrogative clauses (denoting questions) as complements.

(94) Jack asked if Mary liked fish and chips.

(95) Peter wondered what Suzy’s favourite colour was.

(96) Sam investigated whether Kim is hanging out with Bill or with Max.

According to Ginzburg and Sag (2000), factive predicates always take facts as their complement. This is no problem in the case of declarative sentences, since
they can denote either facts, or propositions. Factive verbs predicate of the fact that proves the proposition true. The case of interrogatives, however, is different. Interrogatives always denote questions. At the same time, factive predicates, when they embed interrogative complements, do not take the question denotation as their argument, but the answer to a question (which is, like in the case of declaratives, a fact). This is a clear case of coercion, in which a question (through its answer) gets interpreted as a fact in certain embedded contexts. In addition to some declaratives and embedded questions, the third type of facts is exclamatives. The semantic ontology presented here is thus enlarged compared to the declarative-interrogative dichotomy, denoting propositions and questions, respectively. Declaratives can also denote facts, and interrogatives can also do so in some embedded contexts via coercion. We should conclude that interrogatives complementing factive predicates are not real questions which are to be answered, but facts. In addition to propositions, questions and facts, the fourth important type of message in the ontology is that of outcomes. Outcomes are ’futurate’ events, the denotations of subjunctive, imperative and certain infinitival clauses:

(97) Peter demanded that Sue be included.
(98) John demanded to go out of the room.

Since the main concern of this thesis is the analysis of (multiple) questions, I will not be dealing with this ontology in details any further. We should note, however, that interrogatives embedded in factive predicates are to be treated differently from those in Question Embedding ones.

Let us now turn to the analysis of questions. In this framework, the semantic analysis of questions is akin to the Structured Meaning Approach (Krifka, 2001), in that it does not equate the interpretation of questions with the set of possible/exhaustive answers (which are complete propositions), but relies on the contribution of the wh-word/phrase to the question. Unlike Krifka (2001), where questions are open propositions, in which a variable (represented by the question word) is uninstantiated, Ginzburg and Sag (2000) propose that questions are propositional abstracts, or more precisely, restricted propositional abstracts, which are more complex, but complete semantic units on their own. Proposition abstracts are composed of the body of the abstract, which contain place holders, which indicate where the abstraction occurs. Restriction at the level of the abstract means that the propositional abstract is situationally relativized with respect to the situation the question is relevant to. Consider the following example:

(99) Which foreign students were rejected?

(Ginzburg and Sag, 2000, p. 134, 7)

Without entering into the formal details, the restricted propositional abstract of this question will contain reference to the particular university and its location and thus restricts the set of potential answers to those who are from a different country and applied to that university.
4.7. Possible Answers

*Wh*-expressions play a double role in this system. They enable an abstraction to occur, over the parameter that they associate with the semantic argument role they fill, and introduce certain restrictions over that argument role (personhood in the case of *who*, inanimateness in the case of *what*, etc.).

In her LFG analysis of questions, Mycock (2006) adopts Ginzburg and Sag (2000)’s analysis of question semantics applying glue semantics as the language of logic. Without questioning the adequacy of that kind of analysis, I adopt the Structured Meaning Approach in this thesis. I believe that the two analyses are compatible with each other and can describe the same phenomena with the same adequacy (except for interrogative complementation phenomena, which is not sufficiently accounted for in the Structured Meaning Approach). The Structured Meaning Approach is chosen, because the difference between term answers and full answers on the one hand, and between congruent/underfocused-overfocused/implicative-non-implicative/direct-indirect answers on the other (see section (4.7)) is part of its semantic representation, whereas according to the propositional abstract theory, all questions are treated equally as propositional abstracts and these differences are treated as part of the contextual component (these particular issues have not yet been explored in detail in the latter framework though).

Bearing in mind the observations made in the previous sections (information structure, semantics), in the next section we examine the possible answers that can be given to questions.

### 4.7 Possible Answers

This section is based on Marandin et al. (2009).

#### 4.7.1 Congruent and non-congruent answers

As we have seen, according to the Structured Meaning Approach, in a congruent question-answer pair ”the background of the question and the answer must be the same, and the focus of the answer must be an element of the restriction of the question” (Krifka, 2001, p. 9). Congruence is a purely semantic concept, which by no means predicts that only congruent answers can be given to a question. Non-congruent answers include those that are acceptable as reactions, but do not answer the question:

(100) a. Q: -What did John eat?
    b. A1: -I don’t know.
    c. A2: -I won’t tell you.

Other non-congruent answers can be over-informative or under-informative with respect to the question:

(101) a. Q: -What did John eat?
    b. A1: -John ate pizza and even invited his girlfriend for dinner.
    c. A2: -John’s girlfriend ate pizza.
(101-b) is over-informative, whereas (101-d) is under-informative. In both cases, it is not only the variable in the question and the focus in the answer that are different in the question-answer pair. In addition to the focus, (101-b) contains additional new information, whereas (101-d) does not contain the narrow focus which would qualify as a congruent answer. It seems that (101-c) does not answer the question. However, it is acceptable as an implicative answer, with which the answerer indicates his strategy of answering the question (in this case what we know is that John surely did not eat pizza, but the answerer cannot or does not want to tell what he exactly ate).

4.7.2 Implicative and non-implicative answers

As shown in (101), an answer can be implicative or non-implicative. This dichotomy is rather pragmatic in nature and independent of congruence. As we have seen, implicative answers introduce a complex discourse strategy (Büring, 2003), indicated by a contrastive topic (B-accent in English, C-accent in French, rising tone on the element on the left periphery in Hungarian), in which the answer is partial: it splits the original question into subquestions, and answers one/some of those subquestions, but not necessarily all of them. Partial answers are thus under-informative:

(102) a. Q: -What did the pop stars wear?
   b. A: -The \([CT_{female}]\) pop stars wore \([F] caftans\].

(CT stands for Contrastive topic, F for focus.) In this example, the implicit subquestions are: \{What did the male popstars wear?; What did the female popstars wear?\}, and only one of them gets answered. This answer is implicative and non-congruent.

An answer can be implicative and congruent at the same time. Consider the following example from the previous section:

(103) a. Q: -Did your wife kiss other men?
   b. A: \([-CT_{My wife}] [F couldn’t] kiss other men\].

Here, although the answer is congruent, a complex strategy is introduced, where a global question would be \textit{Whose wife kissed other men?}, which is then split into the following subquestions in the answer: \textit{Did my (the answerer’s) wife kiss other men?} (= the original question); \textit{Did your (the asker’s) wife kiss other men?}; \textit{Did X’s wife kiss other men?}. The pragmatic import of the answer resides in the more universal question implied by the answer, which suggests that not only the answerer’s wife should be taken into consideration, but others’ as well (amongst whom the asker’s).

Furthermore, non-implicative answers can also be congruent (104-b), or non-congruent (105-b):

(104) a. Q: -What did John eat?
   b. A: -(John ate) pizza.
4.7. Possible Answers

(105)  a. Q: -What did John eat in the restaurant?
    b. A: -John never goes to restaurant.

4.7.3 Direct and indirect answers

As a third distinction, an answer can be direct or indirect. In an indirect answer, sometimes new (implicit) questions have to be accommodated in order to understand in which way the answer provides the missing information indicated by the question. Indirect question-answer pairs are shown in the following examples; (106-b) is a general answer, whereas (107-b) is over-informative:

(106)  a. Q: -Did John invite Mary to the party?
    b. A: -John invited all his friends to the party.
(107)  a. Q: -Did John call someone during the holiday?
    b. A: -He called Mary every week.

In (106-b), an implicit, more general question is introduced (Who did John invite to the party?). The answer refers to this question, and the answer to the original question can be inferred from the answer given (Mary is John’s friend, so if all John’s friends were invited, then Mary was also invited). In (107-b), a polar interrogative, the answer expected is yes or no. However, the answer is also provided to probable subsequent questions, such as Who did John call?; How often?. Büring (2003) and Marandin et al. (2009) represent the question-answer structures with the help of discourse trees, which I presented in the previous section, and attempt to integrate into the LFG framework in Chapter 7.

4.7.4 Answers to multiple questions

I showed in Chapter (2), multiple questions are either answered by a single pair or with a pair-list. The direct answer given to a single-pair question contains multiple foci, corresponding to the question words in the question. However, such questions can receive Indirect, or Partial Implicative Answers (see below) as well.

The direct answers to pair-list questions are more complicated. They are answered with a complex strategy, in which the answers contain contrastive topic-focus pairs and the subquestions are implicit. However, all the subquestions get answered:

(108)  a. Q1: - Who ate what at the party?
    b. A: - Julia ate COOKIES, George CHOCOLATE CAKE, and Mary PIZZA.

The answer suggests that there were three people present at the party: Julia, George and Mary. The d-tree representing this discourse is the following (Figure (4.8)):

Such answers differ from what Marandin et al. (2009) call implicative answer in that here all the implicit subquestions get explicitly answered, thus the contrast between them is not only implied, but manifested. In order to distinguish between implicative answers that only imply the presence of other subquestions and answers,
and those in which the answers to the subquestions are all explicitly present, I will call the former *Partial Implicative Answers* and the latter *Complete Implicative Answers*. Note that pair-list answers to multiple questions, although they provide a complete answer to the question, are implicative in the sense that every single answer implies that others (or at least one other) must also exist. In the next chapters, before going on to the analysis, I give an overview of some relevant aspects of Hungarian and French.

### 4.8 Summary

After the presentation of the empirical domain (the Hungarian and French data), and a typological outlook, the aim of this chapter was to provide the necessary background for the analysis with respect to the information structure and semantics. Since the basic notions of information structure (topic, focus, etc.) are applied differently from language to language and from analysis to analysis, and are notoriously difficult to define, after examining a wide range of phenomena, I gave a discourse-based definition of topic, focus, and their subtypes. Topics are considered as the constituents that relate the sentence to the preceding discourse by introducing a subtopic of the discourse topic, whereas foci are the (semantically) prominent elements of replies (answers to questions, corrections, etc.). These are all crucial in the analysis of multiple questions, as we will see in Chapter (7). As semantic descriptions are part of the information structure representation in LFG, in the second part of this chapter, I considered how questions are dealt with in different semantic frameworks and motivated the choice of the Structured Meaning Approach. Finally, I presented three different ways in which questions can be answered, and showed that a further distinction is needed in the *implicative answers* type, differentiating between partial answers and complete, but implicative answers, which constitute the answer to a pair-list multiple question.
5.1 Introduction

In this chapter, I present the basic syntactic structure of Hungarian, with special emphasis laid on the role discourse structure plays in syntax. I discuss the so-called topic- and focus positions in details and argue that although elements bearing these discourse functions usually appear in designated syntactic positions, it is not well-founded to identify these positions (functional projections) with discourse functions (and name them after those). Next, I will go on to the most important aspects of Hungarian prosody. In the second half of the chapter, I will treat interrogatives and examine the formal and interpretational properties of Hungarian single questions,
including syntax, prosody and semantics. The chapter is based on works by Katalin É. Kiss (É. Kiss, 1992b, 1995; É. Kiss et al., 1999; É. Kiss, 2006), a comprehensive grammar of Hungarian (Kiefer, 1992) and on a new descriptive grammar (Kálmán, 2001).

5.2 The Basic Syntactic Structure: a Topology

Schematically, the Hungarian sentence can be divided into two fields: the topic and the comment, and the comment can be further divided into four subfields: the pre-comment, the prominent preverbal position, the finite verb, and the postverbal part. This is illustrated in Figure (5.1).\(^1\)

![Figure 5.1: The structure of the Hungarian sentence](image)

Although the names (topic, comment, prominent preverbal position, etc.) are of semantic/pragmatic nature, there are also syntactic (distributional) and prosodic arguments for this division of the Hungarian sentence into fields and subfields. However, they reflect the assumption that the structure of the Hungarian sentence does not encode grammatical functions, like in configurational languages, but the way the sentence and its parts relate to the discourse in which the sentence is uttered.

5.2.1 Formal aspects: prosody

Hungarian is a syllable-timed language. In content words, stress always falls on the first syllable and there is only one stress in each word. Stressed syllables bear a pitch accent. Function words, on the other hand, are not stressed. According to Kálmán and Nádasdy (1994), although phonetically several degrees of stress can be distinguished in Hungarian (including secondary stress), phonologically only the main stress and the lack of stress (deaccenting) play a role. The main stress can be defined as a type of stress that starts a particular intonation pattern associated with a certain meaning. Secondary stress is the type of main stress that can never start any intonation pattern. Since secondary stress is never obligatory and it does not carry any meaning, it can be eliminated from phonological descriptions. Kálmán and Nádasdy (1994) argue that all (content) words can bear main stress, but...

---

\(^1\)Note that (5.1) is a schematic, topological representation, and not a syntactic structure proposed in a particular framework.
consequently the question to answer is when they are deaccented and why. These factors are partly syntactically determined.

The topic field and the comment can be divided prosodically: the elements in the topic field exhibit a rising intonation pattern, constituting a preparatory section to the first obligatory stress, which falls on the first (content) word of the comment, starting a falling intonation pattern. This main stress can be followed by other main stresses in a sequence of downsteps (in a level-prosody sentence). The verb can lose its stress, cliticizing on the element in the prominent preverbal position (verbal modifier, negative word, etc.):

(1) 'János 'fát vág az 'erdőben.
    János wood chop the forest.INESS
    John is chopping wood in the forest.

(2) 'János 'kirándulni akar.
    János hiking want
    John wants to go hiking.

(3) 'János 'nem akar 'kirándulni.
    János not want hiking
    John doesn’t want to go hiking.

In (1) the finite verb is preceded by a bare nominal modifier, in (2) by an infinitive, and in (3), by a negative word.

The first obligatory stress can also be followed by a so called eradicating stress (a sharp falling pitch accent) on the element in the prominent preverbal position (4), which can only be followed by another eradicating stress (on the right periphery) (5), otherwise the rest of the sentence is deaccented, or all other main stresses are reduced:

(4) "JÁNOS vág fát az erdőben.
    János cut wood the forest-in
    It is JÁNOS who is cutting wood in the forest.

Since the Hungarian sentence can contain more than one focused constituent, all of them can bear an eradicating stress deaccenting the elements that follow them, except for the focused ones. In these structures, more than one constituent is contrasted to parallel ones in another sentence:

(5) A "LÁNYOK nyerték meg tegnap a "KAJAKVERSENYT, a
    the girls won VM yesterday the kayak contest, the
    "FIÚK pedig a "KENUVERSENYT.
    boys and the canoe contest
    It was the girls who won the kayak contest yesterday, and the boys who
    won the canoe contest.
In some cases, the eradicating stress can appear on an element in the topic field as well, but it obligatorily has to be followed by another eradicating stress in the precomment or in the prominent preverbal position. If the prominent preverbal position is filled, the verb is deaccented and cliticized onto the element in that position. If the position is not filled, the verb can carry a main stress or an eradicating stress.

The main difference between the two deaccenting types (the enclisis of the verb, and deaccenting in the presence of an eradicating stress) is related to both syntax and semantics. The former can be found in neutral ("all-focus") sentences (see Section (5.2.3)). It is important to point out that the main stress on the element in the prominent preverbal position does not deaccent the other main stresses in the postverbal sentence part. However, the pitch gradually decreases towards the end of the sentence. The latter type appears in the presence of focus in narrow-focus (non-neutral) sentences. Such sentences have three components from the phonological point of view: material preceding the focus bears lexical stress, the focus eradicating stress, and the postfocal material is deaccented. The two sentence types have different formal and semantic properties.

5.2.2 Distribution

Concerning the distribution of the elements in these fields and positions, we can observe that some positions can be freely occupied by elements, whereas others are more restricted. The topic field is usually reserved for definite or specific indefinite noun phrases, referential (time and place) adverbials (individualizable elements), whose order is free in the topic field. However, the rightmost position of certain sentence adverbials, like *tegnap* (yesterday), *idén* (this year) indicates the right frontier of the topic field itself as well. These adverbials are interpreted as sentence adverbials in the topic field (6), but as referring only to the immediately following constituent in the comment (7).

(6) A *'vonaton 'tegnap sok 'gyerek 'utazott.*
the train.SUPERESS yesterday a lot of child travel.PST
Yesterday, there were a lot of children travelling on the train.

(7) A *vonaton "tegnap utazott sok gyerek.*
the train.SUPERESS yesterday a lot of child travel.PST
It was yesterday that a lot of children were travelling on the train.

Example (6) conveys information about a fact, answering a question like *What happened?*, whereas (7) contrasts *tegnap* (yesterday) with another day (or other days), when supposedly only few children were travelling on the train.

The set of elements that can appear in the topic field carrying an eradicating stress is larger. Carrying an eradicating stress, infinitives, adjectives, bare nouns,
quantifiers, verbal modifiers, and adverbs (other than the ones mentioned above) can appear in the topic field.

In the precomment part, we find the various distributive quantifiers that follow a given order. Kálmán (2001) classifies them based on their order into the IS (also)-group, the MINDEN (all)-field and the SOK (a lot)-position. The precomment does not constitute the main interest of this thesis, but I will refer to it when necessary.

The prominent preverbal position (PPP), which is between the precomment and the finite verb, can also be occupied by a wide range of elements. Some of them appear in the PPP in level-prosody sentences and can receive an eradicating stress 
\textit{in situ}. However, they must follow the verb if there is another element that carries the eradicating stress. The explanation is that only one of them can precede the verb, when there is more than one potential element that can occupy the PPP in a sentence. The others occupy postverbal positions (except for some questions in which there is also a focused constituent).

Kálmán (2001) refers to these elements as \textit{verb carriers}, referring to the fact that the element in that position always bears main stress and the verb following it is destressed and cliticized on the preverbal element (encasis). Since in Kálmán (2001), the term \textit{verb carrier} refers, at the same time, to the position and to the elements that can appear there, the term \textit{verb carriers} is kept as a collective name for elements appearing in the preverbal position, and the position itself will be referred to as \textit{prominent (immediately) preverbal position}. Let us now enumerate the possible elements in that position (based on Kálmán (2001)):

- **Verbal Modifiers (VM)**
  - **Verbal Particles**

   Verbal particles can have an adverbial or a lexicalized aspectual meaning. When there is no other potential verb carrier, they precede the verb, otherwise they follow it:\footnote{In what follows, verbal particles will be referred to as verbal modifiers \textit{VM}. Verbal particles are written as one word with the noun when they precede it, but as two words when they follow it.}

   \begin{enumerate}
   \item \begin{Verbatim}
   János \textit{kiolvasta} a könyvet.
   János . VM.read.pst the book.acc
   John finished the book.
   \end{Verbatim}
   \item \begin{Verbatim}
   János egy hét alatt \textit{olvasta ki} a könyvet.
   János one week under read.pst out the book.acc
   John finished the book in one week.
   \end{Verbatim}
   \end{enumerate}

   About a classification and analysis of verbal particles, see Surányi (2009) and Laczkó and Rákosi (2011).

- **Bare nominal complements** (see also de Swart and Farkas (2003))
Another type of verbal modifiers is bare nominal complements, illustrated by the following example:

(10) János **levelet ír.**
    János letter.ACC writes
    John is letter-writing.

Secondary predicates

Secondary predicates co-occur with some (other) argument of the verb, about which they state something. They often express a goal (11) or a result (12), and appear in the immediately preverbal position:

(11) János **Szegedre utazott.**
    János Szeged.SUBL travel.PST
    John travelled to Szeged.

Goal arguments are argued to be secondary predicates, since they state something about the final state of the theme/patient or the agent, or localize it in time or space.

(12) János **pirosra festette a kerítést.**
    János red.SUBL paint.PST the fence.ACC
    John has painted the fence red.

Infinitives often play the role of such secondary predicates and they can also occupy the immediately preverbal position, for instance when the main verb is an auxiliary (13), or when it expresses the oblique goal (or some other) argument of the main verb (14).

(13) Mari **kirándulni akar.**
    Mari to hike wants
    Mary wants to go hiking.

(14) János **kapálni indult.**
    János to hoe set out.PST
    John set out to go hoeing.

Kálmán C. et al. (1985) examine Verb + infinitive constructions in Hungarian, and argue that the finite verb can be considered as an auxiliary only in those in which the infinitive can appear in the prominent preverbal position, and the verb (deaccented) cliticizes on it even in level-prosody neutral sentences. Compare:
5.2. The Basic Syntactic Structure: a Topology

(15) A vevők utálnak várakozni.
the customers hate to wait
The customers hate waiting.

(Kálmán C. et al., 1985): 50, 10

(16) A vevők perelní akarnak.
the customers to sue want
The customers want to sue.

(Kálmán C. et al., 1985): 50, 11

According to this approach, akar (to want) is an auxiliary, whereas utál (to hate) is a main verb.

Now, the infinitive in the immediately preverbal position can also have its own "verb carrier" (for instance, a verbal modifier, like a particle, a bare nominal complement, or a secondary predicate), and in such cases, the auxiliary appears right after this element.

(17) Laci el akart menni úszni.
Laci VM want.PST.1SG to go to swim
Laci wanted to go swimming.

(18) János fát akar vágni.
János wood.ACC wants to chop
John wants to chop wood.

(19) ?? János fát vágni akar.
János wood.ACC to chop wants
John wants to chop wood.

(20) Jánosnak Szegedre kell mennie.
János.DAT Szeged. must to go.1SG
John must go to Szeged.

The authors explain this phenomenon by the assumption that the finite verb will immediately follow the verb carrier, even if it is part of a more complex structure (an infinitive with its own verb carrier). See also Kenesei (2001) for another possible definition and classification of auxiliaries in Hungarian.

• The Hocus

The hocus (introduced by Kálmán (1985a,b); Kálmán et al. (1986), and also referred to in Kálmán (2001)) is a noun phrase expressing some participant
or circumstance in the event denoted by the predicate. Such noun phrases can bear main stress and appear in the immediately preverbal position, when the event denoted by the verb is not particularly newsworthy, or it is a regular event, apart from the circumstance or participant denoted by the hocus, which expresses something unusual or unexpected. In these cases the main proposition of the sentence is the identification of this participant or circumstance.

(21) János tegnap \textbf{vonattal} utazott haza.
    Yesterday John by train travel.PST home
    Yesterday John took the train to go home.

(22) Ma a \textbf{feleségem} vitte az óvodába a
    today the wife.POSS.ISG take.PST the kindergarten.ILL the
    gyerekeket.
    children.ACC
    Today my wife took the children to the kindergarten.

Example (21) implies that John usually does not take the train, and (22) that usually it is not the wife, but someone else that takes the children to the kindergarten.

In identificational sentences, the subject appears as the hocus, preceding the verb (copula):

(23) \textbf{János} volt az \textbf{írászató}.
    John was the director.

In the mainstream linguistic literature on Hungarian, sentences containing a hocus are not discussed, and they are not clearly distinguished from narrow-focus sentences. This is a problem, since \textit{hocus} and \textit{focus} are clearly different (see below).

- Monotone decreasing quantifiers and negative adverbs

Monotone decreasing quantifiers (24) (opposed to monotone increasing ones that appear in the precomment) and negative adverbs (25) also occupy the immediately preverbal position:

(24) \textbf{Kevesen} jóttek el a bulira.
    few come.PST VM the party.SUBL
    Only few people came to the party.
5.2. The Basic Syntactic Structure: a Topology

(25) János ritkán megy el kirándulni.
    János seldom goes VM to hike
    John seldom goes hiking.

• The Focus

The focused constituent differs from the above mentioned verb carriers in that it bears sharp falling pitch accent, functioning as eradicating stress, referring to the fact that no main stress (only another eradicating stress) can follow it in the rest of the sentence. In Hungarian, the main function of focus is contrast, i.e. it identifies the entities about which the predicate holds and restricts the validity of the predicate to only these entities by excluding the other members of the relevant set. (See Chapter 4 and Section (5.5.1) of this chapter about the exhaustivity of Hungarian focus.) Sentences with focus cannot be uttered out of the blue. In most cases, they are answers to questions (26), reactions or corrections (27) (capitals indicate the constituent carrying a pitch accent):

(26) Answer:
    a. Q: -Ki hívta meg Marit a bulira?  
       who invite.pst VM Mari.acc the party.subl  
       Who invited Mary to the party?

    b. A: -ZOLI hívta meg (Marit a bulira).  
       ZOLI invite.pst VM (Mari.acc the party.subl)  
       It was ZOLI who invited her (to the party).

(27) Correction:
    a. -Mari tegnap kiolvasta a Háború és békét.  
       Mari yesterday VM.read.pst the War and Peace.acc  
       Mary finished yesterday War and Peace.

    b. -Nem, a BÚN ÉS BÚNHÖDÉST olvasta ki.  
       no, the Crime and Punishment.acc read.pst VM  
       No, she finished Crime and Punishment yesterday.

Although focus is defined here at the semantico-pragmatic level, we should note that in Hungarian, it is formally highlighted in syntax and prosody in the way described above. We will see later on in this chapter that if the focus is a universal quantifier, it cannot occupy the immediately preverbal position. Similarly, in the presence of a focus in the PPP, a second one must appear on the right periphery of the sentence. Thus the syntactic position cannot help in the identification of all foci in Hungarian.

Interestingly, in some cases the (unique) focus can follow the finite
verb/auxiliary and precede the infinitive, which is the content verb of the sentence, the finite verb contributing less in the semantic sense to its meaning:

(28) Szeretnék most ÉN elmenni színházba.
    would like this time I VM.to go theatre.
    This time I would like to be the one who goes to the theater.

    (Molnár, 2008, p. 26, 30)

(29) Nem akarod TE felvágni a tortát?
    not want.2SG you VM.to cut the cake.ACC
    Don’t you want to be the one who cuts up the cake?

    (Molnár, 2008, p. 26, 31)

In (28), the speaker (ÉN, (I)) is contrasted to the other people, implying that on other occasions usually the others go to the theatre, and this time s/he would like to be the one who can go. In (29), TE (you) is contrasted to the other potential people who could cut up the cake identifying him/her as the only one who should do it.

However, this is not obligatory:

(30) ÉN szeretnék most elmenni színházba.
    I would like this time to go theatre.
    This time I would like to be the one who goes to the theater.

Molnár (2008) remarks that this construction is only possible with those auxiliaries that agree in person and number with the subject, and thus the identity of the subject is clear from the verb form.

Going back to the hocus-focus distinction, the following examples (based on Kálmán (2001)) illustrate the difference between them:

(31) ’Ezen a héten a ’Mecsekben raboltak
    this.SUPERESS the week.SUPERESS the Mecsek.INESS rob.PST.3PL
    ki egy ’pénsszállító autót.
    VM a money transport car.ACC
    This week it was in the Mecsek (mountains) that a money transport vehicle was robbed.

(32) ’Ezen a héten a ’Mecsekben raboltak
    this.SUPERESS the week.SUPERESS the Mecsek.INESS rob.PST.3PL
ki egy pénzszállító autót.
VM a money transport car.ACC
This week it was in the Mecsek (mountains) that a money transport
vehicle was robbed.

A Mecsekben is hocus in (31) and focus in (32). The difference between
the two sentences can be illustrated with the different contexts. In the first
case, robbing a money transport car counts as a usual event. The hocus
identifies the place where the event happened this week. The location counts
as non-canonical, unusual and surprising at the same time, either because
this happens less often in mountains, or because the Mecsek is not known
for such crimes. In (32), robbing a money transport car is not necessarily a
usual event. The focus identifies the place where it happened, contrasting it
to other locations, where it could have potentially happened, or correcting a
previously proposed other location. In this latter case, the sentence does not
form a true prosodic minimal pair with (31), but would have the following
form:

(33) (Nem!) "A Mecsekben rabolták ki a pénzszállító
(No!) the Mecsek.INESS rob.PST.3PL VM the money transport
autót (és nem a Bakonyban).
car.ACC (and not the Bakony.INESS)
No! It was in the Mecsek that the money transport vehicle was robbed
(and not in the Bakony)!

(33) is about a single event (indicated by the definite article in front of the
noun pénzszállító autót and the definite conjugation of the verb). It identifies
the location of the event, by contrasting it to another location.

We can thus conclude that both the hocus and the focus are identificational
elements, appearing in different discourse contexts. The focus is prosodically
distinguished, carrying a pitch accent (followed by the deaccenting or reduced
stress of the post-focal material), whereas the hocus is not more prominent
prosodically than the other lexical elements of the sentence (except for the
verb which cliticizes on it). In addition, the focused constituent presupposes
that the proposition cannot be true simultaneously with another, in which the
focused element is changed to an alternative to its denotation (the robbery
cannot take place at two locations at the same time). To illustrate this,
consider the possible continuations of (31) and (32):

(34) a. 'Ezen a héten a 'Mecsekben
this.SUPERESS the week.SUPERESS the Mecsek.INESS
raboltak ki egy 'pénzszállító autót.
rob.PST.3PL VM a money transport car.ACC

b. 'Ezen a héten a 'Mecsekben
this.SUPERESS the week.SUPERESS the Mecsek.INESS
raboltak ki egy 'pénzszállító autót.
rob.PST.3PL VM a money transport car.ACC
This week it was in the Mecsek (mountains) that a money transport vehicle was robbed.

b. Nem, nem csak ott. A Bakonyban is kiraboltak
not not only there the Bakony.INESS too VM.rob.PST.3PL
egyet.
one.ACC
No, not only there. One was robbed in the Bakony too.

(35) a. Ezen a héten a Mecsekben raboltak ki egy
this.SUPERESS the week the Mecsek.INESS rob.PST.3PL VM a
pénzszállító autót.
money transport car
This week it was in the Mecsek that a money transport vehicle
was robbed.

b. #Nem, nem csak ott. A Bakonyban is kiraboltak egyet.

- Question words

Finally, question words typically appear immediately in front of the finite verb as well. In the presence of a question word not only verbal modifiers (36) and other verb carriers, but elements of the precomment (37) also occupy postverbal positions:

(36) Kit hívott meg János a bulira?
who.ACC invite.PST VM János the party.SUBL
Who did John invite to the party?

(37) Kire szavazott mindenki?
who.SUBL vote.PST everybody
Who did everybody vote for?

Question words are often argued to constitute a subclass of focus, based on similarities in prosody, syntactic position, semantics and, in some languages, morphology. Despite the apparent similarities, it would be too hasty a generalization to collapse question words into foci in Hungarian. Although foci and question words seem to share the same syntactic position and prosody, some important differences suggest that they belong to different types of objects. For instance, while strictly only one preverbal focus is permitted in Hungarian (if there is a second, it is obligatorily postverbal), two question words can appear preverbally, and they can even be preceded by a focused constituent or followed by a negated focus in some contexts. Furthermore, as we have seen in Chapter (2) two question words can be coordinated, irrespective of
5.2. The Basic Syntactic Structure: a Topology

their grammatical function, whereas this is not possible with two focused constituents. As will be shown in this thesis, their similarities can also be derived from the fact that they play parallel roles in the discourse they occur in: both question words and foci presuppose the rest of the sentence, and foci in the answers correspond to question words in the question.

The elements in the PPP in (8)-(14) are in complementary distribution with each other, i.e. a verb cannot appear simultaneously with a verbal modifier and a secondary predicate, for instance, even if one of those followed the verb. They can all receive an eradicating stress in situ, in the PPP. However, in the presence of the elements in (21)-(37), they have to follow the verb.

5.2.3 The role of discourse structure

Considering the diversity of elements that can occupy the immediately preverbal position, how could we identify what is in common in them? As we have already seen in the case of secondary predicates, these elements contribute to the meaning of the sentence with a secondary/independent proposition that can sometimes modify the proposition formulated by the comment. According to É. Kiss (2006), not only verbal modifiers can be considered as resultative, locative or terminative secondary predicates, but structural focus can be reanalyzed as a specificational predicate (similarly to English cleft sentences) as well. Komlósy (1994) also showed that preverbal bare nominals function as predicates that predicate of an existentially bound variable incorporated into the verb. I will argue that apart from the common grammatical function (secondary predicates), the common properties of some of the elements in the PPP are related to the information structure and the discourse the sentence is uttered in (I formalize this aspect in Chapter 7).

To see this last point more clearly, an important remark is due here. Some of the above mentioned elements can never appear in the same sentence, since the discourse types they can be part of are different. In Hungarian, based on formal, interpretational and discourse factors, two types of sentences can be distinguished: neutral ("all-focus")\(^3\) and narrow-focus ("non-neutral") sentences (see Kálmán (1985a,b)). Formally, narrow-focus sentences contain an eradicating stress (38) in the PPP and possibly also in the topic field, whereas neutral sentences have level-prosody and can contain several main stresses (39):

(38) "Tegnap "MARIT láttam a városban, (nem JÁNOST).
    yesterday Mari.acci see.pst.1sg the city.INESS (not János)

\(^3\)Neutral sentences can answer questions like What happened?. They cannot always be referred to as all-focus sentences, since they can contain topics:

(i) János találkozott Marival a városban.
    János meet.pst Mari.instr the city.INESS
    John met Mary in the city.

In this example János is the topic, and the sentence is not all-focus.
Yesterday I saw MARY in the city, not JOHN.

(39) = (6)

'János 'tegnap 'vonattal utazott ’haza.
János yesterday travel.INSR travel.PST home

Yesterday John took the train to go home.

In (38), the focus (MARI) bears a sharp falling pitch accent, after which the rest of the sentence is deaccented (or bears reduced stress), except for the second focus (JÁNOST). In (39), the topics (János, tegnap) have a slightly rising tone, pointing forward to the comment, whereas the falling tone starts on the first constituent of the comment, the hocus (vonattal).

The two types of sentences are used in different contexts. Neutral sentences, present mostly in narrative contexts, only convey information and continue the narrative, whereas non-neutral sentences are used for asking questions, answering questions, corrections and confirmations, disagreement, and for highlighting parallels. If we analyze the discourse as the hierarchy of topics and subtopics (or, a question under discussion, subquestions and the possible answers), we see that both sentence types contain two prominent preverbal parts (the topic field and the PPP), and a set of (in the sense of Jacobs (1984)) prominent element types that can fill these positions. Elements in the topic field relate to the discourse in a way that they thematize it by selecting the subtopic/subquestion with respect to which the given sentence adds new information to the common ground. On the other hand, elements appearing in the PPP (or possibly in the precomment), constitute the most informative, prominent part of the sentence. In some cases, this can be new information, or the part that answers a question, or the unexpected or unusual part of the meaning (as we have seen in the case of the hocus).

Concerning the topic field, the elements occurring there have different properties in narrow-focus and all-focus sentences. It is common in the two cases that they introduce subtopics/subquestions. In a neutral context, there is no topic in the sentence if the sentence continues the previous subtopic. However, when a sentence changes the subtopic, the element in the linearly first position indicates the topic shift. This is why this type of topic is often called thematic shifter. The other type of topic that we find in narrow-focus sentences (indicated prosodically with eradicating stress and a rising tone) is closely related to the contrastive property of these sentences and is called contrastive topic in the literature. The contrastive topic restricts the domain of the validity of the focused constituent to some element of a set, implying that to other elements of the relevant set the focused constituent does not hold (/ indicates the rising intonation of the contrastive topic):

(40) a. Q: -Mit hoztak a vendégek a bulira?
    what bring.PST the guests the party.SUBL
    What did the guests bring to the party?
According to Büring (2003), in the example (40-b) the contrastive topic (Mari), indicates the strategy of decomposing the set of guests into its elements, the individual guests and associates each of them with an answer (i.e. a focused constituent). This association means at the same time that as opposed to Mary, there is at least someone else who did not bring a chocolate cake. In this respect, the topic types appearing in the two types of sentences have a similar function: they decompose the main question into subquestions, relating the sentence to the discourse in which it is uttered. Thematic shifters are not excluded from non-neutral sentences either.

Since the focus appears only in non-neutral sentences, this part of the information structure cannot be called focus in every sentence. In questions, this element is the question word itself and, in some cases, a focused constituent, in answers and corrections the focus, and in all-focus sentences the hocus, some (bare) nominal complement or a verbal modifier. Our task is then, either to propose a different architecture of information structure for all-focus and narrow-focus sentences, or to propose a general and more abstract structure that can be filled in different ways by the different sentence types, taking into consideration the context as well. I will present such a general architecture, bearing in mind that the other possibility (different information structures) could also derive the expected results.

The two types of sentences are schematically represented below. The square brackets indicate the two main parts of the sentence (the topic, as we mentioned above, is not obligatory, and sentences can even start with the finite verb when there is no quantifier or focus). The round brackets indicate that the position of ordinary topics with respect to the contrastive topic is optional.

(41) Neutral ("all-focus") sentence

[TOPIC] [COMMENT: precomment, hocus/verbal modifiers/secondary predicates, finite verb, other constituents]

(42) Non-neutral sentence

[(TOPIC), CONTRASTIVE TOPIC, (TOPIC)] [COMMENT: precomment, focus, finite verb, (verbal modifier), other constituents]

In the next section, I provide a brief overview of the analyses of the Hungarian syntactic structure, pointing out some problems related to the above issues.

5.3 Syntactic Structure: Previous Analyses

The Hungarian sentence structure has been analyzed and formalized mostly in the transformational literature supposing that the surface structure of the sentence (which is then pronounced at PF (Phonological Form) and interpreted at LF (Logical Form)) is derived, via transformations, from a deep structure. My aim in this
section is not to argue in details against certain aspects of these models, since they are all more or less coherent parts of a transformational grammar. Instead, I will concentrate on certain problems, with respect to which in a non-transformational grammar we could obtain more adequate solutions.

The proposed models can be characterized by the following factors:

- **Syntactocentrism**

  In the transformational analyses, all linguistic aspects are integrated into the syntax. The above-mentioned discourse functions and quantifiers have their own functional projections (FocP, TopP, QP, DistP, etc.), and in order for an element to be interpreted as topic or focus, it has to move into the specifier position of the corresponding functional projection. However, these functional projections have not received an exact and adequate semantic/pragmatic definition, thus their supposed content is often unclear or vague. As we have already seen (and we will see later), the correspondence between syntactic positions and discourse functions is not one-to-one. An element bearing a certain discourse function can appear in more than one position, and the same position can be filled by different elements. Another important example is that elements usually take their scope following (overt) movement and those that are not in their LF scope positions at surface structure are supposed to move there covertly at LF. This is the case of *in situ* question words, which move covertly to the left periphery, where they scope over the whole sentence. However, Mycock (2006) convincingly argues that scope-taking is also possible in prosody, without supposing covert movements. In these analyses, thus, the information structure (the discourse functions) and semantics (the scope of operators) are integrated into the syntax, and prosody cannot play the same role as syntactic operations (for instance, in scope-taking).

- **Lack of discourse and context**

  The derivational analyses, in most cases, take account of isolated sentences without the appropriate context. Nevertheless, the information structure of the sentence and the discourse functions of its constituents reveal the role the sentence plays in the discourse in which it is uttered. It would be therefore worthwhile to include the discourse and the information structure in the formalism of the analysis, but not as part of the syntactic level. LFG is a theoretical framework that makes possible the dissociation of syntax, information structure and discourse at different levels of representation.

