The Logic of Variation
A cross-linguistic account of wh-question formation
The Logic of Variation
A cross-linguistic account of wh-question formation

De Logica van Variatie
Een taalvergelijkende benadering van
wh-vraagzinconstructies
(met een samenvatting in het Nederlands)

Proefschrift

ter verkrijging van de graad van doctor
aan de Universiteit Utrecht
op gezag van de Rector Magnificus, Prof. Dr. W. H. Gispen,
ingevolge het besluit van het College voor Promoties
in het openbaar te verdedigen
op vrijdag 13 januari 2006
des middags te 12.45 uur

doors

Willemien Katrien Vermaat

geboren op 30 januari 1974 te Rotterdam
Promotores: Prof. Dr. N. F. M. Corver
Prof. Dr. M. J. Moorgat
## Table of Contents

1 Introduction .......................... 1

2 Towards a logic of variation ............. 5
  2.1 The type system ...................... 7
      2.1.1 Basic and compound categories .... 7
      2.1.2 Lexicon ........................... 8
  2.2 Base logic ........................... 9
      2.2.1 Sequents ........................... 9
      2.2.2 Rules of inference ................. 10
      2.2.3 Displaying syntactic derivations .. 16
  2.3 The syntax-semantics interface .......... 18
      2.3.1 Semantic types and their interpretation . 19
      2.3.2 Typed lambda calculus .............. 20
      2.3.3 Derivational semantics .......... 22
      2.3.4 Lexical semantics ................. 23
  2.4 Structural module .................... 24
      2.4.1 Global structural reasoning ......... 25
      2.4.2 Controlled structural reasoning .... 26
  2.5 The logic of variation ............... 28
      2.5.1 Higher-order types ................ 30
      2.5.2 Fixed structural module .......... 30
      2.5.3 Strong lexicalism ................ 32

3 A type schema for wh-question formation 37
  3.1 Properties of wh-questions in English .... 37
      3.1.1 Local wh-questions ................ 38
          3.1.1.1 Direct questions .......... 38
          3.1.1.2 Indirect questions .... 39
      3.1.2 Non-local wh-questions ............ 40
      3.1.3 Island constraints ............... 41
      3.1.4 Multiple wh-questions ............ 42
  3.2 Wh-type schema ........................ 43
<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.3.3.2 Non-local wh-questions</td>
<td>122</td>
</tr>
<tr>
<td>4.3.4 Clitic placement</td>
<td>125</td>
</tr>
<tr>
<td>4.4 Partial wh-movement</td>
<td>130</td>
</tr>
<tr>
<td>4.4.1 Basic word order in German</td>
<td>131</td>
</tr>
<tr>
<td>4.4.2 Wh-question formation in German</td>
<td>132</td>
</tr>
<tr>
<td>4.4.3 Basic word order in Hindi</td>
<td>134</td>
</tr>
<tr>
<td>4.4.4 Wh-question formation in Hindi</td>
<td>135</td>
</tr>
<tr>
<td>4.4.5 Scope marking constructions</td>
<td>137</td>
</tr>
<tr>
<td>4.4.5.1 Scope marking constructions in German</td>
<td>139</td>
</tr>
<tr>
<td>4.4.5.2 Scope marking constructions in Hindi</td>
<td>144</td>
</tr>
<tr>
<td>4.4.5.3 A uniform approach</td>
<td>146</td>
</tr>
<tr>
<td>4.4.6 Partial wh-movement in Malay</td>
<td>147</td>
</tr>
<tr>
<td>4.5 Concluding remarks on variation</td>
<td>149</td>
</tr>
<tr>
<td>5 Syntax-semantics interface of wh-questions</td>
<td>153</td>
</tr>
<tr>
<td>5.1 Semantics of questions and answers</td>
<td>154</td>
</tr>
<tr>
<td>5.1.1 Proposition set approach</td>
<td>154</td>
</tr>
<tr>
<td>5.1.2 Structured meaning approach</td>
<td>155</td>
</tr>
<tr>
<td>5.2 Question and answer types</td>
<td>156</td>
</tr>
<tr>
<td>5.2.1 Type definition of wh-questions</td>
<td>157</td>
</tr>
<tr>
<td>5.2.2 Meaning assembly of wh-questions</td>
<td>158</td>
</tr>
<tr>
<td>5.2.3 Derivability patterns of wh-type schemata</td>
<td>160</td>
</tr>
<tr>
<td>5.2.3.1 Semantic derivability</td>
<td>160</td>
</tr>
<tr>
<td>5.2.3.2 Syntactic derivability</td>
<td>161</td>
</tr>
<tr>
<td>5.2.3.3 Overview</td>
<td>165</td>
</tr>
<tr>
<td>5.3 Linguistic application</td>
<td>167</td>
</tr>
<tr>
<td>5.3.1 Single constituent questions</td>
<td>168</td>
</tr>
<tr>
<td>5.3.1.1 Direct questions</td>
<td>168</td>
</tr>
<tr>
<td>5.3.1.2 Indirect wh-questions</td>
<td>171</td>
</tr>
<tr>
<td>5.3.1.3 Non-local wh-questions</td>
<td>171</td>
</tr>
<tr>
<td>5.3.2 Multiple wh-questions</td>
<td>172</td>
</tr>
<tr>
<td>5.3.2.1 Multiple wh-fronting</td>
<td>173</td>
</tr>
<tr>
<td>5.3.2.2 Simple wh-fronting</td>
<td>175</td>
</tr>
<tr>
<td>5.3.2.3 Multiple wh-in-situ</td>
<td>178</td>
</tr>
<tr>
<td>5.3.3 Pied-piping</td>
<td>181</td>
</tr>
<tr>
<td>5.3.3.1 Pied-piping in relative clauses</td>
<td>182</td>
</tr>
<tr>
<td>5.3.3.2 Pied-piping in wh-questions</td>
<td>183</td>
</tr>
<tr>
<td>5.3.3.3 Pied-piping in Japanese</td>
<td>186</td>
</tr>
<tr>
<td>5.3.4 Scope marking constructions</td>
<td>189</td>
</tr>
<tr>
<td>5.3.4.1 German scope marking</td>
<td>189</td>
</tr>
<tr>
<td>5.3.4.2 Hindi scope marking</td>
<td>194</td>
</tr>
<tr>
<td>5.3.4.3 Comparison between German and Hindi</td>
<td>198</td>
</tr>
<tr>
<td>5.4 Concluding remarks</td>
<td>199</td>
</tr>
</tbody>
</table>
6 Evaluation and future perspectives

6.1 A comparison with MG and MMCCG
6.1.1 Minimalist grammars
6.1.2 Multimodal CCG
6.2 Future research

A Overview of rules
A.1 Base logic
A.2 Structural module
A.3 Wh-type schema rules
A.4 Syntax-semantic interface

B Sample sentences
B.1 English
B.1.1 Basic sentences
B.1.2 Direct questions
B.1.3 Embedded interrogatives
B.1.4 Island constructions
B.1.5 Pied-piping
B.1.6 Multiple wh-questions
B.2 Japanese
B.2.1 Basic word order
B.2.2 Argument scrambling
B.2.3 Questions
B.2.4 Ka versus kadooka
B.2.5 Relative clauses
B.2.6 Complex noun phrases
B.2.7 Semantics of questions
B.2.8 Multiple wh-questions
B.3 Serbo-Croatian
B.3.1 Multiple wh-questions
B.3.2 Long-distance extraction
B.4 Bulgarian
B.4.1 Multiple wh-questions
B.4.2 Clitic placement
B.4.3 Long-distance extraction
B.5 German
B.5.1 Basic word order
B.5.2 Basic wh-questions
B.5.3 Scope marking constructions
B.5.4 Constraints on scope marking
B.5.5 Questions and answers
B.6 Hindi
B.6.1 Basic word order
B.6.2 Basic direct and indirect questions
B.6.3 Long-distance wh-questions ........................................ 223
B.6.4 Scope marking constructions ..................................... 223
B.6.5 Questions and answers ............................................. 224
B.7 Minimalist grammar analyses ....................................... 224

Bibliography ................................................................. 225

Samenvatting in het Nederlands ......................................... 231
About a month ago, while I was surfing in France, I was sitting on my surfboard behind the breaking waves. A clear blue sky, a bright warm sun shining on a gently rolling ocean and an empty beach, watching my friends catch some nice waves—a perfect moment. After paddling out through the break and then waiting patiently for the sets to roll in bringing more waves for me to surf on, I realized that surfing and writing a thesis couldn’t be more different. At the same time, they share an important property.

With surfing, you can’t just take your board, paddle out and catch waves at any convenient time. The ocean is different from day to day and even from hour to hour. You have to accept the waves that nature gives you. Sometimes it’s flat for days, but when the swell rolls in and you are at the right spot at the right time, you can catch some of the best waves of your life. While writing this thesis, I’ve had flat days when writing just seemed impossible and great days where ideas and words just kept coming. I have many people to thank for celebrating the great days with me and helping me through difficult times.

First of all, I want to thank my supervisors Michael Moortgat and Norbert Corver. I thank Michael for his continuing enthusiasm, his trust in my abilities and for his efforts in motivating me to set up this research. Norbert became my supervisor only two and a half years ago and my project has greatly benefited from his input. I want to thank him for his advice and for helping me to stay focused on the project’s deadlines. Michael and Norbert complemented each other perfectly.

The research group of Computational Linguistics and Logic provided a good environment to develop my skills as a researcher. For the fruitful discussions during meetings and the nice Italian dinners, I want to thank my former and present colleagues: Dirk Heylen, Esther Kraak, Richard Moot, Raffaella Bernardi, Herman Hendriks, Paola Monachesi and Matteo Capelletti.

I want to express my gratitude to the UiL-OTS, NWO, Fulbright and FoLLI for sponsoring visits to conferences and universities and for creating the right environment for fulfilling my research goals. They made it possible for me to become colleagues and friends with many people among whom I would especially like to mention: Maxime Amblard, Alain Lecomte, Richard Moot and
Christian Retoré in Bordeaux; Ed Stabler, Greg Kobele and Sara Rimensnyder in Los Angeles; as well as Claudia Casadio, Jens Michaelis, Dick Oehrle and Glyn Morrill. My duties as secretary for the ESSLLI summer school have given me the opportunity to cooperate with many enthusiastic researchers of whom I would like to mention a few: Carlos Areces, Patrick Blackburn, Raffaella Bernardi, Balder ten Cate, Antal van den Bosch, Paul Dekker, Judit Gervain, Rosalie Iemhoff, Ivana Kruijff-Korbayova, Geert-Jan Kruijff, Malvina Nissim, Norbert Preining and Maarten de Rijke.

I was in the lucky position to have my office at Achter de Dom and share the building with a wonderful group of researchers. I want to thank my office mates Øystein Nilsen and Nino Grillo for all the laughs and discussions, even though their office hours were opposite to mine. Of the other inhabitants of Achter de Dom, the following people deserve a special word of thanks for their company: Olga Borik, Alexis Dimitriadis, Silke Hamann, Jan de Jong, Jacqueline van Kampen, Joke de Lange, Anna Mlynarczyk, Kakhi Sakhlkhutishvili, Nada Vasic, Evy Vlachou, Frank Wijnen and Carien Wilsenach. Furthermore, I would like to mention some of my colleagues who have made my work a lot more pleasurable with their company during lunch breaks, discussion meetings, drinks, dinners and parties: Anna Ashbury, Fabian Battaglini, Patrick Brandt, Berit Gehrke, Heleen Hoekstra, Annemarie Kerkhoff, Cem Keskun, Oele Koornwinder, Marjo van Koppen, Marija Lekakou, Shakuntala Mahanta, Anne-Marie Mineur, Iris Mulders, Rick Nouwen, Dimitra Papangeli, Brigit van der Pas, Mirjam Rigterink, Oren Sadeh Leicht, Maaike Schoorlemmer, Giorgos Spathas, Krista Szendröi, Anca Sevcenco, Roberta Tedeschi, Mario van der Visser and Judith van Wijk.

A number of people deserve special gratitude: Elise de Bree for helping me through some tough periods; Matteo Capelletti, Berit Gehrke, Alain Lecomte, Richard Moot, Jacqueline van Kampen and Paul Dekker for carefully reading and commenting on early drafts; Sharon Unsworth and Rie von Eyben for proofreading the final version; Anna Ashbury, Svetoslav Marinov, Shakuntala Mahanta, Marijana Marelj, Mana Philip, Berit Gehrke for helping with the data; Céline Chat for allowing me to use one of her beautiful paintings for the front cover; all the neighbors of “Het Feestpaleis” and the “People in Black” for the welcome distraction from writing; all the members of UFO for keeping me physically healthy; my brother Joep and Gert Jan’s parents, Jan and Hennie, for keeping an eye on me; and my paranymphs Sharon Unsworth and Carmen Mecking for all their help until the very end of the project.

There are many more people whose names I have failed to mention. I will thank them personally when possible.

I want to thank my parents, Pim en Leentje who have always stimulated me to stay open-minded and to keep on doing the best I could. Finally, I can’t find the right words for the one person who I like to thank the most, my boyfriend Gert Jan. Without his patience, love and endless support this thesis wouldn’t have been written. Together we are waiting for the next perfect wave. We are curious where it is going to take us.
Languages differ in the way they put phrases together to obtain a meaningful expression. In English, the wh-question ‘Who does John think left?’ has the wh-word (‘who’) in a fronted position while the subject constituent that is being questioned appears as a gap in the embedded clause (‘…left’). In Japanese, a wh-question with the same interpretation has the question word in the position of the embedded subject phrase, “John thinks who left?”. Although the form of the two wh-questions differs, the meaning is the same: ‘who is the person such that John thinks that that person left’. The observation of this contrast between uniformity in meaning assembly and structural diversity naturally leads to the question: What is the logic of this kind of variation?

Linguists have searched for a universal grammar system to explain the variation across languages while accounting for uniformities of interpretation. A theory of grammar must provide the principles that delimit the range of possible variation and that account for the uniformities in semantic composition. Generative grammar, in the recent incarnation of the Minimalist Program (Chomsky, 1995) for example, is an influential theory with these goals. The starting point for linguists working in the Chomskyan tradition is the empirical reality of language. Logicians have their own, more mathematically inclined perspective on natural language. Especially since Lambek (1958) developed his ideas in “The mathematics of sentence structure”, a number of ‘grammar logics’ have been developed in the categorial tradition. The original Lambek systems were not well-equipped to tackle linguistic variation in a principled way. The recent development of so-called Type Logical Grammar (Morrill, 1994; Moortgat, 1997) has changed that situation, and offers logical tools that can provide an understanding of both the constant and the variable aspects of linguistic form and meaning. Our study of variation will be couched within such a type logical framework.

We see then that generative linguists and categorial grammarians have
a common goal, although they will put different emphases in pursuing this goal. The present thesis aims to establish a two-way communication between these approaches. From the generative tradition, we take the broad perspective on cross-linguistic diversity. With such a broad perspective, we can obtain a more solid empirical basis than has been customary in the development of categorial grammar. The type logical tradition, on the other hand, contributes to explanation within generative grammar by providing a rich deductive structure for such explanations, distinguishing between principles which are of a logical nature, and substantive linguistic principles.

Outline of the thesis

Chapter 2 describes the techniques of categorial type logic leading to the formulation of a logic of variation. The different components of the grammatical reasoning system are presented formally and concisely. We introduce different presentation formats to improve the readability of the logical derivations of linguistic expressions. After providing the basic components of the categorial type logical system, we present the restrictions on the structural reasoning system which leads to the formulation of the logic of variation to account for displacement phenomena. The three main factors in the cross-linguistic analysis of wh-question formation are strict lexicalism, restricted structural generalizations and higher-order types.

In chapter 3, we present a uniform account of wh-questions across languages. Focusing on the characteristics of wh-question formation in English, we add a further refinement to our grammar system by introducing wh-type schemata to distinguish different types of wh-phrases. We provide an abbreviated schematic presentation for merging a wh-phrase with a question body. This enables us to provide simple and reader friendly analyses of wh-questions in English.

Chapter 4 consists of a study of languages that structurally differ in the formation of a wh-question. We show that the cross-linguistic variation in wh-question formation can be accounted for in the logic of variation using the wh-type schemata. For each language, we build a grammar fragment to analyze wh-questions in that specific language. Linguistic expressions that illustrate the phenomena of wh-question formation can be parsed with an online parser for categorial grammars, Grail, developed by Moot (2002).

In chapter 5, we concentrate on the syntax-semantics interface of wh-questions. Along with providing the lexical semantics of wh-phrases, we redefine the wh-type schema by incorporating possible answer types into the type assigned to wh-questions. Because of the decomposition of wh-question types and the Curry-Howard interface between syntax and semantics, we can compute the meaning assembly of wh-questions in a more precise manner. Further ‘logical’ generalizations on wh-question formation are drawn on the revised wh-type schemata. We provide empirical evidence for these generalizations by focusing on the syntax-semantics interface of wh-questions.
Chapter 6 summarizes the main results and evaluates our approach by comparing it to two other lexicalist approaches to natural language analysis, the multimodal variant of combinatory categorial grammar (Baldridge, 2002) and minimalist grammars (Stabler, 1997). This leads to a list of open problems and suggestions for further research.

Appendix A contains an overview of rules of the logic of variation for quick reference. In appendix B, we list the linguistic expressions that have been analyzed throughout the thesis. Each expression is a hyperlink to the on-line parser, grail. The reader can check the sentences on-line using the electronic documents on the CD-rom.
Chapter 2

Towards a logic of variation

This chapter presents the version of Categorial Type Logic (CTL) that will be used throughout this thesis. The chapter has a twofold purpose. First, it can serve as a reference for definitions, rules and terminology that will be used in the chapters that follow. Secondly, it is meant as an introduction to CTL for researchers working in the field of generative syntax. In order to make the chapter accessible to that group of readers, we keep the presentation format as simple as possible, and we present formal definitions together with the intuitions behind them and/or graphical illustrations.

A schematic overview of the grammatical reasoning system is given in figure 2.1. Let us comment on this schema. CTL is a strongly lexicalized grammar formalism, which, in the case of a categorial system, means that a derivation is fully driven by the types assigned to lexical elements: these types are the basic declarative units on which the computational system acts. The computational system, or UG (*Universal Grammar*), consists of two parts, a base logic and a structural module. The base logic provides a set of inference rules introducing or eliminating the various type-forming operations for building complex types out of more elementary ones. These inference rules simultaneously determine the assembly of grammatical form (syntax) and meaning (semantics). The structural module characterizes a set of meaning-preserving relations between grammatical forms. UG has two output levels: the form and the meaning of a linguistic expression. The form dimension encodes information on linear order and constituent structure. The meaning encodes the compositional build-up of an expression in the form of a lambda term. Note

---

1Our choice of terminology is meant to highlight correspondences with work in generative syntax. Many researchers in the type-logical grammar tradition, e.g. Cornell (2004); Lecomte (2004); Vermaat (1999) have studied the connections between recent proposals in the Minimalist Program (Chomsky, 1995) and the basic components of type-logical grammar.
that the correspondence between form and meaning is fixed at the level of the base logic; the structural module creates the possibility that a given instruction for meaning assembly has different structural realizations.

In the coming sections, we introduce the different components of the grammatical reasoning system and show how they interact to construct linguistic expressions with both a form and a meaning component. In section 2.1, we present the type system and show how compound types (or categories) are built from atomic types/categories. In section 2.2, we turn to the invariant core of the grammatical reasoning system and present the inference rules of the base logic from a syntactic perspective. In section 2.3, we turn to the syntax-semantics interface, and explain how meanings are put together. In section 2.4, we discuss the role of the structural module and its interaction with the invariant base logic. Finally, in section 2.5, after discussing the different components of the grammatical reasoning system, we are ready to introduce the central theme of this thesis and present our perspective on the *logic of variation*.
2.1 The type system

2.1.1 Basic and compound categories

The basic analytical tool for language analysis in a categorial grammar is the system of types (or categories — we will use these terms interchangeably). The basis for the type system is a set of atomic or basic types. The full set of types is then built out of these basic types by means of a set of type-forming operations. In this thesis, we will consider unary and binary type-forming operations — in principle, one could extend the type system with type-forming operations of any arity. The unary type-forming operations we will use are ♦ (diamond) and □ (box). The binary ones are the two slashes /, \ (forward and backward slash) and • (product). The inductive definition below (in what is known as Extended Backus-Naur grammar notation) characterizes the full set of syntactic categories.

Definition 1 (Syntactic categories):

\[ F ::= A | F/F | F\bullet F | F/F | F/\emptyset | \emptyset F | \diamond F \]

The type system is used to classify groups of expressions with a similar grammatical behavior. Let us look briefly at the intuitive meaning of the type-forming operations from this perspective.

Binary operators  An expression belongs to a certain category depending on its grammatical relation to other expressions. Slash categories express incompleteness with respect to some other expressions. A product category represents the composition of two expressions. An expression of category \( A/B \) is incomplete with respect to an expression of category \( B \) on its right. The composition of these two expressions in the right linear order, i.e. \( (A/B)\bullet B \), classifies complex expressions of category \( A \). Symmetrically, an expression \( B\backslash A \) is incomplete and requests an expression of category \( B \) on its left; the composition \( B\bullet (B\backslash A) \) again classifies complex expressions of category \( A \).

The basic categories are used to classify expressions that are complete in themselves, i.e. expressions for which we have grammaticality judgements that do not depend on their relation with other expressions. For the time being, let us assume we have \( n, np \) and \( s \) as basic categories for nouns, noun phrases and sentences. A category such as \( vp \) for verbs is not needed as a basic category because verbs can be defined in relation to their arguments. In particular, tensed intransitive verbs are characterized as compound categories of type \( np\backslash s \). The type specifies that the verb is incomplete and needs an expression of category \( np \) on its left to form an expression of category \( s \). Exactly which basic categories are needed is an empirical matter. We use Ockham’s Razor to keep the number of basic categories to a minimum, i.e. a new basic category is only introduced if we are faced with a class of expressions with a behavior that cannot be captured in terms of the type-forming operations.
Unary operators The unary operators operate on any kind of category. They may be placed either on basic or compound categories, but also on subcategories of a compound type. Intuitively, unary operators can be understood as features that add additional information to the category. We will refer to the unary operators as features and use the term to decorate to indicate the labeling of a category with a specific feature.

Depending on the type of unary operator and the position in a compound type, the feature may be used to express additional selectional requirements. By decorating the dependent category of a compound category with a box feature: $A/B$, the category encodes that an expression of this category has an additional feature requirement $B$. Besides encoding a feature restriction, the unary features can also be used as parameter setting to control structural reasoning which will be explored further in section 2.4.

These intuitions can be made explicit in terms of a Kripke-style semantics for the type-forming operators. Our base logic can be shown to be sound and complete with respect to such interpretation (see Kurtonina (1995)).

2.1.2 Lexicon

In the lexicon we specify to which categories the smallest elements of a language belong. For the sake of simplicity, we will assume these smallest elements to be words. If desired, one can, of course, extend the categorial approach to the word grammar. We use the symbol ‘::’ to express type-assignment to words. Our lexical entries then have the schematic form given below:

Word :: Category

Later, in section 2.3 after we have discussed the semantical component of the grammatical reasoning system, we will extend the lexical entries with semantic information, but for now, we concentrate on syntax. The following lexical entries constitute a lexicon for a tiny fragment of English; each compound category in this case is constructed with slash operators.

<table>
<thead>
<tr>
<th>Word</th>
<th>Category</th>
</tr>
</thead>
<tbody>
<tr>
<td>John</td>
<td>np</td>
</tr>
<tr>
<td>sleep</td>
<td>np/s</td>
</tr>
<tr>
<td>ate</td>
<td>(np/s)/np</td>
</tr>
<tr>
<td>the, a</td>
<td>np/n</td>
</tr>
<tr>
<td>man, apple</td>
<td>n</td>
</tr>
<tr>
<td>somebody</td>
<td>s/(np/s)</td>
</tr>
</tbody>
</table>

To illustrate the role of the unary connectives as features, we can add number feature information to the following noun phrases and verbs. (For more detailed information on how features can be encoded using the unary operators, see Heylen (1999).)
The unary feature $\square$ is decorated with morphosyntactic properties for number, either singular (= sg) or plural (= pl). The verbs are also decorated with number features. A unary feature may also be underspecified ($\text{num}$) or overspecified ($\text{Num}$) for number. When it is overspecified an expression can be both singular and plural. In Dutch, for instance, the pronoun ‘zij’ (= she or they) is used either as singular or plural, and is therefore overspecified for the number feature ($\text{Num} np$). Furthermore, a transitive verb phrase like ‘eat’ is specified for a plural number feature for the subject noun phrase, but is underspecified for the number feature of the object argument, because it can select both singular and plural object arguments.

### 2.2 Base logic

We now turn to the logical perspective on grammatical composition that distinguishes CTL from other categorial frameworks. The basic insight here, due to Jim Lambek’s pioneering work in the late nineteenfifties, is that one can view the categories as logical formulas and, consequently, the operators that construct compound categories as logical connectives. This is what gives CTL its proof-theoretic flavor, nicely summed up by the slogan ‘parsing as deduction’. The task of the grammatical base logic, in this perspective, is to provide the rules of inference that together characterize the derivability relation, i.e. the rules that specify how one can reason with these formulas in such a way that one derives valid conclusions from valid premises.

#### 2.2.1 Sequents

To present the grammatical base logic as a deductive system, a number of equivalent presentation formats are available. In this thesis, we will use a sequent-style presentation originally due to Gentzen. In this presentation, derivability is cast as a relation between structures and formulas. We already know what the formulas look like. Let us turn then to the structures.

Structures, just like formulas, are built out of elementary structures with the aid of unary or binary structure-building operations. The elementary structures are formulas, the structure-building operations are $\odot$ and unary $\bullet$. A Gentzen structure, in other words, is a tree with formulas at the leaves, and with unary or binary branching internal nodes.

**Definition 2 (structures)**

\[ \text{Struc} ::= \mathcal{F} \mid \text{Struc} \odot \text{Struc} \mid \bullet \text{Struc} \]
In the pages that follow, we will freely switch between the unfolded 2D tree format for structures, and the familiar more compact representation as a bracketed string of formulas. In the tree style presentation, the binary and unary structure-building operations are represented as in the picture below. We will sometimes leave out ◦ to reduce the tree size. We will always explicitly label the unary branching nodes with ♦.

Using \( \{\Gamma, \Delta, \ldots\} \) as structure variables and \( A, B, \ldots \) as formula variables, we are ready now to introduce the concept of a sequent.

**Definition 3 (sequents)** A sequent is a statement \( \Gamma \vdash A \) with \( \Gamma \in \text{Struc} \) and \( A \in F \) expressing the judgment that the structure \( \Gamma \) can be shown to be of type \( A \).

\[ \text{Struc} \vdash F \]

The left-hand side of the turnstile \( \vdash \) is referred to as the *antecedent* and the right-hand as the *succedent*.

### 2.2.2 Rules of inference

Let us turn then to the inference rules that characterize the set of derivable sequents. As we said before, there are several equivalent presentation formats for these rules: Gentzen sequent calculus, natural deduction, categorical combinators, proof nets. In this thesis, we will use Gentzen-style natural deduction because it depicts the simultaneous construction of form and meaning in the most transparent way.

Rules of inference in Gentzen-style natural deduction format have the following schematic form:

\[
\frac{\text{Seq}_1 \ldots \text{Seq}_n}{\text{Seq}} \quad [\text{Label}]
\]

Each sequent \( \text{Seq}_i \), as we saw above, is of the form \( \Gamma \vdash A \). The schematic presentation shows that from a number of sequents \( \text{Seq}_1 \ldots \text{Seq}_n \), a sequent \( \text{Seq} \) is derived. The component above the horizontal line holds the premises of a rule; the lower part has the conclusion. Each rule is identified by a rule name \( \text{Label} \). The rule name indicates which connective the rule operates on, and whether the connective in question is removed from or introduced into the derivation.

The Gentzen format for derivations is very compact. To assist the reader who is not accustomed to prooftheoretic notation, we will also use an alternative presentation that graphically illustrates the operations on tree structures.
in a more intuitively appealing manner. In this format, a sequent is represented as a boxed figure. The antecedent of a sequent $\Gamma \vdash C$ is represented as a tree with syntactic formulas as terminal leaves. We can schematize the tree using the familiar triangle notation. The succedent, or goal formula, is printed in the upper right corner of the box. The inference rules can then be presented as instructions for combining such sequent boxes.

Recall that a sequent expresses the judgment that a structure is of a certain type. Let us inductively define how we can arrive at these judgments.

**Base case**

For the base case of the induction, we look at the smallest possible structure, namely a type formula. Obviously, the one-node tree which consists of just the formula $A$ is of type $A$, which is what the Axiom rule below states.

**Definition 4 (Axiom rule)**

\[
A \vdash A
\]

Because the axiom rule has no premises, we will omit the upper line in later derivations.

In addition to the Axiom rule, for each connective we will need two rules: one rule which tells us how we can use the connective when it occurs as the goal formula of a premise sequent — the so-called Elimination rule for the connective, and one rule allowing us to find the connective as the goal formula of a conclusion sequent — the Introduction rule.

**Binary connectives**

Consider first the elimination rules for the binary connectives / and \. Below we give the Gentzen-style natural deduction rules for the two connectives, and the graphical representation for the /$E$ rule — the \$E$ case is symmetric.

**Definition 5 (elimination rules for / and \)** If the structure $\Gamma$ has type $A/B$ (or $B\backslash A$) and the structure $\Delta$ has type $B$ then the composition of these two structures, $\Gamma \circ \Delta$ (or $\Delta \circ \Gamma$) is of type $A$. 
2. Towards a logic of variation

\[ \Delta \vdash B \quad \Gamma \vdash B \setminus A \quad [\setminus E] \]

\[ \Gamma \vdash A / B \quad \Delta \vdash B \quad [\setminus E] \]

if \( \Delta \vdash B \) and \( \Gamma \vdash B \setminus A \) then \( \Delta \circ \Gamma \vdash A \)

The elimination rule for \( \bullet \) involves the substitution of a structure inside another structure. To represent this substitution operation in the Gentzen style, we use the notation \( \Gamma[\Delta] \) for a structure \( \Gamma \) with a substructure \( \Delta \).

**Definition 6 (elimination rule for \( \bullet \))** Let \( \Delta \) be a structure which has been shown to be of type \( A \bullet B \), and let \( \Gamma \) be a structure with a substructure \( A \circ B \), where \( \Gamma \) has been shown to be of type \( C \). Substituting \( \Delta \) for \( A \circ B \) in the designated position produces a structure which is also of type \( C \).

\[ \Delta \vdash A \bullet B \quad \Gamma[\setminus E] [(A \circ B)] \vdash C \]

if \( \Delta \vdash A \bullet B \) and \( \Gamma[J] \vdash C \) then \( \Gamma[A \circ B] \vdash C \)

The elimination rules above provide instructions for decomposing a complex goal formula in a premise into its immediate subformulas. To make the logic complete, we also need instructions to build a complex formula out of subformulas. These rules are called introduction rules because they introduce a connective in the goal formula of the conclusion of a sequent.
Definition 7 (introduction rules for / and \) Suppose we have shown that a structure $B \circ \Delta$ (or $\Delta \circ B$), i.e. a tree with a leaf formula $B$ as its left (right) daughter, is of type $A$. We can conclude that $\Delta$ by itself is of type $B \setminus A$ (or $A / B$).

\[
\begin{align*}
B \circ \Delta &\vdash A & & \Delta \setminus B \vdash A & & \text{[\ldash I]} \\
\Delta &\vdash B \setminus A & & \Delta \circ B &\vdash A & & \text{[/I]}
\end{align*}
\]

We have presented the introduction rules for / and \ from the perspective that goes from the premises to the conclusion. We may read these rules also from the conclusion to the premise. For that backward-looking perspective, the intuition behind the introduction rule for / is that we assume a hypothesis of type $B$ and check whether $\Gamma \circ B \vdash A$ holds. For this reason, we refer to the introduction rules of / (and \) as hypothetical reasoning.

The Introduction rule for $\bullet$ makes it very clear that the structure-building operation $\circ$ is in fact the structural counterpart of the formula-building operation (connective) $\bullet$.

Definition 8 (introduction rule for $\bullet$) Merging a structure $\Gamma$ of type $A$ and $\Delta$ of type $B$ we obtain a structure $\Gamma \circ \Delta$ of type $A \bullet B$.

\[
\begin{align*}
\Gamma \vdash A & & \Delta \vdash B & & \text{[\ bullet I]} \\
\Gamma \circ \Delta \vdash A \bullet B
\end{align*}
\]
Unary connectives

First, we present the elimination and introduction rules for $\Box$, followed by the rules for $\Diamond$.

**Definition 9 (natural deduction rules for $\Box$) [\[\Box E\]]:** If a structure $\Gamma$ can be shown to be of type $\Box A$, we may conclude that structure $\Diamond(\Gamma)$ is of type $A$. \[\Box I\]: the introduction of the $\Box$ is obtained by switching premise and conclusion of the elimination rule.

$$
\frac{
\Gamma \vdash \Box A
}{\Diamond(\Gamma) \vdash A} \ [\Box E]
\frac{
\Diamond(\Gamma) \vdash A
}{\Gamma \vdash \Box A} \ [\Box I]
$$

The introduction and elimination rules for the $\Box$ connective can be intuitively understood in terms of feature checking. As explored in Vermaat (1999), the elimination rule moves feature information encoded in the lexicon to the structural domain where further structural reasoning may be applied. The introduction rules, in contrast with the elimination rules, transfer feature information from the structural domain to the logical domain. Because of the projection of a $\Diamond$ to the structural side of the sequent the inference rules for $\Box$ are sometimes referred to as ‘key-lock’, as the introduction rule projects a ‘lock’ around the given structure, while the elimination rule for $\Box$ serves as a key that ‘unlocks’ the $\Diamond$ feature projection.
The diamond introduction and elimination rules are similar to the binary
♦ rules. Again, the rules make clear that the ♦ operation is the structural coun-
terpart of the formula-building connective ♦.

**Definition 10 (natural deduction rules for ♦) [♦E]:** Let △ be a structure of type
♦A. If Γ is a structure of type B containing a substructure ♦A, one can substitute △
for ♦A at the designated position in Γ to obtain a structure of type B.

\[
\frac{\triangle \vdash \diamond A \quad \Gamma[\diamond (A)] \vdash B}{\Gamma[\diamond A] \vdash B} [\diamond E]
\]

if and then

![Diagram of ♦I and ♦E rules]

[♦I]: Let Γ be a structure of type A. We can extend Γ at the root with the unary
structure-building operation ♦ to obtain a structure of type ♦A.

\[
\frac{\Gamma \vdash A}{\diamond (\Gamma) \vdash \diamond A} [\diamond I]
\]

if then

![Diagram of ♦I rule]

Summing up, the logical inference rules above, with the complementarity
between elimination rules and introduction rules, give us the tools to deter-
mine for an arbitrarily complex structure whether it is of a certain type A.
Notice that to perform this task it is essential to have both the elimination
rules and the introduction rules. In Vermaat (1999), we discuss the similarity
between slash elimination rules and the MERGE operation as defined in Min-
imalist Grammar (Stabler, 1997). The merge rule is characterized as follows:
“[the] operation takes two syntactic objects and replaces them by a new combined syntactic object” (Chomsky, 1995, Ch.4,p.226).

In contrast to the elimination rules, the introduction rules construct syntactic formulas by introducing a logical connective in the conclusion goal formula of an inference step. The introduction rules, as we saw above, constitute a form of hypothetical reasoning in that they withdraw structural material that is necessary to establish the premise judgment. In the remainder of this thesis, hypothetical reasoning will be a key feature of our analysis of extraction phenomena. Frameworks such as the Minimalist Program or Minimalist Grammars have no support for hypothetical reasoning. As shown in Vermaat (1999), a logical analysis of the MOVE operation crucially involves the slash introduction rules.

2.2.3 Displaying syntactic derivations

To check whether a statement $\Gamma \vdash A$ holds, one uses the inference rules to break up the structure $\Gamma$ in its constituent parts; the attempt to derive $\Gamma \vdash A$ is successful if one obtains a derivation tree with axiom sequents at the leaves. As an example, in figure 2.2 we display the derivation tree showing that structure $[np \circ ((np \circ s)/np \circ (np/n \circ n))]$ is of type $s$. The antecedent structures in this derivation are formatted as bracketed strings of formulas. In figure 2.3, we display the conclusion sequent of this derivation in our boxed tree format.

```
np \vdash np
(np\circ s)/np \vdash np\circ s
 np\circ (np/n \circ n) \vdash np\circ (np/n \circ n) \vdash s
```

Figure 2.2: Natural deduction derivation with syntactic type formulas

```
\vdash np

\vdash np\circ s
\vdash np\circ (np/n \circ n)
\vdash np\circ (np/n \circ n)
\vdash s
```

Figure 2.3: Tree structure format for the conclusion sequent in Fig 2.2
This derivation only needs the slash elimination rules. Illustrations of the applications of the other inference rules will be presented in due course.

**Adding words as terminals** Our sample derivation in figure 2.3 abstracts from lexical items: from the perspective of CF rewriting rules, the axiom leafs of the derivation correspond to the preterminal symbols. To display derivations for particular choices of lexical items, we would like to represent 'lexical insertion' in our derivation trees. Recall that type-assignment in the lexicon is represented in the form of statements \( \text{word} :: \text{A} \). To make our derivation trees fully explicit about lexical insertion, we would have to replace the axiom leafs by the combination of two statements: \( \text{word} :: \text{A} \) (‘\text{word} is an expression of type \( A \)’) and \( A \vdash A \) (‘the one-node tree consisting of the formula \( A \) is a structure of type \( A \)’). We abbreviate this combination thus:

\[
\text{word} \\
\hat{F}
\]

As an illustration, we repeat the derivation of Figure 2.2 with lexical items inserted at the axiom leafs. Figure 2.4 shows the natural deduction style derivation of the sentence ‘Eve ate the apple’.

![Natural deduction derivation of ‘Eve ate the apple’](image)

We can also add the terminals to the boxed tree style presentation. A structure \( \Gamma \) of type \( A \) with preterminal nodes \( B_1 \ldots B_n \) can be extended with lexical items \( w_i \) provided the lexicon has type-assignments \( w_i :: B_i \) for \( 1 \leq i \leq n \).
The tree structure in Figure 2.4 can now be represented as follows:

![Tree structure](image)

2.3 The syntax-semantics interface

In this section, we turn to the syntax-semantics interface. Within CTL, the guiding principle for semantic interpretation is the Curry-Howard correspondence between derivations and instructions for meaning assembly. The Curry-Howard principle is a *derivational view* on how meanings are put together. Put shortly, the idea is the following: we associate each individual inference step (the connective introduction and elimination rules) with an instruction for meaning composition; as a result, a complete derivation for a sequent $A_1, \ldots, A_n \vdash B$ produces a program to compute the meaning of the $B$ conclusion, parametric in the $A_i$ assumptions where we will plug in lexical items with their individual lexical semantics. The Curry-Howard view on meaning composition was introduced in the categorial discussion by van Benthem (1983) (Lambek’s original systems only considered syntax). It offers a particularly attractive realization of the Fregean principle of compositionality that ‘the meaning of a complex expression is a function of the meaning of its parts and the rules that put them together’. These rules, within CTL, are not the syntactic rules of a rewrite grammar, but the rules of inference that tell us how we can reason with categorial types seen as logical formulas. This is what the ‘formulas-as-types’ slogan for the Curry-Howard derivational semantics refers to.

The language we will use to encode the instructions for meaning assembly is the language of the typed lambda calculus. In section 2.3.1, we start by considering the semantic type system. We show how each syntactic type can be mapped to its semantic counterpart and we indicate what the domains of interpretation are for the semantic types. In section 2.3.2, we then present syntax and semantics of the terms of the typed lambda calculus. In section 2.3.3 and 2.3.4, we discuss the derivational and lexical semantics which gives us the compositional interpretation of natural language expressions.
2.3. Semantic types and their interpretation

The set of semantic types has the same structure as the syntactic type set: we have a set of basic semantic types (which for the sake of simplicity will be just type \( e \) for entities and type \( t \) for truth values), and a set of type-forming operations to construct complex semantic types from simples subtypes.

**Definition 11 (Semantic type language)** Let \( \text{Atom}_{\text{sem}} \) be the set \( \{e, t\} \). The set of semantic types, \( \text{Typ}_{\text{sem}} \), is the closure of \( \text{Atom}_{\text{sem}} \) under the type-forming operations \( \rightarrow \) (function) and \( \circ \) (product).

\[
\text{Typ}_{\text{sem}} := \text{Atom}_{\text{sem}} \mid \text{Typ}_{\text{sem}} \rightarrow \text{Typ}_{\text{sem}} \mid \text{Typ}_{\text{sem}} \circ \text{Typ}_{\text{sem}}
\]

**Semantic domains** For each type \( A \) in \( \text{Typ}_{\text{sem}} \), we want to determine its domain of interpretation, \( D_A \). The domain of interpretation of a semantic type determines the kind of semantic objects expressions of that type can denote.

We start from a non-empty set \( E \) (the set of individuals or entities, the things which our language talks about), and a set \( \{0, 1\} \) (the truth values). These are the denotation domains for the basic semantic types \( e \) and \( t \) respectively.

For complex semantic types, the denotation domain is defined inductively as shown below.

**Definition 12 (Denotation domains)**

\[
\begin{align*}
D_e &= E \quad (= \text{the set of individuals}) \\
D_t &= \{0, 1\} \quad (= \text{the set of truth values}) \\
D_A \rightarrow B &= D_B^{D_A} \quad (= \text{the set of functions from } D_A \text{ to } D_B) \\
D_A \circ B &= D_A \times D_B \quad (= \text{the Cartesian product of } D_A \text{ and } D_B)
\end{align*}
\]

The next step in setting up semantic interpretation is to associate each syntactic type with a corresponding semantic type, and, via this mapping, with a denotation domain.

**Definition 13 (Mapping of syntactic types to semantic types)** We define the function \( \text{sem} \) that maps each syntactic type to a semantic type. The base cases of the definition apply to the basic syntactic types:

\[
\begin{align*}
\text{sem}(n) &= e \\
\text{sem}(s) &= t \\
\text{sem}(n) &= e \rightarrow t
\end{align*}
\]

The recursive cases:

\[
\begin{align*}
\text{sem}(A/B) &= \text{sem}(B \setminus A) = \text{sem}(B) \rightarrow \text{sem}(A) \\
\text{sem}(A \bullet B) &= \text{sem}(A) \circ \text{sem}(B) \\
\text{sem}(\Diamond A) &= \text{sem}(\Box A) = \text{sem}(A)
\end{align*}
\]
Notice that types that are atomic syntactically may very well be mapped to complex semantic types, and that information which is relevant only in the syntactic dimension (for example, the directionality of slash types) may get lost in the mapping to semantic types. Additionally, notice that in the recursive cases, the semantic type assignment to unary decorated types map are semantically inert. We use the unary connectives purely as syntactic devices.

**Example 2.1** To illustrate the mapping from syntactic types to semantic types, we give the semantic types for the lexical type assignments that we used in section 2.1.2.

- **John** ⊢ np  ~ e
- **apple** ⊢ n  ~ e → t
- **sleeps** ⊢ np\s  ~ e → t
- **ate** ⊢ (np\s)/np  ~ e → (e → t)
- **somebody** ⊢ s/(np\s)  ~ (e → t) → t
- **the** ⊢ np/n  ~ (e → t) → e

Let us convince ourselves that these lexical entries indeed get associated with the desired domains of interpretation. The expression ‘John’, being associated with semantic type e, will denote an individual. Which particular individual will depend on the model one is considering — it is not the task of D, to determine that, just to make sure that an expression like ‘John’ is interpreted as an individual whatever model one considers. The expressions ‘sleep’ and ‘apple’ are both associated with semantic type e → t, hence they will denote functions from individuals to truth values. These would be the so-called characteristic functions of sets of individuals. Again, which particular set of individuals will constitute the set of sleepers and the set of apples will depend on the model. The type associated with the transitive verb ‘ate’ determines that this expression will be interpreted as a function sending two individuals to a truth value: an appropriate kind of interpretation for a two-place relation. The generalized quantifier expression ‘somebody’, finally, is interpreted as (the characteristic function of) a set of properties, as required.

### 2.3.2 Typed lambda calculus

Let us turn now to the language of the typed lambda calculus. Its role, as we saw at the beginning of this section, will be to provide recipes for meaning assembly that we will be able to read off categorial derivations.

**Definition 14 (λ-terms)** Let C A be a set of constants of type A and V A a (denumerably infinite) set of variables of type A. The full set of typed lambda terms of type A, T A is defined by the following grammar:

\[
\begin{align*}
T^A & ::= \mathcal{A}^A | (T^Barrow A) | (\pi^1 T^A \circ B) | (\pi^2 T^B \circ A) \\
T^Aarrow B & ::= \lambda V^A T^B \\
T^{A\circ B} & ::= (T^A, T^B)
\end{align*}
\]
Definition 15 (Semantic value of lambda terms) Let \( I \) (the interpretation function) and \( g \) (the assignment function) be functions associating constants (respectively variables) of type \( A \) with semantic objects in \( D_A \). We define the semantic value of a term relative to such an interpretation and assignment function as follows.

\[
\begin{align*}
\llbracket x \rrbracket^I_g &= g(x) \quad \text{(where \( x \) is a variable of type \( a \))} \\
\llbracket t \rrbracket^I_g &= I(t) \quad \text{(where \( t \) is a constant of type \( a \))} \\
\llbracket (\lambda x.A \rightarrow B) \rrbracket^I_g &= \llbracket (\lambda x.A \rightarrow B) \rrbracket^I_g \times \llbracket (\psi_B) \rrbracket^I_g \\
\llbracket (\phi, \psi) \rrbracket^I_g &= h_A \rightarrow B \quad \text{such that} \ h(m) = \llbracket (\phi_B) \rrbracket^I_g \end{align*}
\]

Terms in normal form \quad Given the above definition of the semantic value of a lambda term, certain terms will have the same semantic value in all models. We can define a notion of reduction on such equivalent terms, rewriting a syntactically more complex term (the \textit{redex}) to a semantically equivalent simpler term (the \textit{contractum}). These simplification rules are given below.

Definition 16 (Reduction rules) \( \beta \) and \( \eta \) reductions for function and product types in lambda term equations:

\( \lambda \)-conversion: When a semantic (sub)term is of the form \( (\lambda x.A \rightarrow B) \rrbracket^I_g \times \llbracket (\psi_B) \rrbracket^I_g \) or \( (\lambda x.A \rightarrow B) \rrbracket^I_g \times \llbracket (\psi_B) \rrbracket^I_g \), the term at the position which is indicated by the projector \( \pi \) becomes the reduced term:

\[
\begin{align*}
(\pi^1(\lambda x.A \rightarrow B) \rrbracket^I_g) &\sim_{\beta} \llbracket (\pi^1(\lambda x.A \rightarrow B) \rrbracket^I_g \\
(\pi^2(\lambda x.A \rightarrow B) \rrbracket^I_g) &\sim_{\beta} \llbracket (\pi^2(\lambda x.A \rightarrow B) \rrbracket^I_g
\end{align*}
\]

\( \pi \)-conversion: When a semantic (sub)term is of the form \( (\pi^2(t, u) \rrbracket^I_g) \) or \( (\pi^1(t, u) \rrbracket^I_g) \), the term at the position which is indicated by the projector \( \pi \) becomes the reduced term:

\[
\begin{align*}
(\pi^1(t, u) \rrbracket^I_g) &\sim_{\beta} t^A \quad (\pi^2(t, u) \rrbracket^I_g) &\sim_{\beta} u^B
\end{align*}
\]

\( \eta \)-conversion: When a semantic term for a term \( (\lambda x.t) \rrbracket^I_g \), a function type or a term \( (\pi^i(t, u) \rrbracket^I_g) \), a product type, \( \eta \)-reduction yields the following reduced terms

\[
\begin{align*}
(\lambda x.(t x) \rrbracket^B t^A) &\sim_{\eta} t^A \rightarrow B \\
((\pi^1(t \rrbracket^A \rightarrow B), (\pi^2(t \rrbracket^A \rightarrow B)) &\sim_{\eta} t^A \rightarrow B
\end{align*}
\]

\( \eta \)-normal proof terms correspond to proofs with non-atomic axioms (see Moortgat (1997)).
2. Towards a logic of variation

2.3.3 Derivational semantics

With the lambda term language and the Curry-Howard correspondence, we can now determine how meaning assembly can be read off from a syntactic derivation, hence the term *derivational semantics*. Each natural deduction rule is associated with an instruction that computes the term assignment of the conclusions given the term assignments of its premisses. We show for each derivational step how the meaning assembly is built up.

In section 2.2, we have presented the syntactic inference rules for combining syntactic types. We repeat the same rules here and add the semantic term label to each composition or decomposition of a syntactic type formula. In these rules we leave the semantic types of the terms implicit since it is fully determined by the mapping from the syntactic types to the semantic types.

**Definition 17 (Term labeling of the natural deduction rules)** The elimination rules of the two slashes (⁄ and \) correspond with function application. Combining two expressions yields a term where term u, the function, is applied to term v, the argument.

\[
\frac{\Delta \vdash v : B \quad \Gamma \vdash u : B \setminus A}{\Delta \circ \Gamma \vdash (u \circ v) : A} \quad [\smallsetminus E]
\]

\[
\frac{\Delta \vdash u : A/B \quad \Gamma \vdash v : B}{\Delta \circ \Gamma \vdash (u \circ v) : A} \quad [/E]
\]

The introduction rules correspond with abstraction. After abstracting the hypothesis from the antecedent structure, the term variable x assumed for the hypothesis is bound to a lambda operator.

\[
\frac{x : B \circ \Delta \vdash u : A}{\Delta \vdash \lambda x. u : B \setminus A} \quad [\smallsetminus I]
\]

\[
\frac{\Delta \vdash x : B \circ \Gamma \vdash u : A}{\Delta \vdash \lambda x. u : A/B} \quad [/I]
\]

The introduction rule for • corresponds to pairing, while the elimination rule associates with the projection of a term (u) with the already formed expression (t) by substituting the variables in that term.

\[
\frac{\Delta \vdash t : A \quad \Gamma \vdash u : B}{\Delta \circ \Gamma \vdash (\langle t, u \rangle) : A \bullet B} \quad [\bullet I]
\]

\[
\frac{\Delta \vdash u : A \bullet B \quad \Gamma \vdash t : C \quad \Gamma \vdash A \circ y : B}{\Gamma[\Delta] \vdash t[\pi^1 u/x, \pi^2 u/y] : C} \quad [\bullet E]
\]

The syntactic rules in combination with the semantic rules are used to derive a semantic term or *proof term* for each structural expression. The construction of a syntactic proof takes place simultaneously with the construction of a semantic term. The semantic term shows the underlying relations of the basic components.

**Example 2.2** We compute the proof term of the syntactic derivation presented in figure 2.2. To keep the derivation simple, we build up the term from the term variables assigned to the axioms and omit the terms on the antecedent of the sequent.
2.3. The syntax-semantics interface

The syntax-semantics interface is a foundational aspect of natural language processing that bridges the gap between sentence structure and meaning. This interface involves the construction of a meaning assembly that reasons over actual linguistic expressions.

### Structure of Lexical Semantics

Lexical semantics is a crucial component in the syntax-semantics interface, providing a level of detail that bridges abstract terms and concrete linguistic expressions. The lexical semantics of a word is defined by a semantic term label that is associated with each lexical entry. This label is then incorporated into the proof term, which is a formal representation of the sentence structure.

The format of the lexical description is displayed as:

\[ \text{Structure} \vdash \text{Term} : \text{Type} \]

This format indicates that the structure (e.g., sentence, phrase) is mapped to a term that has a specific type, which captures its semantic properties.

#### Example Lexicon

Here is an example of a fully specified lexicon:

- **John**, **Mary**, **Eve** \( \vdash j, m, e : np \)
- **apple** \( \vdash \lambda y.\text{(apple} \ y) : n \)
- **sleeps** \( \vdash \lambda x.\text{(sleep} \ x) : np\s \)
- **ate** \( \vdash \lambda x.\lambda y.\text{(eat} \ x \ y) : (np\s)\snp \)
- **somebody** \( \vdash \lambda P.\exists x.\text{(P} \ x) : s/(np\s) \)
- **the** \( \vdash \lambda Q.\iota\lambda z.(Q \ z) : np/n \)

The sentence presented in example 2.2 can be the derivation of the sentence ‘Eve ate the apple’. The computed proof term contains term variables which can be substituted for the lexical meaning recipes that has been assigned to each individual expression.

### Substitution with Lexical Semantics

We replace each term variable with its corresponding lexical semantics and obtain the following substitutions:
2. Towards a logic of variation

\[ P := \lambda x. \lambda y. ((\text{eat } x) \; y); u := \lambda Q. \iota \lambda z. (Q \; z); v := \lambda y. (\text{apple } y); w := e. \]

After substitution, we get the following semantic term:

\[ ((\lambda x. \lambda y. ((\text{eat } x) \; y) \; (\lambda Q. \iota \lambda z. (Q \; z) \; \lambda y. \text{apple } y)) \; e) \]

The semantic term is not in \( \beta \) normal form. The semantic term still contains a subterm of the form: \([((\lambda x. t) \; u])\]. The term is reduced to a \( \beta \)-normal form using the reduction rules (definition 16, p. 21):

- \( \beta \) \( ((\lambda x. \lambda y. ((\text{eat } x) \; y) \; (\lambda Q. \iota \lambda z. (Q \; z) \; \lambda y. \text{apple } y)) \; e) \)
- \( \sim_\beta \) \( ((\lambda x. \lambda y. ((\text{eat } x) \; y) \; \iota \lambda z. (\lambda y. \text{apple } y) \; z) \; e) \)
- \( \sim_\beta \) \( ((\lambda x. \lambda y. ((\text{eat } x) \; y) \; \iota \lambda z. (\text{apple } z)) \; e) \)
- \( \sim_\beta \) \( (\lambda y. ((\text{eat } \iota \lambda z. (\text{apple } z)) \; y) \; e) \)
- \( \sim_\beta \) \( ((\text{eat } \iota \lambda z. (\text{apple } z)) \; e) \)

Concluding remarks on the syntax-semantics interface

We have presented the semantics of syntactic expressions that gives us a complete and sound system to reason over linguistic expression. Additionally, the Curry-Howard isomorphism between types and terms gives us the means to express the derivational and lexical semantics of expressions. The correspondence ensures a balance between form and meaning. While the form of an expression is led by the syntactic composition of its subexpressions, the meaning of the whole expression is determined by the meaning of its parts. Typed lambda calculus is used to represent the meaning assembly of each composite expression.

2.4 Structural module

In the previous section, we have looked at the interplay between the lexicon and the categorial base logic. In the lexicon, words are assigned one or more types (a finite number). This type-assignment determines how the rules of inference of the base logic will be applicable. If the application of these rules gives rise to a successful derivation, we know how to assemble the words into a well-formed structural configuration of the goal type, and we have a recipe to compositionally put together the meanings of the words in the way dictated by the derivation steps.

Let us turn now to the next component of UG, the structural module. A first question to be answered is: Why do we need such a module? After all, the combination of lexicon and base logic already computes the two output levels of our grammatical architecture: the relation between grammatical forms and meanings. The answer to this question is that the combination of lexicon and base logic in itself is not enough to capture structural generalizations.
Consider what happens when a word can appear in different structural environments. Without a structural module, the only option is to assign it multiple types in the lexicon. Lexical ambiguity as such is not problematic, and it certainly would be possible to enrich the lexicon with facilities to express lexical generalizations over finite sets of types. But in the case of displacement phenomena, the central topic of this thesis, such a lexical strategy will not be satisfactory. Where displacement gives rise to unbounded dependencies between a wh-element and the gap it relates to semantically, one would have to lexically schematize over an infinite set of structural environments. This is inconsistent with the requirement that the lexicon is a finite component. The addition of a structural module to the base logic will make it possible for a word with a given lexical type assignment to appear in different but related structural contexts, thus removing the need for lexical ambiguity.

In contrast with the logical rules of inference, the structural module technically will consist of non-logical axioms or postulates. The postulates make it possible to reconfigure a structure in a meaning-preserving way. Below, we discuss two ways of implementing facilities for structural reasoning. First, in section 2.4.1, we present structural rules that apply in a global fashion. Such global structural reasoning has been considered in the logical literature on CTL, but it is not very useful for linguistic analysis. In section 2.4.2, we discuss how structural operations can be lexically controlled by means of the $\odot$ and $\Box$ connectives. It is this controlled version of structural reasoning that we will further pursue in this thesis in order to obtain a logical perspective on structural variation.

### 2.4.1 Global structural reasoning

One way of realizing a more flexible view on grammatical structure is to attribute global structural properties of associativity and/or commutativity to the composition operator $\bullet$. We present these options as structural rules below, using three presentation formats: as natural deduction sequent rules, reconfiguring a subconfiguration of antecedent structures; in an axiomatic format, which provides a one-line shorthand notation for the sequent rules; and in a graphically explicit tree format.

**Definition 18 (Associativity)**

\[\begin{align*}
N.D. \text{ style:} & \quad \Gamma[(\Delta_1 \circ \Delta_2) \circ \Delta_3] \vdash C & \text{[Ass1]} \\
& \Gamma[\Delta_1 \circ (\Delta_2 \circ \Delta_3)] \vdash C & \text{[Ass2]} \\
& \Gamma[\Delta_1 \circ (\Delta_2 \circ \Delta_3)] \vdash C & \text{[Ass2]} \\
& \Gamma[(\Delta_1 \circ \Delta_2) \circ \Delta_3] \vdash C & \text{[Ass1]} \\
A \bullet (B \bullet C) & \vdash (A \bullet B) \bullet C 
\end{align*}\]
Towards a logic of variation

Definition 19 (Commutativity)

Tree style format:

Definition 19 (Commutativity)

N.D. style: \[ \Gamma[\Delta_1 \odot \Delta_2] \vdash C \]

Axiomatic format: \[ A \bullet B \vdash B \bullet A \]

Tree style format:

Intuitively, the associativity rule says that to determine whether an expression is a well-formed structure of type \( C \), one can freely rebracket this structure. Constituent structure information, in other words, would not be relevant to determine well-formedness. In a similar way, the commutativity rule would imply that well-formedness is not affected by arbitrary changes in linear order. Obviously, such global options are not attractive from a linguistic point of view. But these global structural options have been used in the logical and mathematical literature to unfold a landscape of categorial calculi, starting from a system with no structural rules at all (our base logic, also known as NL, for non-associative Lambek calculus, when one considers just the binary connectives \( /, \bullet, \backslash \)). \( L \) is obtained by adding the structural rule of associativity as global structural option; \( NLP \) is obtained by adding only commutativity as structural option; while \( LP \) is the grammar system which has the two rules. These different categorial calculi are represented in figure 2.5. Their formal properties have been studied by Buszkowski (1997) and van Benthem (1995), among others. Our task in this thesis will be to find out where in this landscape the grammars of natural languages can be situated.

2.4.2 Controlled structural reasoning

Let us see then how we can control the possibilities for structural reasoning in a way that is consistent with our lexicalist principles. Early proposals have

**Definition 20 (multiple modes of composition)**

\[
\mathcal{F} ::= A \mid \mathcal{F} / \mathcal{F} \mid \mathcal{F} \cdot \mathcal{F} \mid \mathcal{F} \backslash \mathcal{F}
\]

\[
\text{Struc} ::= \text{Expr} \mid \text{Struc} \circ \text{Struc}
\]

The families are distinguished by adding an index (referred to as the *mode of composition* (Moortgat, 1997)) to the type-forming operators. The idea is that these different families all share the same logical rules (the elimination and introduction rules of the base logic), but that they can have access to different structural rules, affecting individual families, or the interaction between families. In the terminology of modal logic, this is a shift to a *multimodal perspective*; hence the designation Multimodal Typological Grammar for this line of research.

The multimodal perspective has recently been adopted also in the framework of Combinatory Categorial Grammar (cf. Baldridge (2002) and discussion in chapter 5). A weakness of the multimodal approach is its open-endedness: faced with a problem of grammatical analysis, one would be free to introduce a construction-specific mode distinction, with structural rules applicable to it. The addition of the unary connectives ♦ and □ to the categorial vocabulary creates more attractive possibilities for structural control (Kurtonina and Moortgat, 1997): instead of attributing structural properties to the binary composition operations, one uses the unary operators to trigger or block possibilities for structural reconfiguration.

**Example 2.3 (controlled structural rules)** Restricted commutativity (left) and associativity (right). The product ⋅ in these cases, has no structural possibilities of its own: it is the composition operation of the base logic. The presence of ♦ licenses...
Towards a logic of variation

that a marked structure exchanges position with its right neighbor (restricted commutativity), or that the indicated constituent restructuring can take place (restricted associativity).

\[
\begin{align*}
\Gamma[\Delta_3 \circ \Diamond \Delta_1] & \vdash C \quad \text{[Comm]} \\
\Gamma[\Diamond \Delta_1 \circ \Delta_2] & \vdash C \quad \text{[Comm]} \\
\Gamma[\Diamond \left( \Diamond \Delta_3 \circ \Diamond \left( \Diamond \Delta_2 \circ \Diamond \Delta_1 \right) \right)] & \vdash C \quad \text{[Ass]} \\
\end{align*}
\]

With respect to the categorial landscape of figure 2.5, Kurtonina and Moortgat (1997) show that one can relate the different calculi in this landscape by means of embedding translations using appropriate \( \Diamond, \square \) decorations. These decorations function either as licensors, relaxing a structurally rigid system into a more liberal one, or as constraints, blocking structural options that would otherwise be available. The following diagram schematically presents these two options: the arrows in the inner circle represent the constraining effect of the structural rules, whereas the arrows in the outer circle represents the introduction of structural flexibility.

\[\text{LP} \leftrightarrow \text{L} \leftrightarrow \text{NLP} \leftrightarrow \text{NL} \]

2.5 The logic of variation

At this point, we have all the components in place to address the central issue of this thesis: How can we develop, within the framework of CTL, a logical view on linguistic variation? Before presenting our working hypotheses, we repeat the schematic presentation of our grammar formalism in Figure 2.5, with a summary of the relations between its different components.
The logic of variation

Lexicon

UG

Base logic

+

Structural module

Form

Meaning

Figure 2.6: A model of the grammatical reasoning system in type-logical grammar.

- The lexicon assigns words one or more types. The inference rules of UG (= Universal Grammar) are driven by these types.
- UG consists of a base logic and a structural module.
- The base logic consists of elimination and introduction rules for each binary and unary connective. These rules simultaneously build up grammatical form and meaning.
- The structural module consists of a set of non-logical axioms. They establish meaning-preserving relations between structures.
- UG computations have a syntactic and a semantic output level.
- The syntactic structure of an expression is a tree encoding dominance and precedence relations between subexpressions.
- The composition of meaning of an expression is represented as a lambda term revealing the predicate-argument relations between subexpressions.
Our working hypotheses with respect to the logic of variation can be summed up as follows.

**Definition 21 (The logic of variation)** The combination of the structural variation and the uniform semantic interpretation of wh-question formation cross-linguistically can be accounted for in terms of the following three assumptions:

1. **Higher-order type assignment**: higher-order type assignment to wh-elements, together with the Curry-Howard interpretation that goes with it, accounts for the uniformity in the semantic interpretation of wh-questions.

2. **Fixed structural module**: variation in the structural realization of the uniform interpretation schema can be accounted for in terms of a restricted set of structural rules. This set is fixed by UG; there are no language-specific structural rules.

3. **Strong lexicalism**: as a result of the above, cross-linguistic variation in wh-question formation must be entirely reducible to differences in lexical type-assignment.

Let us briefly discuss these three components of our approach.

### 2.5.1 Higher-order types

Following Vermaat (1999), wh-phrases are assigned higher-order types. Because meaning composition is fully determined by the Curry-Howard of the logical rules of inference, the higher-order type assignment results in a uniform interpretation for wh-question constructions. In a higher-order type, one slash operator is embedded under another slash. The embedded slash enters the derivation by means of a slash introduction rule, which means hypothetical reasoning, and semantically, \( \lambda \)-abstraction over the gap variable in the question body. The governing slash of the higher-order type assignment is used in a slash elimination rule, which means application of the question operator to the lambda abstraction over the gap. As pointed out in Vermaat (1999), this higher-order view demonstrates that the MOVE operation of generative syntax is in fact not a primitive operation: MOVE can be decomposed in a MERGE component (the application part of the meaning composition) and hypothetical reasoning with respect to the gap hypothesis (the \( \lambda \) binding part). Along the same lines, higher-order types account for in-situ wh-phrases which do not move but which nevertheless have a non-local scope construal comparable to that of generalized quantified expressions. Chapter 3 of this thesis is devoted to a detailed discussion of higher-order type-assignment to wh-phrases and the uniform treatment of fronted and in-situ wh-elements.

### 2.5.2 Fixed structural module

In earlier work within Multimodal Type-Logical Grammar, presenting a categorial grammar for a language involved two tasks: giving a lexicon, and giv-
ing the package of structural postulates for the language in question. In this thesis, we move back to the pure lexicalist categorial tradition, in the sense that we fix the structural module as a hard-wired component within UG, and constrain this component to a highly restricted set of structural patterns.

The set of structural postulates that we will use was originally proposed in Moortgat (1999a) in a comparative study of relative clause formation in Dutch and English. The claim of this thesis will be that this set is sufficient to account for the cross-linguistic diversity that we find in wh-question formation. We present the displacement postulates first in natural deduction style, followed by the tree style presentation. The structural rule schemata come in two symmetric groups: left displacement and right displacement postulates.

**Definition 22 (left displacement postulates)**

\[
\Gamma[(\Diamond \Delta_1 \circ \Delta_2) \circ \Delta_3] \vdash C \quad \Gamma[(\Delta_2 \circ (\Diamond \Delta_1 \circ \Delta_3))] \vdash C \quad \Gamma[(\Delta_1 \circ (\Delta_2 \circ \Delta_3))] \vdash C
\]

In contrast with Kayne (1994) who claims that displacement is asymmetric, we propose a symmetry in tree restructuring possibilities.

**Definition 23 (right displacement postulates)**

\[
\Gamma[(\Delta_1 \circ (\Delta_2 \circ \Diamond \Delta_3))] \vdash C \quad \Gamma[(\Delta_1 \circ (\Delta_2 \circ \Delta_3))] \vdash C \quad \Gamma[(\Delta_1 \circ (\Delta_2 \circ \Diamond \Delta_3))] \vdash C \quad \Gamma[(\Delta_1 \circ (\Delta_2 \circ \Delta_3))] \vdash C
\]
For the postulates to be applicable, one of the constituents of the affected structure has to be marked with a licensing ♦ feature. We will see below how this licensing feature is in fact projected from lexical type-assignments decorated with the ♦ and □ connectives.

2.5.3 **Strong lexicalism**

As a result of fixing the structural module at the level of UG, we can maintain a strong lexicalist view on language variation, similar to the view proposed within grammatical frameworks such as Minimalist grammars or LFG. Let us briefly discuss how the binary and unary type-forming operations out of which the lexical type-assignments are built, contribute to this lexicalist view variation.

**Binary connectives encode word order and constituent structure** We assume cross-linguistic variation in underlying word order and phrasal structure to be the result of differences in the selectional requirements of lexical type-assignments, expressed in terms of the two directional slash operations / and \. In this way, we can derive languages which have a SOV underlying word order or a SVO word order. Structural rules will be able to affect this fixed underlying word order and constituent structure only when the licensing control features are present.

**Functions of the unary connectives** Based on the logical properties of the unary connectives (see section 2.2), we can identify three functions of these connectives in accounting for structural variation in natural languages.

1. **Licensing displacement** Only those elements that are decorated with a structural ♦ feature and appear at the appropriate position in the structural configuration of one of the rules in the restricted set of displacement postulates may undergo displacement. The important point to realize here is that the appearance of a structural ♦ feature on a subtree or a leaf in the tree structure
2.5. The logic of variation

is projected from the lexicon. A □ feature on a main type or a ♦ feature as selectional requirement on a subformula projects a structural ♦. The following two logical rules cause a ♦ feature to appear on the structural side of a sequent.

\[ \frac{\Gamma \vdash □A}{\bigodot(\Gamma) \vdash A} \quad [\Box E] \]
\[ \frac{\Gamma \vdash A}{\bigodot(\Gamma) \vdash ▪A} \quad [\Diamond I] \]

The application of the structural rules causes the displacement of a substructure with the corresponding feature decoration. The trigger for the application of the structural rule, similar to the \textsc{move}-operation as formulated in the minimalist program (Chomsky, 1995), is feature checking. A feature is checked when the structural ♦ feature is canceled. A feature can be checked either by selectional requirements of another expression or by feature requirements on the type of the goal formula. In the natural deduction presentation style, the unary feature is checked when the unary feature appears on the right-hand side of the sequent. Therefore, after the application of structural rules, the structural ♦ feature needs to be ‘removed’ from the structural side and ‘moved’ to the categorial side to continue the derivational process. We repeat the two logical rules that remove a ♦ feature from the structural side and replace the structural ♦ occurrence with its logical variant.

\[ \frac{\bigodot(\Gamma) \vdash A}{\Gamma \vdash □A} \quad [\Box I] \]
\[ \frac{\Delta \vdash ▪A \quad \Gamma[\bigodot(A)] \vdash B}{\Gamma[\Delta] \vdash B} \quad [\Diamond E] \]

2. Blocking restructuring

Versmissen (1996); Kurtonina and Moortgat (1997) explore the use of the unary operators for licensing or constraining the structural rules. Depending on the placement of a unary connective on a syntactic type formula, the unary connectives may either allow or block the application of a structural rule. The displacement postulates only apply over purely binary branching trees. If one of the branches is unary branching the displacement postulate cannot be applied.

The best way to explain this function of the unary features is with an illustration.

Example 2.4 (Blocking or licensing of a structural rule) In the following two trees, substructure \( \Delta_2 \) is decorated with a unary feature ♦ which matches the requirement of the left displacement postulate Pl2. In the left tree, the application of the postulates is blocked by an intervening unary feature. In the tree on the right, the application of the structural postulate is licensed, because \( \Delta_2 \) is connected with the
other substructures by binary branching nodes.

**Blocking:**

```
\[ \begin{array}{c}
\Delta_1 \\
\Delta_2 \\
\Delta_3 \\
\end{array} \]
```

**Licensing:**

```
\[ \begin{array}{c}
\Delta_1 \\
\Delta_2 \\
\Delta_3 \\
\end{array} \]
```

In chapter 4, we will show that blocking the application of structural rules can be used to capture islands domains.

Unary control of the application of the structural module results from an interplay between features on lexical elements that occur displaced and features on lexical elements that require those features to be checked. The driving force behind the application of structural rules is feature checking. Heylen (1999) shows that the unary feature can be used to encode morphological features. The logical rules for the unary connectives function as feature checking procedures to cancel, for instance, the features of an argument against the selectional requirements of features of a verb phrase. Feature checking controls the analysis of complex expressions in a language that are either grammatical or ungrammatical due to their morphological properties.

3. Derivability patterns

Finally, there is also a purely logical use of the unary operators. The elimination and introduction rules for these connectives give rise to the following derivability patterns, expressing a notion of 'subtyping':

\[ \Diamond \Box A \vdash A \vdash \Box \Diamond A \]

The derivability relation \( \Diamond \Box A \vdash A \) reveals the use of the \( \Diamond \) operator as 'key' and \( \Box \) operator as 'lock'. The \( \Diamond \) connective on a formula 'unlocks' the \( \Box \) occurrence, but not the other way around. \( \Box \Diamond A \not\vdash A \). Bernardi (2002) exploits these derivability patterns to explore the semantic ordering relation between polarity sensitive items. In chapter 4, we will show that the derivability patterns above can be used to derive strict linear order relations among wh-phrases.

Concluding remarks on the logic of variation

The logic of variation (see definition 21, page 30) hypothesizes that cross-linguistic variation is the result of reasoning with a restricted set of struc-
2.5. The logic of variation

tural postulates and purely lexicalized derivational system, while a uniform interpretation of wh-questions is the result of higher-order type assignment to wh-elements. Our goal in this thesis is to provide empirical support for this hypothesis by presenting a broad cross-linguistic analysis of wh-question formation. We show that only the restricted set of rule schemata is needed to analyze wh-question formation and that cross-linguistic variation is purely lexicalized. We will show that we can account for the syntactic diversity and a uniform interpretation of wh-question formation on the basis of higher-order type assignment of wh-elements.

If our hypotheses on the logic of variation hold, this thesis suggests a number of questions for the formal and computation study of the framework we present. The thesis of Moot (2002) has shown that the complexity of a CTL framework allowing the full class of linear, non-expanding structural rules equals that of context-sensitive grammars. Studies of the computational properties of Tree Adjoining Grammars, Combinatory Categorial Grammars and Minimalist Grammars suggest that for the analysis on natural language phenomena beyond the context-free, the expressivity of so-called weakly context-sensitive grammars would be sufficient. Without going into the technicalities, the class of structural rules for which Moot demonstrates full context-sensitive expressivity only rules out copying and deletion of grammatical material — it allows any restructuring that respects this restriction. The displacement postulates that we consider to be available to UG in this thesis only use a very limited amount of the expressivity of the class Moot considers, which suggests the possibility of accommodating our framework among the weakly context-sensitive formalisms. As indicated before, our focus in this thesis is empirical, so we leave this as an open question for mathematical linguistics.
Chapter 3

A type schema for wh-question formation

The central aim of this chapter is to present a uniform account of wh-questions across languages. First, we discuss some basic characteristics of wh-questions. Then, we will analyze wh-questions by introducing a type schema for wh-phrases that encodes how wh-phrases merge with a question body. We use English as our point of departure, because wh-question formation has been studied extensively\(^1\) in this language.

In section 3.1, we informally present the different properties of wh-questions in English based on data presented in the field of generative linguistics. In section 3.2, we present a type schema for the analysis of wh-question formation in type-logical grammar. The type schema captures the structural variation and semantic uniformity that appears in wh-question formation. After a schematic presentation of the type schema and after showing how the type schema can be decomposed using the standard connectives of the grammatical reasoning system, we use the type schema in section 3.3 to analyze wh-question formation in English.

### 3.1 Properties of wh-questions in English

In this section, we discuss the basic properties of wh-questions and the function of the wh-phrase in a specific context. In a wh-question, a wh-phrase

---

1 The central issues have been raised by Ross (1967) who formulated islands constraints on transformations and by Chomsky (1977) who showed on the basis of a variety of constructions how it is possible to identify the application of wh-movement by means of a set of diagnostic properties.
associates with a constituent\(^2\) that is absent in the clause structure. More intuitively, the wh-phrase fulfills the grammatical requirements in the clause structure that would normally be fulfilled by a constituent phrase. The wh-phrase is sometimes referred to as *filler* and the associated constituent as *gap* or *gap hypothesis*.

We adopt the presentation common in generative syntax to indicate the position of the gap with a trace (\(t\)) and to co-index the wh-phrase and the gap. In main clause questions, the filler and the gap may appear in the same local domain, i.e. local wh-questions.

\[(3.1)\] Who(m) did John see \(t_i\)?

Or the filler and the gap may appear in different domains where the distance between the wh-phrase and its associated gap constituent appears to be unbounded, i.e. non-local wh-questions.

\[(3.2)\] Who(m) did Bill think Mary said John saw \(t_i\)?

Chomsky (1977) has set out diagnostics for recognizing the application of the wh-movement rule in a certain construction type (e.g. comparative construction, relative construction). His work has led to a large amount of syntactic research on wh-questions and related constructions (see Cheng and Corver (To appear)). We limit ourselves here to providing a general overview of different wh-questions and constraints on wh-question formation. We divide the phenomena into three groups. First, we discuss local wh-questions, which are constructions consisting of a wh-phrase and a gap in the same clause; we distinguish between direct questions and indirect questions. Second, we discuss non-local wh-questions, which are constructions where the wh-phrase and the gap appear in different clauses. Third, we discuss locality constraints on the seemingly unboundedness of wh-question formation. These constraints are referred to as island constraints. Finally, we discuss multiple wh-questions.

### 3.1.1 Local wh-questions

We use the term “local wh-questions” to refer to wh-interrogative clauses where a wh-interrogative appears fronted and occurs in the same local domain as the constituent gap with which the wh-phrase associates. We distinguish between two kinds of local wh-questions: direct questions where the wh-phrase appears fronted in the main clause and indirect or embedded questions where the wh-phrase appears fronted in the subordinate clause.

#### 3.1.1.1 Direct questions

Direct questions in English are characterized by the presence of a wh-phrase at the beginning of the sentence. The fronted wh-phrase associates with a gap in

\(^2\)We mainly focus on argument wh-phrases. However, the same line of reasoning applies to adjunct wh-phrases.
the body of the sentence. The following examples illustrate the kind of direct questions that can be formulated on a declarative clause.

\[(3.3)\] John gave a present to Mary (yesterday/because of her birthday).

- a. Who gave a present to Mary?
- b. What did John give to Mary?
- c. Who did John give a present to?  
  or To whom did John give a present?
- d. When/why did John give a present to Mary?

In example 3.3a), the wh-phrase ‘who’ fills the subject argument. In this case, the wh-phrase even occupies the same structural position as the subject noun phrase would have (i.e. John gave a present to Mary).\(^3\) In example b), the wh-phrase associates with the direct object argument position, while in example c) the wh-phrase associates with an indirect object gap in the main clause. Example d) illustrates an adjunct wh-phrase that has the grammatical function of a sentential modifier.

English has the property that main clause direct questions with non-subject interrogative phrases need do-support. Additionally, as illustrated in example c), wh-phrases that are associated with an indirect object gap which is introduced by a preposition can optionally leave the preposition stranded or drag the preposition along with the wh-phrase.

3.1.1.2 Indirect questions

Similar to direct questions, indirect questions or embedded questions are characterized by the presence of a wh-interrogative at the beginning of the subordinate clause or headed by ‘if’ or ‘whether’. Verbs such as ‘know’, ‘wonder’ or ‘forget’ can be merged with an embedded question. Below, we list examples of the embedded interrogatives a verb such as ‘wonder’ can select.

\[(3.4)\] a. John wonders whether Mary saw a bird.
- b. John wonders who saw a bird.
- c. John wonders when Mary saw a bird.

Embedded interrogatives differ from direct questions. The following examples illustrate that no do-support (b), or subject-auxiliary inversion (c) is needed to form an embedded interrogative (a).

\[(3.5)\] a. John wondered what Mary saw.

\(^3\)In early generative syntax, wh-questions were assumed to be derived from a displacement operation (Chomsky, 1977) where the wh-phrase moves from the base position (i.e. the position where the noun phrases gets its \(\theta\) role) to the front of the main clause. In local wh-questions with a subject wh-phrase, the wh-phrase appears at the base position. Therefore, the movement of the subject wh-phrase to the ‘front of the main clause’ has been referred to as vacuous movement.
b. *John wondered what did Mary see.

c. *John wondered what had Mary seen.

3.1.2 Non-local wh-questions

Under some restrictions, a wh-phrase may occur clause-initially in the matrix clause while its associated gap appears in an embedded clause. Such constructions are restricted to embedded clauses that are selected by a verb which belongs to the group of bridging verbs (Erteshik, 1973) such as say, believe, claim or think. These verbs function as a bridge between the wh-phrase and the associated argument position. An unbounded number of embedded clauses may occur between the clause-initial wh-phrase and the bound argument position. We will refer to such constructions featuring a “filler”, the wh-phrase, and a “gap”, the associated argument position appearing in different clauses as long-distance dependencies.

(3.6) Mary saw a bird.

a. Who did Sue believe (that) saw a bird?

b. What did Sue believe (that) Mary saw?

c. What did Sue believe (that) John said (that) Mary saw?

d. What did Sue believe (that) John said (that) Bill claimed (that) Mary saw?