### 5.3.1 Preliminaries

The observation that Hungarian sentence structure reflects information structure, and one of the first analyses about information structure in general, go back to Sánnuel Brassai’s work from the 1850s (presented in the previous chapter), who claimed that both configurational and non-configurational languages have a common structure: some sentence-initial constituent that is discourse-old and constitutes the base for the oncoming information (called the *inchoativum*), which is fol-
5.3. Syntactic Structure: Previous Analyses

followed by the main part (zöm) that conveys information that is new to the hearer. The first part is not obligatory. In configurational languages only arguments in the nominative (i.e. subjects) can be the inchoativum, whereas in non-configurational languages a wide range of arguments can appear in the sentence-initial part.

Brassai’s ideas remained unnoticed in the Hungarian linguistic literature for a long time, until É. Kiss Katalin rediscovered them and based the transformational model of Hungarian syntax on these ideas. Apart from É. Kiss’s, works by Anna Szabolcsi, István Kenesei, Gabi Tóth, Zoltán Bárnéti and Mihály Bródy also greatly contributed to the transformational analysis of Hungarian. Generally, the most important result of the generative transformational approach was that word order in Hungarian is not simply a stylistic question, but it encodes discourse functions and scopal relations. If we change the relative word order of these constituents, the truth conditions of the sentence will also be different, or at least the possible context in which the given sentence can be uttered.

In these analyses, discourse functions are treated as syntactic categories, i.e. discourse notions associated with given syntactic positions. These (later functional) projections are named after one discourse function, after one of the elements that can appear there, and the others are supposed to move there via some supplementary mechanism. Later on, more and more attempts have been made to find the common property of these elements, but then the main challenge has remained to identify the common syntactic position that expresses this common, rather abstract property, and at the same time it is compatible with the current transformational framework.

5.3.2 Early analyses

É. Kiss (1983) defines topic and focus as syntactic positions that can be identified as functions (Top, Foc), and not as categories (VP, NP). The sentence belongs to the non-configurational S category, from which elements move to the left periphery to fulfill the topic or focus function in the corresponding position. The author remarks that elements targeting the focus position constitute a hierarchy: when more than one is present, some of them obligatorily have to appear there, forcing others to occupy postverbal positions. Molnár (2008) notes, however; that focusing is obligatory according to this model (and thus it cannot properly account for neutral sentences that do not contain any focused constituent at all), and that elements raising to the focus position do not have a unified interpretation.

In a latter work (É. Kiss, 1992b), the S node is already configurational. The topic appears on the left periphery as the sister node of the VP, whereas the focus in the Spec,VP position, assuring that it immediately precedes the verb. Focusing is not obligatory, some elements optionally get assigned a focus feature by the verb. The assignment of the [F+] feature is arbitrary. The model does not explain which elements can get it and why (i.e. in which contexts). It is thus still not clear why question words and monotone decreasing quantifiers get assigned a focus feature, due to which they raise to the focus position.

In É. Kiss (1994) the functional projection FocP is already present (the idea
originally comes from Brody (1990)) that dominates the VP and in case if the Spec,FocP is filled, the verb moves into the Foc position, leaving behind the verbal modifiers in the Spec,VP position. Brody proposes an Asp(ectual)P that hosts the verbal modifiers. In both cases, therefore, the focus and the verbal modifiers get dissociated from one another, since they reside in different positions. This approach, nevertheless, still does not shed light on the question, why question words and other elements should move to the Spec,FocP position.

5.3.3 Later developments

Later on, the diversification of elements on the left periphery continued. In É. Kiss (1998b) the whole TopP-FocP sequence can be iterated, in order to account for sentences with multiple focus. The verb moves to the highest Foc position. In Bartos (2000) all morphosyntactic categories have their own functional projection: AgrSP, AgrOP, ModP, FinitP, then in É. Kiss (2008a) PredP and NNP (non-neutral phrase). Concerning the wide range of functional projections, the following problem arises: their head position is, in most cases, empty. Their presence in the structure is motivated only by the single element in their specifier position, thus their explanatory adequacy can be questioned. In addition, as we have already seen above, some of these elements are in complementary distribution with each other, which is not accounted for in a model that hosts all these elements in a separate functional projection. This approach is counter-intuitive in the sense that diverse elements can appear in the same (in this respect neutral) syntactic position, depending on discourse structure.

Another research direction has been to find the common interpretation of elements appearing in the focus position. É. Kiss (2004) argues that verbal modifiers turn the single event denoted by the verb into a complex event. This is on a par with Kálmán (2001) who claims that verb carriers constitute propositions on their own that combine with the proposition in the comment without the verb carrier. Finally, É. Kiss (2005) proposes that the common property of focus and verbal modifiers can be captured in the property of being secondary predicates, and consequently they all reside in the Spec,PredP position.

5.3.4 Problems

Although it is not my aim to criticize the transformational approaches in details, since they have their own internal coherence, I still consider it important to call the reader’s attention to some problems in connection with them. We have already seen that these analyses are syntactocentric. The syntax contains syntactic, morphosyntactic, semantic and discourse information by projecting separate functional projections to almost each of them. This is problematic, because one particular position can be occupied by various kinds of elements and the same element can appear in different positions in the syntactic structure, depending on the other elements in the sentence. However, the particular positions are named after only one element (FocP, TopP). The other drawback of putting only syntax into the foreground is that the theory has to offer syntactic accounts to phenomena that could be more
easily and intuitively accounted for if, for instance, prosody could play a role equal
to syntax. Mycock (2006) argues that in multiple questions *in situ* question words
are focused in prosody and not in syntax. Since both are mapped (in LFG) onto
the information structure, there is no difference between them at that latter level:
both have focus status. In a transformational approach, *in situ* question words are
supposed to move covertly to their scope position. According to Mycock, on the
other hand, the appropriate scope can be signaled prosodically as well. Similarly,
Brody (1990) also suggests that the information structure status of *in situ* question
words is marked intonationally.

The second problem illustrates that not all predictions of the transformational
framework are borne out concerning the syntactic structure. For instance, if the verb
is supposed to move out of the VP to the Foc head, the verbal modifiers are predicted
to follow the verb immediately, which is not always the case (Börjars et al., 1999):

(43) MARI keresett minket fel (és nem Ilona).
     MARI look for.PST us VM (and not Ilona)
     It was Mary who visited us (and not Ilona).

(44) Én csak ezt ismerem körülbelül fel.
     I only this.ACC know approximately VM
     I can only recognize this, approximately.

Finally, it is important to notice that the transformational frameworks do not
account for neutral sentences properly, since the majority of them automatically
include a focus position and projection. Where would the place of the *hocus* be in
such an account? Placing it to some position other than the focus does not seem
to be appropriate either, since they are in complementary distribution depending
on the type of discourse the sentence is part of.

Apart from the problem of the so-called focus position, the topic position is
questionable to the same extent. As we will see in Section (5.5), the topic position
can also be filled with different types of elements depending on the role the particular
sentence plays in the discourse.

5.4 Coordination in Hungarian: the Basic Properties

Since coordination plays an important role in multiple questions, some of its basic
properties are presented here. This section is based on Bánréti (2007), who provides
an extensive study of coordination and ellipsis phenomena in Hungarian.

Bánréti (2007) argues that in a coordination, the conjuncts have to share their
basic grammatical features, such as syntactic category, definiteness, case (nouns);
argument structure, finiteness (verbs), etc. As was pointed out in Chapter (2),
the categorical identity is too strict a requirement, since cross-linguistically, even
conjuncts of different categories can be coordinated provided that they share the same syntactic function.\footnote{About the non-identity of conjuncts, see Dalrymple and Kaplan (2000), who propose set-valued representations for the atomic features such as person, gender, case, etc. (reviewed in Dalrymple et al. (2009), accounting for a wider range of data proposing a feature structure with separate boolean-valued attributes (nom, acc, etc.) for the attribute case, and argue that this analysis can be extended to other features as well), and Sag (2005), who supposes hierarchies of features and their combinations.} This seems to be the case in Hungarian as well:

(45) Péter [beteg vagy nyaral].
Peter is either ill, or on holiday

(46) Péter [festő és bűszke rá].
Peter is a painter and proud of it

(47) ?? Szeretem [a vasárnapot és hogy olyankor semmit se kell csinálni].
I like Sundays and that I don’t have to do anything then.

In (45), an adjective is coordinated with a verb. The adjective (beteg (ill)) forms a predicate without the copula in the third person singular in Hungarian, just like the bare noun (festő (painter)) in (46), where it is coordinated with an adjective (bűszke (proud)). Allegedly, these items can be coordinated based on their common predicative function. In (47), a noun is coordinated with a complement clause, and they share the syntactic function of object.

Bánréti (2007) distinguishes between two groups of coordinative conjunctions: n-ary and binary conjunctions. N-ary conjunctions (és, vagy, etc. (and, or)) form asymmetric coordinate structures, of which they constitute the head. They take the last conjunct as their complement, whereas the others appear as specifier/multiple specifiers in the structure. They can conjoin nominal elements, as well as predicative ones (VPs and clauses). Their function is the unification of the grammatical features of the nominal conjuncts (person, number, case, definiteness), which is reflected by the agreement with the verb form. For instance, in the case of the coordination of a first and a second person subject, the verb exhibits a first person plural form:

(48) [Te és én] elmentünk sétálni.
You and me go.PST.1PL walk
You and I went for a walk.

Needless to say, conjunctions do not have such a unification function if the conjuncts are of predicative nature. In this case, they have a semantic import: conjunction
5.5. Discourse Functions in Hungarian

5.5. Discourse Functions in Hungarian

5.5.1 Focus

The Hungarian focus has been in the center of research on information structure since É. Kiss (1998a) proposed that Hungarian has a structural, called *identificational* focus that can be identified formally by its pitch accent that eradicates all main stresses after it in the rest of the sentence, and by its obligatory syntactic position immediately preceding the finite verb. Concerning its semantics, the identificational focus is claimed to exhaustively identify the individual or individuals about which the predicate holds. At the same time, this means that the other possible candidates of the same set are excluded by this focus operator. Although focus is most commonly defined as the answer to a question, this definition has already been enlarged to corrections and to other types of reactions, claiming that generally the focused constituent is the most prominent part of sentences that are

5 Conventional implicatures refer to those aspects of the meaning of a sentence that reflects the meaning of a specific word, but it is not part of its truth conditions (Matthews, 1997; Kiefer, 2007).
Chapter 5. Relevant Aspects of Hungarian

not uttered out of the blue, but as reactions to some previous sentence(s) in the preceding discourse:

(50) a. Q: -Ki jött el a bulira?
    who come.PST VM the party.SUBL
    Who came along to the party?

    b. A: -MARI (jött el).
        Mari (come.PST VM)
        MARY (came along).

(51) a. S1: -Tegnap látattam Marit és Jánost a városban.
        yesterday see.PST Mari.ACC and János.ACC the city.INESS
        Yesterday I saw Mary and John in the city.

    b. S2: -Az nem lehet! Mari ZOLIVAL volt egész nap.
        'that not possible Mary Zoli.INSTR was whole day
        Impossible! Mary spent the whole day with ZOLI.

As É. Kiss argues, the English equivalent of the Hungarian identificational focus
is the cleft sentence:

(52) It is Mary who came to the party.

We should remark, however, that although cleft sentences also express exhaus-
tive identification, they do not constitute the only natural answers to questions, as
the Hungarian examples do. In other words, there is no need to cleft the focus in
English to answer a question or to correct a previous sentence, the focus can also
be intonationally marked by a pitch accent.

As was already illustrated in (5), in the presence of a preverbal focus, a second
focus is obligatorily postverbal, and in most cases, sentence-final in Hungarian. The
two foci can form a complex focus (see Krifka (1992)), like in (5), or they can be
true multiple foci, like in the following example:

(53) CSAK JÁNOS evett meg CSAK KÉT SÜTEMÉNYT.
    only János eat.PST VM only two cookies.ACC
    It was only John who ate only two cookies.

As discussed in the previous chapter, according to É. Kiss (ibid.), identificational
focus (our information focus) should clearly be distinguished from information fo-
cus (our completive information) that conveys simply unpresupposed information
without exhaustive identification. This type of foci does not appear in the preverbal
position in Hungarian. Compare the following two sentences:

6I slightly modified the glosses with respect to É. Kiss (1998a)
5.5. Discourse Functions in Hungarian

(54) a. Q: -Mit nézett ki magának Mari?
   What look.PST VM for herself Mary
   What did Mary set her eyes on?

   b. A: -Mari EGY KALAPOT nézett ki magának.
      Mary a hat.ACC look.PST VM herself.DAT
      She set her eyes on A HAT./ It was A HAT that she set her eyes on.

      (É. Kiss, 1998a, p. 249, 8a)

(55) a. János és Mari vásárolnak.
      John and Mary are shopping
      John and Mary are shopping.

   b. Mari kinézett magának egy kalapot.
      Mary VM.look.PST herself.DAT a hat.ACC
      Mary has set her eyes on a hat.

      (É. Kiss, 1998a, p. 249, 10)

É. Kiss argues that in example (55) egy kalapot (a hat) conveys new information, without excluding the possibility that Mary picked other things for herself. On the other hand, in (54) the sentence is false if Mary picked out something else as well. To illustrate this, let us have a look at possible continuations of the above examples:

(56) a. S1: -Mari EGY KALAPOT nézett ki magának.
      Mary a hat.ACC look.PST VM herself.DAT
      She set her eyes on A HAT./ It was A HAT that she set her eyes on.
      Correction:

   b. S2: -Nem! Egy SÁLAT is kinézett!
      no a scarf.ACC too VM.look.PST
      No, she set her eyes on a scarf too.

(57) a. S1: -Mari kinézett magának egy kalapot.
      Mary VM.look.PST herself.DAT a hat.ACC
      Mary set her eyes on a hat.

   b. S2: -Igen, és egy sálat is kinézett.
      yes, and a scarf.ACC too VM.look.PST
      Yes, and a scarf too.

In this approach, identificational focus inherently implies some contrast with some contextually determined alternatives. In the case of informational focus such alternatives are not present. É. Kiss proposes to account for structural focus phenomena in different languages with the help of two features: [contrastive +/−] and
[exhaustive +/-] claiming that Hungarian identificational focus is at the same time [contrastive +] and [exhaustive +]. We will come back to these features at some later point of this subsection.

According to Marandin (2006b) (and see also Chapter (4)), all foci that are answers to questions are informational foci. Informational focus can be broad, when larger constituents, or the whole sentence answer the question, and narrow, in answers to constituent questions. If we consider the above examples ((54) and (55)) carefully, we can see that in (55) it would be relatively problematic to find the exact question to which only *egy kalapot* (*the hat*) is the appropriate answer. As we have already seen, such sentences are neutral and answer questions like *Mi történt?* (*What happened?*). However, if this is right, we must also assume that the information focus of this sentence is more than what É. Kiss supposes. It is either the whole sentence, or, if the more general question (the discourse topic, or question under discussion) is *Mari* (*What happened to Mary yesterday?*), it is the rest of the sentence after *Mari*. We can thus conclude that the two types of foci indicated by É. Kiss are, in fact, related to the discourse status of the sentence they appear in. In the previous chapter, we presented examples that illustrate completive information (unpresupposed material that is postverbal in Hungarian, in a narrow-focus sentence). The example is repeated for convenience:

(58)  
| a. | (Gary is an art dealer. Lately he’s been very picky about which museum he deals with; he doesn’t do business with the Metropolitan or the Guggenheim.) |
| b. | Így (csak) a **Louvre-nak** adná ezt az új Modiglianit. |

The constituent *a Louvre-nak* is contrastive focus, contrasted to the other two museums in the context. The painting (*ezt az új Modiglianit*) is, on the other hand, completive information (non-presupposed material in a postverbal position).

Following the terminology introduced in Chapter (4), thus, in Hungarian, contrastive and information foci occupy the prominent preverbal position, whereas completive information is postverbal. Completive information must be distinguished from neutral sentences, in which not only the postverbal, constituent is non-presuppositional, but the whole sentence.

Another example that É. Kiss uses in order to illustrate the difference between the two types of foci is the following:

(59)  
| a. | Q1: *Hol voltatok a nyáron?* |

*Where were you in the summer?*
   go.pst.1pl Italy.iness
   We went to Italy (and nowhere else).

   go.pst.1pl Italy.iness
   We went to Italy (, for instance.)

According to É. Kiss, the first answer exhaustively identifies the place where the answerer went in the summer, whereas the second does not give an exhaustive answer, just mentions an example, and thus, the focus is not preverbal. However, the question-answer structure of this example seems to be more complicated than É. Kiss presents it. In A1, it is presupposed that the answerer (and his family) went somewhere in the summer, and the focus in the answer identifies this place. A2 could be an adequate answer to different questions as well:

(60) a. Q1: -Mit csináltatok a nyáron?
   superess what do.pst.2pl the summer.
   What did you do in the summer?

b. Q2: -Voltatok valahol a nyáron?
   superess be.pst.2pl somewhere the summer.
   Did you go anywhere in the summer?

c. A: -Jártunk Olaszországban.
   go.pst.2pl Italy.iness
   We went to Italy.

As an answer to the first question in (60), Jártunk Olaszországban. contains VP focus, see also Kenesei (1998), identifying what the answerer did in the summer. In this respect, not only the country is new information, but the whole activity. The answer to the second type of question is more complicated. As an answer to Q2, a polar question, A first answers Q2 with a verum focus (this is why the country cannot be preverbal, since in the case of a verum focus, the verb itself is accented), and then, with the help of the postverbal focus, answers the implicit question attached to Q2: Did you go anywhere in the summer; and if so, where? Yes, we did go somewhere and this place was Italy.

An important question that emerges concerning (contrastive/information) focus in Hungarian, is exhaustivity. Should some supplementary mechanism (like the assignment of an exhaustivity feature or an exhaustivity operator) really be assumed to account for the data? Exhaustivity supposes that Hungarian structural focus expresses more than Rooth (1992) or Krifka (2006) assume (focus indicates the presence of alternatives that are relevant for the interpretation of linguistic expressions), in that it exhaustively identifies one of these alternatives by excluding the others from the domain of the predicate. The question is if this exhaustive interpre-
tation comes indeed from some semantic constraint associated with the preverbal position in Hungarian, or is it more linked to pragmatic factors. Onea (2008, 2009a) argues for the latter option. The author remarks that in Hungarian the use of the preverbal focus is more restricted than that of the corresponding prosodic focus in German, in that it typically figures in answers to constituent questions, and as such, answers are expected to be complete. In German, prosodic focus is judged to be less exhaustive than its Hungarian equivalent, since it can be triggered by more factors. In an experiment (Onea, 2009a,b), native speakers of Hungarian were asked to choose from answers to different statements describing exhaustively a non-exhaustive state of affairs on a picture. For instance, in the picture, both John and Mary had a flower. The subjects then heard three types of spoken stimulus:

- containing a focused ‘only x’ subject (i.e. only Mary has a flower)
- containing a preverbally focused subject (expected to express exhaustivity)
- containing a postverbally focused subject (completive information, corresponding to É. Kiss’s information focus)

The task was then to choose one of the following answers:

- Yes, and John has a flower too.
- Yes, but John has a flower too.
- No, John has a flower too.

If exhaustiveness was truth conditional (and not pragmatic), assertions with preverbal focus and with only-subjects would be expected to behave similarly, i.e. the participants would contradict them and use some negative answer (Yes, but … or No, …) to the same extent in both cases. As the results show, stimuli with postverbal (completive information) focus got the least contradicting answers, and those with only-subjects the most. However, answers containing preverbal focus got contradicted at a significantly lower rate than their only-counterparts. The authors conclude that in Hungarian (and also in German) the degree of exhaustivity associated with preverbal focus is much lower than it would be expected in the case of some truth-conditional effect.

Wedgwood also argues against the exhaustivity of preverbal focus in Hungarian (see Wedgwood (2002, 2003, 2004); Wedgwood et al. (2006); Wedgwood (2009)). He assumes that exhaustivity comes from pragmatic factors and that narrow foci (answering questions) are unmarkedly exhaustive. Whenever narrow foci are used non-exhaustively, it has to be indicated by linguistic means, for instance, by expressions that inherently imply non-exhaustivity: among others in English, or jórész (for the most part), legkevésbé (least of all), elsősorban (primarily) in Hungarian, which are all compatible with an element in the prominent preverbal position.

To support the view that interpreting answers exhaustively is a pragmatic phenomenon that can be cancelled, Onea (2008) also mentions the following Hungarian example:
The example explicitly indicates that the focused individual was not the only one in the domain of the predicate.

In sum, based on the Hungarian data, we will mostly be concerned with two types of foci (information and contrastive), which appear in the same syntactic position in all-focus sentences. Completive information is postverbal non-presupposed material.

There is a strong pragmatic tendency to expect the answer denoted by the focused constituent to be exhaustive, since locutors usually expect complete answers to their questions. When it is relevant, I will distinguish between the focus types, referring to them as contrastive and informational focus. Otherwise, when talking about focus, I will refer to the syntactically and prosodically highlighted element found in narrow-focus sentences.

5.5.2 Thematic shifters in Hungarian

In the Hungarian linguistic tradition, the sentence-initial constituent is usually called the topic, which is defined as the part of the sentence that the rest of the sentence is about. In this respect, É. Kiss, in her works, follows the topic as entity approach (Reinhart, 1981). In her view, the topic is a referential and specific entity that has to be known, or at least, identifiable to both speaker and hearer. In this respect, the topic in the Hungarian sentence is always defined within the sentence (sentence topic), without referring to the context in which the sentence is uttered (discourse topic).

The topic can be defined at all linguistic levels. According to É. Kiss, formally the topic is the constituent in Hungarian that precedes (or constituents that precede, since more than one topic is also possible) the first pitch accent in the sentence. Semantically, the topic has to be specific and it is interpreted as the logical subject of the sentence (É. Kiss, 2005).

The specificity condition is illustrated by É. Kiss with the following facts:

- indefinite nouns can appear as topics only if they are specific:

\[(62) \text{ Egy gyerek bekopogott az ajtón.} \]
\[\text{a child knock.pst the door.superess} \]
A child knocked at the door. (One of the children knocked at the door.)

According to É. Kiss, in the above example, a child can only refer to a specific child, i.e. one of the children, where the set of children is already given in the discourse.
indefinite pronouns can appear as topics only if their referent is present in the
universe of the discourse, they do not have to be uniquely identifiable:

(63) Valaki kopog az ajtón.
    somebody knock the door.SUPERESS
    Somebody is knocking at the door.

Similarly to the previous example, the individual that is knocking on the door
is present in the universe of the discourse, thus it is considered to be specific. É.
Kiss compares this meaning of valaki to its Russian equivalents. In Russian, kto-
(to (somebody) refers to an individual that is present in the discourse, whereas
kto-nibud (somebody) to a non-specific individual that is not identifiable for the
speaker and the hearer.

However, various problems emerge in connection with this approach. First of all,
as Gécség and Kiefer (2009) point out, the logical subject/logical predicate and the
topic/comment distinctions belong to different levels of linguistic analysis, the for-
mer being a syntactico-semantic, whereas the latter a pragmatic distinction. From
this it follows that in É. Kiss’s definition (semantically the topic is the logical sub-
ject of the sentence) the two levels overlap. Gécség and Kiefer (2009) show that
there are languages, such as colloquial French, in which the topic/comment dis-
tinction is grammaticalized, and others, in which sentences are structured following
the logical subject/logical predicate distinction, like in the case of Hungarian and
Polish (and possibly other Slavic languages).

According to Gécség and Kiefer (2009), the topic of the sentence denotes the
individual the sentence is about, with respect to a particular context. This is why
the topic referent is observed to be given, or familiar, or, at least identifiable in
the context in question. In contrast to this, the logical subject of the sentence
is context-independent signaling only a particular discourse strategy. The logical
subject does not have to be given or discourse-old, it can denote a brand-new
individual that neither the speaker nor the hearer is familiar with. In addition, the
logical subject is always sentence initial, preceding the logical predicate, whereas the
topic can appear in situ. As the authors show, the initial constituent in Hungarian
sentences corresponds to the above description of logical subjects, and not to that
of topics. The most important argument is, as they demonstrate, that non-specific
individuals can also appear in this position, just like indefinite pronouns denoting
individuals that are not present in the universe of the discourse. Compare the
following examples:

(64) Egy gyerek leesett a villamosról.
    a child VM.fall.PST the tramway.DEL
    A child has fallen off the tram.

(65) Valaki meglokott a buszon.
    somebody VM.knock.PST the bus.SUPERESS
    Somebody pushed me in the bus.
In example (64) the sentence initial constituent (a child) is a non-specific indefinite, still it can appear in the assumed topic position. In (65), the indefinite pronoun denotes an individual that is clearly not present in the universe of the discourse, since the pushing event evoked happened before. The individual denoted by valaki is not existentially presupposed, they merely state the existence of an individual that has not been mentioned before in the discourse.

In addition, they observe that definite constituents that are discourse-new can also occupy the designated topic position. Interestingly, these sentences are equivalent to their topicless variants, from the semantic (truth-conditional) point of view:

(66) A the dog.Poss.2pl bite.PST the daughter.Poss.1sg.Acc
    your dog bit my daughter.

(Gécseg and Kiefer, 2009, p. 605, 49a)

(67) Megharapta a the dog.Poss.2pl a the daughter.Poss.1sg.Acc
    bite.PST your dog bit my daughter.

(Gécseg and Kiefer, 2009, p. 605, 49d)

As can be seen from the examples, the sentence-initial constituent does not correspond to any of the above definition of topics. It does not have to be given or discourse-old, and it can even denote a non-specific entity. These are the reasons why Gécseg and Kiefer (2009) argue that the initial constituent in Hungarian sentences can be identified with the logical subject and not with the topic, which latter, in turn, can appear in situ. However, it is also possible that the logical subject and the topic fall together.

Nevertheless, both Gécseg and Kiefer (2009) and É. Kiss adopt the topic as entity approach, supposing that topics always denote some individual. É. Kiss (2005) mentions that referential temporal and locative adverbials can also be considered as topics, since they can be analyzed as optional arguments of the verb without explaining how these adverbials fit into the topic as entity approach, since they do not denote individuals. Consider the following examples:

(68) Tegnap yesterday VM.bite.PST the dog.Poss.2pl the daughter.Poss.1sg.Acc
    yesterday your dog bit my daughter.

(69) A the trains.Superess VM.multiply.PST the robberies
    more and more robberies on trains.
In order to account for these facts, an alternative analysis of topics in Hungarian is called for. Reconsidering the above examples we can observe that all of them are presented without the discourse that would create the necessary context for them. In all cases, the referent of the sentence-initial constituent can be directly inferred from the discourse in which the sentence is uttered. For instance, examples (58) and (65) can be related to the following context:

(70) Fáj a vállam. (Tegnap utaztam hazafélé a schoulder.POSS.1SG (yesterday travel.PST home the munkából. Tőmeg volt.) Valaki meglokkott a buszon, work.ELAT crowd was) somebody push.PST the bus.SUPERESS, (miközben hazafélé utaztam a munkából). (while home travel.PST the work.ELAT)

My shoulders are aching. (Yesterday I was travelling home from work. There was a lot of people.) Somebody pushed me in the bus, (while I was travelling home from work).

(71) Elképesztő, hogy mit láttam tegnap. Hazafélé mentem, és egy incredible, that I saw yesterday. home go.PST.1SG, and a gyerek leesett a villamosról, de mégsem sérült meg. child VM.fell.PST.3SG the tram.DEL, but not getPST injured VM

Incredible, what I saw yesterday. I was going home and I saw a child fell off the tram, but he didn’t get injured.

Although both the individual pushing the speaker and the child falling off the tram are new in the universe of the discourse, their existence can easily be accommodated by the hearer. It is a usual event that in the rush hours buses in the city are overcrowded (or, if the speaker does not add this detail, it can also be supposed that in the city he often takes the bus, also on that particular day) and that valaki (somebody) refers to a passenger. The same goes for the child falling off the tram. Usually there are also children in the streets taking the tram, the existence of one of them is easily accommodated, even if the child in question is not identified by the speaker or the hearer (it is non-specific). Similarly, in example (66), it is because of the possessive structure that the hearer can accommodate which dog the sentence is about (even if he does not know the friend, nor its dog), this is why it can appear in the initial position. Without the possessive, the sentence is less acceptable:

(72) # Tegnap felhívott a férjem a yesterday VM.call.PST the husband.POSS.1SG the munkahelyemen. A kutya megharapta a workplace.POSS.1SG.SUPERESS the dog bite.PST the lányomat. daughter.POSS.1SG.ACC

Yesterday my husband called me at work. The dog bit my daughter.
In this case, unless the dog is not theirs that they call the dog, the hearer cannot easily accommodate which dog the sentence is about.

The above mentioned examples are all all-focus sentences answering questions like: *Mi történt?* (What’s happened?) or they are part of a narrative-type discourse. What we observed was that the determining factor concerning the sentence-initial element was not givenness or specificity, but rather the extent to which it fits in the discourse the sentence is part of. As we have already seen, in Büring (2003)’s approach the function of the topicalised constituent is to indicate a strategy of discourse organization. Contrary to É. Kiss’s analyses, this conception considers the discourse as a Question under Discussion (QuD), which is divided into subquestions and the answers to them. Büring (2003) shows that contrastive topics indicate such a strategy in non-neutral contexts, *i.e.* in answers to wh-questions, where the answers contain contrastive topics and narrow foci.

Erteschik-Shir (2007), (quoting Vallduví (1992) argues that in Catalan, we find a topic in sentences that do not elaborate further on the previous subquestion (subtopic) of the QuD, but introduces a new subtopic. The case of Hungarian is similar. Sentence-initial topics are all related to the same QuD and define subtopics that structure and thematize the discourse. In the (70) discourse the QuD is *What happened? / What’s the problem?* and since the pure identification of the problem (the aching shoulders) does not seem to be an exhausting answer, the QuD is enlarged to: *What is the problem and why?* and the answers also provide the reason for the problem. This necessitates the introduction of two subtopics: *tegnap* (yes-terday) and *valaki* (somebody). In (71) the first subtopic is the speaker, this is why there is no separate sentence-initial topic in the sentences. The next subtopic is *egy gyerek* (a child), and since this subtopic is continued in the rest of the discourse, again, there is no topicalized constituent thereafter. Notice that in this sense not only individuals can be topicalized, but adverbials as well: *tegnap* (yesterday) indicates that the subtopic of the forthcoming discourse part is the day before. Büring (1997) calls the type of topics that indicate that the answer to a question is not complete, by dividing the main question into subquestions and not answering all of them *S-topics* (Sentence-topics). This topic type is most often referred to in the literature as *contrastive topic* that Büring considers, in turn, as a subtype of S-topics. In this dissertation, I refer to the above described topic of all-focus sentences as *thematic shifter* in order to distinguish it from contrastive topics.

### 5.5.3 Contrastive Topics in Hungarian

The third discourse function often referred to in the Hungarian linguistic literature is the *Contrastive Topic*. Concerning contrastive topics, most analyses point out that they share some features of topics and foci, in that they relate to what is given in the discourse and introduce alternatives, respectively.

From a formal point of view, contrastive topics can be characterized by a fall-rise pitch accent, or B-accent (Jackendoff, 1972) in English, as opposed to the
falling A-accent of focused constituents. In French, contrastive topics bear the so-called \textit{C-Accent} (Beyssade \textit{et al.}, 2004b). In Hungarian, contrastive topics can be distinguished from ordinary topics by their rising intonation.

(73) János, (az/ő) nem jött el.
János, (that/he) not come.PST VM
As for John, he didn’t come along.

A test for the identification of contrastive topics is the insertion of a corresponding resumptive pronoun (az/ő) between the contrastive topic and the rest of the sentence. Contrastive topics, as suggested by the name, imply the presence of alternatives they are in contrast with. As the above example illustrates, János did not come, but the sentence implies at the same time that someone else came indeed. The contrastive topic usually appears on the left periphery of the sentence (although we will see some controversial cases when dealing with multiple questions), and it can be preceded or followed by ordinary topics (thematic shifters).

A sentence containing a contrastive topic cannot be uttered out of the blue. They appear only in narrow-focus sentences that contain a focused constituent (Büring, 1997). This latter constituent is referred to as the \textit{associate} of the contrastive topic (Kálmán, 2001; Gyuris, 2002) (since the term \textit{associate} is often used for the focus in constructions with focus sensitive particles, such as \textit{only}, the term \textit{correlate} will be used here with respect to the contrastive topic). In order to understand the Hungarian data, we should make a clear difference between focus in the semantic/pragmatic and the formal (mostly prosodic and sometimes syntactic highlighting) sense. Recall that we use the term \textit{focus} referring only to the semantic/pragmatic aspects of this notion. The correlate of the contrastive topic in Hungarian is a prosodically prominent constituent, that resides in the preverbal part of the comment, \textit{i.e.} either in the quantifier-field or in the prominent preverbal position. Semantically, the associate has clearly focus status, since it answers a question or constitutes the most prominent part of a correction, highlights a parallelism, etc. This example illustrates the above mentioned fact that information structure constituents (which are semantically defined) do not correspond (in a one-to-one manner) to designated syntactic positions (the associate of the contrastive topic can be a quantifier as well, which clearly does not reside in the immediately preverbal position). The Hungarian focus, thus, when a quantifier is focused, can appear in the quantifier field, and does not have to be in the prominent preverbal position. Let us see now some examples. Among the possible correlates of the contrastive topic we can mention thus quantifiers, negation, verum focus, and other focused constituents:

(74) a. Q: -Ki tudná elénekelni ezt az éneket?
   who could VM.sing this.ACC the song.ACC
   Who could sing this song?
   Verum focus:
   b. -/János EL tudná (énekelni).
      János VM could (to sing).
5.5. Discourse Functions in Hungarian

John could sing it. (but probably some others couldn’t)

Negation:

c. -/János NEM tudná (élénekelni).
   János not could (VM.sing).
   John could not sing it. (but probably some others could)

Focus:

d. -/Mari CSAK BEÉNEKLÉS UTÁN tudná.
   Mari only after voice training could
   Mary could sing it only after voice training (otherwise not).

Quantifier:

e. -/Ezt az éneket MINDENKI el tudná énekelni.
   this.ACC the song.ACC everyone VM could to sing
   Everyone could sing this song (unlike other songs, it is so well-known/easy).

(based on Gyuris (2002))

The bold-faced constituents in the above examples are all to be analyzed as foci, even though not all of them appear in the immediately preverbal position. Their syntactic position results from syntactic constraints, like the one stating that quantifiers precede the preverbal position. Their syntactic position can be identified with respect to the position of the verbal modifier. If the verbal modifier is postverbal, the focus is in the immediately preverbal position, if the verbal modifier is preverbal, the associate cannot possibly occupy the immediately preverbal position.

Concerning the elements that can appear as contrastive topics, they are less restricted than in the case of ordinary topics (thematic shifters). For instance, non-referential noun phrases can be contrastive topics, whereas they are excluded from the topic/thematic shifter category. Such non-referential noun phrases are the ones introduced by universal quantifiers (75), monotone decreasing quantifiers (76), distributive quantifiers (77), and noun phrases containing csak (only) (78) (based on Kálmán (2001)):

(75) /Mindenki nem olvasta el a könyvet.
     everyone not read.PST VM the book.ACC
     Everyone has not read the book (but some of the people have).

(76) /Rítkán csak ezek a gyerekek olvasnak könyveket.
     rarely only these the children read books.ACC
     Only these children read books rarely (others read more frequently).

(77) /Legalább egy könyvet minden gyerek elolvasott.
     at least one book.ACC all children VM.read.PST
     All children have read at least one book (and some might have read even more).
Kenesei (1989) proposed to analyze contrastive topics as a subtype of focus. According to his approach, the focus operator can be exclusive, in which case it has to appear preverbally (or, in transformational terms, move to a preverbal position). However, with the rising intonation the exclusion is weakened. It will not exclude all the other alternatives induced (like the immediately preverbal focus), but it will indicate that there is at least one alternative for which the predicate does not hold. Thus he distinguishes between universally vs. existentially quantified exclusion and calls the contrastive topic kontrafókusz (counterfocus). However, an important difference to mention is that while ordinary preverbal foci has truth conditional meaning, contrastive topics never change the truth conditions of the sentence. Furthermore, contrastive topics are bound to another focused constituent in the sentence, which, again, is not true for ordinary foci.

Molnár (1998) points out that topic and focus are not complementary categories, no matter how they are defined in a given theory (see also the discussion about implicative answers in Chapter (7)). In many cases the topic and the focus overlap in the sentence. Consider the following examples:

(79)  
   a. Q: -What is the news?  
   b. A: -Peter has visited his brother.

In the example, the whole answer is focus (all focus), whereas the topic (Peter) is part of the focus. In other examples, old information that should be topic, is focused:

(80)  
   a. Q: -Who did Felix praise?  
   b. A: -He praised HIMSELF.

(81)  
   a. Q: -Pourquoi tu t’es pas excuse auprès de Jean ?  
      why you cl.aux not excused at Jean  
      Why didn’t you say sorry to John?
   b. A: -Eh ben, c’est LUI qui m’a insulté.  
      Well, it is HIM who cl.aux insulted  
      Well, it was HIM, who insulted ME.

(Nevertheless, if we consider the focus as the answer to a question, it is no longer surprising why discourse-old elements can be focused.) There are cases, in which, according to Molnár (1998), the focus is part of the topic:

(82)  
   Hungarian:
5.6 Basic Facts about Interrogatives in Hungarian

In (82-b) 

5.6 Basic Facts about Interrogatives in Hungarian

In this section, I present some basic facts about interrogatives in Hungarian. The section is based on Kálmán (2001).

The only possible formal identification of polar interrogatives is intonation. They can be derived from all declarative sentences by a rising and then (on the penultimate syllable) sharply falling tone. From the pragmatic point of view, they are the questions that can be answered by -Nem. (-No.). A subtype of polar interrogatives that cannot be answered by yes/no is alternative questions that ask for the choice between one or more items:

(83) Mari kávét vagy teát kér?
Mari coffee.ACC or tea.ACC wants
What would Mary like, tea or coffee?/Does Mary want tea or coffee?

In subordinate interrogatives the intonational differences are neutralized with declaratives. Interrogativity is signaled by an enclitic -e following the finite verb (or the whole predicate in the case of nominal predicates):

(84) Megkérdeztem, hogy Mari üzent-e valamit.
asked.PST.1SG that Mari message leave.PST-CL.INTERR something.ACC
I asked if Mary has left a message or not.

Constituent questions have a falling or a falling-rising tone. They are all constructed with some question word that is obligatorily preverbal. The exact position of the interrogative word is difficult to identify. It can be preceded by any number of topics:

(85) Péter tegnap melyik moziban volt?  
    Péter yesterday which cinema.iness was  
    To which cinema did Peter go yesterday?

Similarly to the focused constituent, interrogative words also immediately precede the finite verb without the verbal particle that follows the verb in the presence of a question word:

(86) Kinek mutatta be János Marit?  
    who.dat introduce.pst vm János Mari.acc  
    To whom did John introduce Mary?

However, quantifiers cannot precede the interrogative phrase in the preverbal field:

(87) *Péter mindenkinek kit mutatott be?  
    Péter everybody.dat who.acc introduce.pst vm  
    Whom did Peter introduce to everybody?