Notice that non-local wh-questions need do-support in the main clause, if a wh-word at the left periphery of the main clause associates with an argument position in an embedded clause. Furthermore, while an embedded clause is normally introduced by the overt complementizer ‘that’, the complementizer can optionally be omitted when it concerns a non-subject wh-phrase (see examples 3.6b-d). The complementizer must be absent when a subject wh-phrase occurs clause-initially.

The main observation for non-local wh-questions is that the matrix verb belongs to the group of bridge verbs (Erteshik, 1973). The examples in 3.6 support this observation. Additional support for the ‘bridging’ function of verbs like ‘believe’ comes from data on relative clause constructions (ex.3.7a) and topicalization (ex.3.7b) which show similar constructions with these verbs.

(3.7) a. A bird which Sue believed Mary saw. (relativization)

b. This bird, I believe Mary saw. (topicalization)
3.1.3 Island constraints

In the previous section, we have explored interrogative clauses where the wh-phrase associates with a gap over a long distance. However, sometimes, a wh-question cannot be formed because the gap is part of an island (Ross, 1967). We present a brief overview of the different types of constructions that form an island: clauses, adjuncts, complex noun phrases, coordinate structures and left branch constructions (see also Szabolcsi (2006) for an overview). In the examples taken from Szabolcsi (2006), the gap position is marked with $i$ and is co-indexed with the wh-phrase. The beginning and the end of an island construction are marked with square brackets.

**Wh-island constraint** A wh-phrase cannot associate with a gap in an embedded wh-interrogative clause.

(3.8)  
\begin{align*}
\text{a. } & \ast \text{Which bird, did John wonder [whether Mary wanted to see } i \text{]?} \\
\text{b. } & \ast \text{Which bird, did John wonder [who wanted to see } i \text{]?}
\end{align*}

**Adjunct island constraint** A wh-phrase cannot associate with a gap in a clausal adjunct.

(3.9) $\ast$Which topic, did John leave [because Mary talked about $i$]?

**Complex noun phrase constraint** A wh-phrase cannot associate with a gap in a complex noun phrase. Complex noun phrases are nouns that are modified either by a restrictive relative clause or a complement clause.

- complex NP built with a restrictive relative clause
  
  (3.10) $\ast$Which kid, must you call [the teacher who punished $i$]?

- complex NP built with a complement clause
  
  (3.11) $\ast$Which man, did you hear [the rumor that my dog bit $i$]?

**Coordinate structure constraint** A wh-phrase cannot associate with a gap in a coordinate structure.

(3.12) $\ast$Which man, did you invite [Mary and $i$]?

Exceptions to the coordinate structure constraint are “across-the-board” constructions consisting of a fronted wh-phrase that associates with a gap in each of the two conjuncts (Ross, 1967).

(3.13) Which man, did you invite [[a friend of $i$] and [a brother of $i$]]?

**Left branch constructions** A wh-phrase cannot associate with a left branch element which enters into a relation with a noun or an adjective, e.g. wh-determiner (‘which’, wh-possessor (‘whose’) or wh-degree word (‘how’).
3. A type schema for wh-question formation

(3.14)  a. *Which, did you see [t, picture]?
     b. *Whose, did you see [t, picture]?
     c. *How, is he [t, tall]?

Notice that the sentences in example 3.1.3 are fine if ‘picture’ or ‘tall’ are pied-piped.

(3.15)  a. Which picture, did you see [t, picture]?
     b. Whose picture, did you see [t, picture]?
     c. How tall, is he [t, tall]?

3.1.4 Multiple wh-questions

Multiple wh-questions in English are characterized by the presence of two or more wh-phrases where only one wh-phrase appears in clause-initial position while the other wh-phrases stays in-situ. The wh-phrase that occurs clause-initially is the wh-phrase whose associated constituent gap is closest to its associated gap. This phenomenon was initially explained by Chomsky (1973) with the superiority condition which stated a constraint on movement of wh-phrases. Pesetsky (1987) illustrates the constraint on the formation of multiple wh-questions with the following examples.

(3.16)  a. Mary asked who, t, read what?
     b. *Mary asked what, who read [tj, t]?

(3.17)  a. Who, did John persuade t, to read what?
     b. ?? What, did John persuade who(m) to read [tj, t]?

In examples 3.16a) and 3.17a) the clause-initial wh-phrase is closer to its associated gap than in the sentences in the b) examples. The superiority condition correctly explains the contrast between the a) and b) sentences. However, as Pesetsky (1987) points out, discourse-linked phrases (D-linked) form an exception to the strict ordering of wh-phrases. Replacing the non-D-linked wh-phrases ‘who’ and ‘what’ with D-linked wh-determiner phrases causes the b) sentences to be grammatical as well, as is illustrated in the following examples.

(3.18)  a. Mary asked which man, t, read which book.
     b. Mary asked which book, which man [tj, t].

(3.19)  a. Which man, did you persuade t, to read which book?
     b. Which book, did you persuade which man to read [tj, t]?

The superiority condition is defined in terms of the c-command relation: the trace of the fronted wh-phrase must c-command the in-situ wh-phrase.
3.2. Wh-type schema

The phrase ‘which book’ may either occur in clause-initial position preceding the other wh-phrase (ex. 3.18a, 3.19a) or in-situ (ex. 3.18b, 3.19b). Thus, the strict ordering between wh-pronouns does not hold for D-linked wh-phrases.5

Summary

In this section, we have presented a brief overview of the characteristics of wh-questions in English. We have mainly focused on the syntactic construction of wh-questions by showing how the position of the wh-phrase relative to its associated gap yields grammatical or ungrammatical sentences. The common order of English direct and indirect questions where the gap appears in the same local domain as the wh-phrase is that the wh-phrase appears clause-initial. In non-local wh-questions, or long-distance wh-questions, whether the wh-phrase occurs in the left periphery is determined by the gap position where the wh-phrase associates with. The gap should not be contained within an island configuration (Ross, 1967). Finally, we have illustrated multiple wh-questions featuring a single wh-phrase in clause-initial position and the other wh-phrases in an in-situ position. The multiple wh-question is subject to the superiority constraint. However, D-linked wh-phrases do not seem to obey this constraint.

3.2 Wh-type schema

To analyze the syntactic and semantic properties of wh-question formation in type-logical grammar, we will use a wh-type schema to lexically identify wh-elements. The selectional requirements of wh-phrases are encoded into this operator type schema and cause a uniform interpretation of wh-questions. Although semantically uniform, we need to express the structural differences between wh-questions. We recognize three structural variants of the wh-type schema that along with the fixed set of structural rules account for structural variation in wh-question formation. The type schemata are additional operators which are decomposable into the usual type-connectives of the base logic.

In section 3.2.1, we introduce a general type schema for encoding the syntactic and semantic properties of wh-phrases. We explain the basic components of the type schema and present a schematic inference rule for the use of the type schema. In section 3.2.2, we show that the semantic interpretation of the wh-type schema maps the syntactic properties to a semantic representation. Although, semantically uniform, structurally wh-phrases may differ. We distinguish between two wh-ex-situ type schema for wh-phrases that occur fronted (section 3.2.3) and a wh-in-situ type schema (section 3.2.4). In section 3.2.3, we explain that for the wh-ex-situ type schema we need two structural

5Pesetsky (1987) explains this on the basis of a semantic difference between D-linked and non-D-linked wh-phrases. Answers to D-linked questions have to be definite. Whereas answers to wh-questions with non-D-linked wh-pronouns, such as ‘who’ or ‘what’ can range from definite, indefinite to non-referring.
variants to encode the structural position of the gap relative to the wh-phrase. We decompose the wh-ex-situ type schema in the base logic and illustrate how the structural differences between the two ex-situ types are accounted for with the fixed set of displacement postulates of the structural module. Along the same lines, in section 3.2.4, we decompose the wh-situ type schema into connectives of the base logic. We show how for wh-in-situ phrases, we can account for the right scope interpretation in the base logic along with the structural module. Finally, in section 3.2.5, we group the syntactic and semantic properties of the different variants of wh-phrases into synthesized inference rules.

3.2.1 A general type schema for wh-phrases

Taking the characteristics of wh-question formation as they have been presented in the previous section into account, we now consider a uniform account of wh-question formation. We incorporate the syntactic and semantic aspects of wh-questions into the types assigned to wh-phrases. Syntactically, the position of the wh-phrase in the sentence may differ. The wh-phrase may occur fronted, in the case of ex-situ wh-phrases, or it may stay embedded, in the case of in-situ wh-phrases. Semantically, regardless of whether the wh-phrase occurs ex-situ or in-situ, the wh-phrase takes scope over the clause that forms the question body.

To account for the distinct syntactic properties but uniform interpretation of wh-questions, we will use a type schema to lexically identify the wh-phrases. For now, we generalize over the different structural variants of wh-phrases and concentrate on the overall syntactic properties of wh-phrases. The choice for a uniform type schema is determined by our goal of deriving a uniform meaning assembly of wh-questions. First, we concentrate on the syntactic properties of the type schema. Then, in section 3.2.2 we show the semantic composition of the type schema.

We adopt the setup of the $q$-type schema which was proposed by Moortgat (1991) to account for in-situ binding of generalized quantifier phrases. Generalized quantifiers share many properties with wh-phrases. In English, quantifier phrases occupy the same position as a noun phrase would, while semantically quantifier phrases are interpreted outside the clause they are embedded in. Consider for instance the sentence, ‘Bill sees someone’ with lambda semantics: $\exists x. ((\text{see } x) \text{ bill})$. Moortgat (1991) shows how the interpretation of quantifier phrases is derived from the syntactic analysis of the $q$-type schema. We use a similar type schema as the $q$-type schema to account for the syntactic and semantic analysis of wh-phrases.

We propose a three-place type schema, $\text{WH}$, ranging over three subtypes: $\text{WH}(A, B, C)$. The three variables indicate subtypes of the grammatical structures where a wh-phrase acts on. $B$ is the type of the body of the wh-question; $A$ is the type of the gap hypothesis contained in the body; $C$ is the type of the result of merging the body of the wh-question with the wh-phrase. Schemat-
3.2. Wh-type schema

ically, the following inference rule defines the merging of an arbitrary wh-phrase (= Γ) with a body of a wh-question (= Δ) which contains a gap hypothesis (= Δ[A]). The result of merging the wh-phrase with the body is a structure where the wh-phrase replaces the gap hypothesis in the structure (= Δ[Γ]).

\[
\frac{\Gamma \vdash WH(A, B, C) \quad \Delta[A] \vdash B}{\Delta[Γ] \vdash C} \quad [WH]
\]

Alternative presentations of the same inference rule are the tree style derivation, which illustrates more intuitively what happens with the structure of a sentence, and an abbreviated natural deduction rule. We will use both presentation styles to illustrate the derivation of a sentence.

- tree style derivation

```
\begin{array}{c}
\Delta
\end{array}
\Gamma \vdash WH(A, B, C)
\Rightarrow
\begin{array}{c}
\Delta
\end{array}
\Delta[A] \vdash B
\Delta[Γ] \vdash C
```

- abbreviated natural deduction style

```
Δ[A] \vdash B
\Rightarrow
Γ \vdash WH(A, B, C)
Δ[Γ] \vdash C
```

The schematic rule provides the necessary type information that is needed to construct a wh-question. To characterize a wh-phrase, we need to define the type of the gap, the type of the structure that the wh-phrase merges with, and the type of the wh-question that is the result of merging the wh-phrase with the body. For now, wh-questions are typed as wh.\(^6\)

As an example, we analyze the direct question ‘Who saw Bill?’. Note that for now we generalize over the distinct structural positions that ‘who’ can occupy. We assign the wh-phrase ‘who’ the wh-type schema, WH(np, s, wh). The following derivation shows the analysis of the wh-question in a natural deduction style with the abbreviated inference rule for merging the wh-phrase. Alternatively, we present the last two steps of the derivation in a tree style presentation.

\(^6\)In chapter 5, we show that the wh type for wh-questions is a type abbreviation for several types of wh-questions which incorporate possible answer types into the type for wh-questions.
The main clause is built as usual, only the subject argument phrase is now filled by a hypothesized np argument. When the main clause is derived, the wh-phrase is merged replacing the np hypothesis, yielding a clause of type wh.

Before we treat the different structural variants of the wh-type schema, we first present the semantic composition.

### 3.2.2 Semantic composition

In chapter 2, section 2.3, we have presented the interface between syntactic and semantic types. Following the Curry-Howard correspondence each syntactic type formula is mapped to a corresponding semantic type. In turn, we interpret the syntactic type by providing a semantic term that matches the semantic type.

Schematically, the semantic type that corresponds to the wh-type schema takes the corresponding semantic types of each subtype in the type schema and arranges them. Wh-type schema $\text{WH}(A, B, C)$ maps to the following semantic type:

$$\frac{A \rightarrow (2) B}{\text{WH}(A, B, C) \rightarrow (1) C}$$

The semantic type reveals the inherent steps encoded in the rule schema. $\rightarrow (1)$ is the application step, merging a wh-phrase with the body. $\rightarrow (2)$ represents abstraction of the hypothesis, withdrawing the gap from the body of the wh-question. In section 3.2.3.2, when we decompose the wh-type schema in the base logic, we see more clearly how the syntactic type matches with the application and abstraction step represented in the semantic type.

The meaning assembly in the lambda style semantics shows the semantic relation between the wh-phrase and its gap. As a binding operator, we assume...
3.2. Wh-type schema

a semantic operator $\omega$. In chapter 5, we show how this semantic operator is interpreted in a lambda style semantics. The semantic operator $\omega$ extends the group of logical constants $\exists$ and $\forall$. $\omega$ binds the gap as a variable in the body of the clause. The semantic term assigned to wh-type schema $WH(A, B, C)$ is:

$$(\omega \lambda x. BODY)\downarrow$$

In this term, $BODY$ is the term computed for the body of the wh-question. The term $BODY$ contains the gap variable $A$.

Incorporating the meaning assembly in the inference rule for the type schema, we obtain the following inference rule:

$$\Gamma \vdash \omega : WH(A, B, C) \quad \Delta[x : A] \vdash BODY : B$$

$$\Delta[I] \vdash \omega \lambda x. BODY : C \quad [WH]$$

The semantic composition of the $WH$-rule reflects the syntactic operation of the wh-type schema. The wh-phrase with term $\omega$ is merged with a structure of type $B$ with the semantic term $BODY$. The body contains a hypothesis of type $A$ with a semantic term variable $x$. After merging the wh-phrase with the question body, the wh-phrase replaces the gap hypothesis $\Delta[I]$ which is reflected in the meaning assembly as a $\lambda$-abstracting over the term variable $x$. The wh-operator $\omega$ in turn applies to the computed semantic term.

As an illustration, we present the meaning assembly of the wh-question 'Who saw Bill?':

```
saw

\[ np \vdash x : np \]\n
\[ saw \circ bill \vdash (see b) : np \]\n
\[ np \circ (saw \circ bill) \vdash ((see b) x) : s \quad [/E] \]

\[ who \vdash \omega : WH(np, s, wh) \quad [WH] \]

\[ who \circ (saw \circ bill) \vdash \omega \lambda x. ((see b) x) : wh \]
```

The general rule schema and the wh-type schemata give the necessary type encoding of the elements involved in the formation of a wh-question. However, we need to add a distinction in the wh-type schema to account for the structural differences between wh-phrases. Structurally, we distinguish between wh-phrases that occur fronted (ex-situ) and wh-phrases that occur in-situ. The distinction is encoded in the type schema by adding a subscript to the type connective: $WH_{ex}$ for ex-situ types and $WH_{in}$ for in-situ types. We will show that due to the isomorphism between types and terms, the semantic composition of the two versions of the wh-type schemata is the same.

### 3.2.3 Wh-ex-situ type schema

Ex-situ wh-phrases occupy a fronted position in a clause. However, we do not only concentrate on the position of a wh-phrase, but also take into account
A type schema for wh-question formation

the underlying position of the associated gap. While the wh-phrase appears clause initially, the gap occurs embedded in the structure that follows the wh-phrase. Depending on the basic word order of the clause structure, the gap may appear on a left or a right branch in the tree structure. We incorporate the directionality of the gap position in the type-assignment of the wh-phrase.

A wh-ex-situ phrase can be associated with an embedded phrase position if the hypothesized element is ‘visible’ for the wh-phrase. A hypothesis is structurally visible if it occurs at the edge of a structure, in which case the wh-phrase may be merged with the clause and replace the occurrence of the hypothesis. The hypothesized element can occur either on a left or on a right edge of a structure depending on the selectional requirements of the grammar. Let us incorporate this structural information about the occurrence of the hypothesis into the syntactic decomposition rule that was proposed in section 3.2.1. We obtain two versions of the syntactic rule which define the position that a hypothesized element has to occupy before merging the wh-element to the structure. In both variants, the wh-phrase is inserted to the left of the structure. The only difference between the two ex-situ type schemata is the requirement on the structural position of the gap.

3.2.3.1 Two variants of wh-ex-situ

We distinguish between a wh-ex-situ type which associates with a gap on a left branch and a wh-ex-situ type which associates with a gap on a right branch. We encode the structural difference between the two ex-situ types by adding a superscript to the type schema.

wh-ex-situ left A wh-ex-situ left is merged with a structure of type $B$ which contains a gap hypothesis of type $A$ on the left edge. The wh-phrase replaces the hypothesis and is inserted at the left edge of the structure.

\[
\Gamma \vdash \text{WH}\_l^l(A, B, C) \quad A \circ \Delta \vdash B
\]

\[
\Gamma \circ \Delta \vdash C
\]

\[
[\text{WH}\_l^l]
\]

wh-ex-situ right A wh-ex-situ right is merged with a structure of type $B$ which contains a gap hypothesis of type $A$ on the right edge. The wh-phrase replaces the hypothesis and is inserted at the left edge of the structure.
3.2. Wh-type schema

\[ \frac{\Gamma \vdash \text{WH}_r(A, B, C) \quad \Delta \circ A \vdash B}{\Gamma \circ \Delta \vdash C} \quad \text{[WH\_r]} \]

3.2.3.2 Decomposition of wh-ex-situ

Although the two ex-situ types differ on the structural position of the gap, semantically the two wh-type schema map to the same semantic type. The semantic representation constructed on the basis of the semantic type as presented in section 3.2.2 is applicable to both ex-situ type schemata. To show that the ex-situ types share a uniform semantic representation, we decompose the type schemata into types in the base logic. Using the decomposed types, we can construct a derivation in the base logic. For these syntactic derivations we can provide the meaning assembly which show that the semantic representation of the two wh-ex-situ type schemata is the same.

The two ex-situ type schemata and the corresponding rules can be written out as:

\[ \text{wh}_l(A, B, C) = C / (A \backslash B) \]

\[ \text{wh}_r(A, B, C) = C / (B / A) \]

3.2.3.3 Semantic composition of wh-ex-situ

To check whether the semantic representation of the left and the right ex-situ types is the same, we give an abstract derivation of the two wh-ex-situ types. We compute the corresponding meaning representation at each step in the derivation. The semantic term labeling corresponds with syntactic derivation: Lambda abstraction matches the introduction step of either / for wh-ex-situ right or \ for wh-ex-situ left, while application matches the elimination step of / or \.
The meaning assembly of the decomposed type for wh-ex-situ wh-phrases is given in the following derivation:

\[
\Gamma \vdash \omega : C/(A \setminus B) \\
\Delta \vdash \lambda x.\text{BODY} : A \setminus B \quad [\setminus I] \\
\Gamma \circ \Delta \vdash \omega \lambda x.\text{BODY} : C \quad [\\/ E]
\]

Figure 3.1: Semantic composition of wh-ex-situ left rule

The same meaning assembly results from the decomposed type for wh-ex-situ right phrases. The semantic representation is derived as follows:

\[
\Gamma \vdash \omega : C/(B/A) \\
\Delta \vdash \lambda x.\text{BODY} : B/A \quad [\setminus I] \\
\Gamma \circ \Delta \vdash \omega \lambda x.\text{BODY} : C \quad [\\/ E]
\]

Figure 3.2: Semantic composition of wh-ex-situ right rule

The syntactic decomposition along with its corresponding semantic representation of the two wh-ex-situ type schemata show that although structurally variant, the meaning assembly for the wh-type schema is uniform. The difference between the two ex-situ schemata is only structural. For the left ex-situ type schema the hypothesis has to occur on a left edge, for the right ex-situ type schema the hypothesis has to occur on a right edge.

3.2.3.4 Displacement postulates

So far, we have assumed that the gap hypothesis of a wh-phrase appears on the edge of a structure. However, the gap may originate more deeply embedded in the structure depending on the underlying clause structure. When the hypothesized element resides deeply embedded in the clause structure, we need to move the hypothesis to the edge of the structure. In short, wh-phrases themselves are not displaced, it is the hypothesized constituent phrase that is displaced from a more deeply embedded position in the structure to the edge of a structure. The restricted set of displacement postulates, as listed in section 2.5 on page 31 and reprinted below, control the displacement of the hypothesis to the edge of a structure.

left displacement postulates Move a feature decorated element on a left branch to a left branch one node higher:
right displacement postulates  Move a feature decorated element on a right branch to a right branch one node higher:

\[
\begin{align*}
\Gamma[\Delta_1 \circ \Delta_2] \vdash C & \quad [P1'] \\
\Gamma[\Delta_1 \circ (\Delta_2 \circ \Delta_3)] \vdash C & \quad [P2']
\end{align*}
\]

Unary features as trigger for displacement  The displacement postulates are triggered when a local tree structure matches the structural rule pattern. Only when the element subject to displacement is decorated with a ♦ feature, can the displacement rule be applied. Therefore, to be able to associate a wh-phrase with a gap hypothesis that occurs more deeply embedded in the tree structure, we need to add features to the gap type of the wh-ex-situ types.

The logical behavior of the ♦ and □ features forces us to add a combination of ♦□ on the gap hypothesis to control the displacement postulates. The following derivation shows how the combination of ♦□ on a hypothesized element projects the occurrence of a ♦. Based on the derivability relation: ♦□A ⊢ A, we can abbreviate these steps and collapse the rule into an abbreviated format.

Natural deduction format:  

\[
\begin{align*}
\Box A & \vdash \Box A & [\Box E] \\
\Diamond(\Box A) & \vdash A \\
\Box \Box A & \vdash \Box A \\
\vdots & \\
\Box \Box \Box A & \vdash A \\
\Gamma[\Box \Box \Box A] & \vdash C & [\Box E]
\end{align*}
\]

Abbreviated format:  

\[
\begin{align*}
\Box A & \vdash A & [\Box E] \\
\Diamond \Box A & \vdash C
\end{align*}
\]

To sum up: The displacement postulates encode the displacement of a constituent on a left or right branch to a left or right branch one node higher up in the structure. Only when a hypothesized argument phrase is decorated with the right features, can the element be displaced. The hypothesis is moved to the left edge of the structure when it occurs on a left branch. If it occurs on a right branch it has to move to the right edge of the structure. Therefore, the underlying position of the gap hypothesis in the body of the question determines whether the wh-phrase can be merged with the question body.

In section 3.3, we use the two ex-situ types to account for local and non-local dependencies in wh-question formation in English. Anticipating the analysis of wh-question formation in English, we schematically present the

---

\(^7\)For instance, with the analysis of non-local dependencies, see section 3.1.2
two instances of ex-situ wh-phrases: wh-ex-situ left requires a hypothesis on a left edge and wh-ex-situ right requires a hypothesis on a right edge. To move the hypothesis from a more deeply embedded position to the edge of a structure, we cyclically (cyclicity is indicated as \( \text{Whr}^* \) or \( \text{Whl}^* \)) apply the displacement postulates. When the hypothesis occurs on a right branch, we apply the right displacement postulates ([Pr*]), and vice versa for a hypothesis on a left branch.

\[
\begin{align*}
\Gamma [\diamond A \circ \Delta] & \vdash B \\
\vdots & [Pl*] \\
\diamond A \circ \Gamma [\Delta] & \vdash B
\end{align*}
\]

With the addition of \( \diamond \) features on the gap type, the wh-ex-situ types are written out using connectives of the base logic as follows.

\[
\text{wh}_{\text{le}}(\diamond A, B, C) = C \backslash (\diamond A \backslash B) \\
\text{wh}_{\text{re}}(\diamond A, B, C) = C / (B / \diamond A)
\]

The following derivation illustrates the analysis of displacing the gap hypothesis for the decomposed types of the wh-ex-situ type schemata.

\[
\begin{align*}
[\diamond A \vdash A] & \quad [\diamond A \vdash A] \\
\vdots & \quad \vdots \\
\Delta [\diamond A \circ \Delta'] & \vdash B \\
\vdots & [Pl*] \\
\diamond A \circ \Gamma [\Delta'] & \vdash B \\
\Delta [\Delta'] & \vdash \diamond A \backslash B
\end{align*}
\]

Blocking restructuring  Besides the directionality of the position of the gap hypothesis, the path along which the gap hypothesis is to be reached also plays a role. The displacement postulates only apply over purely binary branching trees. If one of the branches is unary branching the displacement postulate cannot be applied. In section 2.5.3, we have explained how the unary features can block the application of structural rules. We reprint the two schemata that illustrate blocking or licensing of the application of a structural rule in figure 3.3.

In other words, a gap hypothesis is only to be reached via a pure \( \diamond \) path where no \( \text{\bullet} \) branch blocks the application of structural rules. We show how blocking is used for the analysis of island constructions in section 3.3.4.

3.2.4 Wh-in-situ type schema

The structural position of a wh-in-situ phrase is the same position as the position of the associated constituent. The elimination rule for the wh-in-situ type
3.2. Wh-type schema

Blocking:

\[ \Delta_1 \rightarrow \Delta_2 \]

\[ \Delta_2 \rightarrow \Delta_3 \]

Licensing:

\[ \Delta_1 \rightarrow \Delta_2 \]

\[ \Delta_2 \rightarrow \Delta_3 \]

Figure 3.3: Blocking or licensing the application of a structural rule

encodes that the wh-phrase occupies the same position as the hypothesized constituent.

**wh-in-situ** A wh-in-situ is applied to a structure of type \( C \) which has an embedded hypothesized phrase of type \( A \). The wh-phrase replaces the hypothesized phrase at the same position in the structure.

\[
\Gamma \vdash \text{WH}_{\text{in}}(A, B, C) \quad \Delta[A] \vdash B \quad [\text{WH}_{\text{in}}]
\]

A wh-in-situ merges with a question body of type \( B \) at the structural position of the gap hypothesis and yields a structure of type \( C \). The meaning assembly of the wh-in-situ questions is the same as the meaning assembly of wh-ex-situ questions:
3. A type schema for wh-question formation

\[ \Gamma \vdash \omega : \text{WH}_{\text{in}}(A, B, C) \quad \Delta[x : A] \vdash \text{BODY} : B \]
\[ \Delta[\Gamma] \vdash \omega \lambda x.\text{BODY} : C \]

[\text{WH}_{\text{in}}]

To understand the syntactic and semantic behavior of a wh-in-situ phrase, we decompose the wh-type schema for in-situ wh-phrases into the type connectives of the base logic.

3.2.4.1 Decomposition of wh-in-situ

Wh-in-situ phrases have many resemblances with quantifier phrases. Similar to wh-in-situ phrases, quantifier phrases appear in an embedded position but they can take non-local scope, outside the clause they are embedded in. The \(q\)-type schema, \(q(A, B, C)\), that has been proposed for the analysis of generalized quantifier phrases forms the background of our analysis (Morrill, 1994; Moortgat, 1991). Several proposals have been made to solve the type equation for the \(q\)-type schema in terms of the elementary connectives of the type-logical grammar (see for instance Morrill (1994); Moortgat (1991)). The existing proposals have their drawbacks in that the definition of \(q\) involves construction-specific extra binary mode distinctions, and special-purpose structural postulates. Since in this thesis, we assume the available patterns for structural reasoning are fixed by UG, we cannot make use of these proposals. A new proposal by Moortgat (2005), however, is indeed compatible with our fixed structural module. The proposal is a component of a broader line of research which enriches the categorial vocabulary with duals of the connectives (a co-product, and co-implications) and interactions between the regular and the dual connectives. In this setting, scope construal does not involve structural reasoning at all. The definition of \(q\) that we will use below is presented by Moortgat (2005) as an approximation of what is obtained automatically in the dualized setup. Since this approach available only in the concluding stage of writing this thesis, we will content ourselves with the approximation, and leave the reanalysis of our treatment of in situ phenomena in terms of the richer language of connectives as a subject for further research.

As the type schema for wh-in-situ phrases shows, structurally, the wh-phrase replaces the gap at the position where the associated element is interpreted. However, semantically, the wh-phrase may have scope over the clause where it is embedded in. To compute such a semantic scope reading, the hypothesized element has to raise to the edge of the structure that forms the body of the wh-question. This would be the position at which a ex-situ wh-phrase would be inserted and interpreted; instead a wh-in-situ is inserted at the embedded position where the gap originally resides. We write out the wh-in-situ type schema using the type connectives of the base logic in order to show that the wh-phrase is inserted in the embedded position and gets the right scope interpretation.
3.2. Wh-type schema

\[ \text{WH}_{\text{int}}(A, B, C) = \Diamond \Box (C/B) \bullet A \]

\( (\text{scope marker'}) \quad (\text{gap'}) \)

The type consists of two parts composed with the \( \bullet \) connective. We refer to the left part as \textit{scope marker} as it will mark the position in the clause over which the wh-phrase takes scope. And we refer to the right part as the \textit{gap}. The function of these two parts becomes clear when we inspect how the type is used in a derivation. Figure 3.4 shows a schematic presentation of a natural deduction derivation for merging the decomposed wh-in-situ phrase \( \Gamma \) with a question body \( \Delta \).

\[
\begin{align*}
[A \vdash A] \quad & \quad \text{step 1} \\
[\Diamond \Box (C/B) \vdash C/B] \quad & \quad \Delta[A] \vdash B \quad \text{step 2} \\
\Delta[A] \vdash C \\
\text{step 3} \\
\Delta[\Gamma] \vdash C
\end{align*}
\]

Figure 3.4: Decomposition of wh-in-situ phrase

In the derivation, two hypotheses are assumed: the gap hypothesis \( A \) and the scope marker hypothesis \( \Diamond \Box (C/B) \). First, the body of the question is built up, \( \Delta[A] \) (step 1). When the body of the question of type \( B \) is derived, the hypothesis \( \Diamond \Box (C/B) \) merges with the question body (step 2). The scope marker must move to the in-situ position of the \( A \) gap (step 3). Once the scope marker and the gap are structurally adjacent, the wh-in-situ type is inserted with a \( \bullet E \) step and replaces the two hypotheses.

3.2.4.2 Semantic composition of wh-in-situ

We have demonstrated how the wh-in-situ type can be decomposed using the standard connectives of the base logic. Now we will focus on the semantic composition of the wh-in-situ type. Although the syntactic decomposition of the wh-in-situ types differs from the wh-ex-situ type, we want the meaning assembly of wh-questions with wh-in-situ phrases to be the same as wh-questions with wh-ex-situ phrases. To show that this is indeed the case, we first determine the lambda term that matches the type assignment of the decomposed wh-in-situ type. Then, using this term we compute the meaning assembly for merging a wh-in-situ type to a question body.
The decomposed type consists of the composition of two types \( \diamond \Box (C/B) \) and a type \( A \). These types map to semantic types, \( B \rightarrow C \) and \( A \) respectively.

The semantic type of the whole decomposed wh-in-situ type is the product of the two semantic types: \( (B \rightarrow C) \odot A \). The lambda term of the decomposed wh-type must match this semantic type. We will again use the semantic operator \( \omega \) to bind the term variable of the gap hypothesis in the lambda term of the question body. Recall that \( \omega \) is of type \( (A \rightarrow B) \rightarrow C \). The term that will compute the right meaning assembly is given below. Notice, though, that it is precisely at this point that the proposal of Moortgat (2005) is an approximation: the variable \( x \) is in fact treated as a metavariable. As a proper variable, \( x \) would be bound in the first component of the pair term, and free in the second. Treating \( x \) as a metavariable, we make sure that when terms are standardized alphabetically, the two occurrences of \( x \) in the pair term will be affected simultaneously. To alert the reader to the meta-character of the variable, we have distinguished it typographically.

\[
\langle \lambda y. (\omega \lambda X.y)^{\Gamma}, X^{\Delta} \rangle : \diamond \Box (C/B) \bullet A
\]

In figure 3.5, we illustrate the meaning assembly of the whole wh-question by computing the lambda term for each syntactic step in the derivation.

\[
\begin{array}{c}
\vdash u : \diamond \Box (C/B) \quad \Delta[u : A] \vdash \text{BODY} : B \\
\hline \\
\vdash u : \diamond \Box (C/B) \odot \Delta[u : A] \vdash (u \text{ BODY}) : C
\end{array}
\]

\[
\Gamma \vdash \diamond \Box (C/B) \bullet A \\
\hline \\
\langle \lambda y. (\omega \lambda X.y), X \rangle \\
\Delta[u : \diamond \Box (C/B) \odot \Delta[u : A] \vdash (u \text{ BODY}) : C \\
\hline \\
\Delta[\Gamma] \vdash (\lambda y. (\omega \lambda X.y) \text{ BODY}) \quad [\bullet E]
\]

Figure 3.5: Meaning assembly of decomposed wh-in-situ type
3.2. Wh-type schema

substitutes the term variable of the scope marker hypothesis \((\lambda y . (\omega \lambda x. y))/u\). Finally, the result of these substitutions yields the following term. After \(\beta\)-reduction, we have computed a \(\lambda\)-term which is the same term as for ex-situ wh-questions. (In later derivations, we will use the term variable \(x\) for \(X\).

\[
(\lambda y . (\omega \lambda x. y) \text{ BODY}) \sim_{\beta} (\omega \lambda x. \text{ BODY})
\]

3.2.4.3 Displacement postulates for wh-in-situ

Wh-in-situ phrases occupy the same position as their associated gap hypothesis, so intuitively nothing moves. As the derivation in figure 3.2.4 illustrates, the scope marker hypothesis is inserted at the left edge of the question body. However, only when the scope marker hypothesis is adjacent to the gap hypothesis, the wh-phrase is merged and both hypotheses are replaced. Therefore, the scope marker hypothesis must be displaced in the direction of the gap hypothesis.

To move the scope marker hypothesis from its ‘scope’ position to the in-situ position, we use the exact set of displacement postulates that we have been using for extraction, but in the reverse direction, turning extraction and in-situ binding into structural converses.

\[
\begin{align*}
\Gamma[\Diamond \Delta_1 \circ (\Delta_2 \circ \Delta_3)] & \vdash C & \Gamma[\Diamond \Delta_1 \circ (\Delta_2 \circ \Delta_3)] & \vdash C \quad \text{[PI1]} \\
\Gamma[(\Diamond \Delta_1 \circ \Delta_2) \circ \Delta_3] & \vdash C & \Gamma[\Diamond \Delta_2 \circ (\Diamond \Delta_1 \circ \Delta_3)] & \vdash C \quad \text{[PI2]}
\end{align*}
\]

With this set of postulates we have filled in step 3 of the derivation in figure 3.4 where the scope marker hypothesis is moved from its outer scope position to a position adjacent to the gap hypothesis \(A\).

\[
\begin{align*}
\Diamond \Box (C/B) \circ \Delta[A] & \vdash C \\
\vdash \text{[Pl\*]} \\
\Delta \vdash \Diamond \Box (C/B) \circ A & \vdash C
\end{align*}
\]
3. A type schema for wh-question formation

3.2.5 Synthesized types and rules

We have explored different structural variants of the wh-type schema and have shown how these variants can be written out using the usual connectives of the base logic. The structural differences can be captured in the grammatical reasoning system with only the restricted set of displacement postulates as structural module, whereas the syntax-semantics interface through the Curry-Howard correspondence between syntactic and semantic types results in a uniform meaning assembly for wh-questions with either a wh-ex-situ or a wh-in-situ type.

In the rest of this thesis, we use the following three wh-type schemata to identify the lexical type assignment of a wh-phrase.

\[
\begin{align*}
\text{WH}_{l \text{ex}}(\Box A, B, C) & : \text{wh-ex-situ with a hypothesis on a left branch} \\
\text{WH}_{r \text{ex}}(\Box A, B, C) & : \text{wh-ex-situ with a hypothesis on a right branch} \\
\text{WH}_{in}(A, B, C) & : \text{wh-in-situ}
\end{align*}
\]

To merge each wh-phrase with a question body, we use the following inference rules which synthesize the derivation steps of the decomposed wh-type schemata. We present each inference rule in the abbreviated natural deduction style.

**Wh-ex-situ** The inference rules for wh-ex-situ type schemata apply a wh-phrase \( \Gamma \) of type \( \text{WH}_{l \text{ex}}(\Box A, B, C) \) (or \( \text{WH}_{r \text{ex}}(\Box A, B, C) \)) to a structure of type \( C \) which contains a hypothesized phrase of type \( A \) on a left (or right) branch in the structure. The wh-phrase replaces the hypothesized phrase and is inserted at the leftmost position in the structure.

\[
\begin{align*}
\Delta[\Box A \circ \Delta'] \vdash B & \quad \Delta'[\Box A] \vdash B \\
\frac{}{\Gamma \vdash \text{WH}_{l \text{ex}}(\Box A, B, C)} & \quad \frac{}{\Gamma \vdash \text{WH}_{r \text{ex}}(\Box A, B, C)}
\end{align*}
\]

**Wh-in-situ** The inference rule for a wh-in-situ type schema \( \text{WH}_{in}(A, B, C) \) merges a wh-phrase \( \Gamma \) with a structure of type \( C \) which has an gap hypothesis of type \( A \). The wh-phrase replaces the hypothesized phrase and is inserted at the same position in the structure as the hypothesized phrase.

\[
\begin{align*}
\Delta[A] \vdash B & \quad \Delta[A] \vdash B \\
\frac{}{\Gamma \vdash \text{WH}_{in}(A, B, C)} & \quad \frac{}{\Delta[\Gamma] \vdash C}
\end{align*}
\]

Additionally, we can recognize a null case ex-situ type schema which is derived from the left wh-ex-situ type schema.
3.2. Wh-type schema

**Null case** The null case for a wh-phrase is where the gap hypothesis of type \( A \) already occurs on the outermost left edge of the structure which forms the body of the question. The wh-phrase replaces the hypothesized phrase directly without having to apply any structural rules. The null variant, therefore, does not need additional ♦ ♣ features to the gap hypothesis. Furthermore, we distinguish this type from the ex-situ variant by adding a ♣ as subscript to the type schema, \( WH_{\Delta}^{\text{ex}}(A, B, C) \)

\[
\frac{\Gamma \vdash WH_{\Delta}^{\text{ex}}(A, B, C) \quad A \circ \Delta \vdash B}{\Gamma \circ \Delta \vdash C} \quad [WH_{\Delta}^{\text{ex}}]
\]

The null case is derived from the left wh-ex-situ type schema by means of the derivability relation between unary operators: ♦ ♣ \( A \vdash A \). With this derivability relation, we derive:

\[
WH_{\Delta,\Theta}^{\text{ex}}(\Diamond \Box A, B, C) \vdash WH_{\Delta}^{\text{ex}}(A, B, C)
\]

The benefit of the a null case variant of the wh-ex-situ phrase is a computational matter. Derivations with null case wh-type schemata become simpler because we do not need to apply any structural rules. Moreover, decomposing the null case wh-type schema into the usual type connectives shows that we can remove another syntactic step in the derivation.

Similar to wh-ex-situ left, the null variant of the wh-type schema can be written out as \( C / (A \backslash B) \). Different from the wh-ex-situ type, we can apply this type to a structure without assuming a gap hypothesis for the associated constituent phrase in the question body. The wh-phrase is directly inserted in the structure when the derivation forms a question body which is incomplete for the specified constituent: \( A \backslash B \). The wh-phrase merges with the structure via the usual elimination rule for /.

\[
\frac{\Gamma \vdash C / (A \backslash B) \quad \Delta \vdash A \backslash B}{\Gamma \circ \Delta \vdash C} \quad [/E]
\]

**Intermezzo**

So far, we have mainly concentrated on showing that the wh-type schema is a correct characterization of the syntactic and semantic properties of wh-phrases. The wh-type schema are mainly used as a simplification of the syntactic derivation of the different kinds of wh-questions. With the wh-type schema, the differences and the similarities in wh-question formation within a language and across languages are captured in a uniform way.

Additionally the wh-type schema may contribute to the computational efficiency of wh-question formation. The derivation of a wh-question using the synthesized rules takes fewer steps than a derivation using the usual connectives based on the decomposed wh-type schema. Speculating on the analogy
with human language processing, humans are capable of processing language in linear time. With the addition of the synthesized rules for wh-type schema to the grammatical reasoning system, we can account for a large reduction in processing time. Where the derivation would normally take numerous steps, with the wh-type schema rule the merging of a wh-phrase with a question body reduces to one step.

Thus, the synthesized wh-type schemata and rules for merging a wh-phrase with a question body provide an efficient and useful account of wh-question formation. In the next section, we will illustrate how the wh-type schema account for structural variation in wh-question formation while computing a uniform meaning assembly in the English language.

3.3 **Analysis of English wh-questions**

To illustrate the use of the wh-type schema, we analyze the properties of English wh-question formation that have been addressed in section 3.1. In English direct questions, a wh-phrase occurs standardly at the beginning of the sentence. We have listed most properties and characteristics that have been noted for wh-question formation in English. In this section, we show that we can account for these properties in the grammatical reasoning system of type-logical grammar by using the wh-type schema to identify wh-phrases.

We start in section 3.3.1 by presenting the lexical type-assignment of verbs, nouns and other constituents which we will use to derive the basic word order of sentences. In section 3.3.2, we focus on local wh-questions and identify the basic subtypes of the wh-type schema to analyze local wh-questions. The subtypes of the wh-type schema prescribe in what structural context a subject or non-subject wh-phrase can merge. We illustrate local wh-questions by analyzing single constituent direct and indirect questions. In section 3.3.3, we continue with the analysis of non-local wh-questions. In section 3.3.4, we show how island constraints that block long-distance displacement in non-local questions are accounted for. In section 3.3.5, the properties of multiple wh-questions are captured.

3.3.1 **Basic types and word order**

Before we start analyzing the different syntactic properties, we show the basic types of different sentential clauses. These types indicate the different kind of clauses in English and will be used to instantiate the question body in which a wh-phrase can occur. Additionally, we present a lexicon of English.

We assume basic type $s$ for declarative clauses, $q$ for yes-no questions and $wh$ for wh-questions. The type we will use for yes-no questions ($q$) will be

---

8In chapter 5, we show that the $wh$ type can be further unfolded into a type which reflects the possible answer type that the wh-question requires. For the moment, to concentrate on the syntactic analysis of wh-questions, we use a macro type $wh$ for wh-questions.
3.3. Analysis of English wh-questions

the same type assigned to sentential clauses which are formed by do-support. We are aware that yes-no-questions differ from such sentential clauses. We abstract away from the grammatical properties of inflection in English and use \( q \) and \( s \) to distinguish between sentential clauses with and without do-support. Additionally, we make a distinction for subordinate clauses where we take \( s' \) for embedded declarative clauses, \( q' \) for if/whether clauses and \( wh' \) as type for embedded wh-questions.

<table>
<thead>
<tr>
<th>Type of Clause</th>
<th>Main Type</th>
<th>Subordinate Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Declarative clauses</td>
<td>( s )</td>
<td>( s' )</td>
</tr>
<tr>
<td>Yes-no questions</td>
<td>( q )</td>
<td>( q' )</td>
</tr>
<tr>
<td>Wh-questions</td>
<td>( wh )</td>
<td>( wh' )</td>
</tr>
</tbody>
</table>

Basic lexical entries  The canonical word order of English is SVO. On the basis of this word order we specify the lexical type-assignment of verb phrases and auxiliaries by assigning the relative structural position of arguments to the heads. Some of the verbs and noun phrases that we will use in our example sentences are listed in the lexicon. Lexical entries of words that are involved in specific phenomena of wh-question formation are introduced in the section where the property is being discussed.

<table>
<thead>
<tr>
<th>Lexicon Entry</th>
<th>Type</th>
</tr>
</thead>
<tbody>
<tr>
<td>Mary, John, Bill</td>
<td>( np )</td>
</tr>
<tr>
<td>a</td>
<td>( np/n )</td>
</tr>
<tr>
<td>bird</td>
<td>( n )</td>
</tr>
<tr>
<td>slept</td>
<td>( np\backslash s )</td>
</tr>
<tr>
<td>saw, read</td>
<td>( (np\backslash s)/np )</td>
</tr>
<tr>
<td>whether, if</td>
<td>( q'/s )</td>
</tr>
</tbody>
</table>

With this lexicon we can analyze the basic word order of declarative clauses in English (see appendix B.1.1 for on-line derivations).

3.3.2 Local wh-questions

Let us now look in detail how single constituent wh-questions in English are derived. In section 3.3.2.1, we will illustrate the analyses of direct questions and in section 3.3.2.2 we analyze indirect questions. The basic construction of direct and indirect questions is the same. In both direct questions and indirect questions, the wh-phrase occurs at clause-initial position. However, depending on the type of wh-phrase, the associated constituent may occur on a left or on a right branch.

In section 3.2, we have presented two wh-type schemata to lexically identify wh-phrases which appear in clause-initial position: wh-ex-situ left (\( WH_{ex}^l \)), which associates with a gap hypothesis on a left branch of the structure, and wh-ex-situ right (\( WH_{ex}^r \)), which associates with a gap hypothesis on a right branch. For each wh-phrase we need to determine the wh-type schema that
A type schema for wh-question formation captures the syntactic use of the wh-phrase and leads to a correct analysis of a wh-question. We not only need to specify the left- or right directedness of the associated gap hypothesis, but we also need to specify the subtypes where the wh-type schema ranges over: \( WH(A, B, C) \). Recall that \( B \) is the type of the question body, \( A \) is the type of the associated gap constituent, and \( C \) is the type of the clause that is yielded after the wh-phrases merges with the question body.

With these ‘parameters’, we can now lexically identify wh-phrases. The wh-type schema determines at which point the wh-phrase is merged with a question body. The type of the question body has to correspond to the type encoded in the type schema, and the position and the type of the gap hypothesis in the question body has to correspond to the structural configuration of the wh-phrase. The wh-type schema may vary for wh-phrases. For instance, in English the canonical argument position of a subject argument is on a left branch while non-subject argument phrases or adjunct phrases reside on right branches. Subject wh-phrases are assigned a wh-type schema left and non-subject wh-phrases are assigned a wh-type schema right. Another distinction is encoded in the occurrence of a wh-phrase in a main or in an embedded clause. Each wh-phrase has a wh-type schema that encodes its occurrence. We will present an analysis for the different occurrences of wh-phrases.

For English wh-phrases, we recognize the following instantiations of the wh-type schema to analyze direct and indirect questions:

<table>
<thead>
<tr>
<th></th>
<th>Direct questions</th>
<th>Indirect questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>subjects</td>
<td>( WH'_{ex}(np, s, wh) )</td>
<td>( WH'_{ex}(np, s', wh') )</td>
</tr>
<tr>
<td>non-subjects</td>
<td>( WH'_{ex}(\otimes np, q, wh) )</td>
<td>( WH'_{ex}(\otimes np, s', wh') )</td>
</tr>
<tr>
<td>adjuncts</td>
<td>( WH'_{ex}(\otimes(A \setminus A), q, wh) )</td>
<td>( WH'_{ex}(\otimes(A \setminus A), s', wh') )</td>
</tr>
</tbody>
</table>

Table 3.1: Wh-type schema for English wh-phrases, where \( A \in \{np \setminus s, \text{inf}, s\} \)

### 3.3.2.1 Direct questions

In single constituent direct questions, a wh-phrase occurs in clause-initial position, as is illustrated in the following examples.

\[(3.20)\]

a. Who saw Bill?

b. Who(m) did John see?

c. Where did John see Bill?

On the basis of the wh-type schema in table 3.1, we show how the different wh-questions are derived.
3.3. Analysis of English wh-questions

Subject wh-phrase  As given in table 3.1, a subject wh-phrase gets assigned wh-type schema: \( \text{WH}^\ell_{ch}(np, s, wh) \). This type can be paraphrased in prose as an element that merges with a sentential clause of type \( s \) which contains a gap hypothesis of type \( np \) on a left branch. When the wh-phrase merges with the question body, it replaces the gap hypothesis and is inserted at the front of the structure. The structure that is derived is of type \( wh \). In a natural deduction format, the derivation of the question ‘Who saw Bill?’ can be presented as follows:

\[
\begin{align*}
\text{who} & \quad \text{saw} \quad \text{bill} \\
\text{WH}^\ell_{ch}(np, s, wh) & \quad [np \vdash np] \\
\text{who} \circ (\text{saw} \circ \text{bill}) & \vdash wh \\
(\omega \lambda y.((\text{see bill}) \, y)) &
\end{align*}
\]

In step 1, the basis of the sentence structure is built by merging the noun phrase ‘Bill’ as an object phrase to the right of the verb phrase ‘saw’. The verb phrase requires another noun phrase on the left. The gap hypothesis \( \diamondsuit np \), that can be reduced to type \( np \), merges with the structure in step 2. Because the verb phrase first selects an object phrase to the right and then a subject phrase to the left, the structure is now exactly in the order where the hypothesis is in a left branch position. In step 3 the elimination rule for deconstructing the wh-type schema is applied. The hypothesis on the left branch is replaced by the wh-phrase and the whole clause forms a wh-question of type \( wh \). The meaning assembly of the wh-question shows the abstraction.

In section 3.2.5, we have discussed the option for a null case variant of the wh-type schema. This null case variant is a derived wh-ex-situ left type schema. As the gap hypothesis already occurs on the left edge of a structure, it does not need the \( \diamondsuit \square \) features. The subject wh-phrase merges at the same position in the structure as the gap hypothesis. Therefore, the subject wh-phrase in a local wh-question can be assigned the null case wh-type schema: \( \text{WH}^\ell_{ch}(np, s, wh) \). The analysis of a wh-question with a subject wh-phrase becomes much simpler. The above derivation reduces to the following:

\[
\begin{align*}
\text{who} & \quad \text{saw} \quad \text{bill} \\
\text{WH}^\ell_{ch}(np, s, wh) & \quad [np \vdash np] \\
\text{who} \circ (\text{saw} \circ \text{bill}) & \vdash wh \\
(\omega \lambda y.((\text{see bill}) \, y)) &
\end{align*}
\]

Object wh-phrase  A direct question with an object wh-phrase differs from direct questions with a subject wh-phrase. The sentence is formed by do-support and is typed as \( q \) and the associated object noun phrase resides on
a right branch of a structure. To derive a single constituent wh-question, the object wh-phrase must be identified as \( WH_{\text{ex}}(\Diamond np, q, wh) \) as presented in table 3.1. Figure 3.6 gives the natural deduction derivation of a main clause wh-question ‘Whom did Bill see?’.

\[
\begin{align*}
\text{did} & \quad \text{John} \\
\frac{q/(np \cdot \text{inf})}{\text{see} \circ (\Diamond np)} & \quad \frac{\Diamond np}{\text{inf} / np \quad np} & \quad [\text{E}] \\
\frac{(\text{did} \circ (\text{John} \circ (\text{see} \circ (\Diamond np))))}{q} & \quad [WH_{\text{ex}}] \\
\frac{\text{whom} \circ (\text{did} \circ (\text{John} \circ \text{see}))}{wh}
\end{align*}
\]

\((\omega \lambda x((\text{see} \ x) \ \text{John}))\)

For a clearer view on the structural position of the gap hypothesis, we present the last two steps by displaying the structural part of the derivation in a tree style presentation. In the penultimate step of the derivation, the hypothesis originates in the object position of the verb. The wh-ex-situ wh-phrase merges with the structure on the left side replacing the gap hypothesis.

\[
\begin{align*}
\text{did} & \quad \text{John} \\
\frac{\text{see} \circ (\Diamond np)}{\text{whom} \circ (\text{did} \circ (\text{John} \circ \text{see})) \Rightarrow wh}
\end{align*}
\]

\(\text{whom} \circ (\text{did} \circ (\text{John} \circ \text{see})) \Rightarrow wh\)

Figure 3.6: Natural deduction style derivation of ‘Whom did John see?’

For a clearer view on the structural position of the gap hypothesis, we present the last two steps by displaying the structural part of the derivation in a tree style presentation. In the penultimate step of the derivation, the hypothesis originates in the object position of the verb. The wh-ex-situ wh-phrase merges with the structure on the left side replacing the gap hypothesis.

Figure 3.7: Tree style presentation of ‘Whom did John see?’

In later derivations, we will omit the steps that present the application of the structural rules. We will apply the inference rule for merging the wh-phrase to the body of the questions when the argument appears on the right side of a tree structure. In the above example, in the penultimate step the object argument hypothesis appears on the right branch of a structure. From this configuration, we assume that the hypothesis may be displaced to the right edge.
Direct questions with indirect objects such as ‘Whom did Bill talk to?’ are analyzed similarly to questions with direct objects. The body of the question is a clause that uses do-support and the associated gap hypothesis is the indirect object argument which occurs on a right branch.

**Adjunct wh-phrases** In this thesis, we will mainly focus on argument wh-phrases. However, we want to show that adjunct wh-phrases fit the wh-type schema equally. An adjunct wh-phrase associates with an adjunct gap hypothesis instead of to a noun phrase gap in a question body with do-support. The wh-type type for wh-adjuncts encodes that the wh-phrase ranges over adjunct positions, for instance, q\s\s\s (np\s) or inf\inf. Depending on the kind of adjunct gap, the adjunct wh-phrase associates with a gap hypothesis that modifies either a verb or a clause on the right edge of a structure. Adjunct wh-phrases are typed as: $WH_{ex}(X,X,q,wh)$ where $X$ is the type of constituent that the adjunct wh-phrase modifies.

The natural deduction derivation in figure 3.8 and the tree style derivation in figure 3.9 illustrate the last steps of the natural deduction analysis of ‘Where did John see Bill?’. We present the last steps of the natural deduction derivation. The tree style presentation clearly shows that the adjunct gap hypothesis originates on a right branch in a $q$-typed clause when the wh-phrase merges with the structure.

\[
\text{Figure 3.8: Natural deduction style derivation of ‘Where did John see Bill?’}
\]

\[
\text{Figure 3.9: Tree style presentation of ‘Where did John see Bill?’}
\]
3.3.2.2 Indirect questions

Embedded interrogatives are selected by verbs such as ask, wonder and know. These verbs either select a clause headed by ‘if’ or ‘whether’ or a wh-question while the main clause is not specified as wh-interrogative. The verb phrase requires the embedded clause to be of type $q'$ or $wh'$. The following examples illustrate some of the embedded interrogatives that can be selected by a verb such as ‘wonder’.

\[(3.21)\]
\[\begin{align*}
  & a. \text{John wondered whether Mary saw a bird.} \\
  & b. \text{John wondered what Mary saw.} \\
  & c. \text{John wondered who saw a bird.}
\end{align*}\]

The selectional requirement of the verb phrases are encoded in the lexicon. Because we distinguish between if/whether clauses ($q'$) and embedded wh-questions ($wh'$), we identify two entries for interrogative selecting verbs.

\[
\text{ask} :: \lfloor np/s \rfloor / q' \\
\text{ask} :: \lfloor np/s \rfloor / wh'
\]

As with direct questions, the wh-phrase in indirect questions occurs clause-initially. In table 3.1 on page 62, we have listed the types for wh-phrases. The types for subject and non-subject argument wh-phrases are repeated below.

\[
\begin{align*}
  \text{embedded subject wh-phrases} & :: WH^l_{np} (\Diamond np, s', wh') \\
  \text{embedded non-subject wh-phrases} & :: WH^r_{np} (\Diamond np, s', wh')
\end{align*}
\]

Embedded interrogatives are derived with similar types for wh-phrases as for main clause wh-questions. The only difference is that the question body is an embedded declarative clause type $s'$ and the goal type is an embedded wh-question of type $wh'$. Embedded interrogatives do not need do-support or subject-auxiliary inversion to derive a grammatical ordering. The addition of wh-type schema for wh-phrases to occur in embedded wh-questions is just a variant of wh-type schema that distinguishes between main clause and embedded interrogative clauses. This addition in the lexicon does not affect the derivation of wh-questions in the main clause. The alternatives can co-exist without deriving ungrammatical orderings.

With the given type assignments, we can analyze indirect questions such as the sentences in example 3.3.2.2 (see appendix B.1.3 for on-line derivations).

3.3.3 Non-local wh-questions

There seems to be no limit in how far apart the wh-phrase may appear from their associated argument position. The associated argument is embedded in a complement clause that is selected by a verb, which belongs to the group of bridge verbs. Bridge verbs form no blocking mechanism for associating a wh-phrase in clause-initial position to a hypothesis.
The sentences in example 3.22 show that non-local wh-questions differ from local wh-questions with subject wh-phrases. In local wh-questions the subject wh-phrase merges with a clause of type $s$, while in non-local wh-questions the question body is a verb clause with do-support which are clauses of type $q$. In section 3.3.2, we have shown that subject wh-phrases can be typed with a null variant of the wh-type schema, because the wh-phrase can be merged without hypothesizing an associated gap hypothesis. However, to question a subject argument over a long-distance the question body must be a $q$-typed clause. Furthermore, the gap hypothesis must be decorated with unary features in order to reach the gap hypothesis via the recursive application of the displacement postulates.

We show that for non-local wh-questions, subject and non-subject argument wh-phrases get the same wh-type schema instantiation. The question body is a $q$-typed clause in which the associated gap hypothesis occurs on a right branch. After merging the wh-phrase, the clause becomes of type $wh$.

\[
WH'_{ex}(\Diamond \Box np, q, wh)
\]

Before presenting an analysis of long-distance displacement, we first discuss the matter of 'that'-omission which leads us to the lexical type-assignment of bridge verbs.

**That-omission** The derivation of long-distance displacement is dependent on the occurrence of bridge verbs. As we explained in 3.1.1, bridge verbs function as a bridge between the wh-phrase and the associated argument position. The syntactic behavior of the verb is encoded in its lexical type assignment. The following examples illustrate that the complementizer can sometimes be omitted.

\[(3.23)\]

\begin{enumerate}
\item Who did Sue believe (*that) saw a bird?
\item Who does John say (that) Sue believes (*that) saw a bird?
\end{enumerate}

\[(3.24)\] What did Sue believe (that) Mary saw?

The observation is that both subject and non-subject argument gap hypotheses can be associated with a clause-initial wh-phrase when the that-complementizer is omitted. However, when the subordinate clause contains an overt complementizer 'that', only a non-subject argument gap hypothesis
can be associated with a clause-initial wh-phrase, as illustrated by the ungrammaticality of the that-occurrence in example 3.23a and b) and the optionality of the ‘that’-occurrence in example 3.24. The complementizer is obligatorily omitted when the embedded clause bears the subject gap hypothesis while it is optionally omitted in case of non-subject wh-phrases.

We follow Morrill (1994) who spells out the lexical ambiguity of bridge verbs in the type-assignment.\(^9\)

\[
\begin{align*}
\text{believe}_1 & : (np \backslash s)/s' \text{ (with overt ‘that’)} \\
\text{believe}_2 & : ((np \backslash s)/(np \backslash s))/np \text{ (‘that’ omission)}
\end{align*}
\]

The distinction between the two types is the following. The first type-assignment for bridge verbs is used to derive embedded clauses with an overt complementizer ‘Sue believes that John saw a bird’. It selects an embedded clause of type s’ and, thus, needs an overt complementizer ‘that’ (s’/s). The second syntactic type-assignment is used for embedded clauses where the complementizer is omitted, as for instance the sentence: ‘Sue believes John saw a bird’. In these sentences, the bridge verb ‘believes’ first selects a noun that serves as the subject argument of the embedded clause (john). Subsequently, the bridge verb selects an embedded verb clause (‘saw a bird’). Due to these type assignments, the subject argument phrase occurs on a left or a right branch when the embedded clause is selected either by ‘believe\(_1\)’ or by ‘believe\(_2\)’, respectively.

The grammaticality and ungrammaticality of non-local wh-questions (see appendix B.1.2) can be analyzed on the basis of these type-assignments. Figure 3.10 illustrates that subject wh-phrase cannot associate with a gap hypothesis with an overt complementizer to form the ungrammatical sentence ‘Who did Sue believe that saw a bird?’. The sentence is underviable because the hypothesized noun phrase appears on a left branch and ‘who’ (WH\(_c\)(\(\bigcap\)np, q, wh)) merges with a q-typed clause with the subject hypothesis on a right branch. Figure 3.11 illustrates a structure where the subject wh-phrase can associate with the gap hypothesis to form the grammatical wh-question ‘Who did Sue believe saw a bird?’. This sentence, with the that-complementizer omitted, is derivable because the associated subject argument hypothesis occurs on a right branch.

3.3.4 Island constraints

In section 3.1.3, we have discussed the following different types of islands: wh-clauses, adjunct islands, complex noun phrases and coordinate structures. Left branch constructions will be discussed separately. The common characteristic of islands is that if a gap hypothesis is embedded in an island construction, a fronted wh-phrase cannot associate with it. Following earlier proposals\(^9\)

\[^9\]The analysis is based on earlier research by Gazdar (1981) who has proposed a metarule for the GPSG framework that subsumes all occurrences of the left-branch condition. Morrill (1994) translated this metarule into lexical type-assignments for bridge verbs.
3.3. Analysis of English wh-questions

Figure 3.10: who as $WH_{er}(\diamond \Box np, q, wh)$ cannot associate with its gap hypothesis

Figure 3.11: who as $WH_{ex}(\diamond \Box np, q, wh)$ can associate with its gap hypothesis
by Morrill (1994), we show that the island constraints are the result of feature requirements on the island forming constituent.

In chapter 2, section 2.5.3, we have explored the different uses of unary features. Versmissen (1996); Kurtonina and Moortgat (1997) have shown that the unary features can be used to project a structural domain. The projection of features into a structural domain are the result structures of inference rules of $\Box E$ and $\Diamond I$ as presented in chapter 2 (page 14) and reprinted here:

$$
\frac{\Gamma \vdash \Box A}{\Diamond (I) \vdash A} \text{[} \Box E \text{]} \\
\frac{\Gamma \vdash A}{\Diamond (I) \vdash \Diamond A} \text{[} \Diamond I \text{]}
$$

The island domain is represented in the tree style presentation as a unary branch with a $\Diamond$ feature node. As we have explored in section 3.2 the gap hypothesis needs to be reached in order to merge a wh-phrase to a question body. A deeply embedded gap hypothesis can only be reached via a pure binary branching path. If there is a unary branch, then the $\Diamond$ node blocks the further application of the structural rules. The island forming constituents are elements that project a unary branching $\Diamond$ node. With a unary branch blocking the binary branching path, the embedded gap hypothesis cannot be reached.

We refer to the tree structure that is the daughter of a unary branch as a structural domain. A structural domain is either projected from a $\Diamond$ feature restriction on the selecting subtype requirement of an island forming constituent or it is projected from a $\Box$ feature restriction on the main type. We recognize that wh-clauses and adjunct clauses follow the first strategy, while complex noun phrases and coordinate structures follow the latter strategy. We will present examples for the different types of islands and present an analysis illustrating the blocking of a structural domain.

### 3.3.4.1 Wh-clauses

As described in section 3.1.1.2, embedded interrogatives are recognized as islands. Sentences in which a wh-phrase occurs clause-initially while the associated argument position is embedded in an interrogative clause are marked as ungrammatical. We repeat the ungrammatical sentences as presented in section 3.1.3.

(3.25)  

a. *Which bird, did John wonder [whether Mary wanted to see t_i]?

b. *Which bird, did John wonder [who wanted to see t_i]?

In our analysis, the sentences are ungrammatical because the fronted wh-phrase cannot associate with the gap hypothesis in the embedded interrogative clause. The interrogative verb carries an additional feature which causes the embedded interrogative to be encapsulated in a structural domain which prevents further application of structural rules. An interrogative verb such as 'wonder' has the following type-assignments in the lexicon:
3.3. Analysis of English wh-questions

wonder :: VP/◊q′
wonder :: VP/◊wh′
where VP ∈ \{np\, \mbox{s, inf}\}

The verb selects an embedded interrogative which has feature requirement ◊. The ◊ feature requirement is added to the structure of the embedded interrogative clause with the inference rule, ◊I. The inference rule introduces a ◊ on the syntactic type and imposes an ♦ on the structure, as illustrated by the following derivation of the sentence ‘did John wonder whether Mary wanted to see’.

\[
\begin{align*}
\text{wonder} & \rightarrow \text{inf}/◊q' \\
\text{whether} \circ (\ldots \circ ◊np) & \vdash q' \\
[◊I] & \vdash [E] \\
\text{wonder} \circ 3(\text{whether} \circ (\ldots \circ ◊np))) & \vdash \text{inf}
\end{align*}
\]

The tree style presentation in figure 3.12 illustrates how the unary branch projects over the embedded interrogative clause. The ♦ feature prevents the gap hypothesis to be moved out. Therefore, a wh-phrase such as ‘which bird’ cannot associate with the gap hypothesis and, thus cannot merge with the question body.

![Tree Diagram](image)

Figure 3.12: which bird :: WH\text{\textsubscript{1,2}} (◊dnp, q, wh) cannot associate with its gap hypothesis

A similar analysis applies to the other wh-island construction, ‘*Which bird did John wonder who wanted to see’.
3.3.4.2 Adjunct clauses

Adjunct phrases are verbal or sentential modifiers. Clauses that are headed by function words such as if, although, because form an island construction. We repeat the example presented in section 3.1.3.

(3.26) *Which topic did John leave [because Mary talked about]

The head of an adjunct phrase selects a declarative clause and forms a phrase which acts as a modifier to different types of verbs. In a similar way to the wh-island, the adjunct imposes an additional ♦ feature requirement on the clause type it selects. The creation of an adjunct island is projected from the following type assigned to adjuncts:

although, because, if :: (A \ A) / ♦s where A ∈ {np \ s, inf}

The tree style presentation in figure 3.13 illustrates the blocking of the derivation of the ungrammatical sentence 'Which topic did John leave because Mary addressed?'. The derivation fails at the point where the wh-phrase is inserted. The gap hypothesis in the structural domain is blocked for further restructuring under unary feature decoration (♦Δ). In other words, the wh-phrase cannot merge with the structure because the gap hypothesis is unreachable.

![Figure 3.13: which topic :: WHr(♦ □ np, q, wh) cannot associate with its gap hypothesis](image)

---

10 Notice that 'if' heading a adjunct clause differs from 'if' used in an embedded interrogative: cf. 'I will kick you if you hit my dog.' versus 'I wonder if...'.

3.3. Analysis of English wh-questions

Echo-questions

Echo-questions can be used to question an argument position within an adjunct phrase. In such constructions the wh-phrase is left in-situ. To allow for the derivation of echo-questions, we must assume lexical entries for wh-phrases of type $WH_{in}(np, s, wh)$. Additionally, we remark that in the cases of echo-questions, wh-in-situ phrases are insensitive to the $\bullet$ projection. Whether this would be the case in general is a matter for further research.

The analysis of an echo-question is illustrated by the last step in the derivation where a wh-in-situ phrase ‘which topic’ merges with the question body.

\[
\begin{align*}
\text{John} \circ (\text{left} \circ (\text{because} \circ \bullet (\text{Mary} \circ (\text{addressed} \circ \langle np \rangle)))) & \vdash s \\
\vdash \text{which topic} \vdash WH_{in}((\langle np \rangle, s, wh)) \\
\text{John} \circ (\text{left} \circ (\text{because} \circ \bullet (\text{Mary} \circ (\text{addressed} \circ \langle \text{which topic} \rangle)))) & \vdash \text{wh}
\end{align*}
\]

3.3.4.3 Other island constructions

Other island constructions are formed by complex noun phrases and coordinate constructions. We have presented examples of each construction in section 3.1.3. We repeat two examples: one example of a complex noun phrase (example 3.27a) and one example of a coordinate structure (example 3.27b).

(3.27)  
\begin{align*}
\text{a. } & \ast \text{Which kid, did you see [the teacher who punished t]?} \\
\text{b. } & \ast \text{Which man, did you invite [Mary and t]?}
\end{align*}

The occurrence of a wh-phrase in clause-initial position is blocked due the formation of a structural domain around the clause which embeds the hypothesized argument. Again, the function word ‘who’ and ‘and’ carry additional feature information that causes the clause to be encapsulated in a structural domain. The difference with the constituents that form a wh-clause or an adjunct clause, is that the $\bullet$ feature information is projected from the head category. A $\Box$ feature on the main type projects under the $\Box E$ inference rule to a structure headed by $\bullet$. The function words are encoded in the lexicon with the following type assignments. (Note that the coordinator ‘and’ can range over any arbitrary type $A$.)

\[
\begin{align*}
\text{who} & \:: (n \setminus \Box n)/(np \setminus s) \\
\text{and, or} & \:: (A \setminus \Box A)/A
\end{align*}
\]

With these type-assignments, we can analyze complex noun phrases and coordinate structures (see appendix B.1.4). We illustrate how the derivation of a wh-question with a complex noun phrase is blocked. First, the $\Box$ feature of the function word projects a $\bullet$ unary branch over the tree structure of a complex NP such as ‘teacher who punished $\Box np$’. 
A type schema for wh-question formation

Second, the structural domain blocks the merging of a wh-phrase with a structure whose associated gap hypothesis is embedded in a complex NP or coordinate structure. This blocking is illustrated in the tree structure in figure 3.14. The ♦ feature prevents the wh-phrase from being associated with the gap hypothesis inside the coordinate construction. Therefore, the wh-question ‘Who did you invite Mary and ?’ cannot be derived.

Figure 3.14: who :: WH_{\text{np}}(\Box \square \text{np}, q, wh) cannot associate with its gap hypothesis

Across-the-board An exception to the coordinate structure constraint is the ‘across-the-board’ (ATB) construction where a wh-phrase is associated with a constituent gap hypotheses in both conjuncts of the coordinate structure. In section 3.1.3, we presented the following example of an ATB-construction.

(3.28) Which man, did you invite [[a friend of t] and [a brother of t]]?

We apply the same analysis as Morrill (1994) proposes for ATB-construction in relative clauses. We will provide only a sketch of the proof and refer to Morrill for the complete analysis. The basic idea is that the coordinator can coordinate phrases of any arbitrary type A: (A \setminus \Box A)/A. Therefore, we may coordinate the two structures ‘a friend of’ and ‘a brother of’ which are both np-typed phrases that are incomplete for a noun phrase argument: np/\Box \square np.
3.3. Analysis of English wh-questions

After we have coordinated the two structures, the gap hypothesis for the incomplete noun phrase argument is assumed. The derivation continues with building up the rest of the question body until we have derived a sentence of type \( q \). Nothing prevents the wh-phrase ‘which man’ from merging with the question body and replacing the gap hypothesis. The gap hypothesis occurs free on a right branch in the question body. We illustrate the last step of the analysis of the ATB-construction in a tree style derivation in figure 3.15.

Figure 3.15: which man :: WHr(\( \square np, q, wh \)) can be merged and associate with its gap hypothesis

3.3.4.4 Left branch constructions

A special instance of an island construction is the construction where the gap hypothesis originates on a left branch of a noun phrase. An example of such a noun phrase construction is a possessor phrase featuring a particle ‘s’ connecting two noun phrases such as ‘John’s picture’ and ‘every man’s mother’. As shown in example 3.29, a wh-phrase cannot occur clause-initially and associate with a left branch gap hypothesis in the noun phrase construction.

(3.29)  a. * Which did you see picture?
       b. * Which man did you see ’s picture?
       c. * Which man’s did you see t picture?

The wh-questions in these examples are ruled out, just like the name ‘left branch construction’ says, because the gap hypotheses occur on a left branch. Except for subject argument wh-phrases, all wh-phrases in local and non-local questions have been typed as a wh-ex-situ right type schema \( WH_{ex}^r(A, q, wh) \). Therefore, these wh-phrases cannot associate with a gap hypothesis on a left
branch. We illustrate this by analyzing example 3.29b where ‘which man’ occurs separated from the possessor phrase.

The wh-phrase ‘which man’ is of type $WH_{st}^{r}(\Box \Box np, q, wh)$. The derivation of wh-determiner phrases will be discussed in more detail in section 3.3.5. To analyze the possessor phrase, we assign the particle ‘s’ a category which reflects its function of combining a noun phrase (‘John’) and a noun (‘picture’) into a possessor phrase (‘John’s picture’).

’s :: np\(\backslash(np/n)\)

Figure 3.16 illustrates the tree style presentation of example 3.29a to and shows that ‘which man’ cannot merge with the question body because the gap hypothesis resides in a possessor phrase.

![Figure 3.16: Which man :: WH_{st}^{r}(\Box \Box np, q, wh) cannot be merged and associate with a np gap hypothesis on a left branch](image)

The only way to question the noun phrase constituent in a left branch construction is to pied-pipe the other material along with the wh-phrase.

**Pied-piping** Pied-piping refers to the strategy in which a fronted wh-phrase drags along additional material. The analysis that we present here, follows largely the analysis proposed by Morrill (1994) for pied-piping in relative constructions. With this analysis, we can accounts for at least two instances: The wh-phrase that occurs in a prepositional phrase and the wh-phrase that is part of a noun phrase construction such as those we saw in left-branch constructions.

(3.30)  

a. [To whom], did John give a present $t$?

b. [Which man’s picture], did you see $t$?

Before we present our analysis, we want to point out two properties that will be directly encoded in our analysis. The first property reflects the basic
idea behind pied-piping, namely that the whole phrase that embeds the whphrase functions as a wh-phrase. The second property is that the wh-phrase itself stays in-situ in the phrase in which it is embedded in. Linked to these two properties we present the proposal in two steps. First, we show that by treating the whole construction as a wh-type schema we can derive a wh-question where the construction occurs fronted. Second, we present the lexical type-assignment of the embedded wh-phrase which has the type of a wh-in-situ phrase that yields a wh-type schema after merging with its question body.

To front a noun phrase construction such as ‘Which man’s picture’ or a prepositional phrase such as ‘to whom’, the constructions must be instances of the wh-ex-situ type schema:

\[
\text{pp wh-phrase} :: \text{WH}_{ex}(\Box pp, q, wh) \\
\text{np wh-phrase} :: \text{WH}_{ex}(\Box np, q, wh)
\]

With these types we derive wh-questions where the whole construction acts as a wh-phrase. As an illustration, we present the question body of the wh-question ‘To whom did John give a book’ in figure 3.17.

![Figure 3.17](image)

To derive the whole noun phrase or prepositional phrase as a wh-type schema, the wh-phrases embedded in the constructions must be typed as wh-in-situ phrases. The question body of the embedded wh-phrases can be a noun phrase (which man’s picture) or a prepositional phrase (to whom). Instead of associating to a gap hypothesis in the main clause, the wh-phrase associates with a gap position within the specified body.

\[
pied-piped \text{ wh-phrase} :: \text{WH}_{ex}(np, A, \text{WH}_{ex}(\Box A, q, wh)) \\
\text{where } A \in \{np, pp\}
\]

With this type we derive constructions featuring the whole construction as a wh-phrase. We illustrate the construction with two analyses. We present
an analysis of the noun phrase construction ‘which man’s picture’ and of the prepositional phrase ‘to whom’. Recall that the possessor and the preposition are typed as follows:

\[
\text{to} \;::\; pp/np \\
\text{’s} \;::\; np \backslash (np/n)
\]

The wh-phrase merges with the question body and replaces the gap hypothesis in the construction. The result after merging the wh-phrase is a structure which itself is a wh-type schema. The whole construction merges with a question body and associates with a gap position in the main clause as presented in figure 3.17.

**Preposition stranding** As an alternative to pied-piping, we can also leave the preposition stranded.

(3.31) Whom did you give a book to?

To leave the preposition stranded, we just need to merge a wh-phrase to the question body as usual. The tree style derivation illustrating the merging of a wh-phrase where the preposition is left stranded is illustrated in figure 3.18.

**Meaning assembly** In anticipation of the explanation on the syntax-semantics interface of wh-questions in chapter 5, we show that the meaning assembly of the pied-piping construction is the same as the construction where the preposition is left stranded. For the preposition stranding the wh-phrase merges with the question body as before. The meaning assembly for a wh-question such as ‘Whom did John give a book to?’, as presented in figure 3.18, becomes:

\[
\begin{align*}
\text{did} \circ (\text{john} \circ ((\text{give} \circ (a \circ \text{book}) \circ \text{to}))) \circ \Diamond np \vdash q \\
((\text{give} (a \circ \text{book}) \circ (\text{to} x))) j) \\
\end{align*}
\]

\[
\begin{align*}
\text{whom} \circ \Diamond (\text{give} \circ (a \circ \text{book}) \circ \text{to})) \vdash \text{wh} \\
(\omega \lambda x.(((\text{give} (a \circ \text{book}) \circ (\text{to} x)) j))
\end{align*}
\]
To compute the meaning assembly of the pied-piping analysis, we have to work out the lexical semantics of the wh-phrase 'whom'. Although the type-assignment of the wh-phrase is a wh-in-situ type that yields a wh-ex-situ type after merging with a prepositional phrase, the meaning assembly of the wh-phrase must only project a single ω-operator. Furthermore, the lexical semantics must match the semantic type that maps to the syntactic type of the wh-phrase. The lexical semantics for the pied-piping instance for prepositional phrases must therefore be the following:

\[ \lambda Q (\overline{np} \rightarrow \overline{pp}). \lambda P (\overline{pp} \rightarrow q). (\omega \lambda x^{\overline{pp}}. (P (Q x))) : WH_{\omega} (np, pp, WH_{\omega} (\otimes \square pp, q, wh)) \]

Analyzing the pied-piping construction using this lexical semantics for the wh-type schema for 'whom', we compute the same meaning assembly for 'To whom did John give a book?' as for 'Whom did John give a book to?'. (See appendix B.1.5 for an on-line derivation.) To illustrate how the meaning assembly for the whole construction is computed, we present the derivation in two parts. First, we compute the lambda term for the wh-ex-situ wh-phrase 'to whom'\textsuperscript{11}.

\textsuperscript{11}The meaning assembly, we present here, is based on the synthesized types and rules of the wh-type schema. The actual underlying lexical semantics of the pied-piping type-assignment is determined using the decomposed types of the wh-ex-situ and wh-in-situ type. The synthesized type unfolds into the usual connectives of the base logic as follows (see section 3.2.3.2 for the syntactic decomposition of the ex-situ type and section 3.2.4.1 for the syntactic decomposition of the in-situ type).

whom = WH_{\omega} (np, pp, WH_{\omega} (\otimes \square pp, q, wh))

= \otimes \square ((wh/(q/\otimes \square pp))/pp) \bullet np

The lambda term that matches the semantic type of the unfolded type assignment is the following.
80 3. A type schema for wh-question formation

\[
\begin{align*}
&\text{to} \vdash \text{to} : \text{pp}/\text{np} \quad [u : \text{np} \vdash u : \text{np}] \\
&\quad \to \circ \text{np} \vdash (\text{to} u) : \text{pp} \quad [/E] \\
&\quad \to \circ \text{whom} \vdash \text{WH}_m(np, pp, WH'_r(\square pp, q, wh)) \\
&\quad \lambda Q. \lambda P.(\omega \lambda x.(P (Q x))) \\
&\text{to} \circ \text{whom} \vdash WH'_r(\square pp, q, wh)
\end{align*}
\]

\((\lambda Q. \lambda P.(\omega \lambda x.(P (Q x)))) \lambda u.(\text{to} u))
\sim_{\beta^*} \lambda P.(\omega \lambda x.(P (\text{to} x)))

Subsequently, we use the derivation of ‘to whom’ to derive the complete pied-piping construction ‘To whom did John give a book?’. To simplify the derivation, we abbreviate the question body ‘did John give a book’ to a single entry, \text{BODY}.

\[
\begin{align*}
&\text{to} \circ \text{whom} \vdash WH'_r(\square pp, q, wh) \\
&\quad (\text{BODY} \circ \square pp) \vdash q \\
&\quad \lambda P.(\omega \lambda x.(P (\text{to} x))) \\
&\quad (\text{to} \circ \text{whom}) \circ \text{BODY} \vdash \text{wh} \\
&\quad (\lambda P.(\omega \lambda x.(P (\text{to} x)))) \lambda v.(((\text{give a book}) v) j) \\
&\sim_{\beta^*} (\omega \lambda x(((\text{give a book}) (\text{to} x)) j))
\end{align*}
\]

3.3.5 Multiple wh-questions

In sentences with multiple wh-phrases, one wh-phrase occurs clause-initially while the other wh-phrases stay in-situ, for example ‘Who saw whom?’ or ‘What did John give to whom?’. To analyze multiple wh-questions, we must add lexical ambiguity.