Furthermore, as we have seen in Chapter (2), two interrogative expressions can occur in the preverbal field, which is strictly forbidden in the case of focused constituents:

(88) Péter kinek kit mutatott be?  
    Péter who.dat who.acc introduce.pst vm  
    Whom did Peter introduce to whom?

In addition, interrogative phrases can be preceded by focus in the preverbal field:

(89) Jó, de JÁNOS mit evett?  
    ok, but János what eat.pst  
    Ok, but what about John, what did HE eat?

The sentence in (89) can be used in the same contexts as its equivalent with the focused constituent in a postverbal position:

(90) Jó, de mit evett JÁNOS?  
    ok, but what eat.pst János
Ok, but what about John, what did HE eat?

The status of the prominent constituents in (89) is difficult to determine. Formally (prosodically) it is highlighted, just like foci, bearing a pitch accent that deaccents all postfocal material. From the semantic/pragmatic point of view, however, they can be analyzed as contrastive topics. In her analysis, Géceseg (2001) derives the sentences containing a preverbal contrastive topic from sentences that contain a preverbal and a postverbal prominent element. These latter are, in turn, analyzed by É. Kiss (1998b) as containing two foci. The problem dealt with here concerns two syntactic positions: the one preceding the immediately preverbal position, and the other a postverbal, in most cases sentence-final position, both occupied by a prominent element. This prominence is indicated, on the one hand, formally by a pitch accent and, (since the periphery of the sentence counts most often as prominent), by syntax. Semantically and/or pragmatically, the prominent status of these elements is often related to the implication of some contrast with the exclusion of all (focus) or some (contrastive topic) possible alternatives. The present dilemma reveals an interesting point about the Hungarian linguistic literature, in that it does not make a clear distinction between the syntactic and the semantic/information structure/discourse level of linguistic analysis. Elements that are defined in (sometimes vague) semantic terms, are supposed to appear in designated syntactic positions that bear the name of the semantic category of the given element. As it has already been shown, elements with similar semantics may occupy different syntactic positions, and one position can be filled with several types of elements. For instance, the canonical position of the semantically defined focus is the preverbal FocP, even if the second focus is postverbal in Hungarian. Géceseg (2001) already points towards a solution when assuming ”prominent elements” in the syntax, but continues the previous syntactic labeling by positing a KTopP (Contrastive Topic Phrase) that is filled by the semantically defined contrastive topic element, moving from its postverbal position in the preverbal field of the sentence. The motivation behind this movement is that the elements are not only prominent, but carry a [topic+] feature as well. This (in some sense ad hoc) feature accounts for the fact that these are referring expressions and serves as the common property of contrastive topics and ordinary topics.

In a non-derivational framework, such as LFG, a clear distinction can be made between the syntactic and semantic levels of the analysis. I will propose that the semantically neutral syntactic positions can be targeted by different prominent elements, such as the focus or the contrastive topic. In some cases, like in ((89) and (90)) this will result in ambiguity between a reading where the prominent element is a focus and in another where it is a contrastive topic. In other cases this ambiguity will have certain consequences, for instance in multiple questions, where they have different readings related to the semantic difference between foci and contrastive topics.

Going back to the problem of (89) and (90), the deaccenting of the interrogative word in these sentences can be explained by the fact that it occurs for the second
time, and is, thus, less prominent than it was for the first time. Compare this question with the deaccenting of the NP in the preverbal position in the following example:

(91) Correction:
   a. Csak két gyerek van, aki MARIT választotta. 
      only two children is, who Mari.ACC choose.PST 
      There are only two children who chose Mary.

      No! Everyone Mari.ACC choose.PST 
      No! Everyone chose Mary.

   In the second sentence, the universal quantifier is the focus, and Mari, occurring for the second time, is deaccented. (I thank Zsuzsanna Gécseg for this example.) (See Beaver et al. (2007) and Onea (2009a) on similar considerations in the case of second occurrence foci.) These questions can never start a discourse, they are more like clarification questions in situations, where the locutor does not get a satisfactory answer to his question for the first time. At its second occurrence, the question word is less prominent. As reflected by prosody, the prominent constituent of such sentences is the focus, which delimits the set of congruent answers to those in connection with one particular individual. Since this individual is in contrast with its alternatives in the relevant set, a contrastive topic analysis is also conceivable. Note that according to Beaver and Clark (2008), although second occurrence foci do not carry the same pitch accent as first occurrence ones, there are acoustic correlates that distinguish them from non-focal material. We will come back to these examples in Chapter (7).

   Interrogative words can also be followed by focused negative items. Here the order of the elements is more strict: the negated focus is strictly preverbal.

(92) Ki NEM TEGNAP ment haza? 
    who not yesterday go.PST home 
    Who go.PST home not yesterday (but on some other day)?

   (Kenesei, 2009, p. 569, 9)

(93) Kit NEM SMITH ölt meg? 
    whom not Smith kill.PST VM 
    Who was NOT killed by SMITH (but by someone else)?

   (Kenesei, 2009, p. 569, 10)

(94) Ki NEM A HAMLETET olvasta? 
    who neg the Hamlet.ACC read.PST 
    Who read something else but not the Hamlet?
In these examples, the negated constituent gets the pitch accent, since the other prominent constituent, the question word occurs already for the second time. These questions cannot start a discourse either, but rather point towards the answer expected, compared to previously asked questions.

It also has to be mentioned that according to Kálmán (2001) interrogative words are always local complements of the following finite verb. It follows from this that whenever the question is embedded and the main verb otherwise subcategorises for a non-interrogative complement, interrogativity is indicated by the \textit{wh}-expletive construction (Mycock, 2006). The term \textit{pseudo question word} is used by Kálmán (2001) for the interrogative word in the main clause. Nevertheless, only few verbs permit these constructions (the so-called \textit{bridge-verbs} in Kálmán (2001)’s terminology).

\begin{verbatim}
(95) * Azt mondta, hogy hova megyünk nyaralni?
   that say.PST that where go.PRS.1PL on holiday
   *He said that where are we going on holiday?

(96) Mit mondott, hogy hova megyünk nyaralni?
   what said that where go.PRS.1PL on holiday
   What did he say, where are we going on holiday?
\end{verbatim}

According to Mycock (2006)’s observations, however, it is also possible to extract the question word from the embedded clause, which will then occupy the immediately preverbal position in the main clause:

\begin{verbatim}
(97) Kit mondott János, hogy felvettek az egyetemre?
    whom say.PST János that admit.PST the university.SUBL
    About whom said John that he/she was admitted to the university?
\end{verbatim}

Interrogative words bearing the subject function in the embedded clause go through case-change in these constructions:

\begin{verbatim}
(98) János kit akar, hogy beszédet mondjon?
    whom want that speech give
    Who does John want to give a speech?

(99) Kinek akarod, hogy adjuk ezt az ajándékokat?
    to whom want.2SG that give.PRS.1PL this the present.ACC
    Who do you want us to give this present to?
\end{verbatim}

Mycock (2006) explains the case-change by the fact that two NPs in the nomi-
native case would cause processing difficulties in the main clause. The interrogative words, thus, take the case the main verb assigns to its clausal complement, indicating that it belongs to the subordinate clause. A detailed description of the phenomenon can be found in Horvath (1997); Mycock (2004) and Mycock (2006).

Finally, a short description of echo-questions is due. In echo questions in Hungarian, interrogative words are preferred to be preverbal, but otherwise follow the structure of the previous, not well-understood sentence. They also bear a pitch accent and eradicating stress. The interrogative pronouns *ki* (who) and *mi* (what) are often used in their emphatic forms: *kicsoda, micsoda* (which are usually sentence-initial and "aggressively non-D-linked"):

(100)  
\[
\text{KICSODÁT látott János a moziban?} \\
\text{who.EMPH see.PST János the.cinema.LESS} \\
\text{Who *the hell/on earth* did John see in the cinema?}
\]

## 5.7 Conclusion

I started this chapter with the presentation of the Hungarian sentence structure in a framework-neutral, topological way, dividing it into two fields and one of them further into four subfields. I showed that this division is motivated by formal factors, both on prosodic and on distributional grounds. I pointed out that semantically prominent elements appear in the topic field and in the prominent preverbal position (although focus can also occupy a position in the precomment or on the right periphery). I also showed, concerning the prosodic pattern and the distribution of the various elements that a clear-cut difference exists in Hungarian between sentences that appear mostly in narrative contexts (neutral sentences) and those that are replies to some previous assertion or question. The first can contain thematic shifters in the topic field, and a verbal modifier, a secondary predicate, or a hocus in the prominent preverbal position. In the second, a contrastive topic can appear in the topic field, and the prominent preverbal position is occupied by a focus, a question word, or a negative adverb/quantifier. In the view of these observations, I gave a brief overview of the analyses conducted in the transformational framework and concluded that they cannot take account of both sentence types, since they assume designated syntactic positions for topic and focus. Then I went on to the presentation of discourse functions with respect to Hungarian and showed how information and contrastive focus, completive information, thematic shifter and contrastive topic are manifested in the language. Finally, I presented some basic aspects of interrogatives in Hungarian. I conclude that although the structure of the Hungarian sentence reflects the way the sentence relates to the discourse, discourse functions cannot be assigned to particular syntactic positions. In order to analyze and formalize these assumptions, information and discourse structure are to be investigated separately from syntax. In Chapter (7) I propose an analysis of the information structure that can, at the same time, capture the differences and the similarities between neutral (in some cases all-focus) and narrow-focus (non-
neutral) sentences in Hungarian from a discourse perspective.
It is generally assumed that French is a configurational language, with SVO word order. Similarly to English, the case system disappeared in French, with the exception of personal pronouns, which can still be used in the nominative, accusative and dative. Nevertheless, word order can considerably deviate from the SVO word order in French. In the first part of this chapter, I examine the syntactic structure of French, concentrating on word order variations. I consider two phenomena: the status of pronoun clitics and subject NP inversion, which both support a flat syntactic structure without a VP constituent (elaborated in Chapter 9); and present
Chapter 6. Relevant Aspects of French

the concept of syntactic weight. In the third section, I consider the basic properties of coordination phenomena in French. Then I go on to prosodic and discourse phenomena. In the last section of this chapter, I turn to French interrogatives and examine how the above mentioned phenomena affect question formation.

6.2 Syntactic Structure: Word Order Variation

6.2.1 Clitics as pronominal affixes

The first particularity of French (and, in general, Romance) word order is the excessive use of pronoun clitics, whose position differs from that of the corresponding lexical argument with respect to the verb:

(1) a. Marie voit Jean.
   Marie sees Jean
   Mary sees John.

   b. Marie le voit.
   Marie cl.masc.acc sees
   Mary sees him.

(2) a. Marie donne le livre à Jean.
   Marie gives the book to Jean
   Mary gives the book to John.

   b. Marie le lui donne.
   Marie cl.masc.acc cl.dat gives
   Mary gives it to him.

In the first generative analyses, such items were treated as pronouns that undergo movement to a preverbal position (see Kayne (1975) on French, Perlmutter (1970) on Spanish, introducing surface structure constraints). The main problems to account for include the fact that the pronoun clitics stand in a strict order and cannot combine freely with each other ((3)-(4)). Moreover, in French, they are not always preverbal. In declarative sentences they are, but in imperatives (5) and interrogatives (6) they follow the verb. In negative imperatives (7), again, they are preverbal.

(3) Je te le présente.
   I cl.dat.2sg cl.acc.3sg introduce
   I introduce him to you.

(4) a. *Je te lui présente.
   I cl.acc.2sg cl.dat.3sg introduce
   I introduce you to him.
b. Je te présente à lui.  
I introduce you to him

(5) Donne-le-moi.  
give me  
Give it to me!

(6) Où Pierre est-il allé ?  
where Pierre is-he gone  
Where has Pierre gone?

(7) Ne m’en parle pas !  
PRT CL.DAT.1SG.CL.PART speak not  
Don’t talk to me about it!

Furthermore, in complex predicates (Miller and Sag (1997); Abeillé and Godard (1996, 2003), see also Dalrymple and Zaenen (1996) for an LFG analysis, and Abeillé and Godard (2010) in English), the pronoun clitics are attached to verbs, to which they are not lexically related (Sportiche, 1996). These include auxiliary-infinitive constructions in composed verb tenses, the passive, the causative, as well as other structures in which the clitic is the argument of the infinitive or the predicative noun like in (10), but at the same time appears attached to the finite verb:

(8) Jean l’a vu.  
Jean CL.ACC has seen  
John has seen it.

(9) Marie la fait coudre à Pierre.  
Marie CL.ACC makes sew to Pierre  
Mary has it sewn by Pierre.

(10) Marie en a peur.  
Marie CL.PART has fear  
Mary is afraid of it.

These examples argue against views that suppose that pronoun clitics are not base-generated in the position where they appear.

Miller and Sag (1997) propose to analyze French pronoun clitics as lexically attached inflections, rather than as pronouns or clitics. They present several arguments to support this claim. Let us now enumerate these arguments:

1. Degree of selection with respect to the host
The affixes appear systematically on the verb and not VP-initially if the verb is not VP-initial:

(11) [Tout lui donner] serait une erreur.
    everything CL.DAT to give would be a mistake
    To give her/him everything would be a mistake.

(Miller and Sag, 1997, p. 4, 4c)

2. Arbitrary gaps in the set of combinations

This issue has already been touched upon and illustrated in (4). The authors argue that this is also typical of lexical affixes and not of clitics.

3. Morphophonological idiosyncrasies

Some affix-verb combinations present morphophonological idiosyncrasies that cannot be explained with productive phonological rules. This also reveals the lexicalized nature of these affixed verbs. The authors give the following examples:

    Pierre there goes
    Pierre is going there.

b. Pierre (*y) ira.
    Pierre there go.FUT
    Pierre will go there.

(Miller and Sag, 1997, p. 6, 7a-7b)

(13) a. Chuis (je suis) allé au marché.
    (I am) gone to the market
    I went to the market.

b. Chais pas/Chépa/ Je sais pas.
    I know not
    I don’t know.

(Miller and Sag, 1997, p. 4)

Note that the contraction of je suis into chuis is only possible in the case of the verb être (to be), and not for the homophonous form of suivre (to follow), which also shows that it is a unique idiosyncratic realization and not a regular phonological rule.
4. Rigid and idiosyncratic ordering

   The affixes are ordered rigidly. For instance, accusative and dative first and second person affixes precede third person accusative ones, which in turn, precede third person dative ones. As we can see in (4), however, accusative first and second person pronouns cannot combine with third person dative ones (these latter are represented by a postverbal preposition + personal pronoun construction). This is also typical of affixation and not of cliticization.

5. Pronominal affixes undergo lexical phonological rules

   One example is the obligatory liaison of nasal consonants between the pronominal affix *en* and the following verb stem.

6. Object affixes cannot have wide scope over coordination

   In a coordination, the affixes are present on each verb stem:

       Pierre cl.acc.pl buys and sells
       Pierre buys and sells them.

   b. Pierre les achète et les vend.
      Pierre cl.acc.pl buys and cl.acc.pl sells
      Pierre buys them and sells them.

7. Syntactic explanations for clitic ordering have failed

   See the references in Miller and Sag (1997): 7.

   Based on the criteria mentioned above, these items are considered as pronominal affixes (and not pronouns or clitics) in this thesis.

   Before going on to subject NP inversion, we take a look at subject clitic inversion, or, as it is more adequate to call it now: verb stem + subject pronominal affix verb forms. According to Kampers-Manhe et al. (2004), these are licensed in three contexts: (polar and constituent) questions (15), declaratives with some fronted adverbials (16), and asymmetric coordination of root clauses (17):

   (15) Marie est-elle venue ?
       Marie is she come
       Has Mary come?

   (16) Peut-être viendra-t-elle.
       perhaps come.FUT she
       Perhaps she will come.

   (Kampers-Manhe et al., 2004, p. 554, 2)
(17) Viendrait-elle, (que) je ne changerais pas d’avis.
    come.COND she, (that) I PRT change.COND not opinion
    Should she come, I would not change my mind.

(Kampers-Manhe et al., 2004, p. 554, 3)

Two remarks are due here. Note first that these forms (apart from parentheticals
of the type dit-il (he said)) appear only in root sentences. Secondly, the NP subject
can also be present in these sentences, showing that the pronominal affixes are not
the true subjects of the sentence.

In lexicalist theories (like LFG or HPSG), these pronominal affixes are part of
the lexical entry of the verbs, specifying the contexts in which the given form can
appear.

6.2.2 Subject NP Inversion

Subject NP inversion is a common phenomenon across Romance languages
(Marandin (2000, 2003), see also Marandin (2010) in English). Although
they are not discourse-configurational languages in the sense of É. Kiss (1995),
Kampers-Manhe et al. (2004) and Marandin (2003) argue that subject NP inver-
sion is licensed by semantico-pragmatic/discourse conditions. Now I concentrate
on French. In this section, based on Kampers-Manhe et al. (2004) and Marandin
(2003, 2010), I enumerate the different types of inversion in French and present a
possible account in discourse structure terms. Note that subject clitic inversion,
introduced in the previous subsection is distinct from subject NP inversion, since
it appears only in root sentences and can co-occur with a preverbal subject.

The types of subject NP inversion in French are the following:

1. Inversion in extraction contexts (Stylistic inversion)

   This type of inversion appears (optionally) in interrogatives ((18) and (19)),
   relative clauses (20) and clefts (21):

   (18) Où est allé Pierre?
       where AUX gone Pierre
       Where has Pierre gone?

       (Kampers-Manhe et al., 2004, p. 556, 5a)

   (19) Je me demande où est allé Pierre.
       I me ask where AUX gone Pierre
       I wonder where Pierre has gone.

       (Kampers-Manhe et al., 2004, p. 556, 5b)
6.2. Syntactic Structure: Word Order Variation

(20) la rue où va Pierre
the street where goes Pierre
the street where Pierre goes

(Kampers-Manhe et al., 2004, p. 556, 5c)

(21) C’est dans la cafétéria que déjeune Paul.
It is in the cafeteria that has lunch Paul
It is in the cafeteria that Paul has lunch.

(Kampers-Manhe et al., 2004, p. 556, 5d)

2. Unaccusative inversion

Unaccusative inversion is possible only with ergative and passive verbs, in subjunctive clauses (22), in root clauses introduced by temporal adverbs (24), or temporal adverbial clauses (24), in complement clauses of perception verbs (25):

(22) Je voudrais que soient distribués ces prospectus.
I would like that be distributed those leaflets
I would like those leaflets to be distributed.

(Kampers-Manhe et al., 2004, p. 557, 6)

(23) Alors entrèrent trois soldats.
then enter.PST three soldiers
Then came in three soldiers.

(Kampers-Manhe et al., 2004, p. 558, 11)

(24) Quand sont arrivées les enfants, Pierre travaillait.
when aux arrived the children, Pierre was working
When the children arrived, Pierre was working.

(Marandin, 2010, p. 321, 1a-iii)

(25) Il vit que sortaient deux individus masqués.
he see.PST that were coming out two individuals masked
He saw that came out two men with masks on.

(Kampers-Manhe et al., 2004, p. 558, 12)
Nevertheless, some unergative verbs can also appear in this construction, as far as they denote a usual property of the subject:

(26) Alors sonnèrent les cloches.
    then rang the bells
    Then rang the bells.

(Kampers-Manhe et al., 2004, p. 558, 14a)

3. Elaborative inversion

In this type, the NP is on the right edge of the sentence and requires some sort of semantic plurality:

(27) Paul pense que devront repasser l’examen tous les étudiants
    Paul thinks that must.FUT retake the exam all the students
    qui ont raté le contrôle continu.
    who AUX failed the assessment continuous
    Paul thinks that all the students who failed the continuous assessment have to take the exam again.

(Marandin, 2010, p. 322, 1c)

Note that from a syntactic point of view, the inverted subject is interspersed with the postverbal complements and adjuncts, which would make a hierarchical syntactic structure containing a VP problematic for French (if the VP is supposed to contain only the verb and its complements except for the subject). See (9) for further discussion of the issue.

It is a common assumption in describing inversion in the Romance languages (other than French) that the subject NP is postverbal if it is in narrow focus (in answers to constituent questions). However, as Marandin (2003, 2010) shows, there are focused subjects that cannot appear postverbally (for instance in partial answers), and the subject can be postverbal in all-focus sentences as well (answering questions like What happened?).

Note that in one of the problematic cases that Marandin mentions (referring to É. Kiss (1998a)), although as answer to a question the subject is considered focus (see Chapter (4)), it does not introduce a new discourse referent to the discourse, since the possible referents are introduced in the question and the answer picks out one of those. In this case, the subject can also be preverbal:

(28) Italian:

    a. Chi di voi due ha rotto il vaso?
        who of you two AUX broken the vase
        Which of you two broke the vase?
b. Maria ha rotto il vaso.
   Maria AUX broken the vase.
   It was Maria who broke the vase.

In French questions, contrary to the other Romance languages, the subject NP cannot be indefinite. When it is introduced by an indefinite article, it must be generic. This is why, for some authors (Kayne and Pollock, 2001), the inversion in interrogatives is an instance of topicalization and is by no means related to (semantic) focus. As Drijkoningen and Kampers-Manhe (2001) put it, questions cannot be posed about entities introduced into the discourse at the same time.

(29) Quels romans doit avoir lu un étudiant de français pour être accepté dans le programme?
   Which novels must AUX read a student of French in order to be accepted in the program?
   Which novels must have read a student of French to be admitted in the program?

Kampers-Manhe et al. (2004, p. 561, 19a)

However, the focus-based approach can account for unaccusative inversion in French. According to Kampers-Manhe (1998), unaccusative inversion is possible if the subject is part of the focus, ruling out the ungrammatical examples with narrow focus on the verb. On the other hand, Marandin (2003, 2010) approaches this type, as well as the above mentioned problematic ones from languages other than French, from the perspective of the predicate. According to Marandin, in order for subject inversion to be grammatical, the predicate has to be given (i.e. it has to express the core content of the discourse topic). Whenever the discourse topic is modified, for instance in partial answers, postverbal subject NPs are ungrammatical. Another, marginal possibility is if the predicate denotes a usual property of the subject.

Generally, we can conclude that subject NP inversion occurs in two opposed circumstances. On the one hand, subjects introducing new discourse referents (when the predicate is given) appear postverbally, usually on the right edge of the sentence, which can be considered as a salient position, all the more so that intonationally it is also the right edge of focused constituents that is marked. In this case, the prominence of the subject is emphasized. On the other hand, in questions, and in sentences where the predicate expresses some usual property of the subject, the subject is either given (and not prominent), or predictable from the predicate (ringing-bells). It would thus be a mistake to identify subject NP inversion with focusing; it is rather a mechanism that can be used in different discourse contexts.

### 6.2.3 The concept of syntactic weight

*Weight* as a syntactic feature has been introduced by Anne Abeillé and Danièle Godard in order to account for word order phenomena in French. Its relevance has been demonstrated for other languages, outside the Romance family as well. For a
It is generally observed in the world’s languages that heavy constituents tend to come in the end of the sentence (or clause). E. Kiss (2008a) accounts for the order of postverbal constituents in the Hungarian sentence in terms of their phonological weight, and the same can be illustrated by heavy NP-shift in English:

(30) a. John gave to Mary [the wonderfully illustrated book about animals living in the jungle].
    b. ?? John gave [the wonderfully illustrated book about animals living in the jungle] to Mary.

As opposed to this, there are words that have restricted mobility in the sentence, more precisely, they are in most cases adjacent to a head. In mobility, they are between weak and strong forms, in that unlike weak forms they can appear alone, they can be modified and coordinated, but unlike some strong forms they cannot be detached from their head or extracted (unless they are coordinated, modified, or accented). Abeillé and Godard propose to refer to this class of words as light, which is then opposed to weak forms on the one hand, and non-light forms on the other. The authors assume that unlike non-light forms, light items enter the syntax as words and not as phrases, since they cannot form a phrase alone. Since only phrases can take part in syntactic operations, they cannot undergo extraction or shift to the right, unless they form a phrase with other elements. This feature is supposed to account for word order cross-linguistically: light constituents immediately follow their head, heavy constituents are shifted to the right periphery, and middle heavy ones appear between these two (in head-final languages the order is the mirror image of this).

The feature of weight seems to be relevant across the different categories in French. Let us now have a look at those one by one:

• Nouns

In the nominal domain it is assumed that bare nouns are light, whereas nouns forming a phrase with a determiner and proper nouns are non-light. The former, unlike the latter, cannot be detached from the verb (except for the negative element pas and certain adverbs), and cannot be extracted, unless it is modified or coordinated:

(31) a. Cet endroit fait [peur] [aux enfants].
    this place does fear to the children
    The children are scared of this place.
    
    b. Cet endroit fait [aux enfants] *peur/[une immense peur].
    this place does to the children fear/an enormous fear
    The children are enormously scared of this place.

(32) C’est *peur/une immense peur que cet endroit fait aux
    It is fear/an enormous fear that this place does to the
It is an enormous fear that the children feel at this place.

These words are typical in complex verb structures in which the verb is the syntactic, and the noun the semantic head of the structure.

- **Adverbs**

There are two classes of adverbs that can be classified as light in French: manner and degree adverbs, and adjectives that can be used as adverbs as well (*refuser net* (categorically refuse), *peser lourd* (weigh heavy), *coûter cher* (cost expensive)). Adverbs derived with the suffix *-ment* are non-light. Light and non-light adverbs form doublets (*bien/correctement* (well/correctly), *trop/excessivement* (too (much), excessively), *peu/modérément* (little/moderately), *pas, plus/nullement, aucunement* (not, not anymore), and if a weak form is also available, triplets: *très/beaucoup/énormément* (very/a lot/enormously), *si/tant/tellement* (so/that much), *aussi/autant/pareillement* (as/as much as/similarly). The conditions of use are the same as in the case of nouns:

(33) Paul travaille trop dans le jardin *trop/excessivement.
Paul works too much in the garden too much/excessively
Paul works in the garden too much.

(34) C’est *trop/excessivement que Paul travaille dans le jardin.
it is too much/excessively that Paul works in the garden
It is too much work that Paul does in the garden.

- **Adjectives** (*Abeillé and Godard, 1999a*)

It is well-known that in French, some adjectives are prenominal, while others follow the noun. In some cases, both are possible, but the position of the adjective corresponds to semantic differences. Prenominal adjectives undergo phonological processes like liaison, and some of them change their form according to the gender/first sound (consonant/vowel) of the noun (*beau/bel* (beautiful), *vieux/veil* (old), *nouveau/nouvel* (new)). However, when they are modified, they follow the noun:

(35) a. une (*extrêmement*) longue journée
    a extremely long day
    an extremely long day

b. une journée extrêmement longue
    a day extremely long
Abeillé and Godard propose that preverbal adjectives are light, whereas postverbal ones are non-light.

- Pronouns

Since French personal "pronouns" are analyzed as clitics\(^1\), or rather lexical pronominal affixes, they belong at most to the weight class weak (since they cannot be used in isolation, cannot be modified or coordinated), or, according to the lexical affix analysis, they are part of the verb form, thus they count by no means as independent words (or clitics).

However, English pronouns seem to manifest the light/non-light distinction: personal pronouns can only be detached from the verb if they are coordinated or modified, whereas demonstrative and reflexive pronouns (non-light) enjoy more mobility in the sentence. In example (36) the modified and coordinated personal pronouns follow the particle of the phrasal verb, whereas a personal pronoun alone immediately follows the verb, preceding the particle. Example (37) shows the mobility of non-light pronouns.

(36)  
   a. Paul looked me up (*me) in the telephone book.  
   b. Paul looked up [only me] / [him and her] in the telephone book.

(37)  
   a. He looked himself up/ up himself in the telephone book.  
   b. Paul gave this to John. / Paul gave John this.

(Abeillé and Godard, 2004a, p. 21, 69-70)

Syntactic weight plays a role in interrogatives as well. We come back to the issue of question words que and quoi in Section (6.6.3).

6.3 Coordination in French: the Basic Properties

As Mouret (2007) puts it, two basic properties characterize coordinate structures:

- they contain at least one conjunction
- and they are built with a mechanism of syntactic iteration

Based on these properties, he identifies two canonical structures in French: simple coordination (coordination simple) (38) and correlative coordination (39) (coordination à redoublement):

\(^1\)Nevertheless, there is a class of personal pronouns in French that are non-light and appear without a verb form: moi, toi, lui..., etc.
6.3. Coordination in French: the Basic Properties

(38) a. Paul a appris [l’espagnol, (et) l’italien, et le Portuguese]
Paul learned Spanish, Italian and Portuguese.

b. Paul apprendra [l’espagnol, (ou) l’italien, ou le portugais].
Paul will learn Spanish, Italian or Portuguese.

(Mouret, 2007, p. 15, 1 a, b)

(39) a. Paul a appris [et l’espagnol, *(et) l’italien, et le Portuguese]
Paul learned [and the Spanish, *(and) the Italian, and the Portuguese]

b. Paul apprendra [ou l’espagnol, *(ou) l’italien, ou le Portuguese]
Paul will learn [or the Spanish, *(or) the Italian, or the Portuguese]

(Mouret, 2007, p. 15, 2 a, b)

In simple coordination, the conjunction appears obligatorily before the last conjunct and optionally between the two preceding conjuncts, whereas in a correlative coordination, it appears obligatorily before the first conjunct, and also before all the others.

There are other, non-canonical coordination types, in which one of the above mentioned conditions is not satisfied. These are juxtaposed structures (40), incidental (41-a) and emphatic (41-b) coordinations and fragments (42):

(40) Que Paul apprenne [l’espagnol, l’italien, le portugais],
that Paul learn.SBJ [the Spanish, the Italian, the Portuguese],
peu m’importe.
I don’t care
Whether Paul learns Spanish, Italian, or Portuguese, I don’t care.

(Mouret, 2007, p. 15, 4 b)

(41) a. [Paul, [ou bien j’ai mal compris ce qu’il m’a dit], parle Paul, or LAUX bad understood what he me.AUX told, speaks
dix langues couramment].
ten languages fluently
Paul speaks fluently ten languages, or I misunderstood what he told me.

(Mouret, 2007, p. 15, 5 b)
(42) a. S1: -Paul a appris l’espagnol.
   Paul learned the Spanish
   Paul learned Spanish.

b. S2: -Et Marie a appris l’italien.
   and Marie learned the Italian.
   And Mary learned Italian.

   (Mouret, 2007, p. 16, a)

The example in (40) does not contain any conjunction, whereas (41-a) and (42) cannot be analyzed in terms of syntactic iteration.

Abeillé (2005) argues that the incidental and emphatic conjuncts in examples in (41-a) and (41-b) are to be analyzed as adjuncts and not true coordinate structures. She shows that they are reduced clauses, incompatible with correlative coordination:

(43) Jean est parti, et il est parti vite.
   Jean left, and he left quickly
   John left, and he left quickly.

(44) *Jean est et parti et rapidement.
   Jean AUX and left and quickly

Mouret (2007) shows that although the conjuncts do not have to be identical in all respects, for instance in their syntactic category, they have to be grammatical alone (without the other conjunct(s)) in the same context; i.e. they have to share the same function (subject, object, etc.) (see also Sag (2005)).

(45) Jean est [heureux et de bonne humeur].
   Jean is happy and in a good mood
   John is happy and in a good mood.

(46) Ils veulent [des augmentations et qu’on leur laisse le choix].
    They want an increase and that we leave them the choice.

In (45), an adjective and a prepositional phrase can be coordinated, since both are the complements of the predicate être (to be). Likewise, in (46), a noun phrase is

---

2 This generalization is usually attributed to Thomas Wasow.
coordinated with a complement clause sharing the object function.

As opposed to this, constituent that do not share the same function, cannot be
coordinated. Compare (object and complement of the predicate être in (47), object
and adjunct in (46)):

(47) Jean a trouvé Paul (*et) heureux.
    Jean aux found Paul and happy
    John found Paul (*and) happy.

(48) Jean a vu Marie (*et) rapidement.
    Jean aux seen Marie and quickly
    John saw (*and) Mary quickly.

The examples in (45)-(46) cannot be analyzed as elliptical structures. A convincing
argument against the analysis in terms of ellipsis is the fact that If a correlative coor-
dination is involved, the conjunctions obligatorily stand after the shared predicate,
and not before, as expected if the predicate was included in the first conjunct:

(49) Jean est [et heureux] [et de bonne humeur].
    Jean is and happy and of good mood

(50) Ils veulent [et des augmentations] [et qu’on leur laisse le choix].
    They want and an increase and that we leave them the
    choice.

An interesting structure is the coordination of (non-constituent) clusters:

(51) Paul offrira [un disque à Marie] et [un livre à Jean].
    Paul offer.FUT [a CD to Marie] and [a book to Jean]
    Paul will offer a CD to Mary and a book to John.

(Mouret, 2006, 2a)

It has often been assumed that cluster coordination contains ellipsis (52), just
like structures undergoing gapping (53) or right-node-raising (54):

(52) Paul [offrira un disque à Marie] et [offrira un livre à Jean].
    Paul offer.FUT [a CD to Marie] and [offer a book to Jean]
    Paul will offer a CD to Mary and a book to John.

(53) Jean a acheté un CD et Marie un livre.
    Jean aux bought a CD and Marie a book
    John bought a CD and Mary a book.

(based on Mouret (2006, 1c))
However, Mouret (2006, 2008) shows that non-constituent coordination cannot be analyzed in terms of ellipsis. The arguments he presents are the following:

- syntactic reconstruction

Mouret (2006) argues (citing Abeillé & Godard 1996, 2000) that the syntactic reconstruction of the alleged deleted material is not always grammatical. For instance, there are conjunctions such as *ainsi que* (as well as) that can combine with an argument cluster (55) but not with finite VPs (56) or clauses (57):

(55) Paul offrira un disque à Marie ainsi qu’un livre à Jean.
Paul offers a record to Marie as well as a book to Jean

(Mouret, 2006, 9a)

(56) *Paul écoute la radio ainsi que lit le journal.
Paul listens to the radio as well as reads the paper

(Mouret, 2006, 9b)

(57) *Paul lit le journal ainsi que Marie écoute la radio.
Paul reads the paper as well as Mary listens to the radio

(Mouret, 2006, 9c)

- place of the conjunction

Like above, the conjunctions obligatorily stand after the shared predicate in correlative coordination:

(58) Paul compte offrir [et un disque à Marie] [et un livre à Jean].
Paul plans to offer not only a record to Mary but also a book to John.

(Mouret, 2006, 12a)
6.3. Coordination in French: the Basic Properties 243

(59) *Paul compte [et offrir un disque à Marie] [et un livre à Jean].
Paul plans and offer a record to Marie and a book to Jean
*Paul is planning and to offer a record to Mary and a book to John.

(Mouret, 2006, 13a)

• semantics of adverbs

Additive and restrictive adverbs (for instance aussi (too) and seulement (only)) can take the whole coordinate structure as their semantic associate (60), whereas this is not conceivable under an elliptical analysis in which they reside inside the first conjunct (61):

(60) Paul offrira seulement [un disque à Marie et un livre à Jean]
Paul offer.FUT only a record to Marie and a book to Jean
alors qu’il aurait pu aussi offrir des fleurs à Léa.
while he AUX could also offer ART.PL flowers to Léa
Paul will offer only [a record to Mary and a book to John] while he
could have also offered some flowers to Lea.

(Mouret, 2006, 18a)

(61) Paul compte lire seulement le journal et écouter la radio.
Paul plans read only the paper and listen the radio
Paul is planning to read only the newspaper and listen to the radio.

≠ The only thing Paul is planning to so is both read the paper and listen to
the radio.

(Mouret, 2006, 19a)

• subject inversion and agreement

As was mentioned above, inverted subjects combine freely as sisters with the
two clusters containing inverted subjects are coordinated, two agreement pat-
terns arise. If the conjuncts are treated as independent events (reinforced by
adverbials such as quelques secondes plus tard (a few seconds later)), the verb
is singular (62), agreeing independently with each subject. However, if the
conjuncts denote a complex event (reinforced by the adverbial simultanément
(simultaneously)), the verb is plural (63):

(62) Paul is driving.
Alors surgit/*surgissent d’un buisson une biche, et quelques secondes plus tard d’un champ un renard.
then comes/*come from a bush a doe, and some seconds later from a field a fox

Then comes/come from a bush a doe and a few seconds later from a field a fox.

(Mouret, 2006, 21a)

(63) Alors surgissent/*surgit simultanément d’un buisson une biche, et du champ un renard.
then comes/*come simultaneously from a bush a doe, and from a field a fox

Then come/comes simultaneously from a bush a doe and from a field a fox.

(Mouret, 2006, 21b)

Analyses assuming ellipsis could not account for the plural verb form in (63), since the verb should, in that case, agree with the (singular) subject in the first conjunct. Based on these arguments, Mouret (2006, 2008) adopts a structure without ellipsis, in which non-standard constituents are conjoined in the scope of a shared predicate: Figure (6.1).

![Syntactic structure of non-standard constituent coordination](image)

Figure 6.1: Syntactic structure of non-standard constituent coordination

6.4 Relevant Aspects of French Prosody

This section is based on Beyssade et al. (2004a). Concerning the prosodic units in French, the authors distinguish between those that reflect the syntactic and metrical organization of the sentence and those that reflect the information structure of the utterance. The former contains two types of units: *Rhythmic groups* (RG) and *Major phrases* (MaP), whereas the latter is reflected in terms of a phrasing in *Intonational phrases*. A rhythmic group usually consists of a lexical word and all the functional words on its left side. The last syllable of a rhythmic group is metrically strong (accented). This means that the syllable is lengthened and is sometimes
accompanied by melodic movement. The accent occurring at the right edge of a rhythmic group is called primary accent. Secondary accent may appear at the left edge of a rhythmic group, whenever the distance between two primary accents is too long. It is realized on the first or antepenultimate syllable of a lexical word. Unlike in English, stress or accent is thus assigned postlexically, at the phrasal level in French. Major phrases consist of one or more rhythmic groups and are determined by syntactic factors (phrase structure):

\[(64) \text{Les enfants } R_{RG} \text{ de Jean-François } R_{RG} \ \{ M_{MaP} \text{ sont allés } R_{RG} \text{ au } M_{RG} \text{ en fin } R_{RG} \text{ d’après-midi. } M_{RG} \}\]\n
Jean-François’s children went to the cinema in the late afternoon.

(64) \[(Beyssade et al., 2004a, p. 469, 34)\]

Intonational phrases (IntP) are determined by the ground/focus articulation of the sentence. The position of the first IntP boundary coincides with the end of the focused constituent in the case of narrow focus, and with the end of the sentence in all-focus utterances.