In general, subject wh-phrases in a structure with multiple wh-phrases have the same type as subject wh-phrases in sentences with a single wh-phrase. However, some non-subject arguments occur fronted or in-situ depending on the availability of other wh-phrases in the multiple wh-question. We have seen in previous sections that in a single wh-question a non-subject wh-phrase extracts an \text{np}-typed hypothesis from a \text{q}-typed context. For the analysis of a multiple wh-question, a non-subject wh-phrase may not only merge with a \text{q}-typed sentences, but it may also merge with a \text{wh}-typed question body. A sentence of type \text{wh} is a sentence which already has an occur-

(see section 3.2.4.2):

\[
(\lambda y^\tau. \lambda P(\square pp \vdash (\omega \lambda y^\tau. (P y)), x) : \text{\square pp/(q/\square pp))/pp}) \bullet \text{np}
\]
3.3. Analysis of English wh-questions

tance of a fronted wh-phrase. We first inspect the multiple wh-question ‘Who saw whom?’ where only the subject wh-phrase may occur fronted.

Following the above line of reasoning, the alternative type to assign to non-subject wh-phrases such as ‘whom’ is \( WH_{in}(np, wh, wh) \). Doing so, the wh-in-situ phrase is dependent on the availability of another wh-phrase.

subject wh-phrase :: \( WH_{ex}(np, s, wh) \)
non-subject wh-phrase :: \( WH_{in}(np, wh, wh) \)

The tree style presentation in figure 3.19 illustrates the derivation of the multiple wh-question ‘Who saw whom?’. The natural deduction derivation of ‘Who saw whom?’ is the following:

\[
\begin{align*}
\text{saw} & \quad \vdash \text{saw} \quad [\text{np}] \\
\text{np} \circ (\text{saw} \circ \text{np}) & \quad \vdash \text{saw} \quad [WH_{ex}] \\
\vdash \text{wh} & \quad [WH_{in}] \\
\text{wh} \circ (\text{saw} \circ \text{wh}) & \quad \vdash \text{wh} \\
\end{align*}
\]

\((\omega \lambda y.(\omega \lambda x.(\text{see } y) x)))\)

Without presenting the details of the analyses, we describe how, with the proposed type assignments, we can account for other occurrences of multiple wh-questions in English (see appendix B.1.6 for on-line derivations). For example, multiple wh-questions with only non-subject argument phrases such as the sentences in example 3.32a) and b) can be derived by using the single constituent type \( WH_{ex}(A, q, wh) \) for the fronted wh-phrase and the wh-in-situ type for the embedded wh-phrase. Also multiple wh-questions with more than two wh-phrases can be derived (example 3.32c). Both non-subject wh-phrases are identified as a in-situ wh-type schema \( WH_{in}(np, wh, wh) \).

(3.32) a. What did John give to whom?

b. To whom did John give what?

c. Who gave what to whom?

\(^{12}\)Adding lexical ambiguity is a less elegant solution for the derivation of multiple wh-questions. For the present purpose to show that with the distinct wh-type schema we can analyze the different syntactic occurrences of wh-question formation, we believe that this solution is adequate. In chapter 5, we will show that by decomposing the macro type \( \omega \) assigned to wh-questions, we can derive these ‘multiple’ wh-type schemata from a single occurrence.
3. A type schema for wh-question formation

The ungrammaticality of a multiple wh-questions where the non-subject argument phrase appears fronted while the subject wh-phrase stays in-situ (example 3.33a) or where the indirect object wh-phrase is fronted while the direct object stays in-situ (example 3.33b) are ruled out on a lexical basis. In English, a ‘multiple’ type-assignment for subject wh-phrases does not exist in the lexicon. In chapter 5, when we decompose the macro type for wh, we provide additional evidence for ruling out these wh-questions.

(3.33) a. *Whom did who see?
        b. *Whom did you give what to?

Multiple wh-questions with wh-determiners

An interesting question to ask here is: Why is it not possible to derive the wrong order ‘* Whom did who see?’ or ‘* Whom who see’? If we played the devil’s advocate, we could assign a type to the object wh-phrase such that it can be of an ex-situ type in multiple wh-questions. Then it would get the type $\text{WH}_{ex}(np, wh, wh)$. When we inspect languages where multiple wh-phrases
3.4. Concluding remarks on the wh-type schema

We have proposed a type schema to account for the syntactic and semantic properties of wh-question formation. The general characteristics of wh-question formation have been illustrated with an overview of English wh-questions. These characteristics have formed a background to understand the different properties the wh-type schema has to contain. The general wh-type schema encodes the characteristics of a wh-phrase: firstly, that it associates with a gap in a question body and secondly, that it gives a sentence a question interpretation. The three syntactic components that constitute a wh-question, the gap, the question body and the result, the wh-question itself, are directly encoded in a three place type operator WH. The three argument positions of the wh-type schema identify the syntactic requirements of a wh-phrase.

The wh-type schema provides the basic setup for characterizing wh-question formation. In order to capture the syntactic differences between wh-phrases, we distinguish between two kinds of wh-type schemata: a wh-in-situ
3. A type schema for wh-question formation

Figure 3.20: ‘Subject–Object’-ordering of a multiple wh-question with wh-determiners
3.4. Concluding remarks on the wh-type schema

Figure 3.21: ‘Object→Subject’-ordering of a multiple wh-question with wh-determiners
and a wh-ex-situ type schema. Wh-phrases that occur clause-initially are lexically identified as wh-ex-situ types. While wh-in-situ types are used for wh-phrases that occur in their 'base position'. The ex-situ types have been further subdivided in a left and right variant where the direction indicates the occurrence of the gap in the question body. The decomposition of the three variants of the wh-type schema in terms of the usual connectives in the base logic, resolved a semantic and a syntactic issue on the two variants of the wh-type schema. Semantically, we have shown that a logical constant operator $\omega$ can be used to compute the meaning assembly of wh-questions. Syntactically, both the ex-situ and the in-situ type schema have additional feature requirements in order to apply the restricted set of displacement postulates as encoded in the structural module. For the wh-ex-situ type we need the displacement postulates to access the gap hypothesis when it is more deeply embedded in the question body. For the wh-in-situ type, we need the reversed set of displacement postulates in order to determine the right scope of the wh-phrase in the question body.

After the technical exploration of the wh-type schema, we have analyzed characteristics of English wh-questions. The synthesized types and rules for the wh-type schemata allow us to capture the distinct syntactic and semantic properties of wh-question formation in English. The grammaticality or ungrammaticality of certain sentences are derived on a purely lexical basis specifying the selectional and feature requirements for each syntactic object. Through the interaction between syntactic objects, we can derive local and non-local wh-questions, island constructions and multiple wh-questions.

In the coming chapter, we will show that within the logic of variation we can account for cross-linguistics differences in wh-question formation. With the use of the wh-type schema, we show that we can generalize over the different forms of wh-questions for which, through the syntax-semantics interface between form and meaning, we derive a uniform meaning assembly.
Chapter 4

Structural variation in wh-question formation

In the previous chapter, we have presented a uniform account of wh-question formation. Wh-phrases are typed as instances of a wh-type schema. The different instances of the wh-type schema capture cross-linguistic differences while preserving a uniform semantic interpretation. In this chapter we will explore the syntactic variation of wh-question formation by testing our proposal on languages with a different syntax for wh-question formation. By exploring the phenomena in these languages, we show that the general wh-type schema accounts for wh-question formation cross-linguistically. The syntactic variation of wh-question formation is the result of instantiating the right kind of wh-type schema for a certain wh-phrase. By providing analyses of the wh-question formation in different languages, we provide evidence that syntactic variation among languages is merely a distinction on lexical grounds.

In section 4.1, we present an overview of the different wh-question constructions that occur in natural languages. We illustrate each construction with data from a language that uses the specific construction. In sections 4.2, 4.3 and 4.4, we present each construction in more detail. In section 4.2, we focus on wh-in-situ languages with the main discussion on Japanese and a brief discussion on Chinese. In section 4.3, we consider wh-ex-situ languages and show the differences and similarities between Serbo-Croatian and Bulgarian. In section 4.4, we look at partial wh-movement with scope markers in German and Hindi and briefly mention Malay, a language with partial wh-movement without a scope marker. In section 4.5, we end the chapter with a brief overview of the different ways in which variation in wh-question formation is captured using the wh-type schemata.
4.1 An overview of the data

The following groups describe the structural differences in wh-question formation: wh-in situ languages (Japanese, Chinese), wh-ex-situ languages (Bulgarian, Serbo-Croatian) and languages with partial wh-movement (German, Hindi, Malay). These groups present an interesting typological distinction in the syntactic variation of question formation. A further distinction can be made within these groups. Wh-ex-situ languages can either have simple wh-ex-situ or multiple wh-ex-situ constructions. In multiple wh-ex-situ constructions, more than one wh-phrase can get fronted in a sentence with multiple wh-phrases; whereas in simple wh-fronting only one wh-phrase appears clause-initially. Wh-in-situ languages can be further distinguished by the presence or absence of a question marker. A question marker marks the clause as a question. A similar distinction can be seen in languages with partial wh-movement with the occurrence or absence of a scope marker. As the term indicates, a scope marker determines the scope of a partially moved wh-phrase. Before we present examples for each of the language groups, we list the distinct syntactic realizations of wh-questions in the diagram in figure 4.1.

| i. wh-in-situ | with question marker ⇒ Japanese | without question marker ⇒ Chinese |
| ii. wh-ex-situ | simple wh-ex-situ ⇒ English, Dutch | multiple wh-ex-situ ⇒ Bulgarian, Serbo-Croatian |
| iii. partial wh-move | with scope marker ⇒ Hindi, German | without scope marker ⇒ Malay |

Figure 4.1: Language groups with distinct syntactic realizations of wh-questions

The diagram lists exemplars of languages that allow the specified wh-interrogative construction. Note that although a language is grouped under one type of construction, this does not exclude other wh-constructions from
4.1. An overview of the data

Languages may sometimes use more than one ‘strategy’ for the formation of wh-questions. However, in this chapter, we choose to concentrate on the analysis of the basic constructions and illustrate how we can account for language variation and language particular properties. The analysis of wh-question formation in wh-in-situ languages is presented on the basis of Japanese. We will only briefly extend our proposal to Chinese. In the previous chapter, we have illustrated and analyzed English, which is a simple wh-ex-situ language. For the analysis of multiple wh-ex-situ, we compare Bulgarian and Serbo-Croatian. German and Hindi are used to illustrate the analysis of partial wh-movement. And Malay is used as an example of a language which has a wh-question construction that involves partial wh-movement without the presence scope marker.

4.1.1 Wh-in-situ

Wh-questions in wh-in-situ languages are formed by sentences where the wh-phrase stays in-situ in the base position. Semantically, however, it takes scope over the clause where it is embedded in.

**Wh-in-situ with question marker** In Japanese, questions are indicated by the particle ‘ka’ which appears at the final position in the clause. Thus, the clause that is ‘marked’ by ‘ka’ is interpreted as an interrogative. We refer to the particle ‘ka’ as a question marker and gloss the particle as ‘Q’.

If the clause does not contain any wh-elements, the clause is interpreted as a yes-no question. If the clause does have a wh-phrase the sentence becomes a wh-question. Example 4.1a) illustrates a direct main clause question (Nishi-gauchi, 1990, ex.14,p.6). Example 4.1b) illustrates a non-local direct question where the embedded wh-phrase is interpreted at the main clause (Kim, 1989).

\[(4.1) \]
\[
\begin{align*}
\text{a. } & \text{John-wa nani-o tabe-masita ka?} \\
& \text{John-[top] what-[acc] eat[past] Q} \\
& \text{‘What did John eat?’}
\end{align*}
\]
\[
\begin{align*}
\text{b. } & \text{John-wa [Mary-ga nani-o katta] to] itta ka?} \\
& \text{‘What did John say that Mary bought?’}
\end{align*}
\]

**Wh-in-situ without question marker** In Chinese, wh-questions do not need to be marked with an additional question marker. The wh-phrase appears clause-externally, but the clause, in which the wh-phrase is embedded, is interpreted as a wh-question. Example 4.2a) illustrates a main clause question (Aoun and Li, 1993). Example 4.2b) illustrates that a clause where the matrix clause takes an embedded interrogative, the sentence is interpreted as an indirect question (Huang, 1998, ex.162).
4. Structural variation in wh-question formation

(4.2)  

a. Zhangsan kandao shenme?  
Zhangsan saw what  
‘What did Zhangsan see?’

b. Zhangsan ask wo [shei mai-le shu].  
Zhangsan ask I who buy-[asp] book  
‘Zhangsan asked me who bought books.’

4.1.2 Wh-ex-situ

In wh-ex-situ languages, a wh-phrase appears in fronted position while the gap appears more deeply embedded in the clause. Languages that belong to this group are further divided into simple wh-ex-situ and multiple wh-ex-situ.

Simple wh-ex-situ  Simple wh-ex-situ languages only front a single wh-phrase in sentences with multiple wh-phrases. Examples of languages that have simple wh-ex-situ constructions are English and Dutch. We have presented an analysis of wh-question formation in English in chapter 3. For a comparison with multiple wh-ex-situ language, we repeat the example of a single and a multiple wh-question in English:

(4.3)  

a. What did you eat today?

b. Who ate what?

Multiple wh-ex-situ  Languages with multiple wh-ex-situ constructions share the phenomenon that all wh-phrases get fronted when more than one argument is being questioned. The Slavic languages, Bulgarian and Serbo-Croatian follow this pattern. In section 4.3, we present an account of the similarities and distinctions in wh-question formation in these languages. Example 4.4 and 4.5 illustrate the basic constructions of wh-questions in Serbo-Croatian (SC) and Bulgarian (B).

(4.4) Simple direct question

a. Ko kupuje knjigu? [SC]  
‘Who is buying a book?’

b. Koj otvori vratata? [B]  
Who opened door-the  
‘Who opened the door?’

(4.5) Multiple wh-question
4.1. An overview of the data

Ko kogo vidi? [SC]
Kaj kogo vižda [B]
who whom sees
‘Who does whom see?’

4.1.3 Partial wh-movement

Without a scope marker  German and Hindi belong to the group of languages with a scope marker that allow wh-phrases to be partially moved. Such constructions are sometimes referred to as scope marking constructions, because they depend on the occurrence of a scope marker in the clauses preceding the wh-phrase. However, besides partial wh-movement, both languages have other ways to realize wh-question constructions. Example 4.6a) illustrates a scope marking construction in German (Klepp, 1999, ex.4). Example 4.6b) illustrates a scope marking construction in Hindi with the wh-word occurring in a preverbal position (Dayal, 2000, ex.5, p.160). In section 4.4, we will present an analysis of scope marking constructions in the two languages and discuss the similarities and differences between the two languages.

(4.6) a. Was glaubte Miró welches Bild Picasso tı gemalt
which believed Miró which picture Picasso painted
hatte?
had
‘Which picture did Miró believe that Picasso had painted?’

b. John kyaa socaa hai ki merii kis-se buat karegii?
John what think-[pres] that Mary who-with talk do-[Fut]
‘Who does John think Mary will talk to?’

Without a scope marker  Malay allows both wh-ex-situ question formation, and wh-in-situ question formation, but also wh-questions with the wh-phrase partially moved. However, different from German and Hindi, the clause which is interpreted as a question does not contain a question marker. The following example illustrates such a partially moved wh-phrase (Cole and Hermon, 2000, ex.7a), the wh-phrase ‘ke mana’ (= to where) has moved to the front of the embedded clause, but the whole sentence is interpreted as a direct question.

(4.7) Ali memberitahu kamu tadi apa (yang) Fatimah baca tı ?
A. told you just now what that F. read
‘What did Ali tell you just now that Fatimah was reading?’
Introduction to the analyses

We have illustrated three structurally different strategies for constructing a wh-question. Some languages only allow one kind of construction while other languages have more than one way to form a wh-question. This thesis is not intended to present an extensive overview of the typological diversity of wh-question formation, rather we analyze the cross-linguistic data to provide empirical evidence that the wh-type schema proposed in chapter 3 can be used to capture the syntactic variation in wh-question formation.

We limit the analysis to the languages listed as exemplars in the diagram. For each analysis, we use the following setup. First, we present data that illustrate the basic syntactic properties of a particular language. This is needed to present a proper analysis of wh-question formation in that language. Second, we present an overview of data on wh-questions. Third, we present an analysis of the data by specifying how wh-phrases and other words are encoded in the lexicon. For each set of data, we propose an analysis of the data in the grammar system extended with synthesized types and rules for wh-type schemata, as presented in chapter 3.

The analyses are presented in separate sections in the text that start with a horizontal line with a ♦ on the left edge and end with a horizontal line with a ♠ on the right edge. Derivations of wh-interrogative sentences or parts of a derivation are presented either in a tree style derivation or in a natural deduction format. All sentences discussed in this chapter can be analyzed with an on-line parser. For these on-line derivations, we refer to appendix B for an overview of sample sentences.¹

4.2 Wh-in-situ

Both Japanese and Chinese are wh-in-situ languages. Japanese has an overt question marker that marks a sentence as a question, whereas Chinese does not need such a question marker. Nevertheless, the analyses of Chinese and Japanese are quite similar. We mainly concentrate on Japanese to show the syntactic analysis of wh-question formation in a wh-in-situ language. We only briefly discuss Chinese and indicate the differences with Japanese wh-question formation in section 4.2.7.

Japanese is a wh-in-situ language where the wh-phrase in constituent wh-questions stays in-situ. An overt question marker at the final position of a clause determines whether the clause is interpreted as a question. We analyze Japanese wh-questions using the wh-in-situ type schema. We will show that wh-phrases depend on the occurrence of a question marker to form a wh-question.

First, in section 4.2.1, we look at the basic underlying word order and the selectional requirements of verbs in Japanese. Secondly, in section 4.2.2, we

¹The sample sentences are hyperlinks to an on-line parser for categorial type logics, named 'Grail'. The derivability of a sentence can be checked by clicking on it.
present data that illustrate the different types of interrogative clauses. On the basis of these data, we present an analysis of main clause interrogatives which shows the dependency of a wh-phrase on the occurrence of a question marker. In section 4.2.3 - 4.2.6, we extend our analysis to derive embedded interrogatives, distinctions between question markers, complex noun phrase constructions, scope interaction with quantifier phrases and multiple wh-questions.

4.2.1 Basic word order

Japanese is a head-final language; the basic underlying word order is SOV. The verb appears in a clause-final position and the arguments group to the left of the verb that selects them. The argument phrases are morphologically marked for case features. The case markers, the particles (-wa,-ga,-o,-ni), are post-positionally attached to noun phrases. ‘Wa’ is a topic marker, ‘ga’ indicates the subject, ‘o’ is attached to an object and ‘ni’ indicates the indirect object or dative form of a noun phrase. Because the argument role of the noun phrases is morphologically expressed, the argument order may vary without losing information about the role of the arguments.

The following examples illustrate the underlying word order in declarative clauses. Example 4.8a) illustrates a declarative clause and 4.8b) illustrates a main clause with a subordinate clause. The examples illustrate the underlying word order, but any ordering of the argument phrases, as long as they precede the verb that selects them, is allowed.

(4.8) a. John-ga  mary-ni  konpyuuta-o  ageta
John-[nom] Mary-[dat] computer-[acc] gave
‘John gave a computer to Mary.’

b. John-wa  [Mary-ga  konpyuuta-o  kiratteiru  to]  itta
‘John said Mary hates computers.’

The basic underlying SOV word order will be directly encoded in the type-assignment of verbs in the lexicon. The verbs select arguments that have a certain case feature assigned to them. To express case feature information, we follow Heylen (1999) and use unary operators ♦ to distinguish the different cases of noun phrase elements. To simplify the derivations we will use type abbreviations, NOM, ACC, DAT and TOP to identify noun phrases that are assigned either nominative, accusative, dative or topical case. The type-abbreviations are given in the following table.

<table>
<thead>
<tr>
<th>Type</th>
<th>Abbreviation</th>
</tr>
</thead>
<tbody>
<tr>
<td>NOM</td>
<td>♦□_nomHp</td>
</tr>
<tr>
<td>ACC</td>
<td>♦□_accHp</td>
</tr>
<tr>
<td>TOP</td>
<td>♦□_topHp</td>
</tr>
<tr>
<td>DAT</td>
<td>♦□_datHp</td>
</tr>
</tbody>
</table>
The case markers, which postpositionally attach to noun phrases, get the following type-assignments.

\[
\begin{align*}
wa &: np \to \text{TOP} \\
g\alpha &: np \to \text{NOM} \\
o &: np \to \text{ACC} \\
n\imath &: np \to \text{DAT}
\end{align*}
\]

In Japanese, verbs also have morphological marking. However, to keep things simple and to focus on wh-question formation, we assume verbs to appear fully inflected in the lexicon. Additionally, we add multiple type-assignments for verbs that select a nominative and those that select a topicalized noun phrase as subject argument. Transitive verbs such as ‘tabeta’ (= eat) select both an accusative and a nominative noun phrase. Ditransitive verbs such as ‘ageta’ (= give) also selects an argument with dative case. Verbs like ‘tazuneta’ (= ask), and ‘itta’ (= said) select a sentential complement before selecting a nominative or topicalized noun phrase.

The following list of verbs illustrates the types assigned to the different kinds of verbs. Additionally, we present a selection of the noun phrases that are used in the examples.

\[
\begin{align*}
tabeta &: ACC \setminus (NOM \setminus s) \ (\text{= ate}) \\
ageta &: DAT \setminus (ACC \setminus (NOM \setminus s)) \ (\text{= gave}) \\
itta &: s^\prime \setminus (NOM \setminus s) \ (\text{= said}) \\
tazuneta &: \text{wh}^\prime \setminus (NOM \setminus s) \ (\text{= asked}) \\
\text{John, Mary} &: np \\
\text{hon} &: np \ (\text{= book})
\end{align*}
\]

**scrambling** As we have noted for the declarative clause in example 4.8, arguments in Japanese occur preverbally, but the argument order is free. Arguments may scramble to different positions in the sentence structure. In the type-logical grammar analysis, each case marked noun phrase is decorated with unary features. We propose that the case features in Japanese enable the application of the following displacement postulate. This displacement postulate is an instance of the displacement postulates that we have assumed to be universally available (see chapter 2).

\[
\begin{align*}
\Gamma \Delta_2 \cdot (\langle case \Delta_1 \cdot \Delta_0 \rangle) \vdash C & \quad [\text{Scramble}] \\
\Gamma \langle case \Delta_1 \cdot (\Delta_2 \cdot \Delta_3) \rangle \vdash C & \quad \text{where case} \in \{\text{nom,acc,dat,top}\}
\end{align*}
\]

The **scramble** rule encodes that all arguments decorated with a case feature may scramble to a position further to the left. (see appendix B.2.1 for illustrations of different possible word orders of the sentences in example 4.8). On the basis of the availability of the displacement postulate, we can derive the different argument orders in Japanese.

\[
\square
\]
4.2. Wh-in-situ

4.2.2 Interrogative clauses

A sentence will be interpreted as a question when the sentence contains an interrogative marker, such as ‘ka’ or ‘kadooka’. The function of the interrogative marker is twofold: first, it marks the sentence as interrogative. Second, if there is a wh-phrase, it determines the scope of the embedded wh-phrase. Simply said, the placement of an interrogative marker sets the domain over which the wh-in-situ can take scope.

The following examples illustrate the double role of the question marker. First, in examples 4.9–4.10, we look at sentences without a wh-element, either a main clause or an embedded clause. These sentences are marked by a question marker, ‘ka’ or ‘kadooka’. ‘Kadooka’ can only be used in embedded interrogatives and is interpreted as if or whether, whereas ‘ka’ can mark both the main clause and the subordinate clause. Second, in examples 4.11–4.14, we look at direct and indirect questions where ‘ka’ marks a sentence which contains a wh-phrase.

- When ‘ka’ marks the matrix clause, the main clause is interpreted as a yes-no question (Kim, 1989).

  (4.9) John-wa [Mary-ga hon-o katta to] itta ka?
  ‘Did John say that Mary bought a book?’

- When ‘ka’ or ‘kadooka’ marks an embedded clause, the clause is interpreted as an if/whether clause:

  (4.10) John-wa [Mary-ga hon-o katta ka(dooka)]
  tazuneta asked
  ‘John asked if/whether Mary bought a book.’

When the matrix clause or the embedded clause contains a wh-phrase, the sentence may either be interpreted as a main clause question or a direct question depending on the placement of the interrogative marker.

- When ‘ka’ marks a main clause which contains a wh-in-situ, the sentence is interpreted as a wh-question (Nishigauchi, 1990; Hagstrom, 1999).

  (4.11) John-wa nani-o tabe-ta ka?
  ‘What did John eat?’

- When ‘ka’ marks an embedded clause which contains a wh-phrase, the embedded clause is interpreted as an embedded interrogative clause (Nishigauchi, 1990, ex.13, p.19).
96  4. Structural variation in wh-question formation

(4.12) Sinbun-wa [dare-ga erab-are-ta ka] newspapers-[top] who-[nom] elected-was Q
tutaete-i-nai report-[ing]-[not]

“Newspapers do not report who was elected.”

• When ‘ka’ marks the matrix clause while the embedded clause containing the wh-in-situ phrase is unmarked, the whole sentence is interpreted as a wh-question. The wh-phrase is interpreted at the clause that is marked by the question marker (Nishigauchi, 1990, ex.14, p.19).

(4.13) Sinbun-wa [dare-ga erab-are-ta to] tutaete-iru ka newspaper[top] who[nom] elected-was that report-[ing] Q

“Who do the newspapers report was elected?”

• When the question marker marks both the embedded clause and the matrix clause, a wh-phrase is interpreted under the nearest question marker. The embedded clause is interpreted as an embedded wh-question while the matrix clause is interpreted as a yes-no question (Nishigauchi, 1990, ex.35, p30):

oboete-iru ka? remember-[ing] Q

‘Does Satooo remember what Suzuki ate?’
NOT: “What does Satooo remember that Suzuki ate?”

To sum up, ‘ka’ creates a structural context causing the sentential clause that it marks to be interpreted as a question. When the clause does not contain a wh-phrase, the sentence is interpreted as a yes-no question or an if/whether clause. When the clause does contain a wh-phrase, the clause with the nearest ‘ka’ interrogative marker is interpreted as a wh-question. In the coming section, we will analyze the role of the question marker and the dependency of the wh-phrase on the occurrence of a question marker.

◊ [interrogatives]

We have illustrated how a question marker changes the interpretation of a sentence. In a simplified way, we recognize the following constructions. Example 4.15a illustrates a declarative clause. In example 4.15b, the clause is marked with ‘ka’ and is interpreted as a yes-no question. In example 4.15c, the clause is marked by ‘ka’ and contains the embedded wh-phrase ‘nani’. This clause is interpreted as a direct question.

(4.15)   a. John-ga hon-o katta. (= John bought a book)
4.2. Wh-in-situ

b. John-ga hon-o katta ka? (= Did John buy a book?)
c. John-ga nani-o katta ka? (= What did John buy?)

We will use type $q$ to type yes-no-questions. Based on the function of the question marker to change the type of a declarative clause $s$ into a yes-no-question ($q$), we assign the question marker ‘ka’ the following type-assignment.

$$ka :: s \setminus q$$

Wh-phrases in Japanese are very similar to quantifier phrases. Both wh-phrases and quantifier phrases appear in-situ while they are interpreted outside of the clause where they are embedded in. Following the description of wh-question formation in Japanese, we propose to assign argument wh-phrases a wh-in-situ type schema where the body of the question is $q$ which is the type of structures that are marked by the question marker ‘ka’. The wh-question question that is formed after merging the wh-phrase to the question body is of type $\text{wh}$. Thus, we assign wh-phrases such as ‘dare’ (= who) and ‘nani’ (= what) the following type-assignment.

$$\text{wh-phrase} :: \text{WH}_{\text{in}}(np, q, \text{wh})$$

With the above type-assignment, we can derive the sentences illustrated in example 4.15 (see appendix B.2.3). To illustrate the analysis of a wh-question showing that that the wh-phrase is merged with a question body marked by the question marker ‘ka’, we derive the wh-question ‘John-ga nani-o katta-ka’ (= What did John buy?). To merge the wh-phrase to the question body, we use the wh-in-situ type schema rule $[\text{WH}_{\text{in}}]$ as presented in chapter 3, page 52.

$$\begin{align*}
\text{john} \circ \text{ga} & \quad [np] \quad [\text{ACC}] \quad [\text{E}] \\
\circ & \quad np \setminus \text{ACC} \quad [\text{E}] \\
\circ & \quad \text{katta} \quad \text{ACC} \setminus \text{(NOM \setminus s)} \quad [\text{E}] \\
\circ & \quad \text{ka} \quad s \setminus q \quad [\text{E}] \\
\circ & \quad \text{nani} \quad \text{WH}_{\text{in}}(np, q, \text{wh}) \\
\circ & \quad \text{WH}_{\text{in}}(np, q, \text{wh}) \\
\circ & \quad \omega \, \lambda y . ((\text{buy} \, y))
\end{align*}$$

The derivation shows that the object argument consists of an $np$ hypothesis which is merged with the case marker ‘-o’. The hypothesis-case marker combination and the subject argument ‘john-ga’ fulfill the selectional requirements
of the verb. After the question marker is merged with the main clause, creating a structure of type \( q \), the derivation has reached a structure in which the wh-phrase can be merged. The wh-in-situ rule (see appendix A) inserts a wh-in-situ phrase ‘*nani*’ in a \( q \)-typed structure at the position where the gap hypothesis occurs.

Along with the syntactic merge step, we can compute the semantic term of the interrogative, which is printed in the line below the syntactic derivation. The removed gap hypothesis corresponds to the term variable \( x \) bound by the \( \lambda \) abstractor. And the merging of the wh-in-situ phrase with the question body corresponds to the application of the semantic operator \( \omega \) to the resulting \( \lambda \)-term. The meaning assembly of this wh-interrogative matches the meaning assembly of the wh-interrogative of wh-ex-situ phrases (see section 3.3.2.1 for an analysis of English direct questions).

### 4.2.3 Embedded wh-phrases

An embedded clause may either be an embedded interrogative or an embedded declarative clause. Whether the wh-phrase is interpreted inside the embedded clause or at the main clause, depends on the type of verb that selects the embedded clause. Similarly to English, Japanese shows a distinction between selection properties of verbs such as *ask/wonder* which require an interrogative complement and *think/believe* which take a declarative complement. The Q-morpheme marks complement clauses that are selected by ‘ask’-type verbs. No morpheme or only the complementizer ‘to’ is needed for embedded declarative clause that are selected by ‘think’ type verbs. The following sentences (Kim, 1989, ex. 5,6) illustrate the correct use of question markers with the two types of verbs: ‘*tazuneta*’ (= *asked*) selects an interrogative clause and ‘*sinzita*’ (= *believed*) only selects a declarative clause.

(4.16)  
\[
\begin{array}{c}
\text{John-wa} \quad \text{[Mary-ga hon-o katta ka/to] tazune-ta} \\
\end{array}
\]

‘John asked if/whether Mary bought a book.’

(4.17)  
\[
\begin{array}{c}
\text{John-wa} \quad \text{[Mary-ga hon-o katta to/ka] sinzita} \\
\end{array}
\]

‘John believed that Mary bought a book.’

When the question marker ‘*ka*’ marks the matrix clause, the sentence with an embedded interrogative can only get interpreted as a main clause yes-no-question with an embedded wh-question. Whereas a sentence with a wh-phrase embedded in a declarative complement clause selected by ‘*sinzita*’ (= *believed*) gets wide-scope reading. In both cases, the wh-phrase gets bound to the closest question marker. The following two examples illustrate this distinction.

---

\[\text{Kim (1989) bases his examples on data in Lasnik and Saito (1984)}\]
To capture the distinction between verbs that require an interrogative complement, and verbs that only permit a declarative clause as their complement, we need to encode the selectional requirements of the verbs in the lexicon. As before, we differentiate between types assigned to main clauses (s, q and wh) and subordinate clauses (s', q' and wh'). We take ‘sinzita’ (= believed) as a verb belonging to the class of verbs which only take a declarative complement and ‘asked’ (= tazuneta) as a verb that selects an interrogative complement. Because we use different types for ‘if/whether’ clauses (q') and embedded interrogative clauses (wh'), ‘tazuneta’ comes in two variants. ‘tazuneta’ may either merge with a q-typed clause or it may merge with an embedded interrogative (wh'). The types for the embedded clauses are determined either by the complementizer ‘to’ or the question marker ‘ka’. ‘To’ merges with a s-typed clause and returns the type for embedded declarative clauses (s'). ‘Ka’ or ‘kadooka’ merges with a s-typed clause and yields an embedded if/whether clause type (q'). Additionally, we allow wh-phrases to be merged with embedded q'-typed clauses to yield an embedded wh-question wh'. We extend the lexicon with the following type-assignments.

\[
\begin{align*}
\text{sinzita} &:: s'\backslash (A \backslash s) \quad (= \text{believed}) \\
\text{tazuneta} &:: \text{wh}'\backslash (A \backslash s) \quad (= \text{asked wh...}) \\
\text{tazuneta} &:: q'\backslash (A \backslash s) \quad (= \text{asked if...}) \\
&\quad \text{where } A \in \{\text{TOP, NOM}\}
\end{align*}
\]

\[
\begin{align*}
\text{to} &:: s' \backslash s' \\
\text{ka, kadooka} &:: s' \backslash q' \\
\text{wh-phrases} &:: \text{WH}_p(np, q', \text{wh}')
\end{align*}
\]

With these type-assignments, we can analyze the sentences in the examples 4.16 – 4.19 (see appendix B.2.3). We have now captured the selectional requirement of verbs that do and verbs that do not select an interrogative complement.

In example 4.14, we have shown that a clause with an embedded wh-phrase is always interpreted as a wh-question at the nearest ‘ka’ question marker. However, on the basis of the above type-assignments we can still derive sentence 4.20 where the embedded wh-phrase is interpreted at the matrix clause. We would like to rule out this derivation and only allow the derivation of the sentence in example 4.21.
(4.20) *John-wa [dare-ga hon-o katta ka] tazuneta ka? \( \vdash \) wh

\( (= \text{Who, did John ask whether t bought a book?}) \)

(4.21) John-wa [dare-ga hon-o katta ka] tazuneta ka? \( \vdash \) q

\( (= \text{Did John ask who bought a book?}) \)

To explain what goes wrong with the derivation, we present an analysis of the incorrect example. The following derivation shows that the embedded clause can be merged as a q′ complement to ‘tazuneta’. (See appendix B.2.3 for on-line derivation of both the incorrect and the correct example.) The wh-in-situ phrase is merged with the main clause in the last step of the derivation, replacing its np gap hypothesis and deriving a main clause wh-question.

\[
\vdash \text{\small{\text{tazuneta}}}
\]

\[
\vdash \text{\small{\text{\top \s\[E\]}}}
\]

\[
\vdash \text{\small{\text{\WHin}}}
\]

\[
\vdash \text{\small{\text{\WHin(np, q, wh)} (\omega \lambda \text{\small{x}. (\text{\small{ask (buy book) \ x}})) John)}}
\]

The correct analysis of the sentence should yield a q type with a yes-no question interpretation. Nishigauchi (1990) has indicated that it is due to the properties of ‘ka’ that movement of the wh-phrase beyond the position of the question marker is blocked. According to us, however, it is not due to the properties of the question marker that the derivation is blocked. The only grammatical function of ‘ka’ is that it marks the sentence as a question. Except creating a question body of the right type where the wh-phrase can be merged with, it does not impose any restriction on the scope properties of the wh-phrase. As can be seen in the derivation above, nothing prevents the wh-phrase to be merged with a clause marked by another higher question marker. However, the derivation can be ruled out when the path along which the gap hypothesis is reached, is blocked. Along similar lines we have analyzed wh-islands in English (see chapter 3, section 3.3.4).

Wh-islands in English, such as the sentence in example 4.22 are derived on the basis of the selectional requirements of the verb. The English verb ‘wonder’ has been typed as inf\(\checkmark q\). The \(\checkmark\) feature projects a structural domain around the embedded interrogative clause which prevents a fronted wh-phrase to associate with the gap hypothesis in the embedded interrogative clause.

\[
\vdash \text{\small{\text{\WHin}}} \quad \text{WHin(np, q, wh)}
\]

\[
\vdash \text{\small{\text{\WHin(np, q, wh)}}}
\]

(4.22) *Which bird, did John wonder whether Mary wanted to see t?
4.2. Wh-in-situ

We use the same analysis to project a structural domain around the interrogative clause preventing the wh-phrase to be merged at the main clause, but permitting the wh-phrase to associate with the gap hypothesis in the embedded clause. Thus, the type-assignments of ‘tazuneta’ (= ask) are changed as follows.

\[
\begin{align*}
tazuneta & :: \text{\textbullet}q'( \text{TOP}\_s) \\
tazuneta & :: \text{\textbullet}wh'( \text{TOP}\_s)
\end{align*}
\]

The derivation of the wh-question in example 4.20 is blocked due to additional feature requirements of the verb. The verb ‘tazuneta’ (= ask) selects for an interrogative clause typed \textbullet wh or \textbullet q. The \textbullet feature requirement projects a \textbullet structural domain around the embedded interrogative. The following tree style presentation illustrates that the wh-phrase cannot be merged at the main clause level, because it cannot associate with the gap hypothesis due to the blocking effect of the embedded interrogative, represented as a unary \textbullet branch.

![Tree diagram](image)

**Figure 4.2:** \textbullet dare :: WH_{\text{wh}}(np, q, wh) cannot associate with its gap hypothesis

The sentence can still be derived as a yes-no-question (q). For this interpretation, the wh-phrase is merged with the embedded clause marked by ‘ka’. The \textbullet I rule puts a structural \textbullet domain around the embedded interrogative and adds a \textbullet feature to the clause type. ‘Tazuneta’ selects the embedded interrogative and the main clause is derived as usual. Figure 4.3 presents the crucial parts of the derivation. (For a complete analysis, check the sample sentence in appendix B.2.3.)
4. Structural variation in wh-question formation

4.2.4 ka versus kadooaka

Both ‘ka’ and ‘kadooaka’ can be used as question markers of embedded interrogatives. The difference between the two interrogatives is that ‘ka’ can be used to mark both direct and indirect interrogatives, while ‘kadooaka’ can only be used to mark embedded interrogatives. Furthermore, ‘kadooaka’ can only mark if/whether clauses. The following sentences illustrate the behavior of the use of the interrogative marker ‘kadooaka’ in comparison with ‘ka’.

Both ‘ka’ and ‘kadooaka’ can mark an embedded if/whether clause:

(4.23) John-wa [Mary-ga hon-o katta ka/kadooaka] tazuneta
‘John asked if/whether Mary bought a book.’

‘Kadooaka’ cannot bind a direct question while ‘ka’ can:

(4.24) John-wa [Mary-ga hon-o katta to] itta ka+/kadooaka?
‘Did John say that Mary bought a book.’

3My informant indicates that ‘[verb]-ka doo-ka’ can literally be translated as ‘whether [verb] or not’. Therefore, sentences with ‘kadooaka’ become like coordinating two yes—no questions with a disjunct. Similar to other disjunctive coordination constructions, it is ungrammatical for one part of the coordination to contain a wh-in-situ.

i John-wa [Mary-ga hon-o katta ka [jisho-o katta ka]] tazuneta
ii * John-wa [Mary-ga nani-o katta ka [jisho-o katta ka]] tazuneta
‘Kadooka’ cannot host a wh-in-situ while ‘ka’ can:

       John-[top] Mary-[nom] what bought Q asked Q
‘Did John ask what Mary bought?’

‘Kadooka’ cannot mark a main clause yes-no question:

(4.26) Mary-ga hon-o katta ka/+kadooka]?
       Mary-[nom] book bought Q
‘Did Mary buy a book?’

The examples in 4.23–4.26 indicate that we need to make a distinction between the two interrogative markers ‘ka’ and ‘kadooka’. The type for ‘kadooka’ must encode that the question marker can form an embedded interrogative clause while it cannot host a wh-in-situ or form a direct question. We assign ‘kadooka’ a similar type as ‘ka’, however instead of yielding a type q, it yields a structure of type □q. Recall that □q is the highest type in the derivability pattern □q ⊢ q ⊢ □q. From a q-type clause, we can derive a □q-typed clause, but not vice versa. We need the following lexical type-assignments for question markers that mark a clause as an embedded interrogative.

\[
\begin{align*}
ka & :: s \backslash q' \\
kadooka & :: s \backslash □q' (= \text{whether/if})
\end{align*}
\]

With these type-assignments, we can derive the sentences in examples 4.23–4.26 (see appendix B.2.4). Both types of embedded interrogatives of example 4.23 marked by ‘ka’ or ‘kadooka’ can be selected by ‘tazuneta’, because of the derivability relation, q ⊢ □q. The ungrammaticality of a main clause yes-no question in example 4.26 is derived by the requirement that yes-no questions have to be of type q.

Due to the requirements of wh-phrases, a wh-in-situ cannot be merged with a structure that is hosted by ‘kadooka’ because □q’ ⊨ □q. We show, analyzing the ungrammatical sentence of example 4.27, that the derivation of a subordinate clause marked by ‘kadooka’ with an embedded wh-phrase fails, because nani :: WH(\(np, q', wh'\)) cannot merge with a structure of type □q’.

(4.27) ∗ Mary-ga nani-o katta kadooka?
       Mary-[nom] what-[acc] bought Q

We must change the selectional requirement of verbs selecting an embedded interrogative, because the type of embedded if/whether clauses can be □q. The verb ‘tazuneta’ is now assigned the following type. The type-assignments incorporates both a feature for projecting a wh-island and a selectional feature for a □q-typed clause.

\[
\text{tazuneta} :: □q' (NOM\backslash s) (= \text{asked})
\]
4. Structural variation in wh-question formation

4.2.5 Complex noun phrases

Japanese wh-questions show an interesting contrast with wh-question formation in other languages. Sentences, such as the one presented in example 4.28 (Nishigauchi, 1990, ex. 57, p. 40), in which a wh-expression appears within a relative clause are interpreted as direct questions. In other languages, such as for instance English, a fronted wh-phrase which is bound to an argument position in a relative clause, is perceived as ungrammatical.

The sentence in example 4.29 (Pesetsky, 1987, ex. 38a.) illustrates that in English 'what' cannot move out of a relative clause that is part of a larger noun phrase. Ross (1967) refers to this syntactic structure as complex NP. He explains the ungrammaticality in English in terms of a violation of the complex NP constraint (see section 3.3.4 on wh-islands).

\[ (4.28) \text{Kimi-ua} \quad [\text{dare-ga} \text{ kai-ta}] \quad \text{hon-o} \quad \text{yon-da} \quad \text{ka?} \]

‘You read books that who wrote?’

\[ (4.29) ^* \text{What, did Mary meet [NP the man [S: who gave t to John]]?} \]

In generative syntax, this contrast has been explained in terms of the ‘timing’ of the wh-movement operation. Overt movement is subject to the complex NP constraint, whereas covert movement (i.e. LF movement) is not (Huang, 1982; Lasnik and Saito, 1984). Following this line of reasoning, in languages such as Japanese where the wh-phrase only moves at LF, wh-questions where the wh-in-situ is inside a complex NP can be derived as follows. The wh-phrase moves at LF to the CP node of the matrix clause where it is governed by the question marker ‘ka’. Nishigauchi (1990) illustrates the movement with the following representation which has been slightly adapted.\footnote{Notice that in the GB analysis the wh-phrase moves to the right because Japanese is a head-final language.}
This explanation has been debated by Nishigauchi (1990) and Pesetsky (1987) who have looked more closely at the apparent non-existence of island effects in Japanese. Without going into the details, we will very briefly discuss Nishigauchi’s proposal, because his proposal is related to the approach we will propose in type-logical grammar. Nishigauchi (1990) suggests a pied-piping analysis of wh-expressions inside complex noun phrases. He suggests that in Japanese, as well as in other languages, the entire complex NP where the wh-expression is embedded in moves to matrix clause. The idea essentially is like a pied-piping operation. After the wh-phrase has been moved to the specifier position inside the complex NP, the wh-feature of the wh-phrase percolates to the entire complex NP. Now, the whole complex NP can be treated as a wh-expression. Nishigauchi (1990) illustrates his pied-piping analysis with the following LF representation.

\[(4.31) \text{[CP you t\textsubscript{2} read ][COMP ka [NP [\{t\textsubscript{1} wrote\} who\textsubscript{1} books\}2 ]]}\]

One of the arguments Nishigauchi gives for this analysis comes from analyzing felicitous answers that can be given to these wh-in-situ questions. Besides a full-fledged answer that repeats the entire complex noun phrase the following two short answers are also correct answers.

\[(4.32) \text{Kimi-wa [dare-ga kai-ta] hon-o you-da ka?} \]
\text{‘You read books that who wrote?’}

A. \text{Austen-da}
\text{‘(It’s) Austen'}

B. \text{Austen-ga kai-ta hon da.}
\text{‘(It’s) the book that Austen wrote’}

The argument that the second answer is a possible answer to the wh-question in example 4.28 supports a pied-piping analysis of wh-questions with complex NP’s (Nishigauchi, 1990, p52).

\[\Diamond \text{[complex NP’s]}\]

Along the lines of Nishigauchi (1990), we propose a pied-piping analysis of wh-questions with a wh-phrase embedded in a complex noun phrase. The pied-piping analysis is based on the pied-piping analysis of relative clause constructions in type-logical grammar (Morrill, 1994) (see explanation on pied-piping constructions as a solution for left-branch constructions in English in section 3.3.4.4). The type assigned to Japanese wh-in-situ in a complex NP is a direct extension to the type-assignment proposed for relative pronouns.
Relative clauses  Before we present the analysis for interpreting an embedded wh-phrase at the main clause level, we show how we analyze a relative clause construction in Japanese. Different from English, Japanese has no overt complementizer to mark a restrictive relative clause. Along with the characteristic of Japanese of being head final language, the relative clause precedes the head noun that is modified. The following example illustrates the restrictive relative clause.

(4.33) [[mary-ga kai-ta ] hon]

mary-[nom] write book

“book that Mary wrote”

In English, the relativizer has the function of associating a clause with the noun that is being modified. In a type-logical grammar analysis, a relativizer such as ‘which’ or ‘that’ is typed as \((n\downarrow n)/(np\downarrow s)\). With this type we can derive the complex NP ‘book that Austen wrote’. To derive the complex NP in example 4.33, where ‘hon’ is modified by the relative clause formed by [‘dare-ga kai-ta’], we incorporate the selectional requirements of the relativizer into the type assigned to the noun. Therefore, we allow nouns, such as ‘hon’, to be lexically ambiguous and assign it both a basic type \(np\), and a type that is modified by a relative clause construction. With the following type-assignment for ‘hon’, we can derive the relative construction of example 4.33:

\[
\text{hon} : (\text{case}\downarrow s)\downarrow np \text{ where case} \in \{\text{NOM, ACC, DAT, TOP}\}
\]

Complex noun phrases  Now that we know how to construct relative clauses, we can discuss in more detail how the wh-in-situ in complex NPs can be derived. We repeat here example 4.28 that illustrates the phenomenon.

(4.34) Kimi-wa [[dare-ga kai-ta] hon]-o yon-da ka?


‘You read books that who wrote?’

Two observations can be made: First, the complex \(np\) as a whole is interpreted as a wh-expression and needs to be in the domain of the question marker ‘ka’.
4.2. Wh-in-situ

Secondly, the wh-phrase embedded in the complex np must also take scope outside the domain of the noun phrase. Both observations relate to Nishigau-

chi’s pied-piping analysis (Nishigau-

chi, 1990), namely that the wh-feature of the wh-phrase percolates to the complex NP. The second observation captures the idea that the wh-phrase governs the local np domain in which it is em-

bedded. We look at both observations in isolation and show how these are captured in a type-logical grammar analysis.

To capture the first idea, the complex np needs to be of the same type as a normal wh-in-situ type:

\[
\text{dare-ga kaita hon} \vdash \text{WH}_{int}(np, q, wh) \quad (= \text{book that who wrote})
\]

To simplify structures in later derivations, we abbreviate expressions that belong to the type \(\text{WH}_{int}(np, q, wh)\) to \(\text{whin}\). On the basis of the second obser-

vation, we base the type-assignment of the wh-in-situ phrase ‘dare’. We assign to the wh-phrase a wh-in-situ type that ranges over np-arguments inside a complex np and yields a type which is itself a wh-in-situ type. The following type-assignment matches this description:

\[
dare \vdash \text{WH}_{int}(np, np, \text{WH}_{int}(np, q, wh)) \quad (= \text{abbreviated to } \text{WH}_{int}(np, np, \text{whin}))
\]

Using the type-assignment for wh-phrases we can derive wh-question in example 4.34. To understand the analysis, we will present the derivation of the complex noun phrase ‘dare-ga kaita hon’ (= book that who wrote) which will be of a wh-in-situ type. Then, we will merge that structure in abbreviated form (‘whin’) with the main clause.

\[
\begin{align*}
\text{[np]} & \vdash \text{whin} \\
\text{np} \circ \text{ga} & \vdash \text{nom} \\
\text{ACC} \circ \text{kaita} & \vdash \text{nom} \circ s \\
\text{ACC} \circ (\text{np} \circ \text{ga}) & \vdash \text{np} \circ s \\
\text{ACC} \circ (\text{np} \circ \text{ga}) \circ \text{kaita} & \vdash \text{ACC} \circ s \\
\text{ACC} \circ (\text{np} \circ \text{ga}) \circ \text{kaita} \circ \text{hon} & \vdash \text{np} \\
\text{dare} & \vdash \text{WH}_{int}(np, np, \text{whin}) \\

\langle \omega, \lambda z. ((\text{write book}) z) \rangle
\end{align*}
\]

On the basis of this derivation, we can now merge this complex wh-phrase with the main clause. The following derivation presents the analysis of the matrix wh-question ‘Kimi-wa [dare-ga kaita hon] whin-o yonda ka’ (= You read the
[book that who wrote?] where whin stands for an abbreviation of the complex NP ‘dare-ga kaita hon’ (= who wrote book).

\[
\begin{array}{c}
\text{kimi} \quad \text{wa} \\
\text{TOP} \\
\hline
\text{TOP} \\
\hline
\text{TOP} \\
\hline
\text{TOP} \\
\hline
\text{WH} \quad \text{in}
\end{array}
\]

We will come back to complex noun phrases and treat the semantic interpretation of complex NP constructions in chapter 5, section 5.3.3.

### 4.2.6 Multiple wh-questions

In multiple wh-questions, similar to single constituent questions, wh-phrases receive scope in the clause which is marked by the closest question marker. Thus if two wh-phrases appear in the same clause, the clause is interpreted as a multiple wh-question as illustrated in example 4.35a. If one wh-phrase occurs in the main clause and one wh-phrase is embedded in the subordinate clause where both clauses are marked by ‘ka’, the whole clause is interpreted as a main clause wh-question with an embedded interrogative. The sentences are based on examples presented in Kim (1989).

(4.35) a. John-wa [dare-ga nani-o kiratteiru ka] tazuneta?
    John-[top] who-[nom] what hate Q asked
    ‘John asked who hated what’

b. John-wa dare-ni [Mary-ga nani-o kiratteiru ka]
    tazuneta ka?
    asked Q
    ‘(To) Whom did John ask what Mary hated?’

It is also possible that two wh-phrases are bound to a question marker in the main clause, while one wh-phrase resides in the matrix clause and one wh-phrase appears embedded. Example 4.36 illustrates such a construction. The interpretation of this wh-question can be understood by inspecting possible answers to such a question in comparison to the wh-question in 4.37 where
4.2. Wh-in-situ

the embedded clause has an additional question marker. According to our informant possible answers to the question in example 4.36 are pair-list answers where both wh-phrases need to be filled in. In contrast possible answers to the question in example 4.37 only answer the wh-phrase in the main clause.

(4.36) John-wa dare-ni [Mary-ga nani-o kiratteiru to] itta
Q

'Who did John say/tell what Mary hated?'

Answer: John told Bill that Mary hated cats.

(4.37) John-wa dare-ni [Mary-ga nani-o kiratteiru ka] tazuneta
Q

'Who did John ask what Mary hates?'

Answer: 'John asked Bill what Mary hates.'

♦ (Multiple wh-questions)
Along the same line of the analysis of multiple wh-questions in English (section 3.3.5), we can account for multiple wh-questions as given in example 4.35a and 4.36 by extending the type-assignments to wh-phrases. We need to add a type-assignment in order to merge a wh-phrase with a question body of type wh which is a sentence that already contains another wh-phrase (see appendix B.2.8 for on-line derivations).

multiple wh-phrase :: WH((np, wh, wh))

The sentences in examples 4.35b and 4.37 are derived on the basis of the presented selectional requirements of the verbs, question markers and wh-phrases.

4.2.7 Chinese

Like Japanese, Chinese is a wh-in-situ language. As opposed to Japanese, the basic word order in the Chinese clause is head-initial (Huang, 1998). Another difference with Japanese concerning wh-question formation is that wh-questions do not need an overt question marker. Wh-words in Chinese are for instance, 'shei' (= who) and 'shenme' (= what). The same words that are used for wh-phrases are used for universal quantifier phrases. For instance,
'shen-me' can mean either ‘what’ or ‘anything’ or both depending on the context. Whether a sentence is interpreted as a wh-question is determined in the interaction with other words in the sentence. An ambiguous interpretation is possible in a sentence where ‘shenme’ occurs within the domain of ‘bu’ (= not) (Huang, 1998, 168).

(4.38) ta bu xiang chi shenme?/
he not want eat what/anything

1. ’What didn’t he want to eat?’
2. ’He didn’t want to eat anything.’

As Huang (1998) explains, the grammar determines the use of the word. We concentrate on those sentences that yield a question interpretation as illustrated in the following examples (Huang, 1998, ex.163–165, p.178).

(4.39) a. Zhangsan wen wo [shei mai-le shu].
ask I who buy[asp] book

’Zhangsan asked me who bought books.’
b. Zhangsan xiangxin [shei mai-le shu]?
believe who buy[asp] book

’Who does Zhangsan believe bought books?’
c. Zhangsan zhidao [shei mai-le shu] ?/
know who buy[asp] book

1. ’Who does Zhangsan know bought books?’
2. ’Zhangsan knows who bought books.’

The three sentences in example 4.39 illustrate that the interpretation is largely dependent on the type of matrix verb. In example 4.39a) where the matrix verb is one that selects an interrogative complement, the wh-phrase can only be interpreted locally as an embedded interrogative. Example 4.39b) with the matrix verb ‘xiangxin’ (= believe), the sentence is interpreted as a non-local wh-question. Finally, example 4.39c) illustrates a verb that can optionally select an interrogative complement. The sentence can be interpreted both as a direct and an indirect question.

We treat Chinese wh-question formation on a par with the analysis of wh-question formation in Japanese. Thus, Chinese wh-phrases will be typed using the wh-in-situ type schema: \( WH_{in}(A, B, C) \). For Japanese, we have shown that the dependency of the wh-phrase on the occurrence of a question marker

\(^5\)Huang (1998) indicates that only in a negative context wh words are interpreted as universal quantifiers. In such contexts the wh words behave like negative polarity items.
is incorporated in the type-assignment of wh-phrases: \( WH_{in}(np, q, wh) \). However, in Chinese, the wh-phrase is not dependent on the availability of a question marker and we propose assigning the following types to argument wh-phrases such as ‘shei’ (= who) and ‘shenme’ (= what).

- **main clause interogatives** :: \( WH_{in}(np, s, wh) \)
- **embedded interogatives** :: \( WH_{in}(np, s', wh') \)

A wh-phrase is merged with a question body of type \( s \) which contains a gap hypothesis of type \( np \). After merging, the wh-phrase replaces the gap hypothesis and yields either a main clause of type \( wh \) or a subordinate clause of type \( wh \). The choice for a main clause type or an embedded clause type is determined by selectional requirements of the main clause verbs. For instance, the interpretation of the wh-phrases in the sentences in example 4.39 are solely determined by the selectional requirements of the matrix verbs. We assign ‘wen’ (= ask) and ‘xiangxin’ (= believe) the same type-assignments as the Japanese counterparts. (The verb ‘zhidao’ (= know) which optionally merges with an embedded interrogative or an embedded declarative clause, gets two type-assignments.)

\[
\begin{align*}
\text{wen} &:: IV/wh' (= ask) \\
\text{xiangxin} &:: IV/s' (= believe) \\
\text{zhidao} &:: IV/A (= know) \\
(\text{where } A \in \{wh', s'\})
\end{align*}
\]

With these type-assignments we can account for the different types of questions. The analysis is similar to the analysis of direct and indirect wh-questions in Japanese in section 4.2.2 and 4.2.3, with the difference that there is no question marker. Wh-phrases in Chinese are therefore directly merged with a question body which is a \( s \)-typed or \( s' \)-typed clause instead of a \( q \)-typed question body.

The empirical data on and research related to wh-question formation in Chinese is abundant (Aoun and Li, 1993; Cheng, 1991). Our analysis would surely need further refinement to cover many of the aspects that are known about Chinese wh-question formation. Further research is also needed on the interpretational differences through the interaction of wh-expressions with other expressions in the sentence. For instance, the universal interpretation of wh-expressions in negative contexts, scope alternations between wh-expressions and quantified phrases and multiple wh-questions will help to further develop the analysis of both quantifier phrases and wh-question formation in type-logical grammar. We leave these topics for future research and concentrate on the application of the wh-type schema to account for wh-question formation across languages.
4.2.8 Concluding remarks on wh-in-situ languages

We have presented examples and data from Japanese as an illustration of a wh-in-situ language. The wh-questions in Japanese are analyzed using a wh-in-situ operator type for wh-phrases that only merges with structures that are marked by ‘ka’. The scoping postulates that realize the right interpretation of wh-in-situ elements follow the restricted set of displacement postulates as presented in chapter 3. Additional flexibility in argument order is also captured by the restricted set of displacement postulates. Constraints on the possible interpretations of wh-phrases or the different scope readings a wh-phrase may get, are due to selectional requirements of the matrix verbs. The wide-scope reading of wh-phrases that occur embedded inside a complex np is given a pied-piping account.

4.3 Wh-ex-situ: Bulgarian and Serbo-Croatian

Slavic languages belong to the group of languages where all wh-words are moved to a clause-initial position. These languages share the phenomenon that in sentences which contain multiple wh-words, all wh-phrases appear at the front of the structure. This phenomenon has been studied by several linguists in the generative syntax tradition (Comorovski, 1986; Rudin, 1986, 1988; Bošković, 1997, 1998, e.g.). These analyses either follow the principles laid out in the Government and Binding tradition or the latest developments of the Minimalist Program. In general, the explanations for the extraction phenomena extend early proposals of Chomsky (1973, 1977) which set the stage for a general theory of wh-movement.

We will very briefly sketch the generative syntactic literature on wh-question formation in Slavic languages. Rudin (1988) presents an explanation for the difference between Serbo-Croatian and Bulgarian ordering of wh-phrases. She presents empirical evidence that in Bulgarian the wh-phrases form one constituent which therefore occupies one specifier position, whereas in other Slavic languages (e.g. Serbo-Croatian) the two wh-phrases are separate constituents. Additionally, in Serbo-Croatian wh-phrases are not subject to ordering constraints which results in a free word order of the wh-phrases. Mainly these two points of her analysis have been much discussed in papers that succeeded Rudin’s article (Lambova, 2000; Bošković, 1998, 1997). They have presented new data and analyses to show that there is counter evidence to the analysis that in Bulgarian the wh-words form a constituent. Additionally, Bošković (1997) has provided evidence that wh-phrases in Serbo-Croatian are subject to linear ordering constraints, i.e. the superiority condition.

For our analysis in type-logical grammar, we first lay out the basic underlying word order for both languages. Secondly we will extrapolate the basic analysis proposed for wh-phrases to multiple wh-fronting phenomena. Subsequently, we will present some related phenomena that are argued by generative linguists to provide evidence for locality constraints in Bulgarian and
4.3. Basic word order

The neutral word order of both Bulgarian and Serbo-Croatian is Subject-Verb-Object (SVO).

(4.40) Petar kupuje knjigu. [SC]
Peter-[Nom] buys book-[Acc]
‘Peter buys a book.’

(4.41) Ivan otvori vratata. [B]
Ivan opened door-[the]
‘Ivan opened the door.’

It should be noted that the argument order in Bulgarian and Serbo-Croatian is relatively free. For instance, the following orderings of the sentence in example 4.41 are also permitted. The different permutations are related to shifts in focus.

(4.42)

a. Ivan vratata otvori.
b. vratata Ivan otvori.
c. vratata otvori Ivan.

For an SVO order, we assume the following basic lexical entries with abbreviated type assignments for both languages. The basic np-types are decorated with case features similar to Japanese (Heylen, 1999).

\[
\begin{align*}
\text{NOM} &:: \Diamond \Box_{\text{nom}} np \ (= \text{‘Ivan’, ‘Petar’ [SC,B]}) \\
\text{ACC} &:: \Diamond \Box_{\text{acc}} np \ (= \text{‘knjigu’ [SC], ‘vratata’ [B]}) \\
\text{DAT} &:: \Diamond \Box_{\text{dat}} np \\
\text{IV} &:: \text{NOM} \backslash s \\
\text{TV} &:: (\text{NOM} \backslash s) / \text{ACC} \ (= \text{‘kupuje’ [SC], ‘otvori’ [B]})
\end{align*}
\]

4.3.2 Single wh-questions

Both Bulgarian and Serbo-Croatian are wh-fronting languages. In a wh-interrogative clause which questions only a single argument position the wh-phrase will always appear at the fronted position, independently of the argument role of the wh-phrase. The following examples illustrate single wh-questions in both Serbo-Croatian (SC) and Bulgarian (B).

(4.43) Single wh-question with a subject wh-phrase:
4. Structural variation in wh-question formation

(a) Ko kupuje knjigu? [SC]
   Who-[nom] buys book-[acc]
   ‘Who buys a book?’
(b) Koj otvori vratata? [B]
   Who opened door-the
   ‘Who opened the door?’

(4.44) Single wh-question with an object wh-phrase:

(a) Koga Ivan vidi? [SC]
   whom Ivan-[nom] sees
   ‘Whom did Ivan see?’
(b) Kogo si vidjal? [B]
   whom be-[2sg] see-[2sg]
   ‘Who did you see?’

[Wh-ex-situ phrases]

In section 4.3.1, we have noted that the order in which argument phrases can occur is relatively free, but the basic word order is SVO. A wh-phrase in a single wh-question always occurs fronted. The lexical type-assignment to wh-phrases have to encode the fronted occurrence of the wh-phrase. The interpretation of wh-questions as abstraction over a queried argument position hints at a different analysis with higher-order type assignments.

Similar to wh-phrases in English single constituent questions, wh-phrases have to be assigned an ex-situ wh-type schema. As the underlying word order is SVO, the gap appears either on a left branch for subject wh-phrases or on a right branch for non-subject argument wh-phrases. Wh-phrases in Serbo-Croatian are ex-situ wh-phrases which are lexically typed with the following general type-schema.

subject wh-element :: WH₁^l(◊□NOM, s, wh)
non-subject wh-elements :: WH₂^l(◊□CASE, s, wh)
   (where CASE ∈ \{ACC, DAT, . . . \})

With these type-assignments, we can account for the displacement phenomena presented in the previous paragraph.

Wh-displacement We restrict the analysis of single constituent wh-questions to Serbo-Croatian. We present a step-by-step analysis of the sentence in example 4.44a. The sentence is derived bottom-up, starting from lexical entries, and merging the entries until we obtain the final structure. The wh-phrase will only be inserted at the very end of the derivation. Before the wh-element can be inserted, the verb needs to check its selectional requirements. The verb requires an object np on the right-hand side, and a subject np
on the left. As there is no object np available, because the wh-phrase abstracts over the object position, the computational system assumes a hypothesis. This hypothesis is a basic np-type decorated with unary features ♦□ (which could be seen as special Q-features).6

In order to discard the hypothesized element we need to move the hypothesis to the right edge of the structure. To do so, we must use the restructuring postulates to rearrange the order of elements. We repeat the restricted set of displacement postulates as presented in section 3.2.

\[
\begin{align*}
\Diamond A \cdot (B \cdot C) \vdash (\Diamond A \cdot B) \cdot C & \quad [P11] \\
\Diamond B \cdot (A \cdot C) \vdash A \cdot (\Diamond B \cdot C) & \quad [P12] \\
(A \cdot C) \cdot \Diamond B \vdash (A \cdot \Diamond B) \cdot C & \quad [P2]
\end{align*}
\]

The following natural deduction presentation illustrates part of the derivation of the wh-question ‘koga ived’ where the underlying sentential clause has been built. To move the hypothesized object phrase to the edge of a structure, we apply the right displacement rule [Pr1], because the gap hypothesis appears on a right branch. The resulting structure is a sentential clause with the hypothesized accusative phrase on the right edge of the structure.

\[
\begin{align*}
\text{ivan} & \quad \text{NOM} \\
\text{vidi} & \quad \text{NOM}\setminus s \quad \text{ACC} \\
\text{ivan} \circ (\text{vidi} \circ \Diamond \text{ACC}) & \quad s \\
(\text{ivan} \circ \text{vidi}) \circ \Diamond \text{ACC} & \quad s \\
\vdash \text{ACC} & \quad [P1]
\end{align*}
\]

\(\text{(see x) iven}\)

**Wh-phrase insertion** At this point we can discard the hypothesis and merge the wh-phrase to the structure. The type-assignment of the object wh-phrase \(\text{WH}^C_r(\Diamond \text{ACC}, s, \text{wh})\) requires that it is merged with a question body of type \(s\) which contains a hypothesized element of type \(\Diamond \text{ACC}\) on a right edge. The structure \((\text{ivan} \circ \text{vidi}) \circ \Diamond \text{ACC}\) exactly meets these requirements. The last step of the derivation is that the wh-phrase is merged with the structure with the logical elimination of the WH-operator assigned to ‘koga’.

\[
\begin{align*}
\Diamond \text{ACC} & \vdash \text{ACC} \\
\text{koga} & \quad \vdash \text{wh} \\
\text{WH}^C_r(\Diamond \text{ACC}, s, \text{wh}) & \quad (\text{ivan} \circ \text{vidi}) \circ \Diamond \text{ACC} \vdash s & \quad [\text{WH}^C_r]
\end{align*}
\]

\(\omega \lambda x((\text{see } x) \text{ iven})\)

Notice that the computed meaning assembly shows that the wh-phrase binds the semantic term variable of the hypothesized object gap.

---

6‘Q-features’ is the terminology that is used by syntacticians to distinguish the features that are responsible for the movement of the wh-elements (cf. Chomsky (1995)).
4.3.3 Multiple wh-questions

In Bulgarian and Serbo-Croatian, as well as other Slavic languages, wh-phrases appear fronted in the matrix clause in single wh-questions. In multiple wh-questions, all wh-phrases appear in fronted position. Bulgarian differs from Serbo-Croatian with regards to the possible ordering of the wh-phrases, and Bulgarian shows additional constraints when the interrogative clause contains clitics. First, we present some data to illustrate the difference between the two languages for multiple wh-questions with two wh-phrases. Then, we show how we can account for the possible orderings and the additional restrictions in Bulgarian clitic environments on the basis of a type difference for wh-phrases in Bulgarian.

Most of the data we present here are examples taken from Rudin (1986, 1988) and Bošković (1998), who have analyzed these constructions in a generative grammar framework. The following examples show the subject-object ordering of wh-elements in wh-questions with multiple wh-elements in the two languages.

(4.45) Basic setting

\[
\begin{align*}
Ko & \quad koga \quad vidi? \quad [SC] \\
Koj & \quad kogo \quad vižda? \quad [B] \\
\text{who} & \quad \text{whom} \quad \text{sees} \\
\text{Who sees whom?}
\end{align*}
\]

The word order where the object wh-phrase precedes the subject wh-phrase is allowed in Serbo-Croatian, whereas it is ungrammatical in Bulgarian. In Serbo-Croatian the order of the fronted wh-expressions is relatively free. In a main clause, any ordering of the wh-phrases seems to be possible. The following data shows the contrast between Serbo-Croatian and Bulgarian. The data is taken from Bošković (1998) who compares wh-question formation in Serbo-Croatian with wh-question formation in Bulgarian.

(4.46) Koga \quad ko \quad vidi? \quad [SC]

\[
\text{whom} \quad \text{who} \quad \text{sees} \\
\text{Who sees whom?}
\]

(4.47)* Kogo \quad kaj \quad vižda? \quad [B]

\[
\text{whom} \quad \text{who} \quad \text{sees} \\
\text{Who sees whom?}
\]

In a wh-question with multiple wh--phrases, the body of the question of a wh-phrase that is merged with a question which already contains a wh-phrase
4.3. Wh-ex-situ: Bulgarian and Serbo-Croatian

is of type \textit{wh}.

To specify for \textit{wh}-phrases that they may occur in multiple \textit{wh}-questions, we need the following type-assignments for subject and object argument \textit{wh}-phrases. Notice that the goal type, e.g. \textit{wh}, and the position of the hypothesized elements for both subject and non-subject \textit{wh}-phrases stays the same.

\begin{align*}
\text{subject } \textit{wh-element} & \ := \ WH_{\text{lu}}(\Diamond \Box \text{NOM}, \textit{wh}, \textit{wh}) \\
\text{non-subject } \textit{wh-elements} & \ := \ WH_{\text{ru}}(\Diamond \Box \text{CASE}, \textit{wh}, \textit{wh})
\end{align*}

(\text{where } \text{CASE} \in \{\text{ACC}, \text{DAT}, \ldots\})

Using these type-assignments for \textit{wh}-phrases in addition to the basic type-assignments presented in section 4.3.2, we can analyze the two possible orderings of sentences with two \textit{wh}-phrases.

\begin{align*}
(4.48) \quad & ko \ koga \ vidi \vdash \textit{wh} \\
& (= \text{Who sees whom?}) \\
\hline
(4.49) \quad & koga \ ko \ vidi \vdash \textit{wh} \\
& (= \text{Who was seen by whom?})
\end{align*}

As an example we derive the \textit{wh}-question ‘ko koga vidi?’ (= who whom sees):

\begin{align*}
\Diamond \quad & \vdash \textit{wh} \\
\hline
\vdash \textit{wh} \quad & \vdash \textit{wh}
\end{align*}

\begin{align*}
& (\omega \lambda y. (\omega \lambda x. ((\text{see} \ x) \ y)))
\end{align*}

\begin{itemize}
\item \textit{multiple \textit{wh}-questions in Bulgarian}
\end{itemize}

As example 4.47 illustrates, Bulgarian has a more restricted ordering with respect to the nominative and the accusative \textit{wh}-element than Serbo-Croatian. In line with Serbo-Croatian, we want the \textit{wh}-type schema to apply to structures which are either typed as \textit{s} or as \textit{wh}. So we do not want to rule out the order ‘kogo koz dia’ by disallowing ‘kogo’ to apply to a structure which already contains a \textit{wh}-phrase. In Bulgarian, a \textit{wh}-question like the one given in the following example is also a grammatical sentence.

\footnote{In chapter 5, the \textit{wh}-type schema will be changed due to the syntactic and semantic decomposition of the \textit{wh}-type \textit{wh}. Multiple \textit{wh}-questions will be derived from a single type schema and get an appropriate semantic interpretation due to the decomposition.}
4. Structural variation in wh-question formation

(4.50) Kogo kakvo e pital Ivan?

who what AUX ask Ivan

‘Who(m) did Ivan ask what?’

To account for the correct ordering of the wh-phrases in Bulgarian, we must be more specific about the type specification of the wh-type schema assigned to wh-phrases. By adding feature requirements to the question body and adding features to the question type, the ordering between wh-phrases can be accounted for on the basis of the derivability pattern of feature decorated type formulas (see 2.5.3 on the use of the unary connectives).

**Definition 24 (Derivability pattern with ♦ and □ features)** For arbitrary type formulas $A$, we have the following derivability pattern:

$$\Diamond \Box A \vdash A \vdash \Box \Diamond A$$

In Bulgarian, *koj* precedes the accusative wh-phrase, *kogo*. To derive the ordering relations between these two wh-phrases, *koj* needs a type-assignment such that it can be combined either with a structure of type $s$ (for single wh-questions) or with a structure which already contains a wh-phrase. After we have inserted ‘*koj*’ we do not want any other wh-phrase to be merged with the structure, because no other wh-phrase may precede ‘*koj*’. In order to enforce this, we need to assign such wh-questions the highest type in the derivability pattern: $\Box \Diamond wh$. A wh-phrase such as ‘*kogo*’ may therefore appear in single wh-questions, but also in multiple wh-questions. ‘*kogo*’ is assigned a type which restricts the application of ‘*kogo*’ to structures of a type equal or lower in the hierarchy than $wh$, because it may not be merged with a structure which already contains the wh-phrase ‘*koj*’.

We assign ‘*koj*’ and ‘*kogo*’ the following types to account for single and multiple wh-questions.

<table>
<thead>
<tr>
<th>Single wh-questions:</th>
<th>Multiple wh-questions:</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>koj</em> :: WH$^l_x(\Diamond \Box \Diamond NOM, s, \Box \Diamond wh)$</td>
<td><em>koj</em> :: WH$^l_x(\Diamond \Box \Diamond NOM, \Box \Diamond wh, \Box \Diamond wh)$</td>
</tr>
<tr>
<td><em>kogo</em> :: WH$^l_x(\Diamond \Box \Diamond ACC, s, \Box \Diamond wh)$</td>
<td><em>kogo</em> :: WH$^l_x(\Diamond \Box \Diamond ACC, \Box wh, \Box \Diamond wh)$</td>
</tr>
</tbody>
</table>

The type for wh-questions in Bulgarian is now the feature decorated $\Box \Diamond wh$. With the above type-assignments we can analyze the following wh-questions in Bulgarian (see appendix B.4.1 for on-line derivation).

(4.51) Koj vizda coveka $\vdash \Box \Diamond wh$

(= Who sees the man?)

(4.52) Kogo vizda boris $\vdash \Box \Diamond wh$

(= Who does Boris see?)

(4.53) Koj kogo vizda $\vdash \Box \Diamond wh$

(= Who sees whom?)

(4.54) * Kogo koj vizda $\not\vdash \Box \Diamond wh$

(= Who does whom?)

To understand why the last example is underviable, we will look at the crucial part in the derivation. This is the point where *koj* has already been merged
4.3. Wh-ex-situ: Bulgarian and Serbo-Croatian

with the question body while the structure still contains a gap hypothesis of type \(\Diamond \Box \text{ACC}\). The question body that is formed, ‘\(((\text{k}oj \circ \text{viz}da) \circ \Diamond \Box \text{ACC})\)’, is a structure of type \(\Box \Diamond wh\) which is the highest type in our derivability pattern. However, ‘\(\text{k}oga\)’ has the type requirement to merge with a question body of type \(wh\). Type \(\Box \Diamond wh\) cannot be lowered to type \(wh\) which is the type required by ‘\(\text{k}ogo\)’. Therefore, the derivation fails.

\[
\begin{align*}
\text{koj} & \vdash \text{WH}_{\Box}(\Diamond \Box \text{NOM}, s, \Box \Diamond wh) \\
\text{koj} \circ (\text{viz}da \circ \Diamond \Box \text{ACC}) & \vdash \Box \Diamond wh \\
(\text{k}oj \circ \text{viz}da) \circ \Diamond \Box \text{ACC} & \vdash \Box \Diamond wh \\
\vdash \text{k}ogo \vdash \text{WH}_{\Box}(\Diamond \Box \text{ACC}, wh, \Box \Diamond wh) \\text{[Pr1]} \\
\vdash \text{k}ogo \vdash \text{WH}_{\Box}(\Diamond \Box \text{ACC}, wh, \Box \Diamond wh) \\text{[WH}_{\Box}]) \\
\vdash \text{k}ogo \vdash \text{WH}_{\Box}(\Diamond \Box \text{ACC}, wh, \Box \Diamond wh) \\text{[WH}_{\Box}]) \\
\vdash \text{k}ogo \vdash \text{WH}_{\Box}(\Diamond \Box \text{ACC}, wh, \Box \Diamond wh) \\text{[WH}_{\Box}])
\end{align*}
\]

Figure 4.5: Underivability of ‘\(\text{k}ogo \text{ koj viz}da\)’

4.3.3.1 More than two wh-phrases

Serbo-Croatian While Bulgarian imposes constraints on the linear ordering of the wh-phrases in multiple wh-phrases, the order of wh-elements is completely free in Serbo-Croatian. All possible orders with three wh-words have been judged grammatical. The examples are slightly altered sentences, based on the examples provided by (Rudin, 1988, 473; ex.56a-f).

\[(4.55)\]

\[\begin{align*}
a. & \text{ Ko } \text{\v{c}ta } \text{kome kupuje?} \\
& \text{ who } \text{what } \text{to whom} \text{ buys} \\
& \text{‘Who is buying what to whom?’} \\
b. & \text{ Ko kome \v{c}ta kupuje?} \\
c. & \text{\v{c}ta ko kome kupuje?} \\
d. & \text{\v{c}ta kome ko kupuje?} \\
e. & \text{Kome ko \v{c}ta kupuje?} \\
f. & \text{Kome \v{c}ta ko kupuje?}
\end{align*}\]

The different orders are interpreted with some shift in focus. The first sentence in the examples is the most neutral order, the others have a different focus interpretation.

\[\Diamond \text{[> 2 wh-phrases in Serbo-Croatian]}\]

The ordering in which the wh-elements can occur in Serbo-Croatian follows directly from the type-assignments we have assigned to the lexical entries on
As the wh-movement postulates posit no restrictions on which wh-element has to be moved first, we can move the hypothesized wh-elements in any order to the left or right edge of the structure where they can be extracted. All of the sentences listed in example 4.55 can be derived. As an example, we derive structure 4.55c) ‘stants kome kupuje?’.

First, the hypothesized element with dative case moves to the edge at which point the wh-element ‘kome’ may be inserted. Secondly, the nominative wh-hypothesis moves to the left after which the nominative wh-phrase is merged. Finally, the accusative wh-hypothesis is displaced and ‘sta’ is merged with the structure. Because of the meaning assembly of merging a wh-type schema to a structure, the order in which the wh-phrases are merged is reflected in the proof term.

Bulgarian

As we have seen in the examples in section 4.3.3, the order of wh-phrases in wh-questions with two wh-phrases is restricted. In wh-questions with more than two wh-phrases the linear order again seems strict with respect to the first element, the subject wh-phrase. However, the ordering of the other wh-phrases, ‘kogo’ and ‘kakvo’, that follow the fronted wh-phrase is relatively free, as the following examples illustrate (Bošković, 1998, ex.34; p.13).

(4.56) a. Koj (na) kogo kakvo kupuva?
    who (to) whom what buys
    ‘Who buys what for whom?’

b. Koj kakvo (na) kogo kupuva?
These examples lead to the conclusion that ‘kogo’ and ‘kakvo’ can reorder freely. However, Bošković (1998) points out that ‘kogo’ occurs higher than ‘kakvo’ in sentences without a nominative wh-phrase, as the following sentences illustrate (based on example 33 on page 12 of Bošković (1998)).

(4.57)  
(a) (na) Kogo kakvo Ivan kupuva?  
(to) whom what Ivan buys  
‘For whom does Ivan buy what?’
(b) *Kakvo (na) kogo Ivan kupuva?