The intonational patterns of French are analyzed in terms of boundary tones and pitch accents. Boundary tones, which occur at boundaries of prosodic units can be of two types: demarcative tones occur at the edge of prosodic phrases (rhythmic groups and major phrases) and do not carry any specific meaning; and illocutionary boundary tones that are realized at the right edge of IntPs and indicate the illocutionary force of the utterance. The rising pitch movement (H\%) is prototypically used in questions (65), the falling pitch movement (L\%) in assertions (66) and commands (67), and the rising-falling movement (HL\%) in confirmation requests (68).

\[(65) \text{Est-ce que tu viendras } H\%_{IntP} \text{ ?}\] \(\text{is it that you come.FUT}\) \(\text{Are you coming?}\)

\[(66) \text{Jean est venu. L\%}_{IntP} \text{ }\] \(\text{Jean AUX come}\) \(\text{John has come.}\)

\[(67) \text{Viens tout de suite. L\% }_{IntP} \text{ come at once}\) \(\text{Come here at once!}\)

\[(68) \text{C’est bien pour Chirac HL\%}_{IntP} \text{ que Mathilde a voté HL\%}_{IntP} \text{ que Mathilde has voted Mathilde voted for Chirac, didn’t she?}\]

(65) \[(Beyssade et al., 2004a, p. 469, 34)\]
As for pitch accents, one of them that the authors call \textit{C-accent} (see also Marandin (2006d)) plays a crucial role in the marking of a complex discourse strategy (see Chapter (4)). It is located at the left edge of rhythmic groups and realized on a syllable with lengthened initial consonant and involves a H tone.

Before going on to discussing the manifestations of information structure in French, I consider the meaning of final contours (sequences of tones), since it plays a crucial role in the discourse. Beyssade and Marandin (2007) show that contrary to the common assumption that falling contours characterize assertions and rising contours questions (they are indicators of illocutionary force), in French, contours signal the way the speaker anticipates the reception of her utterance by the addressee. They conclude that falling contours are rather associated with prototypical assertions, questions, or commands, whereas non-falling contours indicate some marked discourse value. This means that the speaker has an image of what the addressee believes about the issue or the content conveyed (belief attribution). At the same time, these assumptions reveal a certain asymmetry between speaker and addressee: the speaker is committed to the propositions/issues conveyed, whereas the addressee may refuse them, or question the content or the relevance of these propositions/issues. The speaker is, however, aware of these possibilities and can anticipate them. This is signaled by final contours in French. The possible final contours and their meaning in French are presented schematically in Figure (6.2).


diagram

If a question is pronounced with a falling intonation, it means that it is added to the questions under discussion without revision. The signaling of anticipated revision is never compulsory. Utterances pronounced with a non-falling contour can always be pronounced with a falling contour as well.

### 6.5 Information Structure Considerations

What are the primitives of information structure is always a question of a given theory, of a given language, of the analysis itself, or of the interplay of more than one of these factors. In this chapter, we consider focus and discourse-thematic phenomena in French.
6.5. Information Structure Considerations

6.5.1 The Focus

Beyssade et al. (2004b) (see also Marandin (2004)) define focus, going back to Jacobs (1984), as the constituent most affected by the illocutionary operator associated with the utterance. (As opposed to this, as we have seen, in the Hungarian linguistic tradition, discourse functions/information structure primitives are defined in terms of syntactic positions.) The advantage of this definition is that it takes into account all utterance types, contrary to information structure analyses working only with declarative sentences (and using interrogatives as a tool for indicating the focus structure of the answer). The authors argue that this approach is reflected by prosody in French: focus cannot be associated with a specific tone or accent, but it is indicated by different illocutionary boundary tones depending on the type of the utterance. I will also make use of this definition when proposing an alternative information structure in LFG (7).

In assertive utterances, focus is marked by the boundary tone L% on its right edge (69). In declarative questions (see Section (6.6.1)), focus (the XP that is specifically questioned in a partial polar interrogative) is marked by H% (70), and in so-called requests for confirmation, by HL% (71), and a copy of this tone is realized at the end of the utterance:

(69) a. Who bought a mandolin?
     b. Mallarmé L% aux acheté une mandoline. L% aux
        Mallarmé aux bought a mandolin
        Mallarmé bought a mandolin.

(70) Tu pars dans la voiture de Jean-Bernard H% aux dimanche prochain
    you leave in the car of Jean-Bernard Sunday next
    H% aux ?
    Are you leaving in Jean-Bernard’s car next Sunday?

(71) Jean-Bernard HL% aux est venu hier ? HL% aux
    Jean-Bernard aux come yesterday
    Jean-Bernard came yesterday, didn’t he?

Example (70) also reveals an interesting factor about the focus status of elements in questions. Focal elements can appear in questions independently of question words. Analyses claiming that question words are uniformly focused (and that all sentences contain some semantic peak/a focal element), could not account for the structure of this example, since the equivalent of the question word of constituent
questions would be (semantically) the yes/no operator, realized sometimes as est-ce que in French, which is by no means the focus of such utterances.

Although focus can be identified in utterances of all illocution types, the authors remark that

“[w]hen Assertion is involved, the propositional content of the focus contributes the update [...] When Question is involved, the focus contributes the type of information the speaker is questioning” (Beyssade et al., 2004b, p. 482).

This means that the information structure concept focus has a different meaning in the different types of utterance. This differentiation is lost if we refer to this with one and the same concept (focus). However, we also miss a generalization if we do not assume that every utterance contains some highlighted element that corresponds to its communicative contribution. Therefore, in this thesis, I use the term prominent (and prominent elements) to refer to the constituents affected by the illocutionary operator. The term focus is reserved for formally highlighted (and semantically prominent) constituents in answers or reactions to previous utterances (which are, in most cases, questions, but can also be assertions in the case of corrections and parallel structures) (see Chapter (4)). Foci can appear in questions, but usually this indicates that the question has already been posed prior to this utterance. Most question words also belong to the set of prominent elements.

The authors argue that in narrow-focus sentences, the focused constituent divides the sentence in a pre-focus and a post-focus domain, which exhibit different prosodic properties: the pre-focus domain is characterized by a non-compressed, whereas the post-focus domain by a compressed pitch register. More importantly, the prosodic realization of the post-focus domain is significant in terms of the information structure associated with the sentence. In the case of short sequences, the sequence can be realized as a low plateau or as a continuous fall. In the first case, there is no new discourse referent introduced in this part of the sentence and it does not modify the discourse topic. In the second, on the other hand, the L% boundary tone at the end of the focused constituent (in assertive sentences) does not reach its target until the end of the sentence, where there is a constituent with a C-accent ($H^c$) on its first syllable. The C-accent indicates that the element modifies the current discourse topic:

(72) a. What does he smoke?
   b. Il fume des cigarettes L% IntP le DI($H^c$) manche. L% IntP
   he smokes cigarettes the Sunday
   On Sundays, he smokes cigarettes.

(Beyssade et al., 2004b, p. 486, 13)

On the other hand, if the sequence is longer, it is usually realized as a succession of downsteps with the copy of the illocutionary boundary tone at the end of each
Major phrase. However, a high continuous boundary tone is also possible at the end of a Major phrase if a new discourse referent is introduced in that part:

(73)  
   a. What did you give for the practice exam?
   b. J’ai donné trois exercices de syntaxe à l’étudiant de licence pour le concours blanc.

(74) C’est à Londres qu’elle a passé son enfance.

Syntactic structures that are usually related to the discourse function focus are cleft sentences:

(75) J’ai ma sœur qui est malade.

Clefts consist of a C’est XP part and a relative clause (called the coda). Another structure used mostly in spoken language is presentational clefts:

(76) Il y a ma sœur qui est malade.

According to Lambrecht (1994), spoken French avoids the co-mapping of the focused constituent and the grammatical subject, which leads to the above illustrated bi-clausal structures.

Clefts are supposed to indicate the focus-ground structure in that the clefted XP corresponds to the focus and the relative clause to the ground. However, Doetjes et al. (2004) show that clefts can be all-focus sentences as well, where the relative clause also belongs to the focus:

(77) C’est avec plaisir que je vous invite à participer à ce séminaire.

(Doetjes et al., 2004, p. 535, 10)

However, what is common in focus-ground and all focus clefts is that the rel-
ative clause (the *coda*) is presuppositional (which can be shown with the famous tests of negation which does not affect the presupposition), and the clefted XP is foregounded, which Doetjes et al. (2004) call *zooming effect*. Compare:

(78) #On dit que c’est pour cette raison qu’elle n’y est plus jamais
we say that it is for this reason that she PRT.CL AUX more
revenue, mais ce n’est pas vrai: elle y est revenue never come.PST back, but this PRT.is not true: she CL AUX
plusieurs fois.
come.PST back several times
They say that it’s for that reason that she never went back, but that’s not true: she went back several times.

(79) On dit que pour cette raison elle n’y est plus jamais
we say that for this reason that she PRT.CL AUX more never
revenue, mais ce n’est pas vrai: elle y est revenue come.PST back, but this PRT.is not true: she CL AUX comePST back
plusieurs fois.
some several times
They say that for that reason that she never went back, but that’s not true: she went back several times.

(78) is anomalous because the presupposed part cannot be affected by negation, whereas (79) is perfectly acceptable.

The authors also argue, contrary to É. Kiss (1998a), that clefted constituents are not necessarily exhaustive (according to É. Kiss (1998a), that clefted constituents are the equivalents of the exhaustive, *identificational* focus in Hungarian):

(80) Mari LONDONBAN tőltötte a gyerekkorát.
Mari London.INESS spent the childhood.POSS.3SG.ACC
It was in London that Mary spent her childhood. / Mary spent her childhood in LONDON.

### 6.5.2 Discourse thematic organization

The other aspect of information structure that is formally (prosodically) marked in French, is the discourse thematic organization. Büring (1997, 2003)’s analysis of the so called B-accent indicating a complex discourse strategy in English (and German) has already been presented in Chapter (4). The French equivalent of the B accent is the so-called *C-accent* in French, referring to a high tone on the first syllable of the element indicating the discourse strategy. The authors observe that elements carrying a C-accent (H*) can appear in pre- or post-focal domain, and even in the focused constituent in French:

(81) a. What did the students study this year?
b. Les étudiants de première année ont fait de la syntaxe générative, les étudiants de deuxième année ont fait de la syntaxe fonctionnelle.

First year students studied generative syntax and second year students functional syntax.

(Beyssade et al., 2004b, p. 490, 17)

They also add that when the C-accent appears in the focused constituent, it is compulsory in the following example, but not in (81-b).

(82) a. Who came?
   b. Ber(H∗c)nard est venu (, pas Marie).
      Bernard AUX come (, not Marie)
      Bernard came, not Mary.

(Beyssade et al., 2004b, p. 491, 20)

It seems that a C-accent is not obligatorily present if the contrast is explicit in the context (the alternatives are enumerated), whereas it is compulsory whenever the focused constituent is implicitly contrasted to some alternative not present in the discourse. If this observation is correct, it clearly goes against the assumption that all foci are contrastive. Contrastiveness means the explicit or implicit presence of alternatives, and the implicit alternatives are indicated by a C-accent in French. If no alternatives are involved, no C-accent not appears on the focused constituent:

(83) a. When are you going to London?
   b. J’y vais dimanche.
      I there go Saturday.
      I am going there on Saturday.

Elements carrying a C-accent are usually referred to as contrastive topics. Contrast is per definitionem included in the definition of a complex discourse strategy, since (even if only implicitly) elements carrying a C-accent are contrasted to other elements in the answers to the other subquestions. In this respect, it is not surprising that this contrasted element can be (part of the) focus, and that contrastive topics are sometimes referred to as topic in focus (Molnár, 1998), or kontrafüksz (Kenesei, 1989).

Based on these facts, I consider that focus and contrastive topic are distinct notions, the former is the most prominent part of a reaction to a previous sentence, whereas the latter indicates a discourse strategy, but they can overlap in some cases, which results in confusion in some analyses.
So far, we have dealt with the (prosodic) manifestation of focus and contrastive topic. Apart from prosody, syntactic processes, such as left- and right dislocation, are sometimes also supposed to correspond to discourse structure categories.\(^3\) French exhibits both left dislocation (with a resumptive element in the remainder of the sentence) and (syntactic) topicalization (without a resumptive element). Both can be used for semantico-pragmatic topicalization, foregrounding the topic (a subtopic of the discourse topic) and indicating its importance in the sentence (or in the subsequent part of the discourse):

(84) **Left dislocation**

a. Q: -Qu’est-ce que tu vas donner à Pierre et à Marie ?
   What you FUT to give to Pierre and to Marie
   What are you going to give to Peter and to Mary?

b. A: -(A) Pierre, je lui donnerai un livre, (à) Marie, je
to Pierre, I (CL).DAT give.FUT a book, to Marie, I
lui offrirai des fleurs.
(CL).DAT offer.FUT flowers
To Peter, I will give him a book, to Mary, I will offer her flowers.

(Delais-Roussarie et al., 2004, p. 508, 18)

(85) **(syntactic) Topicalization:**

Mary a réuni les élèves. Aux filles, elle a donné des
Mary has brought together the students to the girls, she AUX given
exercices d’algèbre. Aux garçons, elle a dicté un problème de
exercises of algebra to the boys, she AUX dictated a problem of
géométrie.
geometry

Mary brought together the pupils. To the girls, she gave algebra exercises.
To the boys, she dictated a geometry problem.

(Delais-Roussarie et al., 2004, p. 508, 19)

The left-dislocated topic constituent must be given (however, according to Laurens (2010), this is not always the case). Lambrecht (1981) distinguishes between *four degrees of givenness: strictly given referents are inferable from the discourse content (these are usually referred to by non-accented pronouns), non-strictly given referents can be textually evoked, situationally evoked (present in the extra-linguistic context or situation), or inferable from the context.*

Unlike contrastive topics, left-dislocated topics do not carry a C-accent, since it does not imply any contrast with parallel thematic shifters in a complex discourse strategy. From the two types of left dislocation, only left dislocation without a

\(^3\)This is, however, a debated issue, which I do not investigate further here
resumptive element can be used for introducing contrastive topics that reshape the discourse topic and indicate a complex discourse strategy (Delais-Roussarie et al., 2004).

Right dislocation is supposed to foreground new information by placing the (semantico-pragmatic) topic in the background.

(86) Je lui ai donné un livre, à Marie.
I CL.DAT.3SG AUX given a book, to Marie
I gave a book to her (, to Mary).

(Delais-Roussarie et al., 2004, p. 515, 30)

Right dislocated constituents are also given (strictly given pronouns, textually or situationally evoked or inferred discourse referents). Contrary to left dislocation without a resumptive pronoun, right dislocation cannot be used as a thematic shifter or to mark an opposition. This is evident if we consider that right dislocation places the topic in the background in order to foreground new information. Right dislocated constituents are referred to as antitopic in Lambrecht (1981, 1994).

6.6 Questions

6.6.1 Variety of French interrogatives

In this section, I present a classification of interrogatives in French. It is a well-known fact that French interrogatives show a variety of different forms, some of them supposedly used more typically in written, or formal language, the others in spoken or informal language. The syntactic difference between the two is the extraction of the wh-phrase in formal French, and their in situ position in informal French. The extracted wh-phrase is often accompanied by the reinforcing sequence est-ce que, which is followed by declarative word order (without inversion). Beyssade (2007) (see also Obenauer (1976)) gives five examples that illustrate ways of asking Who did John invite? in French ((87)-(91)):

(87) Qui (diable) Jean a-t-il invité ?
who (the hell) John AUX he invited

(88) Qui (diable) Jean a invité ?
who (the hell) John AUX invited

(89) Qui est-ce que Jean a invité ?
who is-it that John AUX invited

(90) C’est qui que Jean a invité ?
it is who that John invited

(91) Jean a invité qui ?
John AUX invited who
Beyssade (2007) considers the cases in (87)-(89) as true interrogatives, with a preposed *wh*-phrase, whereas the examples in (90) and (91) as declarative questions (*questions déclaratives*) (clefted (90) or *in situ* (91)), which, although they contain a question word, are similar in word order to the so-called questioning declaratives (*déclaratives questionnantes*):

(92) C’est Catherine que Jean a invitée ?
    it is Catherine that Jean AUX invited
    Was it Catherine that John invited?

(93) Jean a invité Catherine ?
    Jean AUX invited Catherine
    Did Jean invite Catherine?

Beyssade (2007) makes also the following observations. *Pourquoi* (why), just like *que* (discussed in the previous subsection) cannot appear in situ:

(94) *Jean a invité Catherine pourquoi?
    Jean AUX invited Catherine why
    Why did Jean invite Catherine?

In addition, modifiers like *diable* (equivalent of *the hell* in English), cannot appear in situ either:

(95) *Jean a invité qui diable ?
    Jean AUX invited who the hell
    Who the hell did John invite?

According to previous assumptions, (Chang, 1997; Cheng and Rooryck, 2000) (clefted and *in situ*) declarative questions are used in contexts where the event denoted by the question is presupposed, *i.e.*, in the following examples, we presuppose that Marie bought something.

(96) Marie a acheté quoi ?
    Marie AUX bought what
    What did Mary buy?

(97) C’est quoi que Marie a acheté ?
    it is what that Marie AUX bought
    What is it that Mary bought?

However, (Obenauer, 1994; Beyssade, 2007) argue that no such presupposition is associated with the *in situ* questions, since, contrary to cleft questions, they can be answered with *personne* (nobody) or *rien* (nothing):
6.6. Questions

(98)  a. Q: -Tu as fait quoi pendant les vacances ?
     you AUX done what during the holidays
     What did you do during the holiday?

     b. A: -Rien.
     nothing
     Nothing.

According to Beyssade (2007) the difference between questions with preposed and in situ question words is that while the interrogatives with extracted wh-words can introduce a new discourse topic, interrogatives with in situ wh-words continue the current discourse topic\(^4\) (in this respect they are like declarative clauses). As we will see, the possible interpretations of multiple questions do not depend on the fact whether one wh-word is extracted or not. However, we will distinguish them as subclasses of the same type, because of the above mentioned discourse-level difference between them.

6.6.2 Lexical question-formation

It has to be pointed out that unlike English, French interrogatives do not contain subject-auxiliary inversion at all (see also Gazdik (2008)). Instead, French contains interrogative verb forms in its lexicon. According to Huot (1987); Miller and Sag (1997) (as we have seen), French bound pronominal clitics (complement or enclitic) are best analyzed as lexical pronominal affixes forming one single lexical unit (word) with the verb. Similarly, a class of suffixes qualify verbs as interrogative in French, incorporating the subject as well (the hyphenated verb forms in the examples (99) and (102) all illustrate this phenomenon). The clitic/affix status of these elements can be justified by the fact that the subject can be present in the clause as well (thus the clitics/affixes cannot be analyzed as subjects, but sometimes the verb form incorporates the subject as well):

(99)    Paul part-il ? (clitic doubling)
    Paul leaves he
    Paul, is he leaving?

In addition, they cannot take scope over coordinated verbs, unlike ordinary pronouns:

(100)   Il vient et repart aussitôt.
    He comes and leaves soon.
    He comes and leaves soon.

\(^4\)According to Beyssade et al. (2004b) the Discourse Topic can be defined as the question under debate. It is to be distinguished from the notion of topic, which is interpreted at the sentential level, referring to one constituent of the sentence.
In (100), *Il, as a personal pronoun can take scope over coordinated verbs, unlike the homonymous lexical affix, which must be present on each verb (101)-(102):

(101) *Vient et repart-il ?
     comes and leaves he?

(102) Vient-il et repart-il ?
     comes he and leaves he?
     Does he come and leave?

The complementizer in polar questions also contains an interrogative verb (*est-ce*):

(103) **Est-ce que** Paul vient ?
     COMP.INTERR Paul comes
     Is Paul coming?

From the above data we can draw the following conclusion. French simple questions can contain interrogative verbs with enclitics, which is, therefore, a lexical difference to be taken into consideration by any analysis.

### 6.6.3 The role of weight

Not only the information structure, or the formal/informal register, but the concept presented in Section (6.2.3), the feature of syntactic weight, also plays a role in French interrogatives. The items *que* and *quoi* exhibit different properties in finite interrogatives. Consider the following examples:

(104) **Que/** *Quoi* fais-tu ?
     what do you
     What are you doing?

(105) Tu fais *que/quoi ?
     you do what
     What are you doing?

As the examples show, *quoi* and *que* are not interchangeable, since *que* cannot appear in situ, whereas *quoi*, cannot be extracted. However, when modified, or as complement of a preposition *quoi* can be extracted:

(106) **Quoi d’autre** as-tu vu ?
     what else AUX you seen
     What else did you see?

(107) A **quoi** penses-tu ?
     about what think you
6.6. Questions

What are you thinking about?

It was shown above that French exhibits interrogative verb forms with lexical pronominal affixes. According to Abeillé and Godard (1999b,c) not only inverted pronoun clitic subjects are to be treated as lexical affixes, but the interrogative element *que* as well. However, in Abeillé and Godard (2004a) interrogative *que* is analyzed as a weak form, always appearing with an interrogative verb (in finite interrogatives). Even if the exact status of interrogative *que* cannot be decided in this thesis, I conclude, following Abeillé and Godard (2011) that it is by no means an autonomous word, since it cannot be used in isolation and cannot be modified or coordinated. From a diachronic perspective, it is supposed to be on its way to become an affix.

The arguments for analyzing *que* as a weak form (or an affix) are similar to those enumerated in the previous sections:

1. *Que* cannot be separated from the verb:

   \[(108) \] *Que* Jean fera-t-il?
   what Jean do.FUT he
   What will John do?

2. No coordination with other interrogative pronoun

   \[(109) \] Qui ou quoi/*que* te dérange ?
   who or what you disturb
   Who or what disturbs you?

3. Functional restrictions

   *Que* cannot be the complement of a preposition:

   \[(110) \] A qui/quoi/*que* penses-tu ?
   about who/what/what think you
   Who/ What are you thinking about?

4. No scope over a conjunction of verbs or a conjunction of VPs:

   \[(111) \] a. *Que* liras-tu et *(que)* traduiras-tu ?
   what read.FUT you and what translate.FUT you
   What will you read and what will you translate?
   (Abeillé and Godard, 1999b, p. 1, 6b)
b. Il se demandait que choisir et expliquer à ses étudiants la semaine prochaine.

He himself asked what to choose and to explain to his students next week.

(Abeillé and Godard, 1999b, p. 1, 7a)

However, there are attested examples in which *que* can take scope over coordinated infinitives:

(112) Que manger et boire durant un tournoi ?

What should one eat and drink during a tournament?

(Abeillé and Godard, 2011, p. 5, 9c, from tennis-de-table.com/forums, 29/10/2004)

This possibility of scope taking over coordinated infinitives in certain contexts is the only reminiscent of the earlier stage of *que* becoming an affix.

*Que* is obligatorily verb-initial and combines with an interrogative verb form (a subject suffix following the verb stem).

On the other hand, the authors argue that *quoi* is a light pronoun, adjoining to lexical verbs. The arguments supporting this view are the following:

1. Ordering properties

   Bare *quoi* is adjoined to the lexical verb, like light adverbs and quantifiers:

   (113) Paul se demande quoi/*quoi d’autre faire.

   Paul himself asks what else to do

   Paul is wondering what / what else to do.

   (114) Tu donnes quoi à Jean (*quoi) ?

   you give what to John (*what)

   What do you give to John?

2. Extraction properties

   *Quoi*, as referred to above, cannot be fronted unless it is modified or as a complement of a preposition. Like other light forms, it cannot be preposed unless it is made syntactically heavy Abeillé and Godard (1999b).
3. Order of question words in a multiple question

In a multiple question, *quoi* obligatorily follows the verb:

(115) Il a donné *quoi à qui?*
he AUX given what to who
What did he give to whom?

In the reverse order of the question words, the question is ungrammatical:

(116) *Il a donné à qui *quoi*?
he AUX given to who what

In other examples, however, we can find word order variations:

(117) Il a donné *à quel groupe quel exercice de maths*?
he AUX given to which group which exercise of maths
To which group did he give which maths exercises?

(118) Il a donné *quel exercice de maths à quel groupe*?
he AUX given which exercise of maths to which group
Which maths exercise did he give to which group?

As can be seen from the above examples, *quoi* is the *wh*-word in *in situ* questions, but it is ungrammatical with interrogative verb forms. *Que* cannot appear *in situ*, but it always co-occurs with an interrogative verb form.

In some contexts, however, *que* and *quoi* are in competition with each other. This is the case of interrogatives with a verb in the infinitive:

(119) Je ne sais *que/quoi* faire.
I PRT know what to do
I don’t know what to do.

However, *quoi* is preferred in questions embedded by factive verbs:

(120) Il *sait/trouve/dit *quoi/*/que faire.
he knows/finds/says what to do
He knows/finds/says what to do.

According to Abeillé and Godard (2004a), in this case *quoi* is not extracted, but adjoined to the infinitive V, like other light items. Supposedly, this accounts for the fact that it can take scope over the coordination of lexical verbs, but not over coordinated VPs:
a. Je me demande quoi [lire et traduire] au cours la semaine prochaine.
   I wonder what to read and translate at the lesson next week.

b. Je me demande quoi [faire au cours et (*quoi) donner comme devoir pour la semaine prochaine].
   I wonder what to do at the lesson and what to give as homework for next week.

In the next subsection we examine the prosody of French interrogatives.

### 6.6.4 Prosody of French interrogatives

Marandin (2004) and Beyssade et al. (2007) argue that the prosody of interrogatives in French is sensitive to the partition of content to the same extent as assertive utterances. As we have seen above, the illocutionary boundary tone is anchored at the right edge of the focused constituent, and repeated at the end of the sentence, with eventual copies at the end of Major phrases in the post-focus domain.

In most interrogatives, as Beyssade et al. (2007) report, the nuclear contour is falling (non-falling contours reveal an expected disagreement from the locutor). The illocutionary boundary tone (H%) is anchored at the left edge of the part of the question which contributes the restriction. According to Krifka (2001), the meaning of a question can be illustrated as a pair of an open proposition and a restriction on the variable in the proposition (for instance the restriction is [human +] in the case of what). In constituent questions, the illocutionary boundary tone is attracted by the question word. In polar interrogatives, two cases are to be distinguished. First, in partial interrogatives, the tone H% is attracted by the XP questioned:

(122) Est-ce que l’armée(H%) américaine sera là aussi ?
   Will the American army also be there?

(Beyssade et al., 2007, p. 8, 17)

In total polar interrogatives, the H% tone is either anchored on the interrogative marker est-ce que (123), or on the interrogative main verb (124):

(123) Est-ce (H%) qu’il y a un magasin de vélo dans la ville ?
   Is there a bike shop in the city?

(Beyssade et al., 2007, p. 9, 18)
The examples show that the prosody of French is sensitive to the information structure of the sentence. In declaratives, it is the focus/background partition that accounts for the placement of the illocutionary boundary tone, whereas in interrogatives, it is anchored on the element contributing the restriction (the question word), on the interrogative marker, or on the XP especially questioned.

6.7 Summary

The aim of this chapter was to provide the necessary background about French for the analysis of (multiple) questions. In the first section, I examined the syntactic structure of French, with special emphasis on word order phenomena. I considered the status of pronoun clitics/lexical affixes, and showed later that in interrogatives, French exhibits interrogative verb forms and not subject-auxiliary inversion. I also presented the concept of syntactic weight, and showed that it plays a role in the distinction between *que* and *quoi* in interrogatives, and in the order of *in situ* question words in a multiple question. I also considered Subject NP Inversion and right/left dislocation and showed that they were governed by discourse/pragmatic factors. Dislocation with/without a resumptive element also plays a role in the answers to multiple questions. Another formal property I examined was some relevant aspects of French prosody. I illustrated how focus is indicated by different illocutionary boundary tones depending on the type of the utterance, and also presented the C-accent, indicating the reshaping of the current discourse topic. Finally, I turned to some aspects of interrogatives in French. I presented the variety of French interrogatives and showed that the difference between their use is related to the discourse topic. Then I gave a short summary of the prosody of French interrogatives, illustrating that in interrogatives (contrary to assertives) the illocutionary boundary tone (H%) is anchored at the left edge of the part of the question which contributes its restriction.
Part III

Analysis
The Basic Considerations of the Analysis

Contents

7.1 Introduction .............................................. 265
7.2 Characterizing Multiple Questions in the I-structure .... 266
  7.2.1 Question words as foci ................................ 266
  7.2.2 Question words as topics ............................... 268
  7.2.3 Analysis ............................................... 273
7.3 I-structure: an Alternative Analysis ......................... 276
7.4 Discourse Structure Revisited .............................. 288
  7.4.1 A separate level ....................................... 288
  7.4.2 Discourse relations ................................... 291
  7.4.3 The place of d-structure ............................... 306
7.5 Summary ................................................... 307

7.1 Introduction

The essence of what we have seen so far about the languages that are the object of this study and the multiple questions they exhibit is that the phenomena related to multiple questions cannot be understood, analyzed, and formalized without reference to the information structure of the individual sentences, and to the relationship between the individual sentences in the discourse. Information structure has been part of the LFG formalism since the seminal papers of Butt and King (1996) and King (1997). Its architecture is based on Choi (1997, 1999). There have also been attempts to integrate discourse structure into the LFG framework. The present analysis contributes to this direction of research within LFG. The chapter is structured as follows. First, I examine a possible analysis of multiple questions in the mainstream LFG architecture of the information structure, pointing out, nevertheless, some problems. Then I propose an alternative architecture for the i-structure that could capture the already presented facts in a more adequate way. Finally, I present a possible representation of the super-sentential, discourse level in LFG.
Chapter 7. The Basic Considerations of the Analysis

7.2 Characterizing Multiple Questions in the I-structure

Since the analysis of multiple questions is crucially dependent on the information structure, in this section I propose an analysis based on the i-structure architecture of the mainstream LFG-analyses. This structure is based on Butt and King (1996), King (1997) and Choi (1997) and assumes four sets: TOPIC, FOCUS, BACKGROUND INFORMATION and COMPLETIVE INFORMATION. We will consider now the possible i-structure of (multiple) questions. Question words are supposed to be the formally highlighted, prominent elements of questions. Thus, I examine if they are to be analyzed as topics or foci (belonging to one of the [prominent +] sets).

7.2.1 Question words as foci

Analyzing constituent questions cross-linguistically, Mycock (2006) proposes that all question words are focused. The same is advocated by Haida (2007). Mycock argues that an information gap in the speaker’s knowledge has to be highlighted and focusing is the way question words are highlighted. This assumption is supported by the fact that focus and question words tend to occupy the same syntactic position (Hungarian is often cited as a language that supports this view), and exhibit the same prosody (pitch accent). From these it follows that when proposing an i-structure for multiple questions, Mycock places both question words and foci in the focus set at i-structure. Nevertheless, she differentiates between interrogative and non-interrogative foci in order to account for the Hungarian data (despite the ban on more than one preverbal focus, question words can co-occur in the preverbal domain).

We should also note here, that King (1995), still encoding discourse functions in the f-structure, proposes the Q(uestion)-Foc(us) attribute and remarks in King (1997, p. 3) that

"[t]he discourse functions associated with questions are not fully understood. The term Q(uestion)-Foc(us) is used to indicate the role which corresponds to the focus in the answer to the question.”

I assume that this puzzle still persists today.

Haida (2007) also has to confront the problem of Hungarian pair-list questions. He assumes that all question words are obligatorily F(ocus)-marked. After considering the data, he posits that ”[a] wh-word can be left non-F-marked if it is to provide the Sorting key [emphasis mine, G. A.] of a multiple question” (p. 199), but indicates that this constraint will be abandoned later. Later he concludes that
"[i]n Hungarian and Northern Soho, there is a syntactic condition for *wh*-words to be F-marked in simple *wh*-questions, but (for reasons to be explained) this requirement does not hold for all *wh*-words in multiple questions" (p. 202).

Further on, when providing his semantic account, he assumes that in the syntactic structure

"the clause-initial *wh*-phrase is assumed to occupy the specifier position of TopP. [...] However, the following analysis will ignore the interpretive import of the head of TopP (which is therefore not represented above). That is, the non-F-marked *wh*-phrase [...] will not be assigned a topical interpretation, but merely a non-focal interpretation" (pp. 221-222).

Crucially, it seems to be inadequate to ignore the interpretive import of a syntactic position, which is inside a discourse-semantically motivated phrase (TopP). On the contrary, the semantic import of the FocP, where the focused, immediately preverbal *wh*-phrase is supposed to reside, is fully exploited, since the author derives his semantic analysis of preverbal *wh*-words from Szabolcsi (1994)’s analysis of focus in Hungarian. He assumes that the existential presupposition of the *wh*-expression in FocP

"interacts with the denotation of the non-F-marked *wh*-word to give rise to what appears to be an existential presupposition with respect to both *wh*-words. However, it is important to distinguish the true presupposition, the existential presupposition triggered by the F-marked *wh*-word, from the derived presupposition, which is not a presupposition by itself but rather the truth-conditional content of the non-F-marked *wh*-word, interacting with the embedded true presupposition” (p. 223).

The derived existential presupposition on the non-F-marked *wh*-expression in Spec,TopP is then turned into an exhaustiveness presupposition by the interrogative complementizer (demonstrated in dynamic semantics). Let us illustrate this in the following way:

(1) Hungarian:

\[
[\text{CP} \ C^{Q+} \ [\text{TopP} \ \text{ki} \ [\text{FocP} \ \text{mit}^{F+} \ [\text{VP} \ \text{hozott} \ \text{Marinak}]])]
\]

who who what brought Mary.DAT

Who brought what for Mary?
for everyone, there is something that s/he brought for Mary

(2) Hungarian:

\[
[\text{CP } C^{(Q^+)} [\text{TopP mit } [\text{FocP ki}^{(F^+)} [\text{VP hozott } \text{Marinak}]]] [\text{what who brought Mary.DAT}]
\]

What was brought by whom for Mary?

for everything, there is someone who brought it for Mary

Multiple questions (with a pair-list reading) are supposed to carry two presuppositions: for every x and there is an y that.... According to Haida, the second comes from the F-marked wh-phrase and is then transferred onto the other, whereas the first comes from the (abstract interrogative) complementizer.

This account remains silent about the final status of the linearly first question word. It is not originally focused (F-marked); it sits in a Topic position, although it is not analyzed as a topic and the interpretive import of the topic position is ignored. It gets its (existential and exhaustive) presupposition from the interaction with its sister nodes (the abstract complementizer and the preverbal question word), although in order to assume this, one has to suppose the presence of an abstract complementizer, which is at the same time interrogative, and that one question word can overtake a presupposition from another. As we have seen in Chapter (2) (the analysis in Krifka (2001)), and we will see in the next subsections, simpler semantic accounts are also possible that can derive the same interpretive result as Haida’s.

To conclude, both accounts reflect the problem that in spite of the common properties, (all) question words cannot easily be collapsed into the category of focus.

7.2.2 Question words as topics

It has already been suggested in the literature that D-linked, or in some languages, initial interrogative words, share some properties with topics, in that they are both contextually determined, given, salient in the discourse and denote a contextually determined set of entities (see also Gazdik (2008) for an LFG analysis). Let us now have a look at the most important arguments.

7.2.2.1 Further semantico-pragmatic arguments

Surányi (2006) argues that Hungarian high wh-phrases, although not moved by topicalization, are interpreted at interfaces as topics, since, like topics, they invariably quantify over presupposed sets, and constituents that correspond to these wh-phrases appear in the topic, and not in the focus position in Hungarian. Unfortunately, the claim that an element is not moved by topicalization, but is interpreted at interfaces as a topic, is difficult to formalize in a minimalist framework.
that aims to account for all phenomena in syntax, via movement to various functional projections, which, in turn, express discourse or semantic information. It seems, therefore, that an LFG approach, in which information structure is treated independently from syntax, is a more adequate framework to capture these facts.

Another argument supporting this claim is that constituents corresponding to D-linked/Sorting key question words in the answer are contrastive topics (Büring, 2003). As Büring (1997, 2003) shows, the answer to a multiple question necessitates a complex strategy, in which the original question is divided into subquestions, and the congruent set of answers constitute an answer to each of these subquestions. Nevertheless, this does not exclude the possibility of giving incomplete answers (i.e. not providing an answer to all the potential subquestions), which, then, according to this model, imply that the person has more to say about the issue (these are referred to as partial and complete implicative answers in the present analysis). Büring (2003), for instance, illustrates this with the following example (based on Jackendoff (1972) and Roberts (1996)):

\[(3)\] **Who ate what?**

The issue at hand can be approached by the people and by the dishes, depending on the situation. In the first case, the question is the unification of the following subquestions:

\[(4)\] \{What did John eat? What did Mary eat? What did Fred eat? What did Steve eat? etc.\}

Whereas in the second case, the appropriate subquestions would be:

\[(5)\] \{Who ate the beans? Who ate the chicken? Who ate the omelette? Who ate the carrots?, etc.\}

Note that in both cases the question word qualifying as the Sorting key of the original question is replaced by an element of its denotation in the subquestions, whereas the other question word does not change. This shows that while this latter is a true question word in the sense that it does not (usually) denote a contextually given set, the other question word can be paraphrased by: which of the people... ? or which of the things... ?. We should remark here that both (all) question words can denote a contextually given set. In this case, the question is referred to as a matching question, and it has no role of identification, only that of matching the already known elements.

The answers to these subquestions contain a contrastive topic (indicated by a B-accent in English, C-accent in French, rising tone in Hungarian)\(^1\) and a focus (A-accent in English, usually falling tone in French, and pitch accent with a sharp fall in Hungarian), the contrastive topic corresponding to the Sorting key, and the foci corresponding to the ‘true’ question word. Since the focus is analyzed as the

\(^1\)More precisely, note that in Hungarian, in the case of a pair-list answer (the enumeration of answer-parts), the rising tone of the contrastive topic appears at the end of the clause, and the contrastive topic itself exhibits rather a falling tone.
(congruent) answer to a question, or the distinguished part of corrections, etc., it should also be assumed that the status of the two question words is fundamentally different.

Another argument in favour of the view that Sorting key/D-linked question words have a similar interpretation to contrastive topics can be found in Krifka (2001) and Gyuris (2009). Krifka argues that the Structured Meaning Approach can account for the meaning of questions better than Proposition Set Theories, since, among other advantages, it offers a simpler and more elegant account for multiple questions with a pair-list answer. According to this approach, the meaning of such questions is a function, from the denotation of the Sorting key/D-linked question word into the denotation of the other. It follows from the definition of a function that all the members of the first set have to be paired up with a member of the second, but not all members of the second have to be used. This corresponds to the above observation that a (pair-list) multiple question can be considered as a set of single questions, like in (4) and (5), and also explains the intuitive observation that the question is "about" the D-linked question word.

Gyuris (2009) investigates certain quantity-indicating determiners as contrastive topics with verum/falsum focus in Hungarian and observes that certain structures are ill-formed:

(6) Hungarian:

*HATnál több diák ELjött az előadásra.  
hat.ADESS more student VM.come.PST the talk.SUBL

*More than SIX students DID attend the talk.

(Gyuris, 2009, p. 627, 4)

Gyuris argues that the ill-formedness is due to semantic and not to syntactic criteria. She observes that sentences containing contrastive topics cannot be uttered out of the blue (cannot be all-focus sentences), and necessarily contain a focused constituent. A sentence (S) containing a focused constituent implicates that no other sentence (S’) obtained by replacing the focused constituent of S by an alternative from its denotation can be true simultaneously with S. This holds if the background (the non-focused) part of S remains unchanged:

(7) a. What did John bring to the party?  
b. John brought WINE. (S)  
c. #John brought BEER. (S’)  
d. No, John also brought BEER.

Example (7) shows that two sentences that differ only in their focus content cannot be simultaneously true. If the simultaneous truth of the sentences is to be expressed, first the assumption of the truth of (S) has to be cancelled (No!) (see also Szabolcsi (1981)).