As the examples above show, there is a strict order between nominative and accusative wh-elements, and accusative and dative, respectively. As explained in section 4.3.3, p. 117, the order between the occurrences of wh-phrases can be accounted for again by using the derivability patterns between types. We must alter the type-assignments of the wh-phrases in the following way. The types in the wh-type schema that identify the question body or the goal type of the wh-question are decorated with unary features to capture the linear ordering among wh-phrases.

koj :: WH_{ex}^{r} (♦2NOM, □♦wh, □♦wh)
kogo :: WH_{ex}^{r} (♦2DAT, wh, □♦wh)
kakvo :: WH_{ex}^{r} (♦2ACC, wh, wh)

Moreover, the following lexical entry is assigned to the verb ‘kupuva’ (= buys).

kupuva :: (NOM\s/DAT)/ACC (=’buys’)

Based on the given lexical type-assignments, the wh-questions with two and three wh-phrases of example 4.56 and 4.57 can be analyzed (see appendix B.4.1). We account for the fact that koj always precedes kogo and kakvo. The resulting type of koj is □♦wh, and can not be selected by kakvo because ‘kakvo’ needs to be merged with a structure of type wh and □♦wh ⊬ wh. It also immediately follows that kogo occurs higher than kakvo. We could make a linear ordering between the wh-phrases such that the orderings koj ≺ kogo, koj ≺ kakvo, and kogo ≺ kakvo’, but the ordering ‘kogo ≺ koj’ or ‘kakvo ≺ kogo’ can not be derived.

Only the wh-question ‘Koj kakvo kogo kupuva?’ where the order between ‘kakvo’ and ‘kogo’ is changed is undervivable. The reason why this structure is undervivable is because the derivability pattern of wh-types (♦□wh ⊬ wh ⊬ □♦wh) captures a strict ordering among the wh-phrases. Following the arguments presented by Bošković (1998) on similar data, there might be an independent reason why the ordering of kakvo and kogo can be reversed when they follow the subject wh-phrase. Bošković (1998) argues that the word order of the arguments in Bulgarian can be changed when the focus of the sentence is
varied. In sentences with multiple wh-phrases it seems that only the first wh-
phrase is moved into that position for reasons of wh-movement. The other
wh-phrases that follow the fronted wh-phrase are under influence of focus
movement. In this case the order between the dative wh-phrase and the accu-
sative wh-phrase is not strict and does not depend on the derivability pattern
of s-types.

4.3.3.2 Non-local wh-questions

So far, we have discussed multiple wh-questions in a local domain. We have
presented data that show that in Bulgarian there is a strict order among the
wh-phrases while in Serbo-Croatian the ordering is completely free. Bošković
(1998) argues that local wh-questions provide evidence that Bulgarian is sub-
ject to the superiority condition\(^9\) (Chomsky, 1973), while Serbo-Croatian is not. The strict ordering versus free ordering also extends to embedded wh-
questions in Bulgarian (Rudin, 1986) and Serbo-Croatian (Bošković, 1997).

Any variation in the order of wh-phrases may be caused by other interfering
factors, such as focus movement.

Next we will consider whether the ordering constraints also apply to non-
local wh-questions.

Bulgarian In Bulgarian, long-distance displacement with multiple wh-
phrases shows that all wh-phrases appear clause-initially and follow the same
ordering constraints as for local wh-questions (Rudin, 1986, 1988; Richards,
1997). Example 4.58a) illustrates that in non-local wh-questions all wh-
phrases must appear clause-initially. Example 4.58b) shows that a sentence is
ungrammatical when one of the wh-phrases appears in the embedded clause
(Rudin, 1988, p. 450,ex. 6).

(4.58) a. Koj \(\text{i} \ kogo \ j \ misliš \ \ [če \ t_j \ vžda \ t_j \ ]? \)
who whom think-[2s] that sees
‘Who do you think (that) sees whom?’

b. * Koj \(\text{i} \ mišliš \ \ [če \ t_j \ vžda \ kogo \ ]? \)
who think-[2s] that sees whom

Serbo-Croatian In Serbo-Croatian, long-distance extraction seems to follow
the same strategy as local wh-questions. Rudin (1988, p.453; ex.11a,b) presents
the following data on long-distance extraction where the matrix verb is ‘želite’

\(^9\)A simplified description of the superiority condition is defined in terms of c-command rela-
tion: the trace of a fronted wh-phrase must c-command the trace(s) of the wh-phrases that follow
the fronted wh-phrases. The superiority condition as applied to wh-questions is captured in gen-
erative syntax by the following schematic presentation: \([\text{wh,} \ldots \text{wh}, \ [\ldots t_i \ldots t_f] \ldots]\)
4.3. Wh-ex-situ: Bulgarian and Serbo-Croatian

(= want). According to Rudin, either the subject wh-phrase or the object wh-phrase occurs clause-initial while the other wh-phrase stays behind.11

(4.59)  a. Ko Ivan želi da vam šta kupi
who Ivan wants that you what buy-[3s]
‘Who does Ivan want to buy you what?’

b. Šta Ivan želi da vam ko kupi
what Ivan wants that who buy-[3s]
‘What does Ivan want who to buy you?’

These examples illustrate that the free ordering between wh-phrases extends to long-distance displacement. Interestingly, however, Bošković (1997) notices a contrast between local and non-local wh-questions in Serbo-Croatian, which indicates that the ordering between wh-phrases in non-local wh-questions is subject to the superiority condition.12 With the following examples, Bošković illustrates that in long-distance wh-extraction the wh-phrases are strictly ordered (Bošković, 1997, ex.10;p.6).

(4.60)  a. Ko Ivan tvrdi da koga voli?
who Ivan claims that whom love-[3s]
‘Who does Ivan claim loves whom?’

b. * Koga Ivan tvrdi da ko voli?
whom Ivan claims that who love-[3s]
‘Who does Ivan claim is loved by whom?’

Bošković’s explanation that non-local wh-questions are subject to the superiority condition is unsatisfactory in relation to the free ordering in local wh-questions and fails to provide an explanation for the sentences in example 4.59. A better clarification for the contrast between the free order of wh-phrases in example 4.59 and the strict ordering in example 4.60 is given in terms of the difference between the two kinds of matrix verbs. Stjepanovic (2001) shows the distinction between the two types of verbs in relation to clitic climbing. Following Progovac (1993), Stjepanovic takes ‘želite’ (= want) to belong to I-verbs (Indicative selecting verb) and ‘tvrdite’ (= claim) to be an S-verb (Subjunctive selecting verb). Besides the different influence of the two verbs

10In Serbo-Croatian, two constructions are allowed for non-local multiple wh-questions. Bošković (1997, ex.8a,9a) illustrates long-distance wh-constructions with two wh-phrases where either the two elements appear clause-initially in the matrix clause or one of the wh-phrases remains in the embedded clause. My informant approves both constructions, but judges the sentence where only one wh-phrase is fronted as more neutral. We adopt the more neutral ordering as also presented in the literature.

11The examples are slightly altered by my informant to get a better comparison with the examples in 4.60.

12Bošković (1998) provides further empirical support for the idea that also in embedded questions and root questions which contain the question particle ‘li’ the superiority condition holds.
on clitic climbing, licensing of polarity items and topic preposing (Progovac, 1993), it may also explain the difference in long-distance extraction.

The main difference between the two verbs is that I-verbs select opaque complements and S-verbs select transparent complements. ‘ˇZelite’ (= to want) belongs to the group of transparent selecting verbs, i.e. verbs that are sometimes referred to as ‘restructuring verbs’ (Stjepanovic, 2001). These verbs have the property that the complement clause is transparent with respect to the extraction of phrases such as clitics, focus phrases, but also wh-phrases. Progovac (1993) contrasts transparent verbs with opaque verbs like ‘tvrdite’ (= to claim) which block further displacement from the complement clause.

The difference between the sentences in example 4.59 and 4.60 can be explained on the basis of this difference between the two types of verbs. In example 4.59 the gaps of the wh-expression occur in a transparent complement clause which gives the possibility that either of the wh-phrases appears frontal. In example 4.60, on the contrary, the matrix verb selects an ‘opaque’ complement clause and blocks the object wh-phrase from occurring clause-initially.

Non-local wh-questions

In section 4.3.3, we have shown how to account for the difference in linear order for local wh-questions with multiple wh-phrases in Bulgarian and Serbo-Croatian.

For Bulgarian, the analysis for deriving the linear order of the wh-phrases in local wh-questions extends to non-local wh-question formation. With this analysis we can account for the multiple wh-fronting in the wh-questions in example 4.58. Any further restrictions are caused by selectional requirements of matrix clauses or additional feature distribution on syntactic objects in the question body as has been shown for English non-local wh-questions (see section 3.3.3).

For Serbo-Croatian long-distance displacement, we concentrate on the contrast between the sentences in example 4.59 and 4.60. The main difference in the two sentences is that in a sentence with a matrix verb ‘ˇZelite’ either wh-phrase may occur clause-initially, whereas in a sentence with matrix verb ‘tvrdite’ the object wh-phrase is blocked from occurring clause-initially. The contrast between the two types of verbs points at a distinction between the selectional requirements of the verbs. On the one hand, the verb ‘ˇZelite’ selects a complement clause which becomes transparent for the displacement of clitics, focus phrases and wh-phrases. On the other hand, the verb ‘tvrdite’ blocks the extraction of all but one wh-phrase from its complement clause.

We can capture the difference between the two types of verbs and their influence on the occurrence of a specific wh-phrase at clause-initial position using the ◇-feature requirements. The verb ‘tvrdite’ blocks the displacement of the object-argument constituent, but it is transparent for the subject argument constituent. Adding a feature requirement ◇ as a selectional requirement on the sentential complement, IV/◇s, which we have used to capture wh-islands,
4.3. Wh-ex-situ: Bulgarian and Serbo-Croatian

will not work, because the subject wh-expression can occur fronted and associate with its gap hypothesis.

Instead, the analysis we propose is that the sentence that is formed by a verb such as ‘tvrdite’ has additional features $\Box\Diamond s$. With additional $\Box\Diamond$ features on the question body type requirement of subject wh-phrases, $WH_L(\Box, \Box\Diamond s, wh)$, only subject wh-phrases can be merged with clauses with a matrix verb ‘tvrdite’ while non-subject wh-phrases can only stay embedded. Sentences with a matrix verb ‘zeli’ yield a structure of type $s$ with which both subject and non-subject wh-phrases can merge.

Incorporating these changes in the lexicon, we can derive the sentences of examples 4.60 and 4.59 (see appendix B.3.2) in addition to the normal local wh-questions on the basis of the following type-assignments:

\[
\begin{align*}
da &: A/A \\
vidi &: TV \\
zeli &: (NOM\Box)/A \\
tvrdi &: (NOM\Box\Diamond s)/A \\
ko &: WH_L(\Box\Box\Diamond\Box NOM, \Box\Diamond s, wh) \\
koga &: WH_L(\Box\Box\Diamond\Box ACC, s, wh) \\
\end{align*}
\]

We illustrate that ‘koga’ can be merged with structures where the gap hypothesis is embedded in the complement clause selected by the verb ‘zeli’ (= want), while merging ‘koga’ to ‘tvrdi’ (= claim) clauses is underivable. The following two tree style presentations illustrate that the category of the question body for ‘ivan zeli da ko vidi’ matches the required category of the wh-phrase ‘koga’, while the ‘ivan tvrdi da ko vidi’ does not.

Figure 4.6: $koga :: WH_L(\Box\Box\Diamond\Box ACC, s, wh)$ can be merged with a $s$-typed structure, but it cannot be merged with a $\Box\Diamond s$-typed structure.

4.3.4 Clitic placement

Another contrast between Bulgarian and Serbo-Croatian regards the possibility for certain grammatical constructions to interrupt the fronted wh-phrases. In Bulgarian, the wh-cluster can not be interrupted, whereas Serbo-Croatian
is quite flexible. To illustrate the contrast, we first present some examples of grammatical sentences in Serbo-Croatian, and then show that sentences with a similar ordering in Bulgarian are judged as ungrammatical. We present the grammatical ordering of wh-questions with clitics, adverbs and parentheticals in Bulgarian.

**Serbo-Croatian** In Serbo-Croatian, clitics (4.61a) but also some adverbs (4.61b) and parentheticals (4.61c) can appear between the first wh-phrase and the rest of the fronted wh-phrases (Bošković, 1998, ex. 3). Note that in Serbo-Croatian both the nominal pronoun ‘mu’ (= him) and the auxiliary ‘je’ (= is) are regarded as clitics which both appear as a clitic cluster in second position. Similarly the clitic and the adverb form a cluster in example 4.61b.

(4.61) a. Ko mu je šta dao? [SC]
   who him is what given
   ‘Who gave him what?’
b. Ko je prvi koga udario?
   who is first whom hit
   ‘Who hit whom first?’
c. Ko, po tebi, šta piše?
   who, according to you what drinks
   ‘Who, according to you, drinks what?’

**Bulgarian** In Bulgarian, the wh-cluster can not be interrupted. The following example illustrates that the wh-phrases cannot be intervened by the adverbial prˇuv (= first). Similar examples can be constructed with clitics like ‘ti’ (= you) or a parenthetical like ‘spored tebe’ (= according to you) which can not appear between the fronted wh-phrases.

(4.62) * Koj prˇuv kogo e udaril
   who first whom has hit
   ‘Who hit whom first?’

Grammatical orderings of wh-questions with additional syntactic objects are shown in 4.63. Example 4.63a illustrates that the adverb appears left attached to the verb. Example 4.63b shows that a clitic pronoun such as ‘ti’ (= you) are verbal clitics procliticized to the verb phrase ‘e kazal’ (= is told). Example 4.61c illustrates a wh-question with a parenthetical. The sentence is grammatical when the parenthetical appears at the beginning or at the end of a multiple wh-question (Rudin, 1988, p.461-466).

(4.63) a. Koj kogo prˇuv e udaril? [B]
   who whom first has hit
   ‘Who hit whom first?’
b. *Koj* kakvo *ti* is kazal?  
who what you is told  
*‘Who told you what?’*  
c. *Zavisi* ot *tova*, *kog* *prvi* e *udaril.*  
depends on this who whom first has hit  
*‘It depends on who hit whom first.’*  

Based on the above data, Rudin (1988) argues that in Bulgarian a sequence of *wh*-phrases forms one constituent, i.e. each *wh*-phrase adjoins to a previous *wh*-phrase. In Serbo-Croatian fronted expressions are separate constituents, i.e. each *wh*-phrase adjoins to *CP*. The contrast can be captured in the generalization that Serbo-Croatian allows multiple specifier positions in the *CP*-domain whereas Bulgarian only allows one specifier which causes the formation of a *wh*-cluster. The difference between the proposed constructions is presented in the following tree schemata.

**Serbo-Croatian:**

```
CP
  Wh_i
  Wh_j
  IP
    t_i V t_j
```

**Bulgarian:**

```
CP
  Wh_i
  Wh_j
  IP
    t_i V t_j
```

Bošković (1998) claims that Rudin’s conclusion may be made too hasty. He points out that clitics in Bulgarian and Serbo-Croatian act differently. In Serbo-Croatian, clitics are fixed to the second position and put no syntactic requirements on the category of their host. While in Bulgarian the clitic typically attaches to the verb.

---

We assume an analysis of clitics as proposed by Kraak (1998), who has analyzed clitics in Romance languages. We assume that her approach can be extended to Slavic languages. Similar to *wh*-phrases, clitics are assigned higher-order types that capture the relation between the clitic itself, the phrasal body that hosts the clitic and the argument which the clitic associates with. Similar to *wh*-phrases, we could identify clitics with a type schema that could specify for each clitic the argument that it associates with, the host and the type of the phrase after it is merged with the host. Without treating too many details

---

13For a detailed analysis of how clitics are handled in type-logical grammar or in related formal frameworks such as HPSG or pregroups we refer the reader to the work of Kraak (1998), Monachesi (2005), Casadio and Lambek (2001)
of the analysis of clitics, we assign clitics a three-place operator $\text{Cl}(A, B, C)$ which follows similar elimination rules and displacement rules as the wh-type schema.

Below we provide a sketch of the analysis of clitic placement in a wh-question in both Serbo-Croatian and Bulgarian. The precise interaction between an analysis of clitics and wh-question formation is left for further research.

**Clitic placement in Serbo-Croatian** As Bošković (1998) points out, clitics in Serbo-Croatian are second position clitics. Independent of the analysis to merge the clitics with a wh-question, the arguments of the wh-type schema assigned to the wh-phrases are left unchanged. The type specification of a clitic or an adverb must be such that we can derive wh-questions such as the sentence presented in example 4.61a, ‘Ko mu je šta dao?’ (= Who gave what to him?). After the clitics have found their final position, the wh-phrases merge with the structure as usual. Again the linear order of the wh-phrases is unrestricted as pointed out by the following example from Bošković (1998).

(4.64) a. Ko je šta prodao?
   who is what sold
   ’Who sold what?’

   b. Šta je ko prodao?

**Bulgarian** As observed by Bošković (1998), clitics in Bulgarian have different properties from those in Serbo-Croatian. In Bulgarian, these elements cliticize to the verb as illustrated in example 4.63b and repeated here.

(4.65) Koj kakvo ti e kazal?
   who what you is told
   ’Who told you what?’

The type-assignment for the clitic should reflect this property of verbal attachment. As an illustration, we show how a clitic in Bulgarian would be typed in the lexicon. For instance, the Bulgarian clitic ‘ti’ (= you-[dat]) gets a type-assignment that indicates that it merges with a verb phrase which contains a gap hypothesis of type DAT. The gap hypothesis is marked with a special set of clitic features that are needed for the displacement of the associated gap hypothesis. Similar to wh-phrases, to merge a clitic to its ‘clitic’ body, the gap hypothesis in the body must be reachable. After the clitic merges with the verb phrase, it replaces the associated gap hypotheses and returns a verb phrase of type NOM\'$s (abbreviated to IV).

$ti \mathbin{:} \text{Cl}(\Diamond \Box \text{DAT}, \text{IV}, \text{IV})$ (= you)
Figure 4.7 presents a natural deduction derivation of the sentence in example 4.65. To move the hypothesized element from its inner argument position to the right of the verb, we need to apply the same set of displacement postulates as the displacement of gap hypotheses for wh-question formation. The displacement of the clitic gap co-occurs with the displacement of the gap hypothesis of the wh-phrase without blocking the derivation.

Ungrammatical word orders such as ‘Koj ti kakvo e kazal?’ and ‘Koj ti e kakvo kazal?’ are underivable on the basis of the selectional requirements of both the clitic and the wh-phrases. Clitics only merge with structures of type IV and have to be merged before the wh-phrases are merged. Wh-phrases merge only with question bodies with a sentence type s or wh or feature decorated sentence types. Therefore, a clitic must merge before the wh-phrases (check B.4.2 for on-line derivation).

Following Bošković (1998), we have argued that a type-logical grammar analysis of clitics for Bulgarian and Serbo-Croatian must differ. In Serbo-Croatian, clitics are second position clitics, whereas in Bulgarian the clitics host to the verb. Without going into the details of such an analysis, we...
have sketched that as a result of a specific clitic analysis, in Serbo-Croatian, the fronted wh-phrases are interrupted by the second position clitic, whereas in Bulgarian the wh-phrases precede the clitic cluster. The wh-phrases are merged with the structure as before.

**Concluding remarks on wh-ex-situ languages**

We have presented examples and data of Bulgarian and Serbo-Croatian, both multiple wh-fronting languages. To account for wh-question formation in the two languages, wh-phrases are uniformly typed using wh-ex-situ type schema. Due to the underlying SVO word order, both left and right displacement postulates are needed to move the hypothesized argument to the edge of a structure, before the wh-phrase is merged with the structure. The restricted order of wh-phrases in Bulgarian multiple wh-questions has been accounted for on the basis of the derivability patterns between types. Blocking of fronted wh-phrases or other changes in the order of the wh-cluster are due to restrictions on selectional requirements of the verbs, clitics or other phrases involved.

### 4.4 Partial wh-movement

In section 4.1, we have distinguished between languages with partial wh-movement with and without a scope marker. First, we concentrate on German and Hindi which are languages that must have a scope marker to form a wh-question with a partially moved wh-phrase. In the final part of this section, we extend our approach of partial wh-movement to Malay which is a language that has partial wh-movement without the occurrence of a scope marker.

German is like English a single wh-fronting language but it also allows for partial wh-movement of an interrogative. Hindi can have a wh-phrase in fronted position or allows a wh-question with the wh-phrase in preverbal position. Wh-questions that are formed by partial wh-movement are also referred to as scope marking constructions. The term, partial wh-movement is used by McDaniel (1989) to refer to constructions in German and Romani. The phenomenon is sometimes termed scope marking (Dayal, 2000) or was-constructions (van Riemsdijk, 1983) in reference to German. We refer to the construction in both German and Hindi as a scope marking construction.

In order to provide an analysis of scope marking, we will start by presenting an analysis of the basic grammatical constructions in the two languages. Firstly, we deal with the basic word order and the standard construction of wh-question formation in German in section 4.4.1 and section 4.4.2. Then, we do the same for Hindi in section 4.4.3 and 4.4.4. In section 4.4.5, we present a syntactic analysis of a scope marking construction in both Hindi and German by treating it as an instance of the wh-type schema. Finally, in section 4.4.6,
we present a sketch of partial wh-movement in Malay.\footnote{The German examples are taken or derived from examples in Klepp (1999); The Hindi examples are taken from Dayal (2000) and Mahajan (2000); the example of partial wh-movement in Malay is taken from Cole and Hermon (2000). Except for Malay, all examples have been checked by native speakers.}

### 4.4.1 Basic word order in German

German is classified as a verb second language with an underlying SOV word order. However, in main clauses, German has a SVO order due to verb-second. The following examples illustrate the different underlying word orders for declarative clauses (example 4.66a), subordinate clauses (example 4.66b) and yes-no questions (example 4.66c).

\[(4.66)\]

\begin{align*}
a. \text{Mirò \textit{sieht} \textit{ein Bild}.} \\
& \quad \text{Mirò sees a picture} \\

b. \text{Mirò \textit{glaubte}, dass Picasso \textit{ein Bild} \textit{gemalt} \textit{hatte}.} \\
& \quad \text{Mirò believed that Picasso a picture painted had} \\
& \quad \text{‘Mirò believed that Picasso had painted a picture.’} \\

c. \text{Hat Picasso \textit{ein Bild} \textit{gemalt}?} \\
& \quad \text{Has Picasso a picture painted} \\
& \quad \text{‘Did Picasso paint a picture?’}
\end{align*}

\begin{itemize}
\item [\text{basic word order}]\end{itemize}

To concentrate on the phenomenon of wh-question formation, we capture the different argument orderings in main, subordinate and interrogative clauses by assigning multiple types to verbs. Using a categorial type logics account, the different entries may be unified to a single type, while the different orderings are the result of verb movement.\footnote{For a categorial type logics account of a verb second language like German, we refer the reader to Moortgat (1999b) who analyzed verb second constructions in Dutch.}

We assign three different types to verbs. One type yields an SVO order in a main declarative clause. A second type results in an SOV word order for subordinate clauses. To distinguish between main declarative clauses and subordinate clauses, we use the types: \(s\) for main clause types and \(s_q\) for subordinate clauses. Additionally, in German the underlying word order in a yes-no question is VSO which we derive with an additional entry assigned to verbs yielding a clause of type \(q\). To simplify the presentation of the different types assigned to verbs, we will abbreviate the different types of verbs. The abbreviations without a subscript are used for verbs that occur in main clauses, types with subscript \(s\) are used for subordinate clauses and likewise types with a subscript \(q\) are used to derive the argument ordering in yes-no questions. The auxiliary used in the sample fragment is the present or past perfect form of ‘haben’ (= to have) that selects for a past participle, the abbreviated type \(\text{prt}\), such as ‘gemalt’ (= painted).
Thus, in addition to the usual abbreviations for noun phrases (e.g. NOM, ACC, DAT), we distinguish the following types of verbs. With these types we derive the different argument orderings.

**Main clauses:**

\[
\begin{align*}
IV &:: \text{NOM}/s \\
TV &:: IV/\text{ACC} \\
\text{AUX} &:: IV/\text{prt}
\end{align*}
\]

**Subordinate clauses:**

\[
\begin{align*}
IV_s &:: \text{NOM}/s_s \\
TV_s &:: \text{ACC}/IV_s \\
\text{AUX}_s &:: \text{prt}/IV_s
\end{align*}
\]

**Yes-no questions:**

\[
\begin{align*}
IV_q &:: q/\text{NOM} \\
TV_q &:: (q/\text{ACC})/\text{NOM} \\
\text{AUX}_q &:: (q/\text{prt})/\text{NOM}
\end{align*}
\]

In addition to these verbs, we also address bridging verbs such as ‘glauben’ that select a subordinate clause of type \(s^r\) and verbs such as ‘fragen’ that select embedded wh-interrogatives of type \(wh^r\) or ‘if/whether’ clauses of type \(q^r\). The complementizer ‘dass’ gets the type-assignment: \(s^r/s_s\).

The following lexicon presents the type-assignments on the basis of which we can analyze the different types of clauses in example 4.66 (see appendix B.5.1):

- \text{picasso, miro} :: \text{NOM}
- \text{ein} :: \text{ACC}/n
- \text{Bild} :: n
- \text{sieht} :: TV_i where \(i \in \{\emptyset, s, q\}\)
- \text{hatte} :: \text{AUX}_i where \(i \in \{\emptyset, s, q\}\)
- \text{gemalt} :: \text{ACC}/\text{prt}
- \text{glaubt} :: IV_i/s^r or (q/s^r)/\text{NOM}
- \text{fragte} :: IV_i/wh^r or (q/wh^r)/\text{NOM} (where \(j \in \{\emptyset, s\}\))
- \text{dass} :: s^r/s_s

### 4.4.2 Wh-question formation in German

The standard construction of a direct question in German is illustrated in example 4.67, where the interrogative pronoun appears clause-initially in the main clause. Example 4.68 (Klepp, 1999, ex.3) illustrates long-distance wh-movement where the wh-phrase is bound to an argument position in the embedded clause. The wh-phrase can be bound over a long-distance. Similar to English the verbs that select the embedded clauses are bridging verbs that allow for restructuring over several clauses. Finally, example 4.69a illustrates an embedded interrogative that is a complement of the interrogative verb ‘fragen’ (= to ask). Example 4.69b shows that the wh-phrase cannot occur displaced from the embedded interrogative.

*(4.67)* Welches Bild hat Picasso gemalt?

‘Which picture has Picasso painted’

‘Which picture did Picasso paint?’
4.4. Partial wh-movement

(4.68) a. Welches Bild glaubte Mirò dass Picasso gemalt hatte?
   Which picture believed Miró that Picasso painted had
   'Which picture did Miró believe that Picasso had painted?'

   b. Welches Bild glaubte Mirò dass Jacob meinte dass
   Which picture believed Miró that Jacob thought that
   Picasso had painted had
   'Which picture did Miró believe that Jacob thought that Picasso had painted?'

(4.69) a. Mirò fragt welches Bild Picasso gemalt hatte.
   Miró asks which picture Picasso painted had
   'Miró asks which picture Picasso had painted.'

   b. * Welches Bild fragt Mirò dass Picasso gemalt hatte?
   Which picture asks Miró that Picasso painted had?

[wh-question formation in German]

The analysis of basic simple wh-fronting in German is similar to the analysis of
wh-question formation in English. We assume the goal type for wh-questions
to be the macro type wh. The wh-questions are derived by assigning wh-type
schema to the different wh-phrases. Different from English, we do not need
to assign different type schema instances to subject and non-subject argument
wh-phrases. Because German is underlying a SOV language, the wh-phrases
associate with a gap that occurs on the left-hand side of the tree structure.
Therefore, the wh-phrases that occur clause-initially are identified with wh-
type schema $WH_l$. Recall that a wh-phrase of this type merges with a question
body at a clause-initial position replacing a gap hypothesis that occurs on a
left branch in the question body. As a consequence, German only uses left
displacement postulates to derive wh-question formation.

The following type schema account for the analysis of single and multiple
constituent questions in German.

<table>
<thead>
<tr>
<th></th>
<th>direct questions</th>
<th>indirect questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>subject wh-phrase</td>
<td>$WH_{s,\circ}$</td>
<td>$WH_{s,\circ}$</td>
</tr>
<tr>
<td>non-subject wh-ex-siu</td>
<td>$WH_{s,\circ}$</td>
<td>$WH_{s,\circ}$</td>
</tr>
<tr>
<td>non-subject wh-in-situ</td>
<td>$WH_{s,\circ}$</td>
<td>$WH_{s,\circ}$</td>
</tr>
<tr>
<td></td>
<td>where CASE $\in$ {ACC, DAT}</td>
<td>where CASE $\in$ {ACC, DAT}</td>
</tr>
</tbody>
</table>

With these type-assignments, we can analyze the direct and indirect wh-
questions of examples 4.67–4.69 (see appendix B.5.2). To illustrate long-
distance wh-question formation in German, we present a part of the deri-
vation of the sentence 'Welches Bild glaubte Miró, dass Picasso gemalt hatte?'.
The analysis in figure 4.8 illustrates the displacement of the gap hypothesis
from an embedded position on a left branch to the left edge of the structure.
through successive steps of left displacement postulates. The meaning assembly of the analysis reveals that the wh-question has a main clause question interpretation. The wh-operator \( \omega \) takes scope over the main clause with the \( \lambda \) abstractor binding the gap hypothesis in the embedded clause.

\[
\vdash \omega (\lambda \mathbf{y}.(\mathbf{picture} \mathbf{y}) \land (\mathbf{believe} ((\mathbf{painted} \mathbf{y}) \mathbf{p}) \mathbf{m}))
\]

Figure 4.8: Natural deduction derivation of 'Welches Bild glaubte Miró dass Picasso gemalt hatte?'

4.4.3 Basic word order in Hindi

Like German, besides a basic construction for wh-questions, Hindi has wh-questions which follow the pattern of partial wh-movement constructions. Before we present data on wh-question formation in Hindi, we first illustrate the basic grammatical relations underlying Hindi.

Basic selectional requirements Based on examples given in Dayal (1994, 1996), Hindi has an underlying word order SOV. Example 4.70 illustrates a declarative sentence with a basic word order (Dayal, 1996, ex.6).

\[
\begin{align*}
(4.70) \quad \text{Anu-} & \quad \text{ne kitaab khariidii} \\
& \quad \text{Anu-[erg] book buy-[past]} \\
& \quad \text{‘Anu bought a/the book.’}
\end{align*}
\]

The example shows that Hindi has ergative/absolutive morphological marking and is therefore classified an ergative language instead of a nominative/accussative language. Dixon (1994) explains the distinction between
the nominative/accusative system and the ergative system as follows. In lan-
guages with a nominative/accusative system, the argument that fulfills the
subject role of an intransitive verb and the argument that fulfills the agent role
of a transitive verb map to nominative case, while the object maps to accusa-
tive case. In ergative languages, the subject argument of an intransitive verb
and the object argument of a transitive verb are marked as absolutive while
the agent of a transitive verb is marked as ergative.

Despite the ergative/absolutive marking, Dixon (1994, p. 175) points
out that Hindi is a language that has “some ergative morphology but an
entirely accusative syntax”. In Bobaljik (1993), the same observation is
made on the basis of languages other than Hindi. In short, the argument of
the transitive verb that is morphologically marked as ergative (i.e. ‘ne’) in
example 4.70 corresponds to the structural subject. ‘Kitaab’ (= book) which is
unmarked because absolutive case is the default case in Hindi, corresponds
to the structural object. The argument of the intransitive verb which is also
morphologically marked as absolutive follows the syntactic pattern of a
structural subject.

Adopting the line of reasoning of Dixon (1994) and Bobaljik (1993), we treat
the syntax of Hindi as nominative/accusative. To distinguish between the
structural subject and object, we encode the subject as nominative and the
object as accusative. The details on a treatment of morphological erga-
tive/absolutive marking in a type-logical grammar framework are left for fur-
ther research.

With the following lexicon, we can derive sentences with basic word or-
der SOV. The free word order may be derived using the case features to trig-
ger the application of the displacement postulates. Additionally, adopting the
nominative/accusative case system for Hindi helps us to make a comparison
between German and Hindi partial wh-movement constructions.

anu-ne :: NOM
kitaab :: ACC (= book)
khariidii :: ACC\(\text{NOM}\)$s (= buy-[past])

4.4.4 Wh-question formation in Hindi

Local direct questions Wh-elements in direct main clause questions can oc-
cur in two positions. The interrogative clauses in example 4.70 illustrate these
two options. An object wh-phrase occurs either preverbal or fronted. The sem-
antic interpretation of the two direct questions does not differ significantly.

The examples in this section are based on data presented in Dayal (1996, 2000); Mahajan
(2000). The examples or glosses slightly differ from the source examples. For instance, we have
glossed ‘kis-ko’ as ‘who[acc]’ in stead of adopting Mahajan’s 2000 gloss to indicate that the wh-
phrase consists of a stem with a case suffix. Additionally, we have changed the tense or the verb
4. Structural variation in wh-question formation

(4.71)  a. Ravii-ne kis-ko dekhaa?
         Ravi-[erg] who saw
         ‘Who did Ravi see?’

       b. Kis-ko Ravii-ne dekhaa?
         who Ravi-[erg] saw

Embedded interrogatives  The basic argument order of Hindi is SOV. However, additional argument phrases and complement clauses may appear to the right of the selecting verb. However, Dayal (1996) indicates that non-finite complement clauses occur preverbally while finite clauses occur at the right periphery of the clause. We leave the non-finite clauses for further research and restrict the analysis to finite complement clauses.

Like similar verbs in other languages, ‘jaantaa’ (= to know) is a verb that optionally selects an embedded interrogative complement (example 4.72). A wh-phrase in an embedded interrogative can either appear clause-initially in the embedded clause or preverbally. When it appears clause-initial in the embedded clause, it follows the subordinate conjunction ‘ki’.

(4.72)  a. Jaun jaantaa hai ki merii kis-ko dekhaa.
         J. know [pres] that M. who saw
         ‘John knows who Mary saw.’

       b. Jaun jaantaa hai ki kis-ko merii dekhaa.
         J. know [pres] that who M. saw

Non-local wh-questions  A verb such as ‘jaantaa’ (= to know), which optionally selects an embedded interrogative clause can be contrasted with ‘socaa’ (= to think), which only allows declarative complement clauses. Similar to verbs in other languages, ‘socaa’ acts as a bridge verb. The following examples illustrate that wh-phrases may either appear fronted in the matrix clause or occupy a preverbal position within the matrix clause (Mahajan, 2000, ex.4, p.318).

(4.73)  a. Kisko, Sitaa-ne socaa ki ravii-ne ti dekhaa?
         whom Sita-[erg] thought that Ravi-[erg] saw
         ‘Who did Sita think that Ravi saw?’

       b. Sitaa-ne kis-ko, socaa ki ravii-ne ti dekhaa?
         Sita-[erg] who thought that Ravi-[erg] saw

To derive the standard construction of wh-question formation in Hindi where the wh-phrase either appears fronted or preverbally, we encode the following in examples from both Mahajan (2000) and Dayal (1996, 2000) to present a uniform overview of wh-question formation in Hindi. The presented data have been checked with a native speaker of Hindi.
4.4 Partial wh-movement

type-assignments for verbs in the lexicon. The verb 'jaanta' (= know) can optionally select an embedded interrogative clause, while 'socaa' (= know) only selects embedded declarative clauses. Furthermore, Hindi has a complementizer 'ki' that appears clause-initial in embedded clauses.

\[
\begin{align*}
\text{jaanta} &:: ((\text{NOM}\backslash s)/A) (= \text{know}) \\
&\quad \text{where } A \in \{s',wh'\} \\
\text{socaa} &:: ((\text{NOM}\backslash s)/s') (= \text{think}) \\
\text{ki} &:: s'/s (= \text{that})
\end{align*}
\]

Again we use \textit{wh} to type main clause questions and \textit{wh}' to type embedded interrogatives. To derive the ordering of arguments in main clause and indirect questions as illustrated in examples 4.71, 4.72 and 4.73 each \textit{wh}-phrase gets multiple type-assignments. One type for merging with a main clause question body and another type for merging with a subordinate clause. The types are further subdivided in a type that merges at the leftmost position of that clause and a type that merges with the verb phrase at the preverbal position.

\[
\begin{align*}
\text{direct questions} &\quad \text{indirect questions} \\
\text{WH}_1^\text{direct}(\text{CASE}, s, \text{wh}) &\quad \text{WH}_1^\text{indirect}(\text{CASE}, s', \text{wh}') \\
\text{WH}_2^\text{direct}(\text{CASE}, \text{NOM}\backslash s, \text{NOM}\backslash \text{wh}) &\quad \text{WH}_2^\text{indirect}(\text{CASE}, \text{NOM}\backslash s', \text{NOM}\backslash \text{wh}') \\
&\quad \text{where } \text{CASE} \in \{\text{ACC}, \text{DAT}\}
\end{align*}
\]

The sentences in example 4.71, 4.72 and 4.73 can be analyzed on the basis of these type-assignments (see appendix B.6.2 for on-line derivations). We illustrate the point in the derivation where the \textit{wh}-phrase merges with the verb phrase to derive the \textit{wh}-question '(Siitaa-ne) kisko socaa ki raviine dekhaa' (= Who did Siitaa-ne think that Ravi saw?) of example 4.73 with the following tree style presentation.

\[
\begin{align*}
\text{4.4.5 Scope marking constructions}
\end{align*}
\]

Scope marking constructions are characterized by the apparent availability of more \textit{wh}-expressions than corresponding gaps. As the following examples of scope marking constructions in German and Hindi illustrate, both the main clause and the embedded sentence contain a \textit{wh}-phrase.

\[
\begin{align*}
(4.74) \text{Was, glaubt Miró welches Bild, Picasso t, gemalt hatte?} \\
&\quad \text{what believes Miró which picture Picasso painted had} \\
&\quad \text{‘Which picture does Miró believe that Picasso had painted?’}
\end{align*}
\]

\[
\begin{align*}
(4.75) \text{Jaun kyaa socaa hai ki merii kis-se baat karegii?} \\
&\quad \text{J. what think-[pres] that Mary who-[ins] talk do-[Fut]} \\
&\quad \text{‘Who does John think Mary will talk to?’}
\end{align*}
\]
4. Structural variation in wh-question formation

Figure 4.9: \( kisko :: WH^f_{w}(\diamond \Box \text{ACC}, \text{NOM}\backslash s, \text{NOM}\backslash wh) \) associates with its gap hypothesis
In both German and Hindi, the wh-phrase in the main clause is the object wh-phrase. Although the main clause already contains an interrogative pronoun, the embedded sentence also contains an wh-phrase which is partially moved, either to the front of the subordinate clause (German, see example 4.74) or to a preverbal position (Hindi, see example 4.75). Another characteristic of the scope marking construction is that the main clause verb is a verb which normally selects a declarative embedded clause. In the following section, we propose an analysis of scope marking which covers these characteristics.

For both German and Hindi, we show that we can account for these scope marking constructions along similar lines as for standard wh-question formation.

4.4.5.1 Scope marking constructions in German

Apart from the long-distance wh-movement construction (as explored in section 4.4.2), an embedded wh-phrase in German can be marked by a fronted wh-phrase ‘was’ which is referred to as the scope marker. The scope marker ‘was’ is used differently than the morphologically similar object wh-phrase ‘was’. Instead of associating with an argument gap, the scope marker associates with the embedded interrogative clause which contains a displaced wh-phrase. Furthermore, the embedded wh-phrase is partially moved and appears clause-initial in the subordinate clause. Therefore, the construction is sometimes referred to as partial wh-movement.

In example 4.76, we compare a scope-marking construction (4.76a) with a long-distance wh-question (4.76a).

(4.76) a. Welches Bild, glaubt Miró dass Picasso t_i gemalt hatte?
     which picture believe Miró that Picasso painted had

     ‘Which picture does Miró believe that Picasso had painted?’

b. Was, glaubt Miró welches Bild, Picasso t_i gemalt hatte?
     what believe Miró which picture Picasso painted had

     ‘Which picture does Miró believe that Picasso had painted?’