However, S’ and S can be simultaneously true if the background is also different
in the two sentences. Contrastive topics mark that part of the background that has to be different, if $S$ and $S'$ are simultaneously true:

\[(8)\]  
\[a. \] To Jack’s party, John brought wine.  
\[b. \] To Miriam’s party, John brought beer.

Based on this observation, Gyuris proposes to capture the presupposition contributed by contrastive topics by assuming a function that maps the alternatives to the denotation of the contrastive topic into the set of alternatives to the denotation of the focus (including the denotation of the contrastive topic and the focus as well). Furthermore, it is assumed that given a sentence $S$ with a contrastive topic (and a focus), there must be a proposition that provides related information about an alternative of the denotation of the contrastive topic that is not entailed by $S$, i.e. it is mapped on a value of the range of the function (the focus denotation alternatives) that does not correspond to (or is not entailed by) the focus value of $S$.

It is not difficult to see now why (6) is ungrammatical. The possible alternatives to the denotation of the contrastive topic are the cardinalities \{1, 2, 3, 4, 5, 6\}. Since the denotation of verum focus has only one alternative (the falsity of the proposition), if the sentence is well-formed, we would have to show that at least one of the alternatives of the denotation of the contrastive topic can be mapped onto the alternative of the focus denotation and the resulting proposition can be true simultaneously with (6). Such proposition would be the following:

\[(9)\]  
NÉGY diák NEM jött el az előadásra.  
four students not come.PST VM the talk.SUBL  
FOUR students DIDN’T attend the talk.

Obviously, this cannot be true at the same time as (6), since (6) entails the falsity of (9): if more than six students did attend the talk, then this is true for any number of students between one and six as well.

It is thus assumed that both (pair-list) multiple questions and contrastive topics presuppose a function from the alternatives to their denotation into those of the focus/other question word. A sentence containing a contrastive topic implies the existence of a parallel proposition with different contrastive topic and focus values, whereas in the case of pair-list questions, one single answer is not satisfactory; what is expected is a list of pairs, or at least two pairs.

7.2.2.2 Syntactic arguments

Grohman (2006) shows that only topicalizable elements can intervene between two \textit{wh}-phrases in German multiple questions and proposes that therefore all \textit{wh}-phrases undergo topicalization in a German multiple question:

\[(10)\]  
Wer hat viele Bücher/(*wenige Bücher) wo (wenige Bücher)  
who AUX a lot of books/(few books) where (few books)
gekauft?
bought
Who has bought a lot of/few books where?

(Grohman, 2006, p. 16, 17)

He argues that the difference between monotone increasing and decreasing quantifiers is that the former are topicalizable, whereas the latter are not, and this is why the former can appear between the two wh-phrases. He assumes, in addition, that the wh-phrases themselves are also topicalized.

It is not my aim here to contest Grohmann’s analysis (for instance the discourse status and semantics of lower wh-phrases is clearly different from that of initial ones), but to show that the topicality of wh-words has been claimed in connection with various languages, based on a variety of criteria (syntactic position, cliticization, intervention).

Furthermore, in languages in which the order of preverbal question words is not fixed but can change according to the D-linked/non-D-linked status of the question words (Czech, Hungarian), the D-linked question word is the leftmost one in the sequence, appearing on the left periphery of the sentence, just like topics (Lambrecht, 1994, for French) and contrastive topics.

In addition, in Hungarian pair-list questions the question words cannot be preceded by (other) contrastive topics:

(11) *Jánosról, arról mikor mit mondtak?
János.DEL that about when what say.PST.3PL
John, what did they say about him when?

The example illustrates this phenomenon with a resumptive element in the sentence (arról), which is coreferential with the contrastive topic and is one of the tests of contrastive topichood in the language (András Komlósy, p.c.).

7.2.2.3 Prosody

In French, in the answers to pair-list questions, only one XP gets the focus illocationary boundary tone prosodically distinguished, the one corresponding to the non-D-linked question word (Marandin, 2006b). The others bear the already mentioned C-accent, indicating a complex discourse strategy.

(12) a. Quels étudiants étudient quoi dans ce département ?
Which students study what in this department
Which students study what in this department?

b. Les étudiants de première année étudient la syntaxe, ceux de la seconde année la sémantique.
The first year students study syntax, those of the second year semantics.

The illocutionary boundary tone /L% would be impossible on the right edge of les étudiants de première année and ceux de seconde année, since they reshape the discourse topic (les étudiants (the students)) and have the already explained C-accent.

7.2.2.4 Morphology

The topic status of high \(wh\)-phrases has been proposed in connection with other languages as well. According to Jaeger and Gerassimova (2002); Jaeger (2003) topic-fronted objects are clitic-doubled in colloquial Bulgarian. We touched upon this issue in Chapter (2). They also notice that \(wh\)-phrases in multiple questions are also subject to clitic-doubling, in which case superiority effects are cancelled.

\[(13)\quad \text{Kogo kakvo go iznenada?}\]
\[\quad \text{whom what CL.OBJ surprise.PST.3SG}\]

Jaeger and Gerassimova (2002) argue that although the clitics mark topic objects, sometimes the reduplication of fronted focus objects can also be observed in Bulgarian. For them, (13) is one example, the other is that of contrastive topics that can also be doubled. We also observed, both in French and in Hungarian, that contrastive topic and foci share some properties which, in some cases, makes it difficult to distinguish them. However, the authors assume that both question words and contrastive topics are focal elements. This is not a problem for their analysis, since they adopt Choi (1999)'s bi-dimensional information structure, in which the information structure roles differ along two dimensions: [prominent +/-] and [new +/-].

However, the combination of the two features clearly distinguishes between topic and foci, since they are both prominent, but the focus is [new +], whereas the topic is [new −]. In this sense, topic and foci are clearly disjunct sets, consequently it is not clear why interrogative words and contrastive topics are clitic doubled. On the other hand, if we keep ourselves to the assumption that only topical objects are clitic doubled, we should also suppose that the leftmost question word in the sentence is topical and not focal in nature, and clitic-doubling argues for the topicality of high \(wh\)-phrases in Bulgarian. Contrastive topics are also clearly topics in Choi (1999)'s system, since they do not contribute new information.

7.2.3 Analysis

The analysis proposed is thus the following. At the level of information structure, D-linked/Sorting key question words belong to the TOPIC set, but they can be
associated with different syntactic positions, depending on the language. In Czech and Hungarian they precede non-D-linked question words, whereas in English they can be sentence-initial, just like sentence-final. Consider the possible (preliminary) i-structures, and the f-structure of the question in example (3). The curly brackets in the i-structures refer to the fact that the i-structure consists of sets (i.e. there can be more than one topic, focus, background, etc. element in a sentence):

\[
\begin{align*}
\text{TOP} & \quad \{\text{who}\} \\
\text{FOC} & \quad \{\text{what}\} \\
\text{BACKGR} & \quad \{\text{eat}\}
\end{align*}
\]

Figure 7.1: I-structure 1: Who ate what?

\[
\begin{align*}
\text{TOP} & \quad \{\text{what}\} \\
\text{FOC} & \quad \{\text{who}\} \\
\text{BACKGR} & \quad \{\text{eat}\}
\end{align*}
\]

Figure 7.2: I-structure 2: Who ate what?

\[
\begin{align*}
\text{PRED} & \quad \text{’eat} (\uparrow \text{subj},(\uparrow \text{obj})) \\
\text{SUBJ} & \quad \text{PRED} \quad \text{’pro} \\
\text{OBJ} & \quad \text{PRED} \quad \text{’pro} \\
\text{TNS} & \quad \text{PAST}
\end{align*}
\]

Figure 7.3: F-structure: Who ate what?

The question is ambiguous, it can be answered in two ways (it is of course not ambiguous in a given context). The ambiguity is represented by the two i-structures. At the level of the c-structure, the ambiguity is reflected by the different i-structure annotations of the nodes of the question words in Figures (7.4) and (7.5). In these annotations, ∈ can be paraphrased as: "is an element of a set at i-structure”. \(\text{QuesP}\) refers to "question phrase" (see Dalrymple (2001)), which can be defined as the disjunction of other categories (in English):

\[
\text{QuesP} \equiv \{\text{NP}\mid \text{PP}\mid \text{AdvP}\mid \text{AP}\}
\]

Before moving on to the next section, the annotation \(\uparrow_{\alpha_i}\) needs to be clarified. This annotation is proposed by (Dalrymple and Nikolaeva, 2011, Chap. 4) and refers to the discourse function and semantic description of an element at the level of i-structure. The authors assume the following LFG architecture: Figure (7.6).

As this architecture indicates, the information structure projection is linked to the semantic projection via the mapping function \(\iota\). The basic assumption of this
7.2. Characterizing Multiple Questions in the I-structure

Figure 7.4: C-structure 1: Who ate what?

Figure 7.5: C-structure 2: Who ate what?
framework is that the meaning constructors of all the members of a clause are associated with a discourse function (information structure set), represented in the semantic description of their lexical entry. This way, the meaning constructors are categorized according to their information structure role. Importantly, all the f-structures in a clause correspond to the same information structure (in the following figures this is indicated by \[ \sigma_i = \sigma_i \]). The information about the particular i-structure role the meaning constructor takes on can come from various sources: syntactic position (in English, for instance, the Spec.IP is the default topic position), agreement, casemarking, word order, intonation, etc.

I have shown a possible integration of multiple questions into the i-structure architecture of mainstream LFG analyses. One should remark, however, that these assumptions implicitly adopt the view that question words are a subtype of focus, and argue that not all of them share the properties of foci, but that of topics. In the next subsection, I will present an alternative analysis claiming that question words cannot be collapsed into topics or foci.

### 7.3 I-structure: an Alternative Analysis

The architecture of the information structure, as proposed by Butt and King (1996), King (1997) and Choi (1999), contains topic and focus as i-structure primitives. There are a number of problems with this architecture, which are enumerated in this section. Then an alternative architecture is proposed, which is not fundamentally different from the one presented above, but it could be argued to capture the problematic facts better. As I see it, the main problem concerns the fact that the set of elements with different discourse, semantic and prosodic properties is larger than the above architecture could accommodate without simplifying these properties. Let us now present a list of these elements, familiar from previous chapters:

- Thematic shifters:
  I defined this type of topic as the element that links the sentence to the discourse by introducing a new subtopic of the discourse topic. We observed that in Catalan and in Hungarian, a thematic shifter is present in the sentence only...
if it does not continue the previous subtopic. Such sentence topics are typical in narrative contexts (see the formal account in the next section). From this it follows that there are a number of sentence types which do not contain sentence topics, for instance the ones that continue the previous subtopic in narrative contexts, some questions, and answers introducing complex strategies.

• Focus:
The observation I made was that focus appears only in replies and never in out of the blue utterances. Focus is the semantically/prosodically prominent part of answers to questions, corrections, contrastive and parallel structures. There are sentences without a (semantically/prosodically) focused constituent, for instance in narrative contexts. In Hungarian, "neutral" (or all-focus) sentences exhibit level prosody, where no element stands out carrying a pitch accent. The preverbal position is occupied by verbal modifiers or such lexical elements that form a prosodic and lexical unit with the verb (the verb criticizes on them).

• Contrastive Topic:
Contrastive topics are similar to foci, in that they do not appear in out of the blue utterances. Both Büring’s and Gyuris’s model express that contrastive topics carry the presupposition that there is a focus value (different from and not entailed by that of the sentence) associated with an alternative to the denotation of the contrastive topic. This also accounts for the fact that contrastive topics always co-occur with a focused constituent. Contrastive topics appear in answers to subquestions of the main question, linking the partial answers to the discourse topic (modeled as the Question under discussion). Contrastive topics can co-occur with sentence topics, in this case the sentence is linked both to a more general discourse topic and to a more restricted one:

\[(15) \text{ Hungarian:} \]

\[\text{[rJános] [cT a levest] megette(, de a [cT húst] [r nem]).} \]

John the soup.ACC VM.eat.PST but the meat.ACC not

As for the soup, John did eat it (, but he did not eat the meat).

Gyuris (2002, p. 23, 15)

In (15), the thematic shifter is János. The sentence contains a contrastive topic (a levest), which is implicitly or explicitly contrasted to a húst. In the two parallel clauses, the focus values are also different, since different contrastive topic values have to be mapped on different focus values (Gyuris, 2009). The different focus values are verum and falsum foci, respectively.

• Hocus:
The hocus is an argument or adjunct appearing in the preverbal position in neutral (or all-focus) sentences in Hungarian. It lacks the pitch accent and the contrastive-exclusive reading of focused constituents in non-neutral sentences (see Chapter 5).

It follows from the facts presented in Chapter 5 that the hocus cannot be analyzed as a subtype of focus, and thus it would be difficult to integrate it into Butt and King (1996)’s model of information structure.

Question phrases:

In the previous chapters, I discussed the semantic properties of question words, and their syntactic and semantic properties in French and Hungarian. Now we will consider if there is conclusive evidence to claim that question words and foci belong to the same category (since most analyses claim that question words are obligatorily focused). Let us now examine the main considerations of this argumentation:

– Syntax

It has been observed that question words and focused constituents often occupy the same syntactic position in various languages. This seems certainly the case in Hungarian, but not in French, in which focus is not associated with a syntactic position, but with certain prosodic patterns. It is commonly accepted that the preverbal position is a focus-position in Hungarian. Nevertheless, most analyses dealing with the syntax of Hungarian ignore the fact that it is not an exclusive focus position (it can host the hocus, question words, negative adverbs and monotone decreasing quantifiers, and verbal modifiers, just to mention a few), and focused constituents can appear in different positions in the structure as well (on the right periphery, or preverbally, preceding immediately preverbal question words). In addition, we have seen that the cumulation of question words is possible in the preverbal domain in Hungarian, whereas in the case of foci it is strictly forbidden.

Choi (1997) assigns the elements bearing different information structure values (new, prominent, etc.) into specific syntactic positions, showing that [+prominent] and [+new] elements tend to occupy salient syntactic positions. What counts as a salient syntactic position varies from language to language, but on a global scale the (left and right) periphery of the sentence and the preverbal position are argued to be salient. I believe that instead of associating a syntactic position with one particular discourse function (TopP, FocP), we should rather posit the existence of salient or prominent syntactic positions (in the spirit of Choi), and assume that the prominent elements of the information structure appear in those. In Hungarian, the left and the right periphery of the sentence, and the preverbal position qualify as prominent syntactic positions. In French, it is also the periphery of the sentence where prominent constituents appear; right dislocation, for instance, can serve to foreground
non topical material, placing the focus on the right periphery of the core clause:

(16) C’est très BIEN, ce que vous avez fait.
    it is very good what you AUX done
    What you have done, is very good.

Erteschik-Shir (1986) also argues that the

"correlation [of question words and foci occupying the same syntactic position] could follow from the semantic function of question words [i.e. that they are focused], but could equally well follow from the fact that the position of the question-word signals the locus of the answer and the answer (as every one would agree) must necessarily be focused" (p. 119).

The syntactic overlap of focused constituents and question words is thus clearly not complete cross-linguistically, not even within a single language where they are supposed to occupy the same position (Hungarian) (see also Gazdik and Winterstein (2011)).

Prosody
As far as prosody is concerned, I showed in Chapter (2) (based on Mycock (2006)) that in Hungarian, only the immediately preverbal question words have the same prosody (pitch accent) as the focus in the same position. Non-sequence-final question words are pronounced at a higher tone, different from the sequence-final one. This makes them similar to the intonation pattern of thematic shifters and not to foci. In French, the nuclear contour is anchored on the right edge of focused constituents, but on the left edge of question phrases (see Chapter (6)). What we see is that both focus and question words are prosodically highlighted but not necessarily in the same way.

Haida (2007) argues that in German, focusing differentiates between question words and the homonymous indefinite pronouns. Compare:

(17) Wer liest WAS?
    who reads what
    Who reads what?

(18) Wer LIEST was (= etwas)?
    who reads what
    Who reads something?

Haida thus equates prosodic prominence with focusing. As I see it, it is
obvious that when *was* is used as a question word, it has to be distin-
guished from the indefinite pronoun, especially since in a multiple ques-
tion it is not in a sentence-initial position. Pragmatically, it is also plau-
sible that an information gap (which expects an answer) is highlighted
compared to an indefinite pronoun, which, in turn, does not require any
identification from the interlocutor. However, prosodic highlighting does
not necessarily mean that question words are the same type of linguistic
objects as foci.

– Semantics

As we have seen, the common formal properties of focus and the question
words are reflected in their semantics as well. Both define (see Rooth
(1992), Chapter 2) a set of alternatives (that are subject to certain re-
strictions in the case of a congruent question-answer pair). Nevertheless,
this does not prove that interrogative words are a subclass of focus.
Eckardt (2007) observes that some question words can be focused (in
their metalinguistic use):

(19) Hungarian:

(Azt kérdeztem, hogy) MIVEL ment, (nem azt, (that asked, that) what.INSTR went.PST.3SG, (not that,
ho gy HOVA).
that where)

I asked HOW he went there, and not WHERE he went.

In addition, if we considered *wh*-words as a subclass of focus, we could
not distinguish (semantically) multiple questions, and interrogatives con-
taining a focused element:

(20) Hungarian:

JÁNOS mit evett?
János what eat

What did JOHN eat?

(21) Hungarian:

Ki mit evett?
who what eat

Who ate what?

Another problem is the treatment of polar questions. If question words
are supposed to be focused, what would be the focus in polar interrogatives? In French, interrogative verb forms (lexical pronominal affixes) or a complementizer est-ce que can distinguish a polar interrogative from a corresponding declarative, but it is also possible that the difference is only intonational (Beyssade, 2007):²

(22) a. Jean est-il venu ?
    Has John come?

b. Est-ce que Jean est venu ?
    INT.C Jean AUX he come
    Has John come?

c. Jean est venu ?
    Jean is come
    Has John come?

It seems inadequate to argue that the interrogative verb forms (or the lexical pronominal affixes) are the focus of the questions, all the more so that their presence is not even obligatory. In addition, as we have seen, partial polar interrogatives can indeed contain (prosodically/semantically) prominent elements that can be argued to be the focus in the question:

(23) Did John invite MARY? (Was it Mary that John invited?)

Based on these facts, Eckardt (2007) claims that interrogative words and the focus introduce alternatives at different levels (semantic vs. pragmatic). For the present purposes, it is enough to retain that although the interrogative words and the focus have common formal and semantic properties, they do not necessarily belong to the same category. The phenomenon might also be explained by supposing that they play similar roles in the discourse: they constitute the prominent elements in question-answer pairs.

Approaching from the pragmatic side, it is also unclear how question words can introduce new information (which was supposed to be the role of focus in certain approaches). Nevertheless, the inadequacy of this claim concerning focus has already been demonstrated.

Erteschik-Shir (1986) mentions this problem as well:

²The difference between the different types of interrogatives is also present in polar interrogatives. Marandin (2004) shows that only (22-c) is compatible with the question tag n’est-ce pas (John has come, hasn’t he?) indicating that the event denoted by the question is presupposed. On the other hand, an NP subject (Jean) is more natural in (22-a) and (22-b).
"[w]hy is it then that some linguists believe that wh-phrases do function as focus or new information? The main reason seems to be a confusion between the function of the wh-phrase in the question and the function of the constituent which replaces it in the answer.” (p. 119)

In her paper, Erteschik-Shir argues against the analysis of (non-echo) wh-phrases as foci, showing, with the help of the notion of Dominance that the wh-word is not the most dominant part of questions. Informally speaking, Mycock (2006) also argued that

"[i]mpressionistically, a question word is focused in order to highlight the information gap which it represents. Such a lack of a speaker knowledge must ultimately be regarded as semantic rather than syntactic or functional in nature.” (p. 90.) (repeated from Chapter (2))

Mycock thus assumes that foci and question words share the same type of highlighting, referring to a lack of knowledge of the speaker. In this sense, two terms are confused: focus and focusing. Focus, defined the way as I do in this thesis, as the distinguished element of replies, is a highlighted, i.e. focused constituent. However, other elements can also be focused, for instance question words, according to Mycock. Assume that we use highlighting instead of focusing, in order to avoid terminological confusion. Highlighting is a process, as a result of which one element stands out of the others, either formally (prosody or salient syntactic position), or semantically (in the information structure). Prominent elements (in the information structure) are prominent because they are specially affected by the illocutionary operator (Jacobs, 1984), and they are formally (prosodically or syntactically) highlighted as well. Focus (the distinguished element of replies) is highlighted, just like question words, sometimes by the same means, sometimes differently (after all, syntax and prosody have a finite and restricted number of ways of highlighting). Crucially, highlighting does not mean at the same time that question words and foci are the same. Semantically, question words are place-holders indicating a gap, restricting the set of elements that can potentially fill that gap. Focus, on the other hand, (in answers to questions) denotes the element that fills the gap represented by the question word (as such, it has to be part of its restriction). In addition, the two terms are used at different levels of linguistic analysis: the focus is an information structure category that can include many types of elements, whereas the question word status is a lexical unit. Their formal similarities might be due to this parallel in their use. In some languages, focus is placed in the same syntactic position as the corresponding question
word in order to facilitate the recognition of the semantic correspondence (this is the element that fills the gap). Nevertheless, these similarities come from their use (question-answer pairs, highlighting), but they do not entail that they are the same. This is why, instead of the focusing of question words, I prefer to use the highlighting of question words.

- **Complettive information:**

  *Butt and King (1996)* argue (when analyzing the information structure of Hindi/Urdu) that there are elements which are new (not inferable from the discourse-context), but do not appear in the same salient syntactic position as the focus. They call these (from the point of view of the information structure) **Complettive information** and argue that these can be defined by the following combination of features: [new +], [prominent −].

  To sum up, the representation of all the variety of different elements enumerated above in an information structure, which explicitly contains three of them as its primitives (topic, focus and complettive information) seems to be a difficult task. Contrastive topics are different from sentence topics, question words are different from foci, although they share some properties.

  There are three possible ways to solve this problem. Firstly, we can employ *Butt and King (1996)*’s labels (TOPIC, FOCUS, BACKGROUND INFORMATION, COMPLETIVE INFORMATION), with a loose semantic interpretation. Belonging to the topic set, in this case, would mean that an element links the sentence to a discourse topic by introducing a subtopic (which can mean the answer to a subquestion), covering both thematic shifters and contrastive topics. Belonging to the focus set would mean that the element is the highlighted and distinguished constituent of the sentence, covering foci, question words and the hocus. The interpretation of complettive and background information is, in this respect, less problematic. The exact difference between the different types of elements (question words - foci, contrastive topic - sentence topic) would follow from two things: from the semantic description (meaning constructor) of the individual elements included in the i-structure, and from the role the sentence plays in the discourse (question-answer pair, correction, narration, etc., see the next section). This means, for instance, that the element in the focus set would have a different semantic content depending on the role of the sentence in the discourse structure, *i.e.* if it is a question, or an answer. The analysis of multiple questions I presented in the previous section follows this solution. However, why should the sets in the i-structure have exactly these labels (which prove to be only labels) if the semantic content of the elements in them can be different?

  A second solution (László Kálmán, p.c.) would suppose that there is no general i-structure that would suit all sentence types. This means that the i-structure has many different architectures, depending on the discourse-context and the particular sentence. One would be {CONTRASTIVE TOPIC, FOCUS, BACKGROUND INFORMATION}, others would include {THEMATIC SHIFTER,
HOCUS, BACKGROUND INFORMATION}, \{QUESTION, BACKGROUND INFORMATION\}, etc. However, this way the common properties of some groups of elements would not be represented.

The third solution would emphasize the common properties of the different discourse functions. Thus a set would include elements based on a common property, without claiming that these elements must be semantically and discourse-wise identical. The exact semantic and discourse properties would follow, as said above, from the meaning constructors of the individual elements and the discourse structure the sentence appears in. In what follows, I present the proposed i-structure architecture. It keeps some aspects of Choi (1997)'s features, but also deviates from it in others. First of all, we have seen that certain elements are semantically prominent and formally (syntactically or prosodically) highlighted. I will refer to these elements as \(+\text{PROMINENT}\), and to the others as \(-\text{PROMINENT}\). Semantic prominence can be defined, based on Jacobs (1984), with respect to the illocutionary operator associated with the sentence. Prominent elements are the ones specially affected by the illocutionary operator. These elements are different in reactive (focus, contrastive topic) and out of the blue sentences (thematic shifter, hocus, question words), but constitute the prominent set at i-structure. This distinction defines two sets in the i-structure. Furthermore, we saw that among prominent elements we find such that link the sentence to the discourse (by introducing a subtopic of the discourse topic or reshaping the discourse topic), and others which do not. I call the first set D-LINKED, and the second \(\neg\text{D-LINKED}\). This way, I stay neutral with respect to the new status of focus. We have seen that focus is not necessarily new information (in the sense of introducing a new discourse referent). In the \(-\text{PROMINENT}\) set, we also find a D-LINKED and a \(\neg\text{D-LINKED}\) subset, the first corresponding to background, the second to completive information. The proposed architecture hosts the above mentioned elements as shown in Figure (7.7).\(^3\)

\[
\begin{array}{c}
\text{+PROM} \\
\begin{cases}
\neg \text{D-LINKED} & \{\text{FOCUS, QW, HOCUS}\} \\
\text{D-LINKED} & \{\text{THEMATIC SHIFTER, CONTRASTIVE TOPIC, QW}\}
\end{cases} \\
\text{-PROM} \\
\begin{cases}
\neg \text{D-LINKED} & \{\text{COMPLETIVE INFORMATION}\} \\
\text{D-LINKED} & \{\text{BACKGROUND INFORMATION}\}
\end{cases}
\end{array}
\]

Figure 7.7: Proposed i-structure

As can be seen in (7.7), interrogative phrases (QW) can appear both in the D-linked and the non-D-linked subpart of the \(+\text{PROMINENT}\) part. Obviously, this corresponds to the different status of the question words in pair-list questions, in which at least one (the Sorting key) is D-linked. Some question words (which-phrases) are inherently D-linked. This can be specified in their lexical entries in LFG (the abbreviation \(\uparrow \text{VAR} \Rightarrow x\) refers to the fact that question words introduce a variable, the Q attribute refers to a grammatical feature with the possible values

\(^3\)QW stands for Question Word, without specifying its discourse function as topic or focus.
7.3. I-structure: an Alternative Analysis

$+/−$, encoding interrogativity):

\[(24) \quad \text{which: det } (↑\text{PRED}) = 'pro' \]
\[(↑\text{VAR})= x \]
\[(↑\text{Q})= + \]
\[(↑σ \in + \text{PROM D-LINKED}) \]

The different answerhood conditions are related to i-structure differences. Since the meaning of both pair-list questions and answers with contrastive topics is analyzed as a function, I assume that the prerequisite of a question requiring a pair-list answer is at least one non-D-linked question word, the denotation of which is then mapped onto that of the other (in the non-D-linked set). Apart from this type (pair-list answer, one D-linked, one non-D-linked question word), it is possible that both (all) question words belong to the same set (D-linked or non-D-linked). The possible answerhood/ i-structure correspondences are the following:

- Pair-list answer: one D-linked and one non-D-linked question word (Figure (7.8))

\[(25) \quad \begin{align*}
\text{a. Who ordered what in the restaurant?} \\
\text{b. John ordered chicken, Mary turkey, Peter spaghetti…}
\end{align*} \]

\[
\begin{array}{c}
\text{+PROM} \\
\neg \text{D-LINKED } \left\{y \mid (y) \in \{(\text{EATABLE} \text{ THINGS})\}\right\}
\end{array}
\begin{array}{c}
\text{D-LINKED } \left\{x \mid (x) \in \{\text{PEOPLE}\}\right\}
\end{array}
\begin{array}{c}
\neg \text{D-LINKED } \left\{\lambda x.\lambda y.\text{ordered} (x, y)\right\}
\end{array}
\begin{array}{c}
\text{D-LINKED } \left\{\text{in the restaurant}\right\}
\end{array}
\]

Figure 7.8: I-structure: Who ordered what?

A few remarks are due here concerning this representation. Concerning the f-, and i-structure representation of question words, question words are pronouns, thus their semantic content is represented as 'pro' in the f-structure (since they do not contribute an independent referent). Mycock (2006) argues (based on Ginzburg and Sag (2000)’s analysis) that question words also introduce a parameter, whose value is a variable and represents this also at the level of f-structure. However, this variable is a semantic contribution, and, as such, it can be represented only at the i-structure level, containing semantic information as well. I assume that the variable, which is part of the lexical entry of question words, represents the question word at the level of i-structure, along with its restriction (human, non-human, etc.). The semantic representation values are based on Krifka (2001), and expressed in basic lambda calculus terms, keeping in mind that it can be transposed into the glue language, more commonly used in LFG. Items are either represented by their semantic representation, or, when their detailed semantic descrip-
The above mentioned discourse functions (focus, contrastive topic, etc.) do not appear as attributes in the structure. As shown above, since they appear in different contexts, they can be defined in semantic and discursive terms, i.e. they follow from a discourse pattern and from an i-structure role. In turn, this necessitates the formalization of discourse, going beyond the domain of a single sentence.

Another issue pointing into this direction is the interpretation of pair-list questions. Following Krifka (2001), I assume that the meaning of the multiple question in (7.8) is a set of functions $f$ such that every $x$ in the domain of $f$ (people) stands in the relation ordered to $f(x)=y$ (eatable things), which is the range of $f$:

$$\lambda f \forall x \ x \in \{\text{PEOPLE}\} \rightarrow \text{ordered} \ (f(x)), \ (f(x)) \in \{\text{EATABLE} \ 	ext{THINGS}\}$$

This formula contains more than the content of the elements in the question. It formalizes the presupposition that every element of the domain set of the function will be part of one answer in the answer list. In order to account for this, the list of answers would also have to be represented. This representation assumes an ambiguity among the wh-phrases at least in English in that it is not necessarily clear from the syntactic representation which of them is mapped onto the other (this is, however, disambiguated by the information structure roles). We turn to the possible representation of discourse structure in the next section.

- Pair-list answer: matching question (both (all) question words D-linked) (Figure (7.9))

(a) Which student read which book?
(b) John read War and Peace, Mary read Wuthering Heights, Peter read Ulysses...

![Figure 7.9: I-structure: Which student read which book?](image)

The semantic descriptions in Figure (7.9) are not different from Figure (7.8),
and the denotation of the linearly first question word is supposed to be mapped onto that of the second in the same way:

\[
\lambda f \forall x \in \{\text{STUDENTS}\} \rightarrow \text{read } (f(x)), (f(x)) \in \{\text{BOOKS}\}
\]

The formula indicates, at the same time, that the linearly first question word is the *sorting key* of the question. The difference between the two questions is that in Figure (7.9) both question words are D-linked. This means that the sets denoted are both salient in the context and the task of the answer is not their identification, but matching each student with a book. The function is mostly (but not necessarily) bijective (there is one different book for each student, but not more than there are students); nevertheless, it is also possible that more values are mapped onto one (more than one student read the same book). We should note, however, that in French matching questions, it is slightly more natural to start the question with the subject, even if it is not the *sorting key* of the question. In these cases, the context disambiguates the question.

- **Single pair answer:** both (all) question words non-D-linked (Figure (7.10)). This case is demonstrated on a Hungarian example, in which the question words arguably occur in the same clause. In English, they would be analyzed as conjoined questions, the question words appearing in separate clauses.

\[
\begin{align*}
\text{(29) Hungarian} \\
\text{Ki és mikor ment moziba?} \\
\text{who and when go.pst cinema.}
\end{align*}
\]

Who went to the cinema and when?

\[
\begin{align*}
\text{Figure 7.10: I-structure: Ki és mikor ment moziba?}
\end{align*}
\]

In Figure (7.10), both question words appear in the non-D-linked set, they are represented by a variable they introduce. No function interpretation occurs, since it is assumed that in order for such an interpretation to occur at least one of the question words has to be D-linked. Compare:

- **Single pair answer:** directional question; both question words D-linked (Figure
(7.11)

(30) Who fell in love with whom? (John with Mary, or Mary with John)

\[
\begin{align*}
& \text{+PROM} \quad \text{\~D-LINKED } \{\} \\
& \text{D-LINKED } \{ \begin{array}{l}
\forall x, x \in \{\text{Mary, John}\}, \text{fell in love with } (y), \\
\forall y, y \in \{\text{Mary, John}\}, \not= (y)
\end{array} \} \\
& \text{\~PROM} \quad \text{\~D-LINKED } \{\} \\
& \text{D-LINKED } \{ \begin{array}{l}
\forall x, x \in \{\text{Mary, John}\}, \text{fell in love with } (y), \\
\forall y, y \in \{\text{Mary, John}\}, \not= (y)
\end{array} \}
\end{align*}
\]

Figure 7.11: I-structure: Who fell in love with whom?

In this case, it is important that the question words denote the same type of set, otherwise the relation would not be reversible, and the question could not refer to the direction of the relation. Both question words are D-linked, and the mapping between them is bidirectional.

In the next chapters, I will apply the proposed analysis to the French and Hungarian data. However, before going on to the analysis of these languages, I turn to the problem of discourse structure in the next section.

7.4 Discourse Structure Revisited

7.4.1 A separate level

The domain of information structure, as was shown above, is the sentence. However, we have also seen various phenomena, like the adequate formalization of pair-list answers, which cannot be accounted for with reference to a single sentence, only taking into consideration the previous sentences or the immediate discourse context. Let us illustrate this with some examples, which is by no means an exhaustive list of the arguments for the necessity of discourse structure as a level of analysis.

- The definition of topics

When discussing information structure, and the relevant aspects of Hungarian and French (Chapters (4)-(6)), we reached the conclusion that (sentence) topic (thematic shifter or contrastive topic) and discourse topic are related notions. The former usually appear on the left periphery of the Hungarian sentence if the previous subtopic of the discourse topic is not continued by the sentence. Thus, either a new subtopic is introduced, or the discourse topic is altered. The same can be observed in Catalan. Based on Vallduví (1992), Erteschik-Shir (2007) argues that the left dislocated/topicalized constituent in Catalan (that he calls the Link) cannot be a continued topic, only a shifted one, i.e. only elements that are not topic in the previous sentence can be (syntactically) topicalized. In French, left and right dislocated
elements also qualify as (semantic/pragmatic) topics, carrying discourse-old information (although this is not always the case; see Laurens (2010)). But what does discourse-old mean? An informal answer would be to claim that information that is inferable from the discourse topic or relates to an issue already touched upon in the preceding discourse is discourse-old. Consequently, whenever one wants to define (sentence) topics, the discourse topic and the discourse structure also have to be taken into consideration.

• The definition of foci

The same goes for foci. I defined foci in terms of the distinguished part of replies to previous sentences (question-answer pairs, corrections, etc.), and it goes without saying that the logico-semantic relationship of the sentence with the preceding one is crucial. This is clearly reflected by the fact that in analyses dealing with foci, focusing is usually illustrated by question-answer mini-dialogues, since it is always in a context of more than one sentence that the exact semantic description of focus can be provided. Focus can be identified in a single sentence as well, but whether it is contrastive, or information focus, etc. crucially depends on the discourse. Focus is sometimes also defined as new information. The same can be said about this as about the discourse-old status of topics. The information can only be new with respect to the discourse topic or the preceding discourse context. The importance of the preceding sentences with respect to focusing is especially relevant when it has syntactic/grammatical consequences. For instance, in Hungarian, the co-occurrence of a (prosodically/semantically) focused constituent and a question word (in the preverbal position) is only possible if the question is asked for the second time. In this case, it is not the second-occurrence question word that carries the pitch accent, but the focused constituent, which, in turn, cannot occupy the prominent preverbal "focus" position because of the question word. Consider again the following example, familiar from previous chapters:

(31) Hungarian:

a. Q1: -Ki mit evett a bulin?
   who what.pst.3sg the.party.superess
   Who ate what at the party?

b. A: -János pizzát, István spagettit...
   John pizza.acc Stephen spaghetti.acc
   John ate pizza, Steve spaghetti...

c. Q2: -Jó, de BALÁZS mit evett?
   Ok, but BALÁZS what.pst.3sg
   Ok, but what about BALÁZS, what did he eat?

• Questions and contrastive topics
It is usually assumed about sentences containing contrastive topics that they cannot be uttered out of the blue. Sentences containing contrastive topics are partial answers to questions, which is modeled by Büring (2003)’s discourse tree indicating the implicit subquestions and the other potential answers that the answerer is either not aware of or does not want to share with the asker. The use of a contrastive topic cannot be thus understood without making reference to the discourse topic (an initial question) and the hierarchical discourse structure consisting of explicit and implicit moves corresponding to individual sentences.

- Anaphoric binding
  Another set of phenomena relates to anaphora and identifying the antecedent of pronouns. Consider the following example:

(32) John pushed Bill. He fell.

For all speakers of English it is obvious that the pronoun he in the second sentence is coreferential with Bill. The explanation of the coreference comes from world knowledge, i.e. it seems obvious that the one pushed is more likely to fall than the pusher. Without the first sentence, the reference of the pronoun cannot be decided. This is another domain which can be better analyzed if discourse is taken into consideration.

- I-structure
  We have also seen above that the subtle semantic differences between the various elements in the same set at i-structure come from their semantic descriptions, but also from the relation of the sentence with the preceding (or following) ones. In order to illustrate this relation, the i-structures of the individual sentences should be included in a larger structure.

  Discourse structure as a separate level has already been proposed in the LFG framework, although unlike i-structure, it is not yet part of the core architecture. In this thesis, I argue for the existence of such a level of representation and make an attempt to develop some formal aspects of it.

  King and Zaenen (2004) propose to represent discourse-structure as a tree diagram, in which the nodes stand for the Basic Discourse Units (these can be complete sentences or parts of sentences). They also identify some basic relations in which sentences can stand with each other (based on Asher and Lascarides (2003)). One of these is “Subordination”, for example Elaboration, in which a sentence elaborates on the preceding one (for instance, by providing explanation):

(33) Bill fell. John pushed him.

The other relation is “Coordination”, for instance Progression, which is illustrated by (32).

\[4\] Some conjunctions, like because can explicitly indicate the type
of relation, otherwise it has to be deduced from the context. Nevertheless, conjunctions and temporal relations (adverbials, verb tenses, verbal aspect) can help identify a particular relation. These relations are, then, in most cases, identified between clauses (as basic units) and not between complete sentences. The two types of relations in King and Zaenen (2004) reflect the observation that the discourse structure can be modified by a new discourse unit in two ways: either it pushes the structure from left to right, or downwards from the top (these assumptions go back to Polanyi (1985)).

King and Zaenen illustrate this with the following example:

(34)  
(a) The man dove into the pool.  
(b) It was warm and soothing and  
(c) he decided to remain for a little longer than usual.

(34-a) et (34-b) are in a subordinating relation, (34-b) elaborates on (34-a). On the other hand, both (34-a) and (34-b) are in a coordinating relation with (34-c). We illustrate the d-trees later on in this chapter.

On the other hand, O’Connor (2005) proposes an avm representation for the discourse structure (with a similar architecture to the above mentioned information structure), arguing that information structure is the mapping between discourse structure and the other components of the grammar. Thus it does not have to be represented at a separate level.

In this thesis, I follow King and Zaenen (2004) in representing d-structure as a discourse tree. I assume, furthermore, that i-structure is not a simple mapping function between d-structure and the other levels of representation, but it is a level of representation on its own right, with its above defined primitives. Nevertheless, since discourse structure and information structure are closely related, I propose that the nodes in the discourse tree are actually the information structures of the individual sentences or clauses. Since according to Dalrymple (2010), semantic information is also integrated into the information structure, the discourse tree, with the i-structures as its nodes, would contain the meaning (semantics and pragmatics) of an utterance, as opposed to the form, represented by the c-, the f-, the m-, and the p-structures.

### 7.4.2 Discourse relations

The complex d-structure is thus an amalgamation of previous LFG analyses (King, 1997; King and Zaenen, 2004; Dalrymple, 2010), and aims to integrate other approaches as well. The first such approach is Büring’s analysis of contrastive topics as a complex strategy, represented by a discourse tree (Figure 7.12):
How was the concert?

Was the sound good? No, it was awful.
How was the audience? They were enthusiastic.
How was the band?

How was the drummer? Just fantastic.
And what about the singer? Better than ever.
Did they play old songs? Not a single one.
So what did you do after the concert?

Figure 7.12: A Discourse-tree

Nevertheless, the strategy introduced by contrastive topics is only one specific type of the possible relations that can exist between sentences. The second approach that one has to integrate in this framework is a formal account of these possible relations, which we can find in Asher and Lascarides (2003); Lascarides and Asher (2007)’ Segmented Discourse Representation Theory. For the formal account, the reader is referred to these works; its complete presentation would be beyond the scope of this thesis. What I adopt, nevertheless, is the ontology of these relations and offer the first steps of a possible representation in LFG.

Asher (1993); Asher and Lascarides (2003) use rhetorical (discourse) relations to model the semantics/pragmatics interface. They build their analysis on two approaches. First, they consider that dynamic semantics (Kamp, 1981) provide better means of describing the meaning of a sentence in context (they consider it as a relation, called Context Changing Potential (CCP)) than the Montagovian or Davidsonian frameworks (which can only be extended to multi-sentence discourses with certain shortcomings) (see also Heim (1982)). The second approach is Artificial Intelligence (AI) research (Hobbs, 1985), which emphasizes the role of commonsense reasoning with non-linguistic information (commonsense reasoning can play an important role in identifying the antecedent of pronouns, as we have seen above). Integrating dynamic semantics with AI-approaches, Lascarides and Asher (2007) propose a framework called Segmented Discourse Representation Theory (SDRT). I consider that the SDRT is compatible with a representation of d-structure in the LFG framework, although further research is needed to work out the exact
formalism.

Asher (2004); Lascarides and Asher (2007) also assume Coordinating and Subordinating relations and give a detailed ontology of the two types. Note, however, that Asher and Vieu (2005) argue that some discourse relations cannot be classified as coordinating or subordinating, since this does not belong to their intrinsic semantics, but they can be used as coordinating or subordinating in different contexts. Let us now present the basic ontology of discourse relations (the examples are based on Asher and Lascarides (2003)):

- Coordination relations

  - Narration: spatial or temporal progression (then)

    (36)  
    a. Max came in the room.
    b. He sat down.
    c. He lit a cigarette.

  - Background

    (37)  
    a. Max opened the door.
    b. The room was pitch dark.

  - Result

    (38)  
    John pushed Max. He fell.

  - Continuation: the further elaboration of some linguistically explicit or implicit topic (unlike narration, it lacks its spatio-temporal consequences)

    (39)  
    a. The teacher asked the students to look for the lost cat.
    b. John looked under the table.
    c. Mary looked in the garden.

  - Parallel

    (40)  
    a. John loves sport.
    b. Bill loves sport too.

  - Contrast (but)

    (41)  
    John loves sport, but he hates football.

  - Alternation

    (42)  
    Either there is no bathroom, or this is a funny place.

  - Dialogue:
* Correction

(43)  a. John distributed the copies.
b. No, it was Sue who distributed the copies.

• Subordinating relations

– Elaboration (Subordination): description in more detail (see (34))
– Topic: a common discourse topic in a narration (it can be implicit)
– Explanation: elaboration by providing a cause (because)

(44) Max had a lovely evening. He ate a lot of salmon.

– Consequence (if)

(45) Suppose there’s a bathroom/ Then it’s a funny place.

– Dialogue:

* Question Elaboration: answering with a question

(46)  a. Can we meet next weekend?
b. How about Saturday?

* Direct Question-Answer Pair (see (50))

* Indirect Question-Answer Pair: providing enough information to infer an answer

French:

(47) a. Q: Qui a été invité?
    who has been invited
    Who was invited?
    Indirect A:
b. La POSTe est en grève.
    the post is on strike
    The post is on strike.

c. Ok, on annule.
   ok, we cancel
   Ok, we cancel it.

(Asher, 2004): 169, 11

* Partial Question-Answer Pair (referred to as Partial Implicative Question-Answer Pair in the present thesis)

(48)  a. Q: Who’s coming to the party?
b. A: Well, I know Mary isn’t coming.
Based on Chapter (4), I add to this later group *Complete Implicative Question-Answer Pair*, in order to account for the complete answers given to pair-list multiple questions. Recall that in these discourse structures, implicit subquestions are obligatorily present and all of them gets obligatorily answered (unlike partial answers).

Lascarides and Asher (2007) illustrate the *Narration* and *Elaboration* with the following text:

(49) a. John had a great evening last night.
    b. He had a great meal.
    c. He ate salmon.
    d. He devoured lots of cheese.
    e. He won a dancing competition.
    f. ??It was beautiful pink.

The proposed discourse tree is shown in Figure (7.13).

![D-tree](image)

Figure 7.13: D-tree: John had a great evening last night...

The discourse tree reveals why the anaphora is problematic between the pronoun *it* in the last sentence and *salmon*. This is called the *right-frontier constraint* which says that anaphora in the current clause must be bound to an antecedent which is on the right frontier of the structure.

Since my main concern in this thesis is the analysis of multip le questions, I will concentrate on the rhetoric relations that are relevant in this respect. In Chapter (2), I identified (based on Marandin et al. (2009)) three relations that can exist between a question and an answer. Firstly, an answer can be *congruent* or *non-congruent*. This is a purely semantic relation between a question and a potential answer. According to Krifka (2001, p. 9), in a congruent question-answer pair "the background of the question and the answer must be the same, and the focus of the answer must be an element of the restriction of the question". I believe that this approach can easily be integrated into the LFG framework, in which the information structures of the question and the answer appear in the same discourse tree and the focus of the answer must be an element of the restriction of the preceding question.
(Figure (7.14)):

(50) a. Who did Mary see?
    b. (She saw) John.

\[
\begin{align*}
\text{Question} & \\
+\text{PROM} & \neg \text{D-LINKED } \{ x \mid (x) \in \text{PEOPLE} \} \\
\text{D-LINKED} & \{ \} \\
-\text{PROM} & \neg \text{D-LINKED } \{ \} \\
\text{D-LINKED} & \{ \lambda x. \text{MARY saw } (x) \} \\
\hline
\text{Direct Q-A Pair} & \\
+\text{PROM} & \neg \text{D-LINKED } \{ '\text{JOHN}' \in \text{PEOPLE} \} \\
\text{D-LINKED} & \{ \} \\
-\text{PROM} & \neg \text{D-LINKED } \{ \} \\
\text{D-LINKED} & \{ \text{MARY saw } (\text{JOHN}) \} \\
\end{align*}
\]

Figure 7.14: D-tree: Who did Mary see? - John.

Secondly, an answer can be \textit{implicative} or \textit{non-implicative}. An implicative answer (which can be partial or complete) introduces a complex discourse strategy (Büring, 2003), indicated by a contrastive topic (B-accent in English, C-accent in French, rising tone on the element on the left periphery in Hungarian). The implica-
tiveness of an answer is independent of congruence (which is a semantic relation). According to Marandin et al. (2009) etc, implicative answers extend the d-tree hor-
izontally, by adding (implicit) subquestions to the origin al question.\footnote{Note that in an implicative answer the contrastive topic (in (51) ANNA) is usually distinct from the focus (the answer to the question, \textit{French} in (6)). However, in certain cases (see (48)), a constituent (\textit{Mary}) can at the same time answer the question and reshape the discourse topic. Formally, such constituents are contrastive topics. The same can be observed in French (see example (81-b)) in Chapter (6), in which a C-accent appears within a focused constituent). This supports the claim that contrastive topic and focus are not completely disjunct notions.}

The d-tree is schematically represented in Figure (7.15). The letters (A, B, C \ldots) stand for the information structures given below.

(51) a. What did the pupils learn?
    b. /Anna learned FRENCH.
Figure 7.15: D-tree: What did the pupils learn? Anna learnt French.
The bracketed i-structures refer to implicit questions that are not actually pronounced. However, the semantic definition of contrastive topics, as we have seen above, requires the presence of at least one such alternative in the context. Apart from the context of implicative (and indirect) answers, the only context in which implicit i-structures are permitted is that of the discourse topic in narrative contexts.

Thirdly, an answer can be direct or indirect. In an indirect answer, the path between the question and the answer includes supplementary nodes, sometimes new (implicit) questions that have to be accommodated. The authors show that the direct-indirect dichotomy affect the d-tree vertically, i.e. an indirect answer.
extends the d-tree in the vertical dimension. An indirect question-answer pair is shown in the following example:

(52) French
   a. Est-ce que Marie a réussie au bac?
      is it that Marie has passed the school-leaving exam
      Did Mary pass the school-leaving exam?
   b. Tous mes élèves ont réussi cette année.
      all my pupils have passed this year
      All my pupils passed this year.

(Marandin et al., 2009, p. 6, 1)

Marandin et al. (2009) illustrate the example with the following discourse tree (Figure (7.16)), introducing an implicit superquestion, and intermediary nodes:
Figure 7.16: Est-ce que Marie a réussi au bac? - Tous mes élèves ont réussi.
The d-tree supposes a more general, implicit question *Qui a réussi au bac ?* (Who passed the final exam?). Since the answer to all the subquestions (concerning all the individual students) is *yes* (John passed, Pierre passed, Marie passed, George passed...), the answer to the general implicit question (*Tous mes étudiants ont passé cette année.*) implies that Marie also passed, since it is also inferred that Marie belongs to the group of students of the answerer.

Based on these facts, I believe that the different answering strategies correspond to different discourse relations. All these strategies are subordinating relations, since they elaborate on an original question, by answering it in a way. The extension of the discourse tree thus does not correspond to the subordinating/coordinating nature of the relation, since implicative answers introduce subquestions of one single question. The fact that there is more than one subquestion accounts for the fact that the d-tree is also horizontally extended.

The relevant discourse relations are thus the following:

- Direct Question-Answer Pair
- Indirect Question-Answer Pair
- (Partial/Complete) Implicative Question-Answer Pair

Before going on to the further characterization of discourse relations, let us go back to an example presented in Chapter (3). As was shown in that chapter, multiple questions embedded in a question can be ambiguous between the following two readings:

(53)  
  a. Who remembers where Mary keeps which book?
  
  b. Bill remembers, where Mary keeps which book.
  
  c. Joe remembers where Mary keeps Aspects, Max remembers where Mary keeps Syntactic Structures.

*Ginzburg and Sag (2000)* give a compelling account for this example in the HPSG framework, and it would be interesting to see how the present analysis could account for this ambiguity. As a tentative analysis, I would like to claim that (53-a) has one $f$-, but two possible $i$-structures, which is the source of the ambiguity and the two types of answer. These structures are presented below: Figures (7.17)-(7.19).

In (7.18), only *who* counts as a true question word, introducing a variable, the embedded clause is background information. In (7.19), however, *which book* is interpreted as a D-linked question word, which is then mapped onto the set denoted by *who*, making the functional reading available the way illustrated above, summarized by the following formula:

(54)  \[ \lambda f \forall y, (y) \in \{\text{BOOKS}\}, f(y) \in \{\text{PEOPLE}\} \rightarrow \text{remembers where Mary keeps } (y) \]
Figure 7.17: F-structure: Who remembers where Mary keeps which book?

Figure 7.18: I-structure 1

Figure 7.19: I-structure 2
What remains to explain is the status of \textit{where}, \textit{i.e.} the fact that it is not available for the functional interpretation. A conclusion to be drawn from this example is that the basic unit of information structure is arguably the sentence and not the clause (otherwise the functional interpretation in (53-c) could not be derived from the information structure representation.)

Let us now examine the role of the above mentioned elements in this framework (focus, hocus, thematic shifter, etc.). As was pointed out above, they are considered as derived and not as basic categories which follow from the discourse relation the sentence is part of, their information structure role, and from their semantic description.

- Topics (Thematic shifters and Contrastive topics)

As we can see in the above ontology, Asher and Vieu (2005) consider that \textit{topic} is a discourse relation. This assumption corresponds to our definition of \textit{topics}, in that it is the element that relates the sentence to the discourse topic by introducing a subtopic of it, in connection with which the sentence can contribute new information. Asher and Vieu (2005) note that

”[s]ome coordinating relations require a topic; \textit{i.e.}, there must be a simple constituent, a common ”topic”, that summarizes the two related constituents and that is linked to them via the subordinating \textit{Topic} relation. If this third constituent has not been explicitly given in the previous discourse, it must be ”constructed”.” (p. 594-595)

In other words, \textit{topic}, as a discourse relation links coordinated constituents to the discourse topic. This happens typically in narration. I argue, in addition, that the link can be manifest if there is a thematic shifter in a sentence. Discourses that exhibit a strategy (for instance, \textit{Implicative answer}) contain a different type of topic: contrastive topics.

We can conclude, thus that both thematic shifters and contrastive topics are parts of subordinating discourse relations. At the same time, they both extend the discourse horizontally. Contrastive topics indicate a complex discourse strategy (for instance that the main question is divided into subquestions), whereas thematic shifters extend narrations, by integrating the new discourse units into the discourse, linking them to the discourse topic. Although both are +prominent and D-linked, they usually appear in different discourse constellations: contrastive topics presuppose the existence of at least one other contrastive topic in the discourse, which is associated with another focus value. The (at least) two different clauses are part of the strategy indicated by the contrastive topic.

- Focus
Focus, on the other hand, is not a discourse relation in Asher and Vieu (2005). Focus is the distinguished elements of replies in certain discourse relations, like Direct Question-Answer Pair (50), Implicative Question Answer Pair (51), or Indirect Question-Answer Pair, (52) Parallel (55), Contrast (56), Correction (57). The first three have been shown above. Let us now consider examples for the others along with the discourse trees:

- Parallel

  Hungarian:

  (55) ANNA tanult a gimnáziumban OROSZUL és Anna learned the high school.INESS Russian.ESS-FORM and PÉTER FRANCIÁUL.
  Peter French.ESS-FORM
  It was Anna who learned RUSSIAN at high school and Peter FRENCH.

  \[
  \begin{align*}
  &+\text{PROM} \quad \neg \text{D-LINKED} \\
  &\quad \{ '\text{Anna}', '\text{oroszul}' \} \\
  &\text{D-LINKED} \\
  &\quad \{ \} \\
  &-\text{PROM} \quad \neg \text{D-LINKED} \\
  &\quad \{ '\text{tanult a gimnáziumban}' \} \\
  &\text{D-LINKED} \\
  &\quad \{ \}
  \end{align*}
  \]
  \hspace{1cm} \text{Parallel}

  \[
  \begin{align*}
  &+\text{PROM} \quad \neg \text{D-LINKED} \\
  &\quad \{ '\text{Péter}', '\text{franciául}' \} \\
  &\text{D-LINKED} \\
  &\quad \{ \} \\
  &-\text{PROM} \quad \neg \text{D-LINKED} \\
  &\quad \{ '\text{tanult a gimnáziumban}' \} \\
  &\text{D-LINKED} \\
  &\quad \{ \}
  \end{align*}
  \]
  
  Figure 7.20: D-tree: Anna tanult a gimnáziumban oroszul és Péter franciául.

  This example presents two states of affairs, which are parallel and partly identical. The difference between them is focused.

- Contrast

  Hungarian:

  (56) Anna nem OROSZUL, hanem FRANCIÁUL
  Anna not Russian.ESS-FORM, but French.ESS-FORM
  tanult.
  learnt
  It wasn’t Russian that Anna learnt at high school, but French.
Example (56) is about one state of affairs that is seen differently by two people. The difference between the way they see it is focused (*i.e.* what exactly Anna learned) (Figure (7.21)).

\[-\text{Correction}\]

**Hungarian:**

(57) Nem! ANNA tanult franciául!

\[\text{No! ANNA learned French.}\]

\[-\text{Correction}\]

**Figure 7.21:** D-tree: Anna nem oroszul, hanem franciául tanult.

In a correction, the focus of the previous sentence is changed to one of its alternatives (Figure (7.22)).
• Hocus

The hocus is in complementary distribution with the focus, and appears as the +prominent, non-D-linked element of narrations:

(58) El kellett utaznom. Így tegnap a feleségem vitte
vm had to leave so yesterday the wife.Poss1sg take.Pst
óvodába a gyerekek.
kindergarten.ill the children.acc
I had to leave, so yesterday it was my wife who took the children to
the kindergarten.

The d- and i-structures are shown in Figure (7.23).

```
+PROM [− D-LINKED {el kellett utaznom}] ]
−PROM [− D-LINKED {a feleségem}] ]
```

Figure 7.23: D-tree: Narration-Hocus

• Question phrases

Question words also appear in the +prominent part of the information structure, in discourse relations containing questions. Most of them are, like foci in the answers, non-D-linked, whereas some of them in pair-list multiple questions, D-linked (see the examples above).

• Background information and Completive information

Background information, −prominent and non-D-linked, can be part of any discourse relation and any clause, just like completive information, which is, as opposed to background information, non-D-linked (see the examples above).

The following table provides a summary of these considerations:

Since the information structure status (attributes) and the discourse structure identify the particular discourse functions (information structure roles), they are not part of the representations as i-structure attributes.

7.4.3 The place of d-structure

In LFG, all levels of representation are related to some others via correspondence functions, although there is not necessarily a mapping function from every single level into all the others (see Mycock (2006)'s architecture in Chapter (2)). D-structure is particular in the sense that its basic unit is not the individual sentence, but the larger discourse, comprising more than one sentence. Discourse functions
are identifiable in the individual sentence, based on formal criteria (syntactic position, prosody), and this is mirrored by the mapping functions between syntax, prosody, and the information structure. However, the exact semantic content of these information structure units is highly context-dependent. I propose, thus, that the d-structure is accessible from the i-structure (or i-structures of the individual clauses) via the correspondence function $\delta$. The definition of a function makes it possible that several i-structures (clauses) are mapped onto one d-structure, constituting the discourse. The exact formalization of this correspondence function is, however, left for future research.

7.5 Summary

In this chapter, I proposed an analysis of multiple questions in the mainstream LFG framework, emphasizing the crucial role of the i-structure. However, I identified some problems with respect to the i-structure architecture, notably that it has to accommodate a larger set of elements than its basic units. This is why I posited an alternative architecture, which is based on the common properties of these units (prominence, D-linkedness) and can thus host all of them. I also assume a level of discourse-structure in the LFG framework, represented as a tree-diagram with the information structures of the individual sentences in its nodes. The d-structure encodes discourse-relations, based on Büring (2003) and Asher and Lascarides (2003), which also specify (via the correspondence function $\delta$) the set of i-structures and the semantic content of the i-structure units the d-tree contains. In the next two chapters, I apply the proposed analysis to the Hungarian and French data.
Chapter 8

Analysis of Multiple Questions in Hungarian

Contents

8.1 Introduction ................................................................. 309
8.2 Syntactic Structure (C-structure) ................................. 310
8.3 Multiple Questions ....................................................... 316
  8.3.1 All question words in the preverbal domain ............... 317
  8.3.2 One left-peripheral and one right-peripheral question word 331
  8.3.3 Coordinated question words .................................... 339
8.4 Beyond Multiple Questions ........................................... 346
8.5 Summary ................................................................. 351

8.1 Introduction

In the previous chapters on Hungarian, I presented the syntactic structures proposed in the most influential analyses in the movement-based frameworks. It was shown that discourse functions are usually supposed to reside in the specifier position of (discourse)-functional projections, such as TopP, FocP, or even CTOPP. In a more recent approach, the elements occupying the preverbal position (in complementary distribution) are assigned to Spec,PredP, reflecting the assumption that they are analyzed as secondary/specifical predicate (É. Kiss, 2006).

In the modular architecture of Lexical-Functional Grammar, there is no need to suppose that discourse functions are encoded in the syntactic structure (i.e. in the c-structure, more precisely). For a long time, they appeared in the f-structure and were identified with a grammatical function via functional uncertainty (see Bresnan and Mchombo (1987) on grammaticized discourse functions). At the same time, functional annotations indicated in the c-structure which syntactic position a particular discourse function was associated with, in case syntactic positions were argued to reflect discourse functions (see for instance, King (1995) on Russian). However, an independent level of information structure has been more and more advocated and assumed in the LFG framework (see Butt and King (1996), King (1997), Mycock (2006), Dalrymple and Nikolaeva (2011), etc.), which means that discourse functions are no longer encoded in the f-structure, but in a separate i-structure (however, certain discourse functions are argued to be syntacticized and
thus, represented also in the f-structure). In the previous chapter, I discussed in
details the possible architectures of this structure, and opted for one that does not
contain topic or focus as its primitives, but consists of more general sets that can
accommodate a larger variety of elements. In this and the following chapter, I will
make use of the alternative architecture proposed, bearing in mind that the analysis
can be compatible with the i-structure of the mainstream LFG analyses. However,
first I propose a c-structure for Hungarian.

8.2 Syntactic Structure (C-structure)

In LFG, constituent structure corresponds to a flexible X-bar theory representation,
in which no node, not even the head is obligatory, and exocentric constituents are
permitted (there is no binary-branching constraint). The question is, what kind
of c-structure should be associated with Hungarian. To my knowledge, there have
been two proposals in the LFG literature for the c-structure in Hungarian, but they
concentrated mostly on the problem of the preverbal position and the elements it
can host: focus and question words.

In the first analysis (Börjars et al., 1999), the immediately preverbal constituent
is sister to the verb in an extended verbal projection, which is supposed to host also
all the elements of the preverbal domain (topics and quantifiers). The discourse
functions are associated with syntactic positions via functional annotations. This
analysis does away with the set of functional projections (TopP, CTopP, DistP/QP),
whose head position is usually empty, since they are only postulated for accommod-
ating one type of element in their specifier position. FocP is an exception to this,
since the verb is supposed to move into its head position, leaving behind the verbal
modifier. However, according to Börjars et al. (1999), even a FocP is superfluous
in a theory in which no Foc feature is supposed to be assigned or checked. The
authors assume OT-type constraints as well, which account for word-order and the
immediately preverbal position of the focus. The second analysis to be mentioned
here is that of Mycock (2006), who assumes that the focus and the question words
are in Spec,VP, thus obligatorily sister to the verb.¹

According to Dalrymple (2001), functional categories vary from language to
language, and each of them has to be motivated for each language. According to
this, the I head position can be occupied by a finite verb or an auxiliary, like the C
position (in inversion contexts). Thus King (1995) assumes that in Russian, only
non-finite verbs reside in the VP, finite verbs occupy the I position, the topic and
the contrastive focus the Spec,IP and interrogative words the Spec,CP position.
Dalrymple (2001) also mentions that positing a VP projection is motivated only if
it contains only the verb and its complements (except for the subject) and these
constituents can appear together at other parts of the sentence as well. On the other
hand, if the subject can appear as sister to the V, the VP projection is unmotivated.
Now, the syntactic structure of non-configurational languages is represented with
the help of the non-configurational S node, which does not necessarily contain a CP

¹Laczkó and Rákosi (2011) also assume a VP projection in Hungarian, in which the verbal
modifiers occupy the specifier position.
or an IP projection. It is also possible that one part of the sentence is hierarchical and the other exhibits a free word order, flat structure, in which case the tree diagram contains both CP/IP and S nodes. Such languages are Warlpiri and Welsh.

These considerations about the VP undermine Mycock (2006)’s (and Laczkó and Rákosi (2011)’s) c-structure, since in Hungarian, the subject can be postverbal, appearing as sister to the verb, between the verb and the direct object:

\[(1)\quad \text{Marinak} \quad \text{adta} \quad \text{oda} \quad \text{János} \quad \text{a könyvet.} \]

\[\text{Mari} . \text{DAT} \quad \text{give} . \text{PST} \quad \text{VM} \quad \text{János} \quad \text{the book} . \text{ACC} \]

John gave the book to MARY.

Moreover, Mycock assumes that two question words (interrogative foci in her analysis) can jointly occupy the Spec,VP position, which is (presumably) not possible in the case of non-interrogative foci. (I pointed out that Mycock posits a distinction between interrogative and non-interrogative foci based on the Hungarian data in order to account for the very same data). On the other hand, Börjars et al. (1999)’s architecture does not deal with the postverbal section in details, and neither of the analyses account for the fact, why only one focus can precede the verb. Since neither of the structures proposed so far can account for all the necessary data, a new structure is proposed in this section, which aims to capture these data and to correspond to the above mentioned LFG assumptions better than the previous ones.

In Hungarian, as we have seen, the preverbal and postverbal parts of the sentence differ, in that in the preverbal section, the position and the order of the elements depend on their role in the information structure. This can be directly represented in LFG via the functional annotations. The question is now, if a hierarchical preverbal section is motivated even in the LFG framework. In the transformational frameworks, two factors motivated the hierarchical preverbal structure: the obligatory binary branching in the tree diagrams and the fact that the linear order of the elements determines their relative scope as well. As opposed to this, the postverbal part of the sentence exhibits free word-order (obeying, supposedly, certain phonological factors, such as heavy elements tend to follow lighter ones). According to András Komlósy (p.c.), in LFG, neither of these factors necessitate a hierarchical architecture, since the linear order of elements can in itself reflect the scopal relations, thus there is no reason for positing a hierarchical sentence structure in Hungarian. As was pointed out above, a VP projection is not motivated. In the c-structure, annotations indicate the grammatical and discourse functions. In addition, they can express other constraints as well, for instance, the obligatory presence of other elements, which is needed in the case of sentence-final focus (which can appear on condition if there is a preverbal focus in the sentence). With the annotations, thus, we can express and formalize the observation that the preverbal part of the Hungarian sentence is determined by the information structure. The flat, annotated c-structure is shown in Figure (8.1).
Figure 8.1: C-structure of Hungarian
A few remarks are due here concerning this structure:

- The annotation \( \uparrow \sigma \in (\uparrow \sigma, \text{PROM D-LINKED}/\neg \text{D-LINKED}) \) refers to the language-specific fact that in Hungarian the (left) peripheral and the preverbal positions are prominent and also specifies if the given position hosts D-linked or non-D-linked elements. This has to be indicated, since syntactic highlighting in this language means placing the element into one of these positions, depending on other factors. The set of prominent positions is constant in a given language, thus it is not indicated in all the c-structures of this chapter.

- The structure reflects that the preverbal part of the sentence is determined by the i-structure. The annotation \( \uparrow \sigma \) refers to this. As pointed out in the previous chapter, the signs \( \sigma \) and \( \iota \) represent mapping functions, mapping semantic and discourse information, respectively, onto the i-structure (see Dalrymple and Nikolaeva (2011)).

- According to this, the first block of constituents can be thematic shifters and contrastive topics (or, eventually, interrogative phrases). They come in a block, since more than one topic is possible in a sentence and they constitute an indivisible unit. These discourse functions do not appear in the annotations, because in the framework proposed in this thesis, they are the obvious ”intersection” of a given set in the information structure (for instance, +PROMINENT and D-LINKED) and the discourse relation of a given sentence with the preceding (or following) one.

- Quantifiers are best assigned to a position via annotations with the help of their lexical properties, i.e. that they are, for instance, universal quantifiers (\( \forall \)). The attribute \( \forall \) refers to the lexical property of universal quantifiers; its values can be + or −. Just like in the case of topics, there can be more than one preverbal quantifier in the sentence.

- In the postverbal part, there are only annotations referring to the f-structure. As we have seen, those elements can fill any grammatical function (even that of the subject). Later on we will add specifications to the sentence-final position, since it can host interrogative words, focus and contrastive topic. It is important to note that postverbal prominent elements are also possible in Hungarian, but they are prosodically, and not syntactically highlighted (for a detailed description of prosodic representations and prosodic highlighting, see Mycock (2006)). Such elements are right peripheral foci, contrastive topics and question words, and such question words and foci that precede infinitives following an auxiliary.

- Contrary to the usual analyses, what we see here is that there are two prominent preverbal positions, one D-linked and the other \( \neg \)D-linked. As we have seen in Chapter (5), there are (marginally) sentences in Hungarian in which we can find two prominent preverbal elements. Interestingly, one of these is
always an interrogative word: a focus and an interrogative word (2) and an interrogative word and a negative focus (3) combinations are permitted, but the combination of two foci is strictly prohibited:

(2) -Jó, de JÁNOS mit evett?
     ok, but János what eat.pst
     Ok, but what about John, what did he eat?

(3) -Ki NEM A HAMLETET olvasta?
     who not the Hamlet.acc read.pst
     Who didn’t read Hamlet (but something else)?

Gazdik and Winterstein (2011) investigate the relationship between discourse relations and syntactic positions in Hungarian. The authors argue that those discourse relations that imply that a particular element of their semantics is distinguished will use (one of) the prominent preverbal position(s) to localize this element. Thus, the focus will be in the Prominent Preverbal Position (PPP) for relations such as contrast, parallel, and question-answer pair, since these make an explicit reference to the focus. In case the focus must be on the verb (VP focus, verum focus), the PPP will be empty, because it is the verb that is highlighted. For questions, the PPP will mostly be occupied by the question word, which is central to the semantics of the question.

Other discourse relations have no such requirement and often impose an all-focus reading of their right argument. These will allow the variety observed in Chapter (5): verbal particles, secondary predicates, hocus, etc. Since the sentence must be all-focus, a single sub-constituent cannot act as narrow-focus. A good example of such a discourse relation is the case of narration.

If we consider that the PPP is a semantically privileged position, it is no surprise that it is the locus of the hocus since the hocus or verbal particles, carrying a specific semantic meaning (identification and aspect, respectively). More generally, the elements present in the PPP are a way to underline some aspects of what the speaker wants to convey.

Let us now turn to the problem of the preverbal position(s). I assume, thus, that the immediately preverbal position is non-D-linked (i.e. the element it hosts appears in the non-D-linked subset of the prominent set at i-structure). However, there is an optional d-linked position between the previous one and the verb, hosting second occurrence prominent elements. The model has to rule out multiple preverbal foci, but permit the co-occurrence of foci and question words. In order to achieve this, we could use disjunction (Figure (8.2)).

The disjunction in the annotations refers to the following: the element in the linearly first preverbal position is either a interrogative word, or a focus, in the second, either an interrogative word or a negative focus. Thus, the disjunction permits that both positions be filled by an interrogative word, or that one is filled
8.2. Syntactic Structure (C-structure)

by an interrogative word and the other by a focus or negative focus. The problem is
that these annotations would make it possible that one position is filled by a focus
and the other by a negative focus, and these sentences would be ungrammatical in
Hungarian.

Another possible formalization would be if the annotations could not refer only
to a given node and to the one dominating it, but to sister nodes as well. Dalrymple
(2001) (citing Nordlinger 1998) argues that it is definitely possible to refer to sister
nodes in LFG, and this is symbolized with arrows pointing to the left or to the
right. Let us see the structure with these annotations (Figure (8.3)).

\[
\begin{array}{c}
S \\
\ldots ZP \\
\downarrow------------------------
\end{array}
\quad
\begin{array}{c}
(ZP)\ldots \\
\downarrow------------------------
\end{array}
\]

Figure 8.3: The ”focus” position 2

The annotation under the node A is the following:\footnote{In the following annotations, \(\implies\) stands for implication, whereas \(\rightarrow\) and \(\leftarrow\) refer to the constituent immediately to the right or to the left.}

\[\downarrow \in (\uparrow \sigma_t \ +PROM \ \neg D\text{-LINKED}) \implies \neg[\leftarrow \in (\uparrow \sigma_t \ PROM \ \neg D\text{-LINKED}) \ (\leftarrow Q) = +] \]

Node B is annotated as follows:

\[\downarrow \neg (\neg \text{NEG} = +) \implies \neg[\rightarrow \in (\uparrow \sigma_t \ +PROM \ \neg D\text{-LINKED}) \ (\rightarrow Q) = +] \]

Node B is optional, it is present in the structure only if there is a second +PROM
element in the immediately preverbal part of the sentence. The annotations on these
nodes express bans on the environment of these nodes. The annotation under A
means: if the current constituent is +PROM and \(\neg D\text{-LINKED}, then there can be
no +PROM and \(\neg D\text{-LINKED} element on its left, which is not an interrogative
pronoun.

The annotation under B reads as follows: if the current constituent is a negative
expression, there can be no +PROM and \(\neg D\text{-LINKED} element on its left, which is
not an interrogative pronoun.

These annotations describe the facts without providing an explanation or a useful generalization about them. We should remark, concerning these structures, that they are marginal with respect to those sentences in which a focus and an interrogative word do not co-occur. They are possible only in narrow-focus sentences, more precisely, in questions which do not start a discourse, but repeat a previously asked question. In fact, the question word in (2), and the (negative) focus in (3) are already known, and the real prominent element of the sentence is the focus and the question word, respectively. This contrast can be illustrated by the following paraphrases:

(6) a. -Jó, de JÁNOS mit evett?
   OK, but János what eat.PST
   Yes, OK, but what about John, what did he eat?

   b. -Jó, de János az, akiről kérdezem, hogy mit evett.
      OK, but János that who.DEL ask.PRS.1SG, that what eat.PST
      Ok, but it is John, about who I am asking what he ate.

(7) a. -Ki nem a Hamletet olvasta?
   who not the Hamlet.ACC read.PST
   Who DIDN’T read Hamlet?

   b. -Ki az, aki nem a Hamletet olvasta?
      who that, who not the Hamlet.ACC read.PST
      Who is it that DIDN’T read Hamlet?

The exact treatment of second occurrence foci and question words in the present framework awaits further research.

An advantage of the present analysis is that it automatically rules out multiple foci in the preverbal domain: since one of the positions hosts D-linked and the other non-D-linked elements, multiple foci cannot occupy these positions. What seems to be forbidden in Hungarian, is thus the presence of two preverbal, prominent non-D-linked elements. The cumulation of prominent elements in the positions between the quantifier and the verb is only possible when a question is asked for the second time, or it has some antecedent in the preceding discourse. This is the case both in (2) and (3).

8.3 Multiple Questions

In this section, we examine the different levels of representation associated with multiple questions in Hungarian. We will always analyze mini-dialogues, since answers can prove to be revelatory with respect to the representation of the question itself. The i-structures of the question and the answer will be integrated into discourse trees, which specify the discourse relation between the question and the answer. We
consider first multiple questions with more than one left-peripheral question word:

8.3.1 All question words in the preverbal domain

8.3.1.1 Left-peripheral question words - pair-list reading

Let us first have a look at the example of questions expecting pair-list answers.

(8) a. Q: **Ki mit hozott a bulira?**
   Who what bring.PST the party.SUBL
   Who brought what to the party?

   b. A: János bort, Péter sört, Mari pedig süțiket
   János wine.ACC, Péter beer.ACC, Mari and cookies.ACC
   hozott. bring.PST
   John brought wine, Peter beer, and Mary cookies./ as for Mary, she
   brought cookies.

As was discussed in Chapter (2), the question contains a non-D-linked and a D-linked interrogative word, whereas the answer contains contrastive topic-focus pairs.

In the preceding chapters we could also observe that the Hungarian examples contradict the assumption that question words constitute a subtype of focus. In pair-list answers, the information structure status of the constituents clearly differs (contrastive topic + focus pairs), the question remaining if the interrogative words themselves have different status at the level of information structure or not. The answer is most probably yes, since otherwise the interlocutor would have no clue about how to structure the information in the answer, in other words, which word is the Sorting key.

Since semantic and discourse information is represented at a separate level, the question is if syntactically discourse functions should be represented at the level of f-structure. Hungarian would be a typical example of a language in which discourse functions are associated with syntactic positions. However, as we have seen at the beginning of this chapter, the preverbal part of the Hungarian sentence is determined by information structure constraints, and this can be accounted for via the c-structure-i-structure mapping, without referring to the f-structure (Butt and King (1996, 2000) also assume that discourse functions are associated with particular syntactic positions in Hindi/Urdu). One could argue that some discourse-function related processes are syntactic in nature, for instance, topicalization (left dislocation without a resumptive element). For example, Dalrymple (2001) (citing Bresnan 1994) assumes that a sentential argument undergoes topicalization when it is the subject of the sentence, *i.e.* it is identified with both the subject and the topic in the f-structure:

(9) That David yawned surprised Chris.

(Dalrymple, 2001, p. 115, 89c)
Preposed constituents have often been associated with discourse functions in English (and other languages, like French), but as Abeillé et al. (2008) demonstrate, NP-preposing cannot be exclusively analyzed as topicalization or focalization in French. The question is what the syntactic import of analyzing extraction structures as topicalization or focalization really is. Since topics and foci (as discourse functions or syntacticized discourse functions) have been associated with (mostly left peripheral) syntactic positions, the fact that certain constituents undergo topicalization or focalization does not only have certain semantic or discourse-related aspects with respect to a certain constituent, but also indicates, indirectly, that the constituent is extracted into a left-peripheral position. If we consider that the discourse-semantic aspects are encoded in the i-structure, what remains is that (certain) interrogative phrases, relative pronouns, some clauses or constituents are extracted. Extraction may seem to be a reminiscent of movement-based approaches, since it assumes that a constituent is not in its canonical (argument) position, but is fronted/preposed. However, there is another possible approach to extraction. It can also be assumed that extracted is also a function, like subject or object (see also the filler-gap constructions in HPSG). Alsina (2008) refers to this function as \( OP(erator) \), and Asudeh (2010) as \( UDF \) (Unbounded Dependency Function), arguing that OP has semantic connotations that are not appropriate for all instances of this function (see also Asudeh (2004)).

Being the subject of the sentence means being associated with a certain position in configurational languages, being in the nominative case, and showing (person/number/gender) agreement with the finite verb. Being \( OP/UDF \) in a sentence means appearing in a non-canonical, sentence-initial position. Extracted constituents can play a discourse function in the i-structure, and can be associated with a grammatical function via functional uncertainty at the level of f-structure. After all, this is what the syntactic aspect of topicalization or focalization means. In addition, certain constituents can behave differently if they are extracted. For instance, Skrabalova (2006) argues that in Czech, question words with different grammatical functions can only be coordinated if they are extracted. This means that extracted question words have a double function: \( OP/UDF \) and another grammatical function. If the function UDF is assumed, the generalization about coordination can be maintained: coordinated elements must share the same function. From now on, I will represent the UDF function for question words in the f-structures (and c-structure annotations). A problem emerging from this assumption is the analysis of cases where there is more than one extracted question word (without coordination). In order to avoid problems of consistency (two elements having the same function), we analyze UDF as a set (like ADJuncts and coordinate structures).

The c-structures of the question and the answer are shown in Figures (8.4) and (8.5).

---

3 I thank Mary Dalrymple for calling my attention to these proposals.
4 I leave for future research the investigation of the role of the UDF function in the whole preverbal domain in Hungarian.
8.3. Multiple Questions

Figure 8.4: C-structure: Ki mit hozott a bulira?

Figure 8.5: C-structure: János bort hozott...
The f-structures are presented in the Figures (8.6) and (8.7).

Let us consider now the discourse and i-structures of the examples above. The i-structures of the question and the answer are organized in the same discourse tree, and the relation between them is elaborative: a Complete Implicative Question-Answer Pair (in which the original question is split up into subquestions, but all the subquestions get answered). The d-structure containing the i-structures is illustrated schematically in Figure (8.8).
8.3. Multiple Questions

A: Question

\[
\begin{align*}
+ \text{PROM} & \quad \neg \text{D-LINKED} \left\{ y \mid (y) \in \{\text{THINGS}\} \right\} \\
& \quad \text{D-LINKED} \left\{ x \mid (x) \in \{\text{PEOPLE (János, Péter, Mari)}\} \right\} \\
- \text{PROM} & \quad \neg \text{D-LINKED} \left\{ \right\} \\
& \quad \text{D-LINKED} \left\{ \lambda x. \lambda y. \text{HOZOTT A BULIRA} \ (x, y) \right\}
\end{align*}
\]

B: Subquestion

\[
\begin{align*}
+ \text{PROM} & \quad \neg \text{D-LINKED} \left\{ y \mid (y) \in \{\text{THINGS}\} \right\} \\
& \quad \text{D-LINKED} \left\{ \text{János} \in \{\text{PEOPLE}\} \right\} \\
- \text{PROM} & \quad \neg \text{D-LINKED} \left\{ \right\} \\
& \quad \text{D-LINKED} \left\{ \lambda y. \text{János HOZOTT A BULIRA} \ (y) \right\}
\end{align*}
\]

C: Answer

\[
\begin{align*}
+ \text{PROM} & \quad \neg \text{D-LINKED} \left\{ \text{bort} \in \{\text{THINGS}\} \right\} \\
& \quad \text{D-LINKED} \left\{ \text{János} \in \{\text{PEOPLE}\} \right\} \\
- \text{PROM} & \quad \neg \text{D-LINKED} \left\{ \right\} \\
& \quad \text{D-LINKED} \left\{ \text{János HOZOTT A BULIRA} \ (\text{bort}) \right\}
\end{align*}
\]

D: Subquestion

\[
\begin{align*}
+ \text{PROM} & \quad \neg \text{D-LINKED} \left\{ y \mid (y) \in \{\text{THINGS}\} \right\} \\
& \quad \text{D-LINKED} \left\{ \text{Péter} \in \{\text{PEOPLE}\} \right\} \\
- \text{PROM} & \quad \neg \text{D-LINKED} \left\{ \right\} \\
& \quad \text{D-LINKED} \left\{ \lambda y. \text{Péter HOZOTT A BULIRA} \ (y) \right\}
\end{align*}
\]

E: Answer

\[
\begin{align*}
+ \text{PROM} & \quad \neg \text{D-LINKED} \left\{ \text{sört} \in \{\text{THINGS}\} \right\} \\
& \quad \text{D-LINKED} \left\{ \text{Péter} \in \{\text{PEOPLE}\} \right\} \\
- \text{PROM} & \quad \neg \text{D-LINKED} \left\{ \right\} \\
& \quad \text{D-LINKED} \left\{ \text{Péter HOZOTT A BULIRA} \ (\text{sört}) \right\}
\end{align*}
\]
The whole d-structure representation leads to the following semantic interpretation, in which the subquestion gets answered with respect to all the elements in the domain set of the question:

\[
\lambda f \forall x \in \{\text{PEOPLE}\} \rightarrow \text{hozott a bulira } (y), \ (y) \in \{\text{THINGS}\}
\]

Note that this semantic representation is not part of the i- or d-structure, but it follows from this particular organization of the i-structures in this discourse context.

The i-structures reflect the assumption that clauses can have separate i-structures, since they answer different subquestions of the original question. As was illustrated previously, in Hungarian, the order of the question words indicates how the information is to be organized in the answer. In the above analyzed example, the question word qualifying as the Sorting key was at the same time the D-linked one and the leftmost one in the sentence. I assume this is so in most cases (Sorting keys are D-linked), and they are leftmost in the \textit{wh}-sequence. Let us now compare the above example with the LFG representation of its counterpart, in which the order of the question words is the inverse:

\[
\begin{align*}
\text{Q: } & -\text{Mit ki hozott a bulira?} \\
& \text{what who bring.PST the party.SUBL} \\
& \text{What was brought by who to the party?}
\end{align*}
\]

\[
\begin{align*}
\text{A: } & -\text{A bort János, a sör Péter, a sütiket pedig the wine.ACC János, the beer.ACC Péter, the cookies.ACC and}
\end{align*}
\]

\footnote{This contradicts what we saw in the previous chapter, where we concluded that the basic unit of the information structure was the sentence. There might be a difference, in this respect, between the coordination of answers to a multiple question and the relationship between a main and a subordinate clause, both from the grammatical and from the discourse perspective.}
Mari hozta.
Mari bring.pst
The wine was brought by John, the beer by Peter, and the cookies by Mary.

The c-structures are shown in Figures (8.9) and (8.10), and the d-structure in Figure (8.11) (the f-structures are similar to those of the previous question-answer pair).
Figure 8.9: C-structure: Mit ki hozott a bulira?

Figure 8.10: C-structure: A bort János hozta...
Figure 8.11: D-structure: Mit ki hozott a bulira? A bort János, ...
The semantic interpretation following from the above structure is the following:

(12) \( \lambda f \forall y \in \{\text{THINGS}\} \rightarrow \text{hozta a bulira} (x), (x) \in \{\text{PEOPLE}\} \)

8.3.1.2 Contrastive topic on the right periphery?

There is another way of answering (8), in which the constituents corresponding to the D-linked question word are on the right periphery of the sentence (I thank Zsuzsanna Gécseg for pointing this out to me):
8.3. Multiple Questions

(13) a. Q: -Ki mit hozott a bulira?
   who what bring.PST the party.SUBJ
   Who brought what to the party?

   b. A: -Bort hozott János, sört Péter, és sütiket
      wine.ACC bring.PST János, beer.ACC Péter, and cookies.ACC
      Mari.
      Mari
      John brought wine, Peter beer and Mary cookies.

The same word-order variations can be observed in single questions with Implicative Answers, indicated by contrastive topics:

(14) a. Q: -Mit tanultak idén a diákok?
   what study.PST this year the students
   What did the students study this year?

   b. A1: -Az elsőévesek szintaxist, a másodévesek
      the first year students syntax.ACC, the second year students
      pedig szemantikát tanultak.
      as for semantics.ACC study.PST
      The first year students studied syntax, the second year students
      semantics.

   c. A2: -Szintaxist tanultak az elsőévesek, és szemantikát
      syntax.ACC study.PST the first year students and semantics.ACC
      a másodévesek.
      the second year students
      The first year students studied syntax, the second year students
      semantics.

However, the contrastive topic cannot be sentence-final in the case of Partial Implicative Answers:

(15) a. Q: -Mit tanultak idén a diákok?
   what study.PST this year the students
   What did the students study this year?

   b. A: # -Szintaxist tanultak az elsőévesek.
      syntax.ACC study.PST the first year students
      The first year students studied syntax.

This difference also motivates a distinction between Complete and Partial Implicative Answers.
I conclude thus that the sentence-final/right-peripheral position of the Hungarian sentence can be occupied by the following prominent elements: question words, foci and contrastive topics, which are prosodically highlighted (see Mycock (2006)) (the position is unspecified with respect to D-linkedness). Consider the c-structure of the answer in (13) (Figure (8.12)).

\[
\text{S} \quad \text{NP} \quad \text{V} \quad \text{NP} \\
\uparrow \sigma \in (\uparrow \sigma + \text{PROM} \sim \text{D-LINKED}) \quad \uparrow \sigma \in (\uparrow \sigma + \text{PROM}) \\
(\uparrow \text{OBJ})=\downarrow \quad \uparrow = \downarrow \quad (\uparrow \text{SUBJ})=\downarrow \\
\text{Bort} \quad \text{hozott} \quad \text{János}
\]

Figure 8.12: C-structure: Bort hozott János...

The question is now, if right peripheral question words can also be D-linked or not. The data suggest that multiple questions with one preverbal and one right-peripheral question word can (marginally) also trigger pair-list answers. We examine this in Section (8.3.2).

8.3.1.3 Left-peripheral question words - single-pair reading

As we have seen in Chapter (2), some multiple questions containing preverbal question words in Hungarian do not license a pair-list answer, but expect a single pair answer, like in the following example:

(16) a. Q: **Honnan hová** utazik János?  
    From where to where travels János  
    From where to where is John travelling? 

b. A: -Budapestről Szegedre utazik.  
    Budapest.DEL Szeged.SUBL travels  
    He is travelling from Budapest to Szeged.

I argued that the explanation for the single-pair reading is that the question words refer to the starting and final points of a process (travelling, in this case). The question is synonymous with the following one:

(17) **Honnan** utazik János **hová**?  
    From where travels János to where  
    From where to where is John travelling?

The position of the second question word is optional. In (17), it is postverbal, like other single-pair interrogative structures, whereas in (16), it occupies the immediately preverbal position, which can be accounted for by the fact that the
corresponding constituent in the answer, like secondary predicates, typically also appears in that position, forming a prosodic and lexical unit with the verb (the verb cliticizes onto it). Since both question words are non-D-linked, they cannot occupy the separate positions preceding the verb, since one of them hosts D-linked elements. Based on cross-linguistic data (see similar cases in French in Chapter (2)), I propose that the question words constitute a complex question word and occupy the non-D-linked preverbal position together. This case, thus, corresponds to Mycock (2006)'s analysis of multiple preverbal question words in Hungarian. I assume that it is applicable to this specific case.

In the answer, the constituents corresponding to the question words appear both in the preverbal position, which contradicts the strict unity of preverbal focus in Hungarian. A possible analysis is to argue that these focus items do not constitute true multiple foci, but a complex focus (Krifka, 1992) that can occupy the preverbal position. A remarkable difference compared to pair-list answers is also the fact that the linearly first focused constituent lacks the rising intonation of contrastive topics, which characterizes the constituent corresponding to the linearly first question word in pair-list answers. This is not surprising, since this first constituent does not indicate any answering strategy. Parallelly to this, we can consider the two interrogative words as a complex question word that does not expect any pair-list answer.

The complex question word and ordinary pair-list questions illustrate the claim that sentences sharing the same linear order of words can have different information structures. (Whether the two questions have exactly the same syntactic structure or not is difficult to determine. Nevertheless, in both cases, the question words belong to the same intonation unit, and they are, in most cases, inseparable by intervening material. Even if we suppose that complex question words occupy one single position and that in pair-list questions, the question words are in two different positions, the syntax would provide no clue concerning the expected answer and the interpretation of the question.)

Another approach would be to claim that only the linearly first question word is extracted, and the second is in fact in situ. However, in situ means, in this case, that it is in the immediately preverbal position. If we consider the constituents corresponding to the immediately preverbal question word, they are all verb carriers/secondary predicates. This, however, does not contradict the assumption that the two question words constitute a complex interrogative word.

Consider the proposed c-structures of (16) (Figures (8.13) and (8.14)).
330 Chapter 8. Analysis of Multiple Questions in Hungarian

![Diagram](image-url)

**Figure 8.13:** C-structure: Honnan hová utazik János?

![Diagram](image-url)

**Figure 8.14:** C-structure: Budapestről Szegedre utazik.
The f-structures are illustrated in Figures (8.15) and (8.16).

Figure 8.15: F-structure: Honnan hová utazik János?

Figure 8.16: F-structure: Budapestről Szegedre utazik.

The d-structure of (16), representing a Direct Q-A Pair is the following (Figure (8.17)).

The analysis of this question differs from that of the previous, pair-list ones in that here there is no function application between the question words, since neither of them is D-linked. The answer is, thus, a complex focus (see Krifka (2006)), containing one element from the set denoted by each question word.

8.3.2 One left-peripheral and one right-peripheral question word

In Hungarian, this type is mostly restricted to question words denoting the same type of set, since often the question does not refer to the denotation of the question words itself, but to the direction of their relation:

(18)  
Ki hívott meg moziba kit?
who invite.PST VM cinema.ILL who.ACC
Who invited who to the cinema?

The rule states that in structures containing one preverbal and one sentence-final question phrase, if the two question words are D-linked, they have to have the same restriction (denote the same type of set), symbolized with α.
These questions are supposed to contain multiple foci (É. Kiss, 1992a; Surányi, 2007). However, approaches differ in explaining how the sentence final interrogative word is focused. In the transformational literature, it is argued that these question words go through LF/covered movement and attach to the item in the Spec,CP or Spec,FocP position.

Surányi (2006) proposes different, movement-based mechanisms in order to derive the different interpretations of these structures. He argues that sentence final interrogative words in 'English-type' questions can optionally carry a \([F+]\) focus feature. If they do, they move covertly to the left periphery and behave like 'multiple fronting' questions, whereas if they do not, they stay in situ and the question is answered by a single pair.

This proposal reflects the approach, according to which taking a certain scope is only possible in syntax, as a result of syntactic processes. On the other hand, as has already been mentioned, Mycock (2006) argues that focusing is also possible in prosody, syntax being another, but not the exclusive domain where an element can be focused. In this thesis, I follow Mycock’s approach, with the remark that the syntactic and prosodic prominence (highlighting) of question words does not necessarily mean that they should be considered as a subclass of focus.

We have already seen in the previous chapters that discourse functions, such as topic, focus, or contrastive topic, cannot be exclusively assigned to designated syntactic positions, but a more adequate approach is to claim that all languages have formal (syntactic and prosodic) means of highlighting prominent elements, and these are exploited differently in different structures. For instance, we have seen that contrastive topics, which typically indicate an answering strategy to a
question, usually appear on the left periphery of the Hungarian sentence; however, this is not always the case, since in Implicative answers to single questions they can be right-peripheral (14-c), like in answers to pair-list multiple questions (13).

Recall also the case of questions (asked for the second time) containing a focus. We also showed that the prominent constituent of such sentences is prosodically highlighted in the same way as foci, but from the discourse point of view, these constituents can also be analyzed as contrastive topics in that they continue the strategy started by the (partial) answers. Note that this constituent can precede the preverbal question word, but it can also appear sentence-finally:

(19) Jó, de JÁNOS mit evett?
    OK, but János what eat.PST
    Ok, but what about John; what did HE eat?

(20) Jó, de mit evett JÁNOS?
    OK, but what eat.PST János
    Ok, but what about John; what did HE eat?

Interestingly, Gécseg (2001), in a movement-based framework, derives the left-peripheral contrastive topic from a postverbal position. This supports the hypothesis that the contrastive topic can also appear postverbally in the sentence, and the left-peripheral position is rather a tendency and not a rule.

In addition, Surányi (2006) shows that when an English-type question licenses a pair-list answer, the list has to include an individual as a second member of the pair. We have already seen that in multiple questions with all question words in the preverbal domain, the linearly first question word(s) must be individualizable (since their denotation is mapped onto that of the other). This is why such question words as hogy (how) and miért (why) are excluded from that position. This observation supports the assumption that sequence-initial question words in multiple extraction questions and sentence-final ones in the English-type can behave similarly, and this corresponds to the claim that D-linkedness is the property that contrastive topics and the question words pointed out above share. Thus, in interrogatives with multiple preverbal question words, the non-final question word has to be D-linked, whereas in the English-type, the sentence-final one can be D-linked (in that case licensing a pair-list answer), but in most cases it is non-D-linked, licensing a single pair answer.

The question in (18) is in fact ambiguous between two readings. In one of them, the participants of the inviting event are known. What is not known is which of them was the inviter and which was the invitee. In the other, in this case less probable interpretation, neither of them is known, and the question asks for the identification of both of them. In the first case, the question can be paraphrased by two subquestions, and the answer to one of those is enough to answer the original question (the relation between the two subquestions being expressed by the exclusive or):
(21) János hívta meg moziba Mari? vagy Mari hívta meg Jánost?

Did John invite Mary to the cinema? or Did Mary invite John to the cinema?

The c-structures of the question (referring to the direction of the relation) and the answer are shown in Figures (8.18) and (8.19). The other meaning would be represented by non-D-linked question words in the question and (non-D-linked) foci in the answer.
8.3. Multiple Questions

Figure 8.18: C-structure: Ki hívott meg moziba kit?

Figure 8.19: C-structure: Mari hívta meg moziba Jánost.
The f-structures are shown in Figures (8.20) and (8.21), and the d-tree, including the subquestions indicated above, is illustrated in Figure (8.22).

Figure 8.20: F-structure: Ki hívott meg moziba kit?

Figure 8.21: F-structure: Mari hívta meg moziba Jánost.
Multiple Questions

Question

\[
\begin{align*}
&\text{INT} + \\
&\text{+PROM} \quad \neg \text{D-LINKED} \{ \text{Mari Jánost} \in \{ \text{János, Mari} \} \} \\
&\text{D-LINKED} \{ \} \\
&\text{¬PROM} \\
&\text{D-LINKED} \{ \} \\
&\text{D-LINKED} \{ \text{MEGHÍVTA MOZIBA' (Mari, János)} \} \\
\end{align*}
\]

Subquestion

\[
\begin{align*}
&\text{INT} + \\
&\text{+PROM} \\
&\text{D-LINKED} \{ \} \\
&\text{¬PROM} \\
&\text{D-LINKED} \{ \text{MEGHÍVTA MOZIBA' (Mari, János)} \} \\
\end{align*}
\]

Subquestion

\[
\begin{align*}
&\text{INT} + \\
&\text{+PROM} \\
&\text{D-LINKED} \{ \} \\
&\text{¬PROM} \\
&\text{D-LINKED} \{ \text{MEGHÍVTA MOZIBA' (János, Mari)} \} \\
\end{align*}
\]

Answer

\[
\begin{align*}
&\text{INT} + \\
&\text{+PROM} \\
&\text{D-LINKED} \{ \} \\
&\text{¬PROM} \\
&\text{D-LINKED} \{ \text{MEGHÍVTA MOZIBA' (Mari, János)} \} \\
\end{align*}
\]

Figure 8.22: D-structure: Ki hívott meg moziba kit?
The structure of the subquestions and the answer is the same in Hungarian, since polar questions are indicated only by prosodic means. The representation of prosody (the p-structure) is beyond the scope of this chapter; however, we assume, following Mycock (2006), that there is a mapping function from the p- onto the i-structure. Since the only difference between declaratives and polar questions is intonational in Hungarian, the information about questionhood comes only from the p-structure.\(^6\) I indicate this with \textit{INT +} in the i-structure, distinguishing declaratives and polar interrogatives.\(^7\)

Concerning the semantic representation in the original question in the d-structure (8.22), the denotation of the question words is restricted to the same set of two people (both questions are D-linked). The two subquestions contain a +PROM, −D-LINKED part, which is not the individuals themselves, but the relationship between them. The answer is the choice of one of these relations, and thus, this part of the subquestions and the answer is non-D-linked.

The d-structure is different in the other reading of example (18), in that the denotation is again the same (the set of people (relevant in the context)), but it is not restricted to two individuals. The question is not split into subquestions. Consider Figure (8.23).

As was suggested above, this multiple question structure can also trigger pair-list answers, since one question word can be D-linked (in that case, the set denoted by the D-linked question word is not restricted to two individuals and the same functional interpretation applies as in Figure 8.8). Although such answers are mostly triggered by multiple questions with preverbal question words, this is rather a tendency and not a general rule. This means that such answers are not completely excluded:

\[(22)\]

\[
\text{a. Ki nézett rá kire?}
\]

\[
\text{who glanced.PST at who.SUBL}
\]

\[
\text{Who glanced at who?}
\]

\[
\text{b. János nézett rá Marira, Péter Julira, és Pista}
\]

\[
\text{János glance.PST at Mari.SUBL, Péter Juli.SUBL and Pista}
\]

\[
\text{Annára.}
\]

\[
\text{Anna.SUBL}
\]

\[
\text{John glanced at Mary, Peter at Julia and Steve at Anna.}
\]

\(^6\)In the case of constituent questions, questionhood is indicated by the presence of question words. Intonation is not enough to differentiate between declaratives and interrogatives, since question words and foci can have the same prosodic highlighting (pitch accent). Nevertheless, prosody certainly plays a role, for instance in distinguishing between true interrogatives and echo questions.

\(^7\)However, this representation does not account for \textit{which}-questions, which do not necessarily denote the same type of set. Another way of formalizing this property in LFG would be to posit a rule that states that the question words must share some feature: \textit{+ANIM(ate)} (in the case of \textit{ki} (who) and its inflected forms), \textit{−ANIM(ate)} (in the case of \textit{mi} (what) and inflected forms), or \textit{+LOC(ations)} (in the case of \textit{hol}, \textit{honnan}, \textit{hová}, etc. (where, from where, to where)); or they are both \textit{+[SPEC(ific)]} (in the case of \textit{melyik} (which) phrases). Specificity is assumed to be part of the lexical entries of \textit{melyik} (which) phrases, just like interrogativity (Q) and D-linkedness. The exact formalization of this rule awaits further research.
These examples are analyzed like the ones we have seen so far that license a pair-list reading. The D-linked question word denotes a contextually denoted set (this is possible in one of the preverbal positions and in the sentence final one too), and it is mapped onto the denotation of the other question word.

Now we turn to multiple questions with coordinated question words.

8.3.3 Coordinated question words

As we have seen in Chapter (2), Hungarian allows the coordination of question words, which is highly problematic, in that the question words that are coordinated can have different functions. However, first we look at the coordination of adjunct question words (i.e. question word sharing the same function).

(23) **Hol és milyen feltételekkel lehet és szabad**
where and on which conditions be possible and be allowed to
**diákmunkát vállalni?**
student job to undertake

Where and on which conditions can one undertake a student job?

In such examples, the question words share the both of their functions both in the f-structure (UDF and ADJ). In the c-structure, they occupy the same preverbal position.
Coordination is analyzed in LFG as a set (Dalrymple, 2001; Peterson, 2004), in which the elements have to share certain features. The general rule of coordination is the following:

\[(24) \quad \text{Rule:} \quad \begin{array}{c} \text{XP} \\ \downarrow \in \uparrow \end{array} \rightarrow \begin{array}{c} \text{XP} \\ \downarrow \in \uparrow \end{array} \text{Conj} \begin{array}{c} \text{XP} \\ \downarrow \in \uparrow \end{array} \]

Let us consider the c- and f-structures first (Figures (8.24)\(^8\) and (8.25)):

---

\(^8\)Arguably, the only construction that necessitates a VP constituent in Hungarian is infinitives with a complement. Such sequences can appear at different parts of the sentence, showing that they form constituents.
Figure 8.24: C-structure: Hol ... lehet és szabad munkát vállalni?
The c-structure in (8.24) shows that the coordinated question words occupy together the preverbal, prominent, non-D-linked position (and license a single-pair answer). Note that they can also appear in the D-linked preverbal position. In this case the question receives a pair-list answer, although, as we have seen in Chapter (2), this is rare.

The d-structure represents a **Direct Question-Answer Pair** (Figure (8.26)):

**Question**

\[
\begin{align*}
+\text{PROM} & \quad \neg \text{D-LINKED} \left\{ \begin{array}{l}
\neg \text{D-LINKED} \\
\neg \text{D-LINKED} \\
\lambda x.\lambda y. \text{‘LEHET ÉS SZABAD DIÁKMUNKÁT VÁLLALNI’} \ (x, \ y) \\
\end{array} \right\} \\
\neg \text{D-LINKED} & \quad \neg \text{D-LINKED} \\
\end{align*}
\]

**Direct Q-A Pair**

**Answer**

\[
\begin{align*}
+\text{PROM} & \quad \neg \text{D-LINKED} \left\{ \begin{array}{l}
\text{‘AT COMPANIES’} \in \{ \text{PLACES} \} \\
\text{HAVING A STUDENT CARD} \in \{ P_1 \ldots P_n \ \text{(PRECONDITION)} \} \\
\end{array} \right\} \\
\neg \text{D-LINKED} & \quad \neg \text{D-LINKED} \\
\end{align*}
\]

Figure 8.25: F-structure: Hol ... lehet és szabad munkát vállalni?

Figure 8.26: D-structure: Hol ... lehet és szabad diák munkát vállalni?
Let us now consider multiple questions with coordinated question words having different functions. This structure complements the one dealt with in the previous section, since the question words coordinated in the preverbal position cannot denote the same type of set:

(25) ??/* Ki és kit hívott meg moziba?
    who and who.ACC invite.PST VM cinema.ILL

Derivational and non-derivational theories choose different ways to account for the coordination of unlike categories. Derivational frameworks usually assume the coordination of identical categories with ellipsis. Some examples still prove to be difficult to account for:

(26) I will always remember my professor and that languages are easy to learn.

In (26), an NP is coordinated with a clause and it would be difficult to reconstruct the part undergoing ellipsis in the first conjunct.

Non-derivational analyses, on the other hand (Sag, 2005; Dalrymple and Kaplan, 2000; Dalrymple et al., 2009), build their analyses on the under- or overspecification of the features of the conjuncts, assuming that what they share is the same syntactic function.

In the case of coordinated question words, it has been claimed (see Lipták (2001) on Hungarian and Skrabalova (2006) on Czech) that they can be coordinated if they are semantically/prosodically marked, i.e. focused. If focus is no longer assumed to be an f-structure attribute, the common function of the conjuncts can be the UDF. Cross-linguistically, the following generalization can be posited: on the one hand, in languages in which only constituents sharing the same function can be coordinated (like English and French), they must share all their functions in the f-structure (for instance, UDF and ADJunct); on the other hand, in Slavic languages and Hungarian, the conjuncts have to share at least one function, which can be the UDF.

However, this still would not explain the ungrammaticality of (25), as opposed to the coordination of ([ki és mikor] (who and when); [mit és hogyan] (what and how)). Moreover, the answers to multiple questions with coordinated question words are ungrammatical if the corresponding constituents (the foci) are coordinated:

(27) a. Mikor és mit ettél?
    when and what eat.PST.2SG
    What did you eat and when?

b. *Spagettit és tegnap ettem.
    spaghetti.ACC and yesterday eat.PST.1SG
    I ate spaghetti yesterday.

(based on Kálmán (2001, p. 126, 107))

The grammatical answer is an elliptical structure with forward ellipsis:
Thus, the assumption that question words can be coordinated in Hungarian because they share the function UDF has to be completed with further specifications.

In order to account for the issue at hand, we must first conclude that the basis of the coordination is not some categorical identity, or not only the functional identity of the conjuncts. Informally speaking, speakers use this construction as a last resort. They would like to indicate that they expect a single-pair answer (the cumulation of question words in the preverbal domain without coordination would require a pair-list answer), but since the question words do not denote the same type of set, they cannot use the structure in which one question word is preverbal and the other postverbal. This structure is thus licensed by two factors. First, both conjuncts have the function UDF, second, they do NOT denote the same type of set. Note that this is the case even if both question words are adjuncts (they do not denote the same type of set either). The first criterion is formalized by the c- and i-structures themselves. The second is purely semantic. The question words must not denote the same type of set.

From the data we have seen above, it seems (in LFG terms) that in Hungarian (unlike French and also Czech) sharing all the functions or at least one at the level of f-structure is not enough for the coordination to be grammatical in this specific case of interrogative words. We also need some semantic restrictions with respect to the denotation of the question words: the denotation of the question words (represented again at the level of i-structure) cannot be the same. The exact formalization of this aspect is beyond the scope of the present analysis and, according to Mary Dalrymple (p.c.), it necessitates further research, possibly in the framework of linear logic (glue semantics).

The above observations lead us to the following structures (Figures (8.27)-(8.29)).

(29) a. Q: -Ki és miért bántotta a sértettet?
   who and why maltreat.PST the aggrieved.ACC
   Who maltreated the aggrieved and why?

b. A: -János bántotta, mert részeg volt.
   János maltreat.PST, because drunk was
   John maltreated him, because he was drunk.

The c-structure (8.27) shows that the coordinated interrogative words occupy the preverbal, prominent, non-D-linked position together. This reflects that prominence is also a syntactic notion: in a given language, syntactic positions can be prominent. In the f-structure (Figure (8.28)), the question words share the function UDF, but not their other grammatical function: subject and adjunct.
8.3. Multiple Questions

Figure 8.27: C-structure: Ki és miért bántalmazta a sértettet?

Figure 8.28: F-structure: Ki és miért bántalmazta a sértettet?
The D-structure is again a Direct Question-Answer Pair (Figure (8.29)).

Question

\[
\begin{align*}
+\text{PROM} & \quad \neg \text{D-LINKED} \quad \left\{ x \mid (x) \in \{\text{PEOPLE}\} \right\} \\
\text{D-LINKED} & \quad \left\{ y \mid (y) \in \{P_1 \ldots P_n \text{ (REASON)}\} \right\} \\
-\text{PROM} & \quad \neg \text{D-LINKED} \\
\text{D-LINKED} & \quad \left\{ \lambda x, \lambda y. \text{‘BÁNTALMAZTA A SÉRTETTET’} (x, y) \right\}
\end{align*}
\]

Direct Q-A Pair

Answer

\[
\begin{align*}
+\text{PROM} & \quad \neg \text{D-LINKED} \quad \left\{ \text{‘JÁNOS’} \in \{\text{PEOPLE}\} \right\} \\
\text{D-LINKED} & \quad \left\{ \text{‘RÉSZEG VOLT’} \in \{P_1 \ldots P_n \text{ (REASON)}\} \right\} \\
-\text{PROM} & \quad \neg \text{D-LINKED} \\
\text{D-LINKED} & \quad \left\{ \text{‘BÁNTALMAZTA A SÉRTETTET’} (\text{JÁNOS, részeg volt}) \right\}
\end{align*}
\]

Figure 8.29: D-structure: Ki és miért bántalmazta a sértettet?

8.4 Beyond Multiple Questions

In this section, we take a look at questions that contain a focus and a question word. In the introduction, I presented two types, which both contain two prominent preverbal constituents. We will consider those types in this section. The first example is the following:

(30) -Jó, de JÁNOS mit hozott?
    ok, but János what bring\text{,\text{PST}}
    Ok, but what about John, what did he bring?

The c-structure of this sentence is the following: Figure (8.30).

The f-structure is represented in (8.31).

As was explained in Chapter (4), this question originates in a pair-list multiple question like (13), which has not been exhaustively answered. The question in (30) asks the original question again and continues the strategy employed by the interlocutor at the same time. This is why the question word appears in the preverbal position, and the contrasted constituent is focused (highlighted). This can be illustrated by the D-structure in Figure (8.32).
Figure 8.30: C-structure: JÁNOS mit hozott?

Figure 8.31: F-structure: JÁNOS mit hozott?
Figure 8.32: D-structure: Jó, de JÁNOS mit hozott?
Like in the previous cases of pair-list answers, the intended meaning of the question is the set of functions, from all the elements of the domain set into the range:

\[(31) \quad \lambda f \forall x \in \text{PEOPLE} \rightarrow \text{hozott a bulira } (f(x)), \ (f(x))\in(y)\]

The second example with a complex discourse structure was the following:

\[(32) \quad -\text{Ki NEM a Hamletet olvasta? who not the Hamlet.ACC read.PST Who didn’t read Hamlet (but something else)?}\]

This question can be asked in a context, in which every student in a class was supposed to read a Shakespeare play. After asking ten of them, who all read Hamlet, the teacher wants to know if there is someone who has read another play, not Hamlet. The original question can sound as:

\[(33) \quad \text{Ki melyik Shakespeare darabot olvasta? who which Shakespeare play.ACC read.PST Who read which Shakespeare play?}\]

The f-structure belonging to (32) is shown in Figure (8.33).

![Figure 8.33: F-structure: Ki NEM A HAMLETET olvasta?](image-url)

The d-structure is illustrated in Figure (8.34).
Figure 8.34: D-structure: Ki NEM A HAMLETET olvasta?
8.5. Summary

The intended meaning of the superquestion is, like before, a set of functions:

\( \lambda f \forall x \in \{ \text{STUDENTS} \} \rightarrow \text{olvasta} \ (y), \ (y) \in \{ \text{SHAKESPEARE PLAYS} \} \)

8.5 Summary

In this chapter, we looked at the LFG analysis of all multiple question types in Hungarian in the framework adopted in this thesis. First, I proposed a flat c-structure for the Hungarian sentence, dominated by the non-configurational S-node. I argued that there is no need to posit a VP constituent in Hungarian (except for infinitives), and that linear precedence reflects scope relations as well. In addition, contrary to previous analyses I assumed two prominent positions between quantifiers and the verb, one of them D-linked, the other non-D-linked, which can host a focus and a question word, but not two (or more) foci (which are both non-D-linked). Multiple questions containing preverbal question words either license a pair-list reading, represented as an implicative answer in the d-structure, or, when the question words form a complex question word, occupying the non-D-linked preverbal position, they are answered by a single pair (Direct Q-A Pair in the d-structure). Multiple questions containing one preverbal and one right-peripheral question word are usually answered by a single pair (Direct Q-A Pair), although the pair-list answer is not excluded either. In this latter case, I assume that the D-linked question word resides in the D-linked preverbal position or on the right periphery (this prominent position can host D-linked and non-D-linked elements as well, like contrastive topics or foci). Finally, in multiple questions with coordinated question words, the question words are supposed to occupy the same (mostly the non-D-linked) preverbal position. The question words can be of different categories or functions, the constraint on them is semantic: they must not denote the same type of set. In the next chapter, we go on to the analysis of the French data.
Chapter 9

Analysis of Multiple Questions in French

Contents

9.1 Introduction ............................................. 353
9.2 Syntactic Structure (C-structure) ..................... 353
9.3 Multiple Questions ..................................... 357
   9.3.1 One left peripheral/in situ question word(s) ...... 357
   9.3.2 Multiple questions with coordination ............... 375
9.4 Summary .................................................. 381

9.1 Introduction

This chapter is about the application of the LFG analysis to the French data. Like in the case of Hungarian, first I consider a possible c-structure for French. Then I proceed to the LFG representation of the different multiple question types.

9.2 Syntactic Structure (C-structure)

The syntactic structure of French is a long debated issue. Although hierarchical structures have been assumed, following the transformational analyses worked out for the English language, it has also been argued that a VP projection (if it is supposed to contain the verb and its complements but not the subject) is difficult to motivate in French (see Abeillé (1997, 2002, 2007), citing also Tesnière 1959, Gross 1975). The most important reasons are the positions of sentence adverbs, and the free permutation of the complements with adjuncts and inverted subjects in the postverbal domain. Sentence adverbs can appear at the beginning of the sentence, and between the subject and the verb, or between the verb and its complements as well:

(1) a. Généralement, Paul prend le métro.
   generally, Paul takes the subway
   Paul usually takes the subway.

   b. Paul, généralement, prend le métro.
c. Paul prend généralement le métro.  

(Abeillé, 2007, p. 108)

In the postverbal domain, complements can be interspersed with inverted subjects and adjuncts:

(2)  

a. Je sais quand parlera Jean à Marie.  
   I know when talk.FUT Jean to Marie  
   I know when John talks to Mary.

b. Je sais quand Jean parlera à Marie.

(Abeillé, 2007, p. 77)

(3)  

a. Le directeur parlera demain à Marie.  
   the director talk.FUT tomorrow to Marie  
   The director will talk to Mary tomorrow.

b. Le directeur parlera à Marie demain.

(Abeillé, 2007, p. 109)

Based on these criteria and on the flexible c-structure of the LFG framework, I conclude that a hierarchical structure comprising a VP constituent is not necessary in French either and follow Abeillé (2007)’s proposal of a flat structure. Nevertheless, a V’ node is assumed to host verb forms with lexical pronominal affixes (clitics), and the VP node is only used in the case of non-finite verb forms with a non-realized subject (infinitives, participles).

The proposed c-structure of example (3-a) is illustrated in Figure (9.1).

Figure 9.1: C-structure: Le directeur parlera demain à Marie.

Left- and right dislocated constituents are adjoined to the sentence (see also Abeillé (2007)):

(4) Left dislocation:
9.2. Syntactic Structure (C-structure)

Cette femme, Marie n’a pas confiance en elle.
This woman, Marie PRT has not confidence in her

This woman, Mary does not trust her.

(5) Right dislocation:
Marie ne le connaît pas, cet homme.
Marie PRT CL..ACC know not, this man

Mary doesn’t know him, this man.

The c-structures are shown in Figures (9.2) and (9.3), respectively.

Figure 9.2: C-structure: Cette femme, Marie n’a pas confiance en elle.

Figure 9.3: C-structure: Marie ne le connaît pas, cet homme.

In (5), the object is incorporated into the verbal form: le connaît. This can be represented in the lexical entry of the verb:
(6) \( \text{le connaît:} \quad (\uparrow \text{PRED}) = \text{‘know} < (\uparrow \text{SUBJ}), (\uparrow \text{OBJ})> \)  
\( (\uparrow \text{OBJ PRED}) = \text{‘pro} \)  
\( (\uparrow \text{OBJ NUM}) = \text{SG} \)  
\( (\uparrow \text{OBJ PERS}) = 3 \)  
\( (\uparrow \text{OBJ GEND}) = \text{MASC} \)

I assume that the S-adjoined positions both in left- and right dislocation are mapped onto the +PROM, (in these cases) D-linked set at the level of i-structure. As Abeillé et al. (2008) argue, dislocated constituents cannot always be identified with a discourse function. As far as question phrases are also analyzed as extracted and S-adjoined, they are mapped onto the non-D-linked set.

As far as interrogatives are concerned, I assume that left-peripheral question words are also adjoined to S. Contrary to dislocated constituents, question phrases bear a grammatical function within the clause. In addition, when they are extracted to the left periphery they are also assigned the already introduced \textit{Unbounded Dependency Function} (UDF) (see Asudeh (2010) and the explanations in Chapter (8)).

Let us have a look at an example:

(7) Quel livre donneras-tu à Marie?
Which book give.cl.nom2sg to Marie
Which book will you give to Mary?

The c-structure is illustrated in Figure (9.4).

![Figure 9.4: C-structure: Quel livre Jean donne-t-il à Marie?](image-url)

As discussed in Chapter (6), French exhibits interrogative verb forms with pronominal lexical affixes. In this case, the subject is incorporated in the verbal form, like the object in (4). Unlike Hungarian, syntactic positions cannot be unanimously associated with i-structure sets and subsets in French. However, there are also prominent syntactic positions in French, such as those adjoined to S, (see
Figure (5)). I assume, furthermore, that in situ question words are highlighted prosodically and it is due to the prosody-information structure mapping that they are represented in the + PROMINENT set at the level of i-structure.

In the next section, we go on to the discussion of multiple questions in French.

9.3  Multiple Questions

As we have seen in Chapter (2) about French multiple questions, if both question words are arguments, both the pair-list and the single-pair readings are possible. If one of them is an adjunct, the pair-list reading is preferred, most probably because there is an alternative, bi-clausal structure, in which the adjunct is coordinated sentence-finally, which licenses only a single-pair reading. This latter is not considered as a true multiple question. In addition to the argument-adjunct dichotomy, the extracted/in situ position of the question words also has to be considered, although this does not influence the type of answer expected. In the first type of multiple questions, either one question word is extracted and the other is in situ, or (mostly in spoken language) both are in situ. In the second type of multiple questions, the question words are coordinated, either in an extracted, or in an in situ position. The coordinated question words have to share the same function. If the question words are adjuncts coordinated sentence-finally, the question is difficult to distinguish from conjoined single questions. Multiple questions with coordinated question words can also license both answer-types.

9.3.1  One left-peripheral and one in situ question word/all question words in situ

As was pointed out in Chapter (2), the extracted or in situ position of the question words in a multiple question does not influence the type of the answer expected. On the difference between them, see Section (6.6.1) in Chapter (6). Let us first consider the case in which both question words are arguments.

9.3.1.1  Arguments

If both question words are arguments, both a pair-list and a single-pair reading are possible:

(8)  a. Q:  "Qu’est-ce que tu as donné à qui ?
    What (is it that) you give to whom
    Single-pair:
    I gave a book to Mary.
    Pair-list:
    c. A2:  J’ai donné un livre à Marie, et un stylo à Paul...
    I gave a book to Mary, and a pen to Paul...
I gave a book to Mary, and a pen to Paul.

What we see in this case is that the position of the question words is not revelatory with respect to the answer expected. As we have seen above, the existence of a pair-list reading depend on the presence of a D-linked question word in the question, the denotation of which can be mapped onto the set denoted by the other. I assume that this is also the case in French. The indirect evidence for this assumption is that, in structures where the question words appear within the same clause, the pair-list reading is always available, whereas only the single-pair reading is possible in bi-clausal structures (except for multiple questions containing sentential adverbs in French, see Chapter (2)), indicating the fact that the mapping function between two question words is a local relation (see also Krifka (2001)) restricted to clauses.

The other interesting factor is that contrary to Hungarian, where the D-linked or Sorting key question word is identifiable from the syntactic position, this difference is not manifested in the syntax in French. Since word order is more rigid in French, the order, or the syntactic position of the question words, does not change systematically according to the discourse-status of the question words. Consider the following examples (repeated from Chapter (2)):

(9) a. Q: -Quel groupe a visité quel monument ?
   which group AUX visited which monument
   Which group visited which monument?

   b. A1: -Les linguistes ont visité la Tour Eiffel, les psychologues
      the linguists AUX visited the Tower Eiffel, the psychologists
      l’Arc de Triomphe...
      the Arc de Triomphe
      The linguists visited the Eiffel Tower, the psychologists the Arc de Triomphe...

   c. A2: -La Tour Eiffel, c’est les linguistes qui l’ont visitée,
      the tower Eiffel, it is the linguists who cl.acc.aux visited,
      l’Arc de Triomphe, c’est les psychologues...
      the Arc de Triomphe, it is the psychologists
      The Eiffel tower, it was visited by the linguists, the Arc de Triomphe,
      it was visited by the psychologists.

Both the question and the answer (A1) are ambiguous between a contextually determined set of tourists and a contextually determined set of monuments, whereas (A2) refers to a contextually given set of monuments. However, the interlocutor is aware of the fact that s/he has to enumerate exhaustively all the tourist groups or all the monuments and pair them up with an element of the other set. In the case in which the monuments are D-linked, left dislocation can also be used, but it is not obligatory (see A2). What helps, in this case, is the context, and
not syntactic or prosodic information. The contextual difference means that the difference between the question words is encoded in the information structure and not at the other levels of linguistic analysis. The parallel architecture of Lexical-Functional Grammar makes it possible to account for these differences, since it dissociates discourse functions from syntax and prosody.

Let us illustrate the analysis on the example (9). The c-structure of the question (in which the groups are D-linked) is illustrated in Figure (9.5), the f-structure in Figure (9.6). For an analysis based on Butt and King (1996)'s i-structure, see Gazdik (2010c).

\[
\begin{align*}
\text{S} & \quad \text{QuesP} \\
\uparrow \sigma \in (\uparrow \sigma + \text{PROM D-LINKED}) & \quad \uparrow = \downarrow \\
(\uparrow \text{SUBJ}) = & \quad \uparrow = \downarrow \quad \uparrow \sigma \in (\uparrow \sigma - \text{D-LINKED}) \\
\text{Quel groupe} & \quad \text{a} \quad \text{visité} \quad \text{quel monument?}
\end{align*}
\]

Figure 9.5: C-structure: Quel groupe a visité quel monument?

\[
\begin{align*}
\text{[PRED} & \quad \text{‘visiter}(\uparrow \text{SUBJ})(\uparrow \text{OBJ})]\text{]}
\text{SUBJ} & \quad \text{PRED} \quad \text{‘pro} \\
\text{OBJ} & \quad \text{PRED} \quad \text{‘pro} \\
\text{TENSE} & \quad \text{PAST}
\end{align*}
\]

Figure 9.6: F-structure: Quel groupe a visité quel monument?

In questions in which neither of the question words refer to the subject the f-structure (Figure (9.8)) belonging to the question in which one question word is extracted (10) would be the same as the one belonging to the question in which both question phrases are in situ (11). The difference between them can be demonstrated by the UDF attribute. Recall that these varieties of interrogatives in French suppose different discourse contexts, which also motivate an explicit distinction: Figures (9.7) and (9.8).

(10) Quels exercices as-tu donné à quel groupe?  
which exercises AUX you given to which group  
Which exercises did you give to which group?

(11) Tu as donné quels exercices à quel groupe?  
you AUX given which exercises to which group
Figure 9.7: F-structure: Quels exercices as-tu donné à quel groupe?

Figure 9.8: F-structure: Tu as donné quels exercices à quel groupe?
Figure 9.9: D-structure: Quel groupe a visité quel monument?
As Figure (9.9) reflects, both groups are paired up with a monument, which is summed up in the following formula:

\[(12) \quad \lambda f \forall x \in \{\text{GROUPS}\} \rightarrow \text{visité} \ (y), \ (y) \in \{\text{MONUMENTS}\}\]

The question has another reading as well, in which the monuments are D-linked. The f-structure is, in this case, the same as in Figure (9.6); however, the c-structure annotations and the d-structure are different. Let us now consider these structures: Figures (9.10) and (9.11).

![Figure 9.10: C-structure: Quel groupe a visité quel monument?](image)
Figure 9.11: D-structure: Quel groupe a visité quel monument? La Tour Eiffel, ...
The semantic representation following from figure (9.11) is given below:

(13) \( \lambda f \; \forall x \; x \in \text{MONUMENTS} \rightarrow 'a \; \text{été \; visité \; par}' \ldots (f(x)), \; (f(x)) \in (y) \)

The discourse structure and the information structures reveal that the question phrase *quel groupe* is D-linked in this context, thus it is divided into two partitions (the linguists and the psychologists) and answered exhaustively in the answer (implying that the Eiffel Tower and the Arc de Triomphe are not the only monuments to see in the program offered by a tourist agency in Paris).

On the other hand, I illustrate the single pair answer with the following example.

(14) a. **Qui** chante **quoi** à la télévision ?
   who sings what on the television
   Who is singing what on TV?

   b. Yves Montand chante une chanson.
   Yves Montand sings a song
   Yves Montand is singing a song.

The c-, f-, and d-structures are shown in Figures (9.12), (9.13) and (9.14) respectively.

Figure 9.12: C-structure: Qui chante quoi à la télévision?

Figure 9.13: F-structure: Qui chante quoi à la télévision?

No we go on to argument-adjunct multiple questions.
In multiple questions containing an argument and an adjunct, both readings are possible. Since the single-pair answer has a bi-clausal alternative, the pair-list reading can be preferred over the single-pair one, especially if it is reinforced by D-linked question words, such as *quel* (17). Again, either one question word is extracted and the other is *in situ*, or both are *in situ*.

(15) Pair-list:

a. Q: -En famille, *qui* dort *où* ?
   in family, who sleeps where
   In the family (home) who sleeps where?
   (http://generation-blogueurs.blogs.lavoixdunord.fr)

b. A: -Les parents dorment au premier étage et les enfants
   the parents sleep on the first floor and the children
   au deuxième.
   on the second
   The parents sleep on the first floor and the children on the second.

(16) Single-pair:
a. Q: **Quand est-ce que** tu as vu **qui** ?
   When is it that you seen whom
   When did you see whom?

b. A: -J’ai vu Jean ce matin.
   I saw John this morning.

(17) Pair-list:

a. Q: -Tu rencontres **qui à quelle heure** ?
   you meet whom at which hour
   With whom are you meeting when?

b. A: -Je rencontre Max à 17 heures et Léo à 18 heures.
   I meet Max at 17 hours and Léo at 18 hours
   I’m meeting Max at 5 pm and Léo at 6 pm.

The c-, f-, and d-structures of example (15) are shown in Figures (9.15), (9.16) and (9.17), respectively.

![C-structure](image)

Figure 9.15: C-structure: En famille qui dort où?

![F-structure](image)

Figure 9.16: F-structure: En famille qui dort où?
Figure 9.17: D-structure: En famille qui dort où ? Les parents . . . , les enfants . . .
Example (16) with a single-pair reading is illustrated in Figures (9.18)-(9.20).

\[
\begin{array}{c}
\text{QuesP} \\
\uparrow_{\sigma} \in (\uparrow_{\alpha} + \text{PROM} \& \neg \text{D-LINKED}) \\
\downarrow \in (\uparrow_{\text{ADJ}}) \\
\downarrow = (\uparrow_{\text{UDF}}) \\
\end{array}
\]

\[
\begin{array}{c}
S \\
\text{Quand} \\
\text{est-ce que tu as vu qui ?} \\
\end{array}
\]

Figure 9.18: C-structure: Quand est-ce que tu as vu qui?

The c-structure in (9.18) illustrates another instance of incorporation. The subject-clitic (affix) is part of the verb form, and thus no other subject is needed in the sentence. The lexical entry of the verb form is the following:

\[
\begin{align*}
(18) \quad & \text{tu as:} \quad (\uparrow \text{ASP}) = \text{PERF} \\
& (\uparrow \text{SUBJ PRED}) = \text{'pro'} \\
& (\uparrow \text{SUBJ NUM}) = \text{SG} \\
& (\uparrow \text{SUBJ PERS}) = 2
\end{align*}
\]

\[
\begin{array}{c}
\text{voir} \\
\text{pro} \left\langle \left( \text{SUBJ} \right) \left( \text{OBL} \right) \right\rangle \\
\text{pro} \\
\text{pro} \\
\text{PAST}
\end{array}
\]

Figure 9.19: F-structure: Quand est-ce que tu as vu qui?

The d-structure (Figure (9.20)) illustrates a Direct Question-Answer Pair. Example (17) is illustrated in Figures (9.21)-(9.23).

The verb form \textit{tu rencontres} is analyzed as follows:

\[
\begin{align*}
(19) \quad & \text{tu vas:} \quad (\uparrow \text{SUBJ PRED}) = \text{'pro'} \\
& (\uparrow \text{SUBJ NUM}) = \text{SG} \\
& (\uparrow \text{SUBJ PERS}) = 2
\end{align*}
\]
9.3. Multiple Questions

Direct Q-A Pair

Figure 9.20: D-structure: Quand est-ce que tu as vu qui ? J’ai vu Jean ce matin.

Figure 9.21: C-structure: Tu rencontres qui à quelle heure?

Figure 9.22: F-structure: Tu rencontres qui à quelle heure?
Figure 9.23: D-structure: Tu rencontres qui à quelle heure?
Figure (9.23) shows that the question is split into two subquestions, which are answered separately in the answer. The interpretation is, again, a set of functions:

\[ \lambda f \forall x \in \{\text{PEOPLE}\} \rightarrow \text{tu vas chercher (} y \text{), } (y) \in \{\text{POINTS IN TIME}\} \]

In the next subsection, we look at questions containing adjuncts.

### 9.3.1.3 Adjuncts

In multiple questions containing adjuncts, again, both readings are possible. Some question phrases (quel) can reinforce the pair-list reading, like in (21) when they stand with a countable noun, but they can also be answered by a single pair (23), when the noun they modify is abstract or uncountable. One question word can be preverbal ((21) and ((22)), or both question words can appear in situ (23).

### Pair-list:

a. Q: -Où Jean a dormi quel jour ?
   where Jean AUX slept which day
   Where did John sleep on which day?

b. A: -Lundi, Jean a dormi à Rome, mardi à Nice, mercredi Monday Jean AUX slept in Rome, Tuesday in Nice, Wednesday
   in Cannes.
   On Monday, John slept in Rome, on Tuesday in Nice, on Wednesday
   in Cannes.

### Single-pair:

a. Q: -Quand est-ce qu’il est arrivé avec quelle
   when is it that he AUX arrived with what kind of
   intention?
   When did he arrive and what was his intention?

b. A: -Il est arrivé lundi pour rencontrer Jean.
   he AUX arrived Monday to meet Jean
   He arrived on Monday to meet John.

### Single-pair:

a. Q: -Il est arrivé quand avec quelle intention ?
   he AUX arrived when with what kind of intention
   When did he arrive and what was his intention?

b. A: -Il est arrivé lundi pour rencontrer Jean.
   he AUX arrived Monday to meet Jean
   He arrived on Monday to meet John.
First, we examine the LFG representation of (21). The c-, f-, and (schematic) d-structures are illustrated in Figures (9.24)-(9.26).

![Diagram of C-structure](image)

Figure 9.24: C-structure: Où Jean a dormi quel jour?

![Diagram of F-structure](image)

Figure 9.25: F-structure: Où Jean a dormi quel jour?

A: Question

\[
\begin{align*}
&\text{A: Question} \\
&\text{+PROM} \quad \neg \text{D-LINKED} \left\{ y \mid (y) \in \{\text{PLACES}\} \right\} \\
&\text{D-LINKED} \quad \left\{ x \mid (x) \in \{\text{DAYS}\} \right\} \\
&\text{−PROM} \quad \neg \text{D-LINKED} \\
&\text{D-LINKED} \quad \lambda x.\lambda y.'\text{JEAN A DORMI} (x, y)'
\end{align*}
\]
9.3. Multiple Questions

A: Question

Implicative Q-A Pairs

(B: Subquestion) (D: Subquestion) (F: Subquestion)

C: Answer E: Answer G: Answer

Figure 9.26: Où Jean a dormi quel jour ? Lundi, Jean a dormi à Rome . . . .

B: Subquestion

\[
\begin{align*}
+\text{PROM} & \neg \text{D-LINKED} \left\{ y \mid (y) \in \{\text{PLACES}\} \right\} \\
\text{D-LINKED} & \left\{ \text{"lundi"} \in \{\text{DAYS}\} \right\} \\
-\text{PROM} & \neg \text{D-LINKED} \\
\text{D-LINKED} & \left\{ \lambda y.\text{"LUNDI JEAN A DORMI"} (y) \right\}
\end{align*}
\]

C: Answer

\[
\begin{align*}
+\text{PROM} & \neg \text{D-LINKED} \left\{ \text{"ROME"} \in \{\text{PLACES}\} \right\} \\
\text{D-LINKED} & \left\{ \text{"LUNDI"} \in \{\text{DAYS}\} \right\} \\
-\text{PROM} & \neg \text{D-LINKED} \\
\text{D-LINKED} & \left\{ \text{"LUNDI JEAN A DORMI"} \left( \text{Rome} \right) \right\}
\end{align*}
\]
The question and the answer together give as a result the mapping function between the three days when Jean was travelling, and the the cities in which he slept each day:

\[(24) \quad \lambda f \forall x \in \{\text{DAYS}\} \rightarrow \text{Jean a dormi} (y), (y) \in \{\text{PLACES}\}\]

The structures also illustrate the fact that in French, D-linked question words have no constant position like in Hungarian. They can be sentence-final, following the other question word. However, the answer is organized with respect to this question word, since in (21), the question is split according to the relevant days.
9.3. Multiple Questions

(Monday, Tuesday and Thursday), which appear sentence-initially in the answers, followed by the cities.

Now we examine the representation of (22) with a single-pair answer (Figures (9.27)-(9.29)).

![Diagram](image)

Figure 9.27: C-structure: Quand est-ce qu’il est arrivé avec quelle intention?

![Diagram](image)

Figure 9.28: F-structure: Quand est-ce qu’il est arrivé avec quelle intention?

9.3.2 Multiple questions with coordination

Unlike in Hungarian, coordinated question words must share the same function in French at the level of f-structure. The shared function can be for instance OBL (25) or ADJ ((26) and (27)). The coordinated question words can appear in situ or in a preverbal position, and like in the previous cases, both readings are possible, depending on the context.

OBL questions words:

(25) Single-pair:
Figure 9.29: D-structure: Quand est-ce qu’il est arrivé avec quelle intention? ... 

a. Q: -La conférence a eu lieu où et quand ?
   the conference AUX had place where and when
   When and where did the conference take place?

   she AUX had place in Paris in 2007
   It took place in Paris in 2007.

ADJ question words:

(26) Single-pair:

   a. Q: -Pourquoi et quand avez-vous décidé d’arrêter vos études
      why and when AUX-you decided to stop your studies
      university
      Why and when did you decide to stop your academic studies?
      (Rochefort, Christiane (1978) Ma vie revue et corrigée par l’auteur à

   b. A: -J’ai décidé de les arrêter en 2001, parce que je n’en
      I AUX decided to them stop in 2001, because I PRT-CL.PART
      voyais plus aucune perspective.
      saw anymore no perspective
      I decided to stop them in 2001, because I couldn’t see any perspective
9.3. Multiple Questions

thereof anymore.

ADJ question words:

(27) Pair-list:

a. Q: -Quand et pourquoi voit-on circuler des trains sans voyageurs? (SNCF)
   When and why can we see trains without passengers?
   (http://www.infolignes.com/article.php3?id_article=3505)

b. A: -Par mesure de sécurité, chaque matin avant les premières circulations commerciales, un TGV-balai effectue un aller-retour trip sur toutes les lignes à grande vitesse; après un incident, une rame peut repartir à vide quand les voyageurs ont été orientés vers une rame de substitution; en période de vacances, la SNCF fait venir un grand nombre de trains de province dans la capitale pour assurer tous les départs, etc.
   For security reasons, every morning before the first return trip on all the lines with high speed; after an accident, a train can leave empty when the passengers have been oriented towards a train of substitution; in period of holidays, the SNCF makes come a big number of trains of country into the capital in order to assure all the departures.

In what follows, we will consider the LFG representations of examples (25) and (27). The c-, f- and d-structures of (25) are shown in Figures (9.30), (9.31), and (9.32), respectively. For an LFG analysis based on Butt and King (1996)’s information structure, see Gazdik (2010b).

The structures belonging to (27) are shown in Figures (9.33)-(9.35).

A: Question

\[
\begin{align*}
\text{A: Question} & \\
\text{\{d-linked \{y | (y)\in\{p_1...p_n\text{(reason)}\}\}\}} & \\
\text{\{d-linked \{x | (x)\in\{\text{POINTS IN TIME}\}\}\}} & \\
\text{\{\}} & \\
\text{\{d-linked \{\lambda x,\lambda y,'on voit circuler des trains sans voyageurs' (x, y)\}\}} & \\
\end{align*}
\]
La conférence a eu lieu où et quand?

Figure 9.30: C-structure: La conférence a eu lieu où et quand?

```
Figure 9.31: F-structure: La conférence a eu lieu où et quand?

```

```
Figure 9.32: D-structure: La conférence a eu lieu où et quand?

```
9.3. Multiple Questions

Figure 9.33: C-structure: Quand et pourquoi et voit-on circuler des trains . . . ?

Figure 9.34: F-structure: Quand et pourquoi voit-on circuler des trains . . . ?

Figure 9.35: Quand et pourquoi voit-on circuler des trains . . . ?
B: Subquestion

\[
\begin{align*}
\text{+PROM} & \quad \neg \text{D-LINKED } \{ y \mid (y) \in \{ p_1 \ldots p_n \} \} \\
\text{D-LINKED} & \quad \{ \text{CHAQUE MATIN} \in \{ \text{P. IN TIME} \} \} \\
\text{PROM} & \quad \neg \text{D-LINKED } \{ y \mid \text{\'CHAQUE MATIN, ON VOIT CIRCULER DES TRAINS...'} (y) \} \\
\end{align*}
\]

C: Answer

\[
\begin{align*}
\text{+PROM} & \quad \neg \text{D-LINKED } \{ \text{\'SÉCURITÉ'} \in \{ p_1 \ldots p_n \} \} \\
\text{D-LINKED} & \quad \{ \text{CHAQUE MATIN} \in \{ \text{P. IN TIME} \} \} \\
\text{PROM} & \quad \neg \text{D-LINKED } \{ y \mid \text{\'CHAQUE MATIN, ON VOIT CIRCULER DES TRAINS...'} (\text{sécurité}) \} \\
\end{align*}
\]

D: Subquestion

\[
\begin{align*}
\text{+PROM} & \quad \neg \text{D-LINKED } \{ y \mid (y) \in \{ p_1 \ldots p_n \} \} \\
\text{D-LINKED} & \quad \{ \text{\'APRÈS UN INCIDENT} \in \{ \text{P. IN TIME} \} \} \\
\text{PROM} & \quad \neg \text{D-LINKED } \{ y \mid \text{\'APRÈS UN INCIDENT ON VOIT CIRCULER DES TRAINS...'} (y) \} \\
\end{align*}
\]

E: Answer

\[
\begin{align*}
\text{+PROM} & \quad \neg \text{D-LINKED } \{ \text{\'rame de substitution'} \in \{ p_1 \ldots p_n \} \} \\
\text{D-LINKED} & \quad \{ \text{\'APRÈS UN INCIDENT} \in \{ \text{P. IN TIME} \} \} \\
\text{PROM} & \quad \neg \text{D-LINKED } \{ y \mid \text{\'APRÈS UN INCIDENT ON VOIT CIRCULER DES TRAINS...'} (\text{rame de substitution}) \} \\
\end{align*}
\]
In this chapter, I applied the analysis to the French data. First, I proposed a flat c-structure for French, without a VP node. Then I showed, with the help of the c-, f-, and d-structures, how the structural and interpretive differences between the different multiple question types can be demonstrated. In multiple questions requiring pair-list answers, at least one of the question words is D-linked, but it cannot be identified via syntax in French, since it can appear on the left or on the right periphery. The clue is, in this case, the information structure of the question, which determines the type and structure of the answer. In multiple questions containing an adjunct, the pair-list reading is preferred, since the alternative, bi-clausal structure with final coordination licenses only single pair answers. In multiple questions containing coordination, the question words must share the same function at the level of f-structure. They usually belong to the prominent, non-D-linked set at i-structure, but it is also possible that one of them is D-linked and distributes on the other, in which case they appear in different subsets in the i-structure.
10.1 Main Results

The overarching goal of this thesis was to offer a comprehensive analysis of multiple questions in Hungarian and French in the framework of Lexical-Functional Grammar. Since no such analysis can be conducted without reference to the information- and discourse-structure, I also proposed an LFG representation of the syntax-discourse interface that takes into consideration a wider range of data in the two languages than multiple questions.

In the first part of this thesis I presented, with the help of a large number of attested examples, all the possible structures of multiple questions in the two languages, along with the type of answer they license. In Hungarian, three types of (monoclausal) multiple questions were identified. The first, in which all question words appear in a preverbal position, is usually answered with a pair-list, and explicitly indicates which question word is to be mapped on the other in the answer. In the second, we find one preverbal and one postverbal question word, which (in most cases) have to denote the same type of set. The third type contains coordinated question words in a preverbal position, and even question words with different functions can be coordinated in this structure.

In French, four structures can be distinguished: the question words can be coordinated or not, and one of them (or the coordinated question words) can appear extracted or in situ. The argument/adjunct status of the question words influences the expected reading: multiple questions containing two arguments can be answered both by a pair-list or a single pair depending on the context, whereas those containing an adjunct question word are usually answered with a pair-list, since another structure (the clause-final coordination of the adjunct) can only be answered with a single pair. An important exception identified was that of sentence adverbs that must appear coordinated in a second (elliptical) clause, but the multiple question can be answered with a pair-list.

We concluded that the prerequisite of a pair-list reading is the possibility of thematizing one of the elements of the question in the subparts of the answer. This possibility is usually but not exclusively due to the presence of a \textit{D-linked} question.
word in the same clause as the other question word. Since the pair-list reading is available also in coordinated structures, it was admitted that multiple questions containing coordinated question words are not always biclausal and elliptical, but the coordination can apply at the level of the lexical items.

In order to proceed in the analysis, a definition of discourse functions and their subtypes had to be provided, as they are used in this thesis. Focus was defined as the semantically prominent (Jacobs, 1984) element of replies (answers, corrections, etc.). It has three subtypes: contrastive focus (in case the alternatives the focus is contrasted to are explicitly present in the context), information focus, which is an answer to a constituent question, and completive information,\(^1\) which is un presupposed, but not highlighted material. Focusing is understood only in the semantic-pragmatic sense, and does not refer to formal means of highlighting: placing in a salient syntactic position, or salience lending prosody.

The discourse function topic is defined as the part of the sentence that links it to the current Discourse Topic. Two subtypes have been adopted, Thematic Shifters and Contrastive Topics. The former introduces a subtopic of the Discourse Topic, mostly but not exclusively in narrative contexts, whereas the latter indicates the presence of a complex discourse strategy, mostly in answers to questions (Büring, 2003).

The highlighting of these discourse functions is related to syntactic positions in Hungarian (although prosodic highlighting also plays a role with respect to contrastive topics, foci, and constituents on the right periphery), whereas in French, the role of prosody is more significant (illocutionary boundary tones, C-accent). However, spoken/informal French also uses syntactic means: dislocation with/without a resumptive element.

The LFG representation of the two languages is then based on these assumptions. Since the main focus of this thesis is on the syntax-discourse interface, the representation of prosody is set aside, but a detailed description can be found in Mycock (2006). The most important assumptions of the proposed analysis are the following:

- Syntactic structure (both c- and f-structure) is discourse-neutral in the sense that it does not contain discourse functions. However, discourse functions can be associated with certain syntactic positions via functional annotations, which is more characteristic of discourse-configurational languages, such as Hungarian, than configurational languages, like French.

- As mentioned above, discourse functions are defined as semantically and pragmatically prominent, which is clearly distinguished from formal highlighting. Syntax and prosody have a limited number of means for highlighting. This is why some syntactic positions, intonation patterns or accent types cannot be exclusively associated with a certain discourse function.

\(^1\)Completive information is not prominent in the same sense as contrastive focus and information focus, but it is considered as a subtype of focus, since it is unpresupposed and cannot be collapsed into background information.
• The information structure architecture proposed builds on the fundamental properties of the elements it has to accommodate. These are prominence and D-linkedness. Prominence is semantically defined, but it is usually accompanied by syntactic and/or prosodic highlighting. This is formalized by the mapping functions that exist between the c-structure and the i-structure, and between the p-structure and the i-structure, respectively. Foci, the hokus, and non-D-linked question words appear in the non-D-linked subset of the +PROMINENT set, whereas thematic shifters, contrastive topics, and D-linked question words in the D-linked subset. The −PROMINENT set contains completive information (non-D-linked), and background information (non-D-linked).

• Multiple questions licensing pair-list answers contain at least one D-linked question word. It is demonstrated that these question words occupy syntactic positions associated with this i-structure subset in Hungarian. No such correspondence is established for French. The semantic mapping from the denotation of the D-linked question word into the denotation of the other is represented at the i-structure level. No such mapping is available if both question words are non-D-linked (the answer to the question is, in this case, a single pair).

• It is also argued that in an analysis of multiple questions and discourse functions in general, the discourse level also has to be taken into consideration, since the factors that come into play go beyond the level of the individual sentences. This is what we observed in the definitions of discourse functions, or in the different answer types given to multiple questions (pair-list answers follow strategies). D-structure is proposed to consist of the information structures of the individual sentences, constituting a tree-diagram.

• With respect to the coordination of question words, it is assumed that in French, they must share all their functions at the level of f-structure (usually ADJ and UDF), whereas in Hungarian and Czech, it suffices if they share only one (UDF). This is supported by the fact that in those languages, only extracted question words can be coordinated. Moreover, it is argued that further semantic restrictions apply in Hungarian, according to which coordinated question words must not denote the same type of set (this is true for the coordination of two adjuncts and for the coordination of an argument and an adjunct). This way, this question type complements the one in which the question words (preferably) denote the same type of set.

• The d-structure representation encodes the possible discourse relations between the individual sentences it contains. These relations are listed and formalized in Asher and Lascarides (2003)’s SDRT. As far as questions are concerned, the relevant discourse relations are Direct Question-Answer Pair, Indirect Question-Answer Pair, and Partial/Complete Implicative Q-A Pair. Pair-list questions are usually answered with a Complete Implicative Answer, referring to the strategy indicated by the contrastive topics in the answer,
which can imply the presence of implicit sub- or superquestions as well. Single-pair questions can be part of a *Direct Question-Answer Pair*, an *Indirect Question-Answer pair*, or *Partial Implicative Question-Answer Pair*.

- Considering now the other discourse relations, it has been argued (*Gazdik and Winterstein, 2011*) that the *focus* corresponds to the distinguished element of the semantics of those relations that are built upon such a specific element. It is obviously placed in the *prominent preverbal position* (PPP) in Hungarian (*parallel, contrast*, etc.). If the semantics of a relation does not rely on a distinguished element, the PPP will host an element with some specific semantic meaning (identification in the case of the *hocus*, and *aspect* in the case of verbal particles). However, the issue awaits some further research.

The proposed framework can account for the syntactic and interpretive properties of all multiple question types in French and Hungarian. However, some aspects still necessitate further research.

### 10.2 Future Perspectives

There are at least three areas in which research can be fruitfully continued. One of them is the coordination of question words with different functions. This phenomenon, as pointed out earlier, is not restricted to Hungarian, but appears in Slavic languages (Russian) and in Romanian as well. On the one hand, it should be determined why this is possible in certain languages, but forbidden in others. On the other hand, it also has to be investigated what the cross-linguistic and language-specific properties of this unusual coordination are.

Further research is to be done with respect to French sentence adverbs, in order to give a precise account of why they cannot appear in the same clause as the other question word in pair-list questions.

Finally, the exact formalization of the relationship between the PPP in Hungarian and the various discourse-relations, especially in all-focus ("neutral") sentences also has to be investigated. Such an analysis could finally solve the long-standing problem of this position in the Hungarian sentence.
Bibliography


Anne Abeillé and Danièle Godard. Interrogative verbs in French. Going Romance Conference, Utrecht, 1999c. 122, 257


Anne Abeillé and Danièle Godard. Les interrogatives du français. Presented at the HPSG Reading seminar, April 2002. 122


Bibliography


Michael Brody and Kriszta Szendrői. Hungarian focus is an answer, hence exhaustive. To appear, 2011. 159


Mary Dalrymple. Glue and Information Structure. Paper presented at the LFG10 Conference, Carleton University, Ottawa, Canada, June 2010. 4, 291


Katalin É. Kiss. A magyar mondat szerkezet generatív leírása [The Generative Description of the Hungarian Sentence]. Akadémiai Kiadó, Budapest, 1983. 201


Anna Gazdik and Grégoire Winterstein. A Discursive Analysis of Discourse Functions in Hungarian. work in progress, 2011. 279, 314, 386


Andreas Haida and Sophie Repp. Monoclausal question word coordination across languages. to appear, 2011. 97


Konstantin I. Kazenin. On Coordination of Wh-Phrases in Russian. unpublished manuscript, Universität Tübingen, 2010. 50, 97


Tracy Holloway King. Focus domains and information structure. In Butt Miriam and Tracy Holloway King, editors, LFG97, University of California, San Diego, 1997. CSLI Publications. 6, 124, 126, 149, 151, 265, 266, 276, 291, 309.


Tibor Laczkó and György Rákosi. On Locative Dependencies Involving Particle Verbs in Hungarian. Talk presented at the Thematic Workshop on Spatial and Temporal Relations in LFG, April 2011. 189, 310, 311


Frédéric Laurens. French dislocates as incidentals licensed by anaphoric expressions. Paper presented at the HPSG 10 Conference, Université Paris 7, July 2010. 252, 289


International Conference on Head-Driven Phrase Structure Grammar, July 2010. 151


Jason Merchant. Spurious coordination in Vlach multiple wh-questions. in progress, 2009. 97


François Mouret. Les coordinations de termes dissemblables sont-elles elliptiques ? In *Actes du 1er Congrès Mondial de Linguistique Française (CMFL 08)*, 2008. 242, 244


Mark Newson and Vili Maunula. Word order in Finnish: whose side is the focus on? *The Even Yearbook*, 7, 2006. 119


Hana Skrabalova. Parataxe apparente et coordination des interrogatifs en tchèque. *Faits de langue*, 28, 2006. 4, 45, 49, 50, 97, 318, 343


Résumé

L’objectif de cette thèse est de proposer une analyse des questions multiples (Qui fait quoi ? Où et quand a eu lieu le dernier concert ?) dans le cadre formel de la Grammaire Lexicale-Fonctionnelle (LFG). L’analyse est centrée sur les aspects syntaxiques, discursifs, et sur leur interface. L’étude des questions multiples est englobée dans une approche plus générale des fonctions de discours et de la structure discursive qui toutes les deux jouent un rôle important dans l’analyse.

Les structures syntaxiques proposées sont « neutres » par rapport au discours dans le sens où elles ne contiennent pas de projections fonctionnelles étiquetées TopP, FocP, même dans des langues discours-configurationnelles, comme le hongrois. Les fonctions de discours sont traitées dans les structures informationnelle et discursive, distinguées systématiquement du soulignement syntaxique et prosodique.

Les deux niveaux représentationnels principaux sont la structure en constituants et la structure informationnelle. De plus, une possible représentation de la structure discursive est également esquissée. La structure informationnelle proposée ne contiennent pas les fonctions de discours comme unités de base, mais leurs propriétés fondamentales : la prémonience et l’ancrage discursif. Bien que prémonience soit considérée ici comme une notion sémantique, il est aussi admis que les constituants prémonients sont formellement soulignés. Ces observations sont formalisées à l’aide des fonctions de correspondance entre la structure en constituants et la structure prosodique d’un côté, et la structure informationnelle de l’autre. La première de ces fonctions de correspondance se manifeste dans le fait que certaines positions syntaxiques peuvent être systématiquement associées à ces propriétés informationnelles.

L’interprétation des questions multiples est dérivée de deux propriétés sémantico-pragmatiques des mots interrogatifs : la prémonience et l’ancrage discursif. En hongrois, les différents types de réponse attendus sont directement liés à la correspondance structure en constituants – structure informationnelle (des positions syntaxiques sont associées avec ces propriétés), tandis que la syntaxe n’est pas révélatrice en français de ce point de vue.

Un avantage important de cette approche c’est qu’elle prend en compte également les phrases dites "neutres" (ou all-focus), qui ne contiennent pas nécessairement un focus étroit ou un topique, mais d’autres éléments partageant certaines propriétés avec les topiques et les focus. En plus, l’analyse explicite la différence entre les mots qu discursivement ancrés et non-ancrés dans les questions multiples qui attendent une réponse de liste de paires.

En ce qui concerne la structure discursive, son architecture proposée contient les structures informationnelles des phrases individuelles qui constituent le discours en question. En plus, les relations discursives entre les énoncés individuels sont également intégrés dans le cadre formel.

La thèse propose non seulement une analyse compréhensive de tous les types de questions multiples en hongrois et en français, mais elles contribue à la recherche sur la structure informationnelle et discursive.
Összefoglalás

A dolgozat célja a magyar és francia többszörös kérdések elemzése a Lexikai-Funkcionális Grammatika (LFG) eszköztárával felhasználva. Az elemzés központjában a szintaktikai és diskurzussal kapcsolatos aspektusok és azok interélésével áll. A többszörös kérdések tanulmányozása részét képezi a diskurzusfunkciók és a diskurzus szerkezet egy átfogóbb megközelítésének, hiszen ezek fontos szerepet játszanak az elemzésben.

Amellett érveik, hogy a szintaktikai szerkezet diskurzus szempontjából semleges abban az értelemben, hogy nem tartalmaz diskurzusfunkciók (pl. topik, főkusz) számára kijelölt pozíciókat, még az ún. diskurzusconfigurációs nyelvekben sem. A diskurzusfunkciókat kizárólag szemantikai és pragmatikai alapon definiálom, és szisztematikusan megkülönböztetem a formai kiemelési technikáktól.

Két LFG-s reprezentációs szintet vizsgálunk részletesen, az összetevős szerkezetet és az információs szerkezetet. Ezeknek kívül kísérletet teszünk a diskurzus szerkezet egy lehetséges formalizálására is. A javasolt információs szerkezet alapegységei nem diskurzusfunkciók, hanem azok két alapvető tulajdonságára épülnek, a prominencia és az adott diskurzusot kötőtség (D-linkedness). Bár a prominenciát szemantikai fogalomként kezelem, a szerkezetek közötti leképezési függvények segítsével azt is formalizálni próbálom, hogy a prominens elemek formailag is ki vannak emelve. Ilyen leképezést feltételez az összetevős és az információs szerkezet, valamint a prozódia szerkezet és az információs szerkezet között. Az előbbi arra utal, hogy egyértelmű leképezés jöhet létre egyes nyelvekben bizonyos szintaktikai pozíciók és az említtet i-struktúra szerepek között.

A többszörös kérdések interpretációja (a rájuk adható válasz típusa) a magyarban egyértelműen levezethető az összetevős szerkezet és az információs szerkezet közti leképezésből. A párlista választ váró kérdések tartalmaznak legalább egy diskurzushoz kötött kérdőszót, amely szisztematikusan az elhez kapcsolódó pozícióban jelenik meg. A specifikus pár olvasatú kérdésekben nem találunk ilyen kérdőszót. A magyarral szemben a franciaiban a szintaxis kevésbé fejezi ki ezeket az információkat.

Egy fontos előnye ennek a megközelítésnek az, hogy számtud adni az un. semleges mondatokról, amelyek nem feltétlenül tartalmaznak (szük) főkusz összetevőt, sőt, előfordul, hogy topikot sem. Tartalmazhatnak azonban egy sor másfajta elemet, amelyek bizonyos tulajdonságai közötti osztalnának a topikkal és a fókussal. Ezen kívül az elemzés egyértelműsíti a diskurzushoz kötött és nem kötött kérdőszavak viszonyát a párlista olvasatú többszörös kérdésekben.

A javasolt diskurzus szerkezet az adott diskurzust alkotó mondatokból, pontosabban azok információs szerkezeiteiből épül fel. Javaslatot tesz továbbá a mondatok közötti fennálló diskurzus relációk formalizálására az LFG-ben.

A dolgozat tehát nemcsak egy átfogó elemzés a többszörös kérdésekről a magyarban és a franciaiban, de az információ és diskurzus szerkezet kutatásához is hozzájárul, mind az LFG elméleten belül, mind pedig annál általánosabban.