When more subordinate clauses intervene between the matrix clause and the embedded interrogative, the scope marker appears in a clause-initial position in each subordinate clause (example 4.77a). Alternatively, the embedded wh-phrase may appear clause-initially in one clause higher up. In those cases the preceding clauses are marked by ‘was’, while the following clauses are marked by the complementizer phrase ‘dass’ (example 4.77b).
Before we discuss how the wh-scope marker finds its position at the front of the matrix clause, we first discuss the grammatical role of the scope marker. To understand the role of the scope marker, we compare the use of the scope marker to the syntactic use of the complementizer. In CTL, the complementizer, ‘dass’ (= that), is categorized as $s'/s_s$. The complementizer changes the type of the subordinate clause such that it can be selected by the matrix verb. In scope marking constructions, the matrix verb is a bridge verb which selects an embedded declarative clause ($\text{glaubt } \circ \text{ miro} \vdash q/s'$). Note that in a scope marking construction (see example 4.76b), the embedded clause is a wh-interrogative clause and not a declarative clause. We hypothesize, for now, that the scope marker merges with the embedded interrogative clause and returns the category of an embedded declarative ($s'$). This category can be selected by the bridge verb.

The following derivation illustrates how the scope marker ($s'/wh'$) takes the embedded interrogative and returns the type of the embedded declarative clause. \text{wbpgh} is an abbreviation for an embedded interrogative clause, ‘welches Bild Picasso gemalt hatte’ (= wh')

\[
\text{glaubt } \vdash \text{ IV}_s/s' \\
\text{glaubt } \circ \text{ miro } wbpgh \vdash \text{ wh'} \\
\text{wbpgh } \vdash \text{ wh} \\
\]

However, ‘was’ always occurs at a distance from the embedded clause at the front of the main clause which in turn is interpreted as a wh-question. The proposed type-assignment of the scope marker as $s'/wh'$ does not account for the structural position of the scope marker at the front of the main clause. Thus, the scope marker must have a type-assignment which accounts for 1) its structural position at the front of the main clause, 2) its function to change the type of the embedded interrogative clause, and 3) its function to type the main clause as a wh-interrogative clause. These three characteristics can be incorporated in the wh-type schema.
The scope marker is typed using the wh-ex-situ type schema: the scope marker associates with a gap hypothesis of category s′/wh′ on a left branch, it merges with a question body of type q and it yields a wh-question of type wh. Thus, the instantiation of the type schema for the German scope marker ‘was’ becomes:

\[ \text{was} :: \text{wh}/(\Diamond \Box (s'/\text{wh}'), q) = \text{WH}_{\lambda_2}(\Diamond \Box (s'/\text{wh}'), q, \text{wh}) \]

On the basis of this type the scope marking construction can be derived. Figure 4.10 displays the last steps in the derivation where ‘wbpgk’ is an abbreviation for an embedded interrogative clause, ‘welches Bild Picasso gemalt hatte’ (= wh)

\[
\begin{align*}
glaubt & \qquad \text{miro} \\
\frac{(q/s')/\text{NOM} \quad \text{NOM}}{glaubt \circ \text{miro} \vdash q/s'} & \quad [E] \\
\frac{[\Diamond \Box (s'/\text{wh}')] \quad \text{wbpgh} \vdash s'/\text{wh}'} {s'/\text{wh}'} & \quad [E] \\
\frac{\Diamond \Box (s'/\text{wh}') \circ (glaubt \circ \text{miro}) \circ \text{wbpgh} \vdash q} {\vdash q} & \quad [P/2] \\
\frac{\vdash \text{was} \vdash \text{WH}_{\lambda_2}(\Diamond \Box (s'/\text{wh}'), q, \text{wh})} {\vdash \text{was} \circ ((glaubt \circ \text{miro}) \circ \text{wbpgh}) \vdash \text{wh}} \\
\end{align*}
\]

\[ \omega \lambda y.((\text{believe } y (\omega \lambda x((\text{picture } x) \land ((\text{has } (\text{painted } x)) \ p))))) \ p) \]

Figure 4.10: Natural deduction derivation of a scope marking construction in German

The derivation of the scope marker construction can be paraphrased in prose as follows. The embedded interrogative (‘welches Bild Picasso gemalt hatte’ = ‘wbpgk’) is built up as usual. The scope marker hypothesis (\(\Diamond \Box (s'/\text{wh}')\)) merges with the embedded interrogative clause and yields a construction of type s′. After the main clause glaubt Miró \(\vdash q/s'\) merges with the subordinate clause, the scope marker hypothesis is displaced to the front of the matrix clause. At this point the scope marker can merge with the question body and replace its hypothesis. The complete expression becomes of type \text{wh}.

The lambda term that corresponds to this sentence construction shows how the wh-scope marker, as a higher order type, binds the complete embedded wh-phrase. With this type-assignment, we can derive the sentences in example 4.76 and 4.77, page 139 (see appendix B.5.3 for on-line derivation). In chapter 5, in which we discuss the syntax-semantics interface of wh-question, we will show that the meaning assembly of a scope-marking construction is
equal to the meaning assembly of a wh-interrogative with long-distance displacement.

**Multiple wh-scope marking** Constructions with multiple scope markers are derived by successively merging a scope marker hypothesis with an embedded interrogative clause. Each embedded clause that intervenes with the embedded interrogative and the main clause must contain another scope marker. The intervening scope markers play the same role as the scope marker in the main clause, i.e. the scope marker hypothesis merges with the embedded interrogative clause and returns a declarative clause type. The type of the intervening scope marker differs slightly from the main clause scope marker. The wh-question body of the intervening scope markers is the embedded declarative clause type $s'$ while it yields an embedded interrogative clause type $wh'$. The intervening scope marker has the following type-assignment:

\[
\text{intervening scope marker } '\text{was}' :: \text{WH}_{ix}(\Diamond \Box(s'/wh'), s_s, wh')
\]

Multiple scope marker construction have been illustrated in example 4.77a and b. To illustrate the use of the intermediate scope marker, we derive the intermediate clause ‘*was Hans meint welches Bild Picasso gemalt hatte*’ of example 4.77a. Notice we abbreviate ‘welches Bild Picasso gemalt hatte’ again to ‘wbpgh’.

\[
\begin{align*}
\text{hans } \text{NOM} & \vdash \text{meint } \Box(s'/wh') \text{ wbpg} \vdash s' \\
\text{hans } \text{NOM} \circ \text{wbpg} & \vdash \text{NOM}' s_s \\
\text{hans } \text{meint } \text{wbpg} & \vdash s_s \\
\text{hans } \text{meint } \text{wbpg} & \vdash \text{wh'} \\
\text{was } \text{wbpg} & \vdash \text{wh'} \\
\omega \lambda y.((\text{think } (y \omega \lambda x((\text{picture } x) \land ((\text{has } (\text{painted } x)) p)))))) h
\end{align*}
\]

Figure 4.11: Natural deduction derivation of the intermediate clause in a multiple scope marking construction

To derive the complete construction, the derivation continues along the same lines as the scope marking construction in figure 4.10.
The scope marking construction in example 4.77b is derived from the partial movement of the wh-phrase to the clause-initial position of the preceding embedded clause. The scope marker hypothesis is merged with the embedded interrogative clause as before and moves to the edge of the main clause. The verb in the intermediate clause, *Hans meint* selects the embedded declarative clause where the gap hypothesis of the wh-phrase has been displaced. The scope marker hypothesis merges with the intermediate subordinate clause and moves to the edge of the main clause as presented before in figure 4.10. (See appendix B.5.3 for an on-line derivation.)

**Constraints on partial wh-movement**

In the previous section we have looked at scope marking constructions. The verbs that allow partial movement constitute more or less the class of bridge verbs, for instance verbs such as: ‘*sagen*’ (= to say), ‘*meinen*’ (= to think/mean) and ‘*glauben*’ (= to believe). These verbs select a declarative subordinate clause. In contrast with the bridge verbs, Reis (1996) observes that preference verbs and (strong) factive verbs do not license partial movement. Examples of preference verbs are ‘*mögen*’ (= to wish) and ‘*wollen*’ (= to want). Examples of factive verbs are ‘*ärgern*’ (= to anger) and ‘*berücksichtigen*’ (= to regard). The sentences in example 4.78 illustrate that partial movement with a matrix clause which contains a preference verb such as ‘*mögen*’ is ungrammatical, while long-distance displacement is grammatical.

(4.78)  

\[\begin{align*}
(4.78) & \quad \text{a. } * & & \text{Was } & & \text{möchte/will } & & \text{Anna, welche Schokolade } & & \text{Felix} & & \text{mitbringt?} \\
& & & \text{what wants/wishes } & & \text{Anna } & & \text{which chocolate } & & \text{Felix} & & \text{with brings} \\
& & & \text{b. } & & \text{Welche Schokolade } & & \text{möchte/will } & & \text{Anna, dass Felix} & & \text{mitbringt?} \\
& & & & & \text{which chocolate } & & \text{wants/wishes } & & \text{Anna } & & \text{that Felix} & & \text{with brings} \\
& & & & & \text{‘Which chocolate does Anna want/wish Felix to bring?’}
\end{align*}\]

That certain verbs do and certain other verbs do not allow for partial movement is the result of the selectional requirements of the verbs. We will account for the difference in selectional requirements on the basis of the derivability relation between sentence types (\(\bigcirc\Delta s \vdash s \vdash \bigcirc\Delta s\)). Bridge verbs such as ‘*glauben*’ (= believe) allow for both long-distance displacement and scope marking constructions, while factive verbs and preference verbs only allow for long-distance displacement. Thus, it should be possible for ‘*glaubt*’ to merge with both subordinate clauses headed by ‘*dass*’ or with subordinate clauses...
headed by the scope marker hypothesis. However, for the factive or preference verbs it should only be possible to merge with embedded clause headed by the complementizer ‘dass’.

Taking the different possible selectional requirements into account, we obtain the following lexicon.

\[
\begin{align*}
\text{möchte} &:: (q/\Diamond s')/\text{NOM} \\
\text{glaubt} &:: (q/s')/\text{NOM} \\
\text{dass} &:: \Diamond s'/s_i \\
w\text{was} &:: \text{Wh}^{l,3}(\Diamond (s'/\text{wh}'), q, \text{wh})
\end{align*}
\]

As the type-assignment shows, a preference verb such as ‘mögen’ imposes an additional feature requirement (\(\Diamond s'\)) on its complement clause. The bridge verb ‘glaubt’ and the scope marker are left unchanged. However, complementizer ‘dass’ carries the same features as requested by the preference verb. With these type-assignments we account for the ungrammaticality of the sentence in example 4.78a), while we can still derive long-distance displacement of example 4.78b), and long-distance displacement and scope marking constructions with bridge verbs such as ‘glauben’. (see appendix B.5.4 for online derivation)

Bridge verbs can still merge with subordinate clauses headed by ‘dass’ because the derivability relation holds: \(\Diamond s \vdash s\). The following derivation illustrates that sentence in 4.78a) is underivable, because the preference verb cannot merge with the construction of the scope marker hypothesis and the embedded interrogative clause (wsfm is used as an abbreviation for ‘welches Schokolade Felix mitbringt’).

\[
\begin{align*}
\text{möchte} &:: (s/\Diamond s')/\text{NOM} \quad \text{anna} &:: \text{NOM} \\
\text{möchte} \circ \text{anna} &\vdash s/\Diamond s' \quad \Diamond (s'/\text{wh}') \quad \text{wsfm} &:: s'/\text{wh'} \quad \text{wh'}
\end{align*}
\]

4.4.5.2 Scope marking constructions in Hindi

In Hindi, similar to German, the scope marker ‘kyaa’ is morphologically similar to the wh-phrase for ‘what’. Scope marking constructions in Hindi differ from German with respect to the position of the scope marker. In scope marking constructions, ‘kyaa’ always occurs preverbally. The wh-phrase in the embedded interrogative clause may occur clause-initially or preverbally in the subordinate clause. Semantically, the scope marking constructions is interpreted as a direct question. The following example illustrates a scope marking construction in Hindi. The gloss [ins] to the morphological marking
4.4. Partial wh-movement

‘-se’ of the wh-word ‘kis’ is an instrumental marking which we abbreviate to [ins].

(4.79) (Dayal, 2000, ex.5, p.160)

Jaun kyaa socaa hai ki merii kis-se baat karegii?
J. what think-[pres] that Mary who-[ins] talk do-[Fut]

‘Who does John think Mary will talk to?’

The following example illustrates that a sentence with an embedded interrogative clause selected by a bridge verb is ungrammatical. The wh-phrase must either occur clause-initially or the sentence needs a scope marker to mark the sentence as a wh-interrogative.

(4.80) * Jaun socaa hai ki merii kis-se baat karegii?
J. think-[pres] that M. who-with talk do-[Fut]

Multiple scope marking Similar to German, the scope marker ‘kyaa’ may be distributed over multiple embedded clauses. Mahajan (2000) and Dayal (1996), both indicate that a scope marker must appear in each intermediate clause between the matrix clause and the embedded interrogative clause (ex. 4.81a). Alternatively, the embedded interrogative may occur in an embedded clause closer to the matrix clause while associating to the gap in the lower clause as illustrated in example 4.81b. The examples are based on examples presented in Dayal (1996, ex. 26, p. 70) and Mahajan (2000, ex. 9–11, p. 320).

(4.81) a. Jaun kyaa socaa ki anu kyaa kahaa ki merii-ne
J. KYAA thought that Anu KYAA said that M.-[erg]
kis-ko dekhaa?
who saw

Lit: ‘Who did John think that Anu said that Mary saw?’

b. Jaun kyaa socaa ki anu-ne kis-ko kahaa ki
J. KYAA thought that Anu-[erg] who said that
merii-ne dekhaa?
M.-[erg] saw

After having explored the data that illustrates scope marking constructions, we now summarize the following points on wh-scope markers in Hindi:

- ‘kyaa’ appears in a preverbal position in the clauses preceding the clause containing the embedded wh-phrase
- ‘kyaa’ functions as a scope marker for the embedded wh-phrase, i.e. the embedded wh-phrase is interpreted at the main clause
4. Structural variation in wh-question formation

- 'kyaa' marks the main clause as a wh-interrogative clause

On the basis of these points, we propose to assign scope marker 'kyaa' the following type in the lexicon.

\[ kyaa :: WH_l \left( s'/wh' \right), IVs, IVwh \]
where \( it_s \) and \( iv_{wh} \) are abbreviations of \( iv_s :: NOM\backslash s \)
\( iv_{wh} :: NOM\backslash wh \)

With this type-assignment of 'kyaa', we can derive scope marking constructions in Hindi along the same lines as scope marking constructions in German. The derivation in figure 4.12 illustrates the analysis of the scope marking construction:

(4.82) Jaun kyaa socaa ki kis-ko merii-ne dekhaa?
John KYAA thought that who mary-[erg] saw

'Who did John think that Mary saw?'

\[ kjaa \vdash WH_l \left( s'/wh' \right), IVs, IVwh \]
\[ socaa o (\diamond(s'/wh')) o (ki o (kis-ko o (meri-ne o dekhaa))) \vdash wh \]
\[ \vdash NOM\backslash s \]
\[ jaun o (socaa o (ki o (kis-ko o (meri-ne o dekhaa))) \vdash NOM\backslash wh \]
\[ \vdash \lambda z. \left( (\text{think} (z ((\lambda x. (\text{see} x) m)))) \right) \]

Figure 4.12: Natural deduction derivation of scope marking construction in Hindi

For more examples of grammatical and ungrammatical scope marking constructions in Hindi, look at the sample sentences listed in appendix B.6.4.

4.4.5.3 A uniform approach to scope marking constructions

The debate on the syntactic role and the semantic properties of the wh-scope marker in Hindi and other languages mainly relates to the interpretation of
4.4. Partial wh-movement

scope marking constructions in comparison to direct main clause questions. Based on the interpretation of the construction in German, some linguists have claimed that the scope-marker is an expletive (McDaniel, 1989; Mahajan, 2000), whereas the data in Hindi points into the direction of the scope marker as a ‘real’ wh-phrase (Dayal, 1994). These distinct views form the basis of two approaches, the direct dependency approach and the indirect dependency approach. In the direct dependency approach (McDaniel, 1989; van Riemsdijk, 1983) the wh-particle and the wh-phrase are linked to the same chain; the semantic scope is obtained in a similar way as with the long-distance movement constructions. This approach provides an explanation for German was-constructions, but the account does not provide a sufficient explanation for Hindi. In the indirect dependency approach (Dayal, 1994), the scope marker is seen as a wh-phrase which is indirectly linked to the complement clause. Therefore, the whole embedded interrogative is associated with the scope marker.

Recently, proposals have been made to reconcile the two approaches (Horvath, 1997; Mahajan, 2000) in a unified treatment of wh-scope markers in Hindi and German. We adhere to this unified treatment and have presented a uniform analysis for handling scope marking constructions. In both German and Hindi the scope markers are typed using the wh-type schemata. The difference with normal wh-phrases is that instead of hypothesizing over an argument position they associate to a position which is usually occupied by the complementizer ‘dass’. The role of the scope marker hypothesis is comparable with the complementizer. The scope marker hypothesis merges with an embedded interrogative clause and returns an embedded declarative clause which can be selected by the bridge verb. Depending on the structural position of the scope marker, the scope marker is merged to the structure clause-initially (German) or preverbally (Hindi).

4.4.6 Partial wh-movement in Malay

As we have illustrated in the overview of languages in section 4.1, Malay is a language that allows wh-phrases to be partially moved. Different from German and Hindi, there is no need for a scope marker in the matrix clause to mark the clause as a wh-question.

Besides partial wh-movement as a wh-question construction, Cole and Hermon (2000) provide data showing that Malay allows three kinds of constructions to form a wh-question. The wh-phrase can be left in-situ (ex. 4.83), the wh-phrase may appear in the left-most position in the matrix clause (ex. 4.84) or it may partially move (ex. 4.85) (Cole and Hermon, 2000, ex.5a,6a,7a). The square brackets in the example indicate the embedded clause.

(4.83) Ali memberitahu kamu tadi [(yang) Fatimah baca apa ]?
    A. told you just now that F. read what
    ‘What did Ali tell you just now that Fatimah was reading?’
(4.84) Siapa, Bill harap [(yang) ti akan membeli baju untuknya]?­
who B. hope that will buy clothes for him

‘Who does Bill hope will buy clothes for him?’

(4.85) Ali memberitahu kamu tadi [apa, (yang) Fatimah baca ti]?­
A. told you just now what that F. read

‘What did Ali tell you just now that Fatimah was reading?’

♦

[Sketch of the analysis]

To provide an analysis that captures the different ways to form a wh-question in Malay, thorough research should be done on the basic syntax of Malay. It is beyond the scope of this dissertation to provide such a complete analysis of Malay. Therefore, we only sketch a possible analysis for partial wh-movement in Malay based on a similar approach to German and Hindi.

The sentences in example 4.83-4.85 show that the verbs that select the embedded clause are all bridge verbs. We have shown for other languages that bridge verbs are lexically typed as verbs which select for an $s'$ typed clause. Nevertheless, similar to German and Hindi, the sentence in example 4.85 contains an embedded interrogative. The gloss of the sentence reveals that the embedded clause is not interpreted as a wh-question, but that the whole clause is interpreted as a main clause question. The sentences in examples 4.83 and 4.84 illustrate that the wh-phrase may occupy different positions in the wh-question.

To account for the different positions of the wh-phrase in the question body, we will assign multiple type-assignments to the wh-phrase. We suspect that the different types of wh-phrases may be derivable from a single type-assignment. This is left for further research.

The in-situ position of the wh-phrase in the wh-question in example 4.83 can be derived from the assignment of a wh-in-situ type. Similarly, the clause-initial position of the wh-phrase in the wh-question in example 4.84 can be derived from the assignment of a wh-ex-situ type. The more interesting case is where the wh-phrase occupies the clause-initial position in a subordinate clause as presented in example 4.85. By assigning the wh-phrase ‘apa’ ($=$ what) the following ‘pied-piping’ type, we can account for the derivation of a wh-question as presented in example 4.7.

\[
\begin{align*}
\text{wh-ex-situ ‘apa’} & : WH_X(s',s,wh) \\
\text{wh-ex-situ ‘apa’} & : WH_{X}(np,s,wh) \\
\text{partially moved ‘apa’} & : WH_C(s',s,wh)
\end{align*}
\]

We present a sketch of the derivation. Instead of using the actual Malay words, we take the glosses and shorten the sentence to the essential parts and derive ‘Ali told you what that Fatima read?’. The wh-phrase ‘what’ is first merged in clause-initial position with the embedded sentence. It turns the embedded interrogative clause into a wh-in-situ type which associates with a $s'$ hypothesis in the main clause. The subordinate clause merges with the main clause.
4.5 Concluding remarks on variation

at the position of the $s'$ gap hypothesis. After merging, the whole sentence is typed as a $wh$.

$$
\vdash_{WH_{ex}} \left( \text{that} \circ (\text{fatima} \circ \text{read}) \circ \square np \right) \vdash s' \quad [WH_{ex}']
$$

Conclusion remarks on partial wh-movement

In this section, we have concentrated on the syntactic analysis of scope marking constructions. We have shown that we can account for partial wh-movement constructions in both German and Hindi in addition to the standard account of wh-question formation in the two languages. Wh-phrases in the two languages are typed as wh-ex-situ type schemata that associate with gap hypotheses on a left branch, because the word order in both languages is assumed to be SOV. The scope markers in both German and Hindi are both typed as wh-ex-situ left type schema, but instead of reasoning over argument phrases, these elements reason over scope marker hypotheses that merge with an embedded interrogative clause. After merging, the embedded clause becomes a declarative clause type that can be selected by a bridge verb. The scope marker associates with the scope marker hypothesis and merges with the main clause at clause initial position (German) or in preverbal position (Hindi).

Additionally, we have provided a meaning assembly for each derivation using the semantic operator $\omega$. In chapter 5, we refine the meaning assembly by decomposing the macro type for wh-questions. The syntactic decomposition maps to the decomposition of the semantic $\omega$-operator. On the basis of this decomposition, we derive a meaning assembly of scope marking constructions which is the same as the meaning assembly that is computed for main clause questions.

4.5 Concluding remarks on variation

With the wh-type schemata proposed in chapter 3, we have accounted for the syntactic variation in wh-question formation. We have recognized three patterns for wh-question formation. First, wh-ex-situ where the wh-phrase appears in fronted position. Second, wh-in-situ where the wh-phrase occupies the position of the questioned constituent. Third, partial wh-movement where
the wh-phrase appears between the fronted position and the position of the gap.

To account for these three patterns, we have introduced a wh-type schema that yields a uniform semantic analysis for these different patterns. Syntactic variation is reduced to lexical variation. We will briefly summarize how the variation in wh-question formation in the three groups is captured in terms of the various instances of the wh-type schema. The different options for instantiating the wh-type schema are the distinction between wh-ex-situ and wh-in-situ type schemata, the encoding of the gap hypothesis on a left or a right branch, determining the type of the question body and the goal type of the question and finally, the categorial type of the gap hypothesis, i.e. \( np, pp, s / s \). Determining the right instantiations for these four parameters in the lexicon leads to the correct encoding of syntactic variation.

**in-situ versus ex-situ** The wh-type schema \( WH(A, B, C) \) comes in two structural variants, a wh-ex-situ type and a wh-in-situ type. Recall that the wh-ex-situ type is prefixed to a question body of type \( B \), while the wh-in-situ replaces its gap hypothesis at the same position in the question. Languages such as Japanese and Chinese, which leave their wh-phrases in-situ only use the wh-in-situ type schema. Languages, such as Serbo-Croatian and Bulgarian, that allow the wh-phrase to be fronted follow the wh-ex-situ type schema. Additionally, some languages only allow one kind of construction while other languages allow constructions to co-occur; these languages have both wh-ex-situ and wh-in-situ occurring in the lexicon.

<table>
<thead>
<tr>
<th>Wh-in-situ languages</th>
<th>( WH_{\text{in}}(A, B, C) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wh-ex-situ languages</td>
<td>( WH_{\text{ex}}(A, B, C) )</td>
</tr>
</tbody>
</table>

**left versus right branching** A second parameter for variation is encoded in the wh-ex-situ types. Wh-ex-situ types come in two variants: one type variant encodes wh-phrases whose associated gap hypothesis occurs on a left branch; the other encodes wh-phrase whose gap hypothesis occurs on a right branch. Notice again that the constituent affected by displacement is the gap hypothesis: the wh-question word itself is base generated in the fronted position.

The choice between the two types is related to the basic word order of a language. In an SOV language such as German, arguments occur on left branches. Wh-phrases in German are identified as wh-ex-situ left type schema. In SVO languages, such as Bulgarian and Serbo-Croatian, on the contrary, the subject wh-phrase is a left-ex-situ type, while the object wh-phrase is identified as a right ex-situ type.

<table>
<thead>
<tr>
<th>Wh-ex-situ</th>
<th>Subject</th>
<th>Non-subject</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bulgarian/SC</td>
<td>( WH_{\text{ex}}^{\text{SC}}(\Box \text{NOM}, B, C) )</td>
<td>( WH_{\text{ex}}^{\text{SC}}(\Box \text{CASE}, B, C) )</td>
</tr>
<tr>
<td>German/Hindi</td>
<td>( WH_{\text{ex}}^{\text{H}}(\Box \text{NOM}, B, C) )</td>
<td>( WH_{\text{ex}}^{\text{H}}(\Box \text{CASE}, B, C) )</td>
</tr>
</tbody>
</table>
4.5. Concluding remarks on variation

The left and right direction of the wh-type schema interacts with the restricted set of displacement postulates. Wh-question formation uses these universal postulates to associate the wh-phrase with its gap hypothesis.

question body type and question goal type A further type of variation that is expressed by the wh-type schema is the type of the body of the question that the wh-phrase merges with and the type of question that it yields after replacing its gap hypothesis. This kind of variation expresses the possible ways for forming a wh-question in a language. The following table illustrates how this variation is captured by specifying the type of the question body $B$ and the type of the wh-question $C$ that is yielded after the wh-phrase is merged to the question body. We generalize over the different wh-type schema that are possible.

<table>
<thead>
<tr>
<th>Question type</th>
<th>Direct questions</th>
<th>Indirect questions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Single constituent question</td>
<td>$\text{WH}(A, s, wh)$</td>
<td>$\text{WH}(A, s', wh')$</td>
</tr>
<tr>
<td>without question marker</td>
<td>$\text{WH}(A, q, wh)$</td>
<td>$\text{WH}(A, q', wh')$</td>
</tr>
<tr>
<td>with question marker</td>
<td>$\text{WH}(A, wh, wh)$</td>
<td>$\text{WH}(A, wh', wh')$</td>
</tr>
<tr>
<td>Multiple wh-question</td>
<td>$\text{WH}(A, np, \text{WH}(np, s, wh))$</td>
<td></td>
</tr>
</tbody>
</table>

Multiple wh-questions are formed through wh-phrase that have a wh-type schema whose question body type is set to $wh$. Syntactically, this wh-type schema forces the derivation of multiple wh-questions. To force multiple wh-questions in this way is unsatisfactory for two reasons. First, because we need multiple type-assignments; second, because the question goal type $wh$ does not encode a syntactic and semantic difference between single and multiple wh-questions. In chapter 5, we present a proposal that solves both problems. We show that the wh-type schema to derive single and multiple wh-questions can be derived from a single wh-type schema. The goal type for single wh-questions and multiple wh-questions will encode the difference between the two kinds of questions.

A special instantiation of the question body and the question type are formed by wh-phrases that are embedded inside a noun phrase, but as a whole behave and are interpreted as a wh-phrase. Such wh-phrases have been type as wh-phrases that occur in an $np$-typed body but yield a wh-type schema. With this type we accounted for complex noun phrase constructions in Japanese.

Additionally, we have shown that further restrictions can be added to the type of the question body. By using the derivability pattern of feature decorated types ($\Box \Diamond A \vdash A \Diamond \Box A$) we can restrict the merging of a specific wh-phrase. We have illustrated the use of the derivability pattern to account for the strict ordering relation between wh-phrases in multiple wh-questions in Bulgarian.
gap hypothesis type Another option for variation is linked to the type of the gap hypothesis. In this chapter, we have mainly concentrated on argument wh-phrases whose gap hypothesis in the wh-type schema is instantiated by an *np* type with additional features for case marking. But languages have wh-phrases that can associate with any type of gap. For instance, adjunct wh-phrases range over *A*/A or A\A gaps where *A* is any type of phrase which the wh-phrase is modifying. In other words, by instantiating the gap hypothesis in the wh-type schema one captures the syntactic role that the wh-phrase fulfills. In this way, we have accounted for the phenomenon of partial wh-movement in German and Hindi, solely by instantiating the right gap hypothesis for the scope marker in a wh-ex-situ type schema.

Concluding remarks

In this chapter, we have shown how the wh-type schemata account for the syntactic variation of wh-question formation. The specific construction of wh-question formation in a language is captured by giving the right wh-type schema for wh-phrases in the lexicon. The different instantiations of wh-type schema determine how a wh-question in a certain language or in a certain context is construed. However, additional variation in wh-question formation is due to the selectional requirements of other words that are present in the sentence. Words interact with the wh-phrases and cause variation in word order or even prevent a wh-question from being derived. For instance, we have shown that clitics may intervene the wh-cluster in Serbo-Croatian (section 4.3.4) or that verbs that select an interrogative complement create a wh-island preventing a wh-phrase from associating with its gap hypothesis as in English (see section 3.3.4.1). By studying the specific properties of constituent constructions in interaction with wh-phrases we can account for the syntactic variation across languages.

So far, we have concentrated on the syntax of wh-question formation and used a common type *wh* for all types of wh-questions along with a wh-operator *ω* to determine the lambda term of a wh-question. In the coming chapter, we will work out the semantic type for wh-questions and show that we can account for a uniform interpretation of wh-questions. We decompose the syntactic type for wh-questions *wh* to reflect its relation towards the type of answers that are requested. Along with the syntactic decomposition of the wh-question type, we also work out how the wh-operator *ω* is decomposed for each wh-phrase. With the decomposed type and a matching semantic term label for each wh-phrase, we reanalyze some of the data revealing the syntax-semantics interface of wh-questions.
Chapter 5

Syntax-semantics interface of wh-questions

In the previous chapter, we concentrated on the syntactic aspects of wh-questions. We showed how to account for cross-linguistic differences in terms of the wh-type schema proposed in chapter 3. So far, we have used \( wh \) as a temporary abbreviation for categorizing wh-questions.

In this chapter, we propose to spell out the \( wh \) type by incorporating the types of possible answers into the type assigned to wh-questions. Based on the Curry-Howard interpretation of the decomposed type, we provide the semantic term decomposition of the semantic operator \( \omega \). With the decomposed type and term, we can determine the lexical semantics of each wh-phrase. Interestingly, the syntactic and semantic decomposition of wh-phrases leads to derivability patterns between instances of wh-type schema. Syntactically, the derivability patterns between wh-type schemata make it possible to derive different types of wh-questions from a single wh-type schema. Semantically, the semantic term assigned to a wh-type schema results in a uniform meaning assembly of wh-questions. Using these patterns, we can account for multiple wh-questions and different types of answers from a single instantiation of the wh-type schema. Additionally, we provide syntactic analyses of different types of wh-questions, e.g. multiple wh-questions, embedded wh-questions, scope-marking constructions.

In section 5.1, we discuss two approaches to the semantics of questions: the proposition set approach and the structured meaning approach. We will argue in favor of a structured meaning approach. In section 5.2, after providing the type definition of wh-questions, we show how the decomposition of the macro type \( wh \) for wh-questions leads to derivability schemata between instances of wh-type schema. Finally, in section 5.3, we explore the polymorphic use of the
wh-type schema by reanalyzing some of the cross-linguistic data from chapter 4 and we show that we can account for the correct syntactic type and meaning assembly of wh-questions.

5.1 Semantics of questions and answers

The type-logical grammar approach imposes a strong correlation between the syntactic construction of an expression and its semantic interpretation. This correlation is formulated in the Curry-Howard correspondence between syntax and semantics. The correspondence stems from Frege’s principle of compositionality: The meaning of the words (= lexical semantics) and the way they are put together (= derivational semantics) determines the meaning of the whole.

Many theories that account for the semantics of questions relate the meaning of a question to its possible answers (for an overview of those theories, see Groenendijk and Stokhof (1997)). We present two approaches of relating questions and answers: the proposition set approach (Hamblin, 1958; Karttunen, 1977; Groenendijk and Stokhof, 1984) in which questions represent propositions; and the approach which Krifka (2001) named the structured meaning approach, also referred to as the functional or categorial approach (Groenendijk and Stokhof, 1984). In this latter approach, the interrogative in combination with its answer forms a statement.

5.1.1 Proposition set approach

The proposition set approach was formulated by Hamblin (1958), who wrote a set of principles that largely influenced the logical approach to the semantics of questions (Karttunen, 1977; Groenendijk and Stokhof, 1984). With these principles Hamblin (1958) captures the main idea behind a logical approach to interrogatives, which states that to determine the meaning of an interrogative one has to inspect what kind of statement can serve as a response. The first of Hamblin’s principles was paraphrased by Groenendijk and Stokhof (1997) as “an answer to a question is a sentence, or statement”. The principle implements the idea that the semantic status of an answer is a proposition and that the syntactic form of an answer is irrelevant.

Answers as statements  Constituent questions such as ‘What did Mary read?’ or ‘Who read The Minimalist Program?’ receive a full statement as an answer. A likely answer to either question could be ‘Mary read The Minimalist Program’. Answers which only give a single constituent as an answer (‘The Minimalist program’ or ‘Mary’) are said to be derived from the full statement. Thus, the underlying structure of an answer has to be a complete statement with the requested argument embedded in it.
The semantics of questions in the proposition set approach are represented in variants of predicate logic. Adopting the presentation style of Krifka (2001), we illustrate the semantics of a question ‘Who read The Minimalist Program?’ in predicate logic.

**Question:** ‘Who read The Minimalist Program?’

\[
\{ p \mid \exists x. (\text{person } x) \land p = ((\text{read mp }) x)\}
\]

\[
= \{(\text{read mp }) x | (\text{person } x)\}
\]

\[
= \{(\text{read mp }) m, (\text{read mp }) j, \ldots \}
\]

**Answer:** ‘John read The Minimalist Program.’

\[
= (\text{read mp }) j
\]

The answer to such a question is represented by the set of propositions \( p \) that are restricted by the predicates that have to hold for this proposition. For example, there is a person for which it is true that this person read the Minimalist Program. By analyzing the restrictions as part of the proposition, the answer yields all interpretations that make such a proposition true.

The main critique of the proposition set approach is that it lacks expressibility. Because Hamblin (1958) chooses to ignore the syntactic form of the answer and treats all answers as full statements, the proposition set approach cannot directly relate a certain type of answer to a certain type of question. Krifka (2001) notes that questions with alternative answers or multiple constituent questions cannot be handled easily in the proposition set approach. The dominant view is that questions belong to a single category and should therefore be mapped to one type. This stands in contrast to the view which is reflected by the structured meaning approach, which takes a polymorphic stance towards the categorial nature of questions.

### 5.1.2 Structured meaning approach

The structured meaning approach is sometimes referred to as the *functional* or *categorial* approach. The approach is developed by logicians and semanticists and supports the idea that the meaning of a question is dependent on the meaning of the answer and vice versa. Along similar lines, Hız (1978) points out that questions and their answers are not autonomous sentences, but that they form a semantic unit — a question-answer pair. We briefly discuss the structured meaning approach and its syntactic consequences.

**Constituent answers** An appropriate answer to a single constituent question may be any type of syntactic object. This might be a generalized quantifier phrase or a verb phrase, as well as a noun phrase or prepositional phrase. Additionally, in multiple wh-questions, different combinations of syntactic objects can be used as an answer. The wh-question directs the kind of answers that can be expected.

(5.1)  

a. ‘Who saw Mary?’ John, nobody, John’s sister, …

b. ‘When did John see Mary?’ In the afternoon,
5. Syntax-semantics interface of wh-questions

c. ‘Which man did John see?’ His father, the neighbor, …

d. ‘Why did John see Mary?’ Because …

e. ‘Who saw whom?’
   pair list reading: John (saw) Bill, Mary (saw) Sue, …
   functional reading: every professor/his student, John/his sister

As the sentences illustrate, the answers have a direct relation to the interrogative phrase in the question. To capture the relation between the question and its possible answer type, the structured meaning approach formulates the idea that the question and answer form a unit, both syntactically and semantically. Syntactically, the interrogative in combination with its answer forms an indicative sentence or a question-answer sequence. This syntactic unit is reflected in the semantics where the question meaning is a function that yields a proposition when applied to the meaning of answer (Krifka, 2001).

In the proposition set approach, answers belong to a uniform type of syntactic object viz. s, whereas the answers in the structured meaning approach may fall into any kind of category. A syntactic analysis must formulate a way to account for the possible variation in question-answer pairs. Along similar lines, Groenendijk and Stokhof (1984) have argued that a proper theory of the semantics of questions should not focus on one type for wh-questions, but define a multitude of types and determine the semantic relationship between them. Groenendijk and Stokhof (1997) refer to this view as “the polymorphic stance”.

Within the type-logical grammar framework, a polymorphic view on question and answer types comes quite naturally, as shown in work of Hausser (1983) and more recently in Bernardi and Moot (2003). These works show that the diversity in answer types can also be derived from uniformly typed wh-phrases. We will follow this line of “derivational polymorphism” and show that by incorporating the answer type into the type assigned to wh-phrases, we can account for different question-answer sequences on the basis of a single type for wh-questions.

5.2 Question and answer types

In chapter 3, we assigned type wh to wh-questions as a temporary abbreviation. In a structured meaning approach, questions are expected to be functions that, when applied to an answer, yield a proposition. In this section, we will spell out the wh abbreviation in such a way that it reflects the functor-argument relation between a wh-question and its response. In section 5.2.1, we determine how this relation is captured in the syntactic and semantic type definition of wh-questions. The next step is to determine the lexical semantics of wh-phrases. In section 5.2.2, we work out the lambda term for the ω operator that we have used so far as a semantic operator for wh-questions. In section 5.2.3, we will show that spelling out the wh-question type along
with the meaning assembly for the $\omega$-operator leads to derivability patterns between instances of the wh-type schema. These patterns follow from characteristic type-shifting laws that hold in the semantic type language.

### 5.2.1 Type definition of wh-questions

So far, we have assumed wh-questions to be expressions of type $wh$. Adopting a structured meaning approach of questions, we will incorporate the type of possible answers into the type of the wh-question. Generalizing over the possible types of answers and questions, we decompose the type abbreviation $wh$ of wh-questions into the following type:

**Definition 25** *Decomposition of wh-question type wh*

<table>
<thead>
<tr>
<th>Syntactic type</th>
<th>Semantic type</th>
</tr>
</thead>
<tbody>
<tr>
<td>$B/\gamma A$</td>
<td>$A \rightarrow B$</td>
</tr>
</tbody>
</table>

The semantic type $A \rightarrow B$ is a direct mapping from the components of the syntactic type $B/\gamma A$. $A$ is the semantic type of category $A$ which is the type of the expected answer. $B$ is the semantic type of category $B$ which is the type of the question-answer sequence.

Notice that the type connective has an additional index $\gamma$. We use this index to capture a compositional difference between predicates and arguments on a sentential level (structural composition relation: $\circ$) and between questions and answers on a dialogue level (structural composition relation: $\circ_\gamma$). Following the structured meaning approach, we assume question-answer sequences to form a syntactic and semantic unit. Syntactically, we assume the question-answer sequence to belong to category $s$.\(^1\) Semantically, the question-answer sentence is a proposition which has a certain truth value, similar to declarative clauses. Now that we have defined the general type schema for wh-questions, we continue to illustrate how such a type for wh-questions can account for different types of question-answer combinations.

In most languages, wh-questions can range over any type of answer depending on the kind of wh-phrase. For instance, in English, a single constituent question with ‘who’ or ‘what’ requires a referential noun phrase (*John, the man, him*) or a quantified noun phrase as an answer (*everyone, every man*). Whereas wh-questions with adjunct wh-phrases such as ‘where’ or ‘when’ require a locative or temporal modifier as response. Additionally, we can also have multiple wh-questions that require pair-list answers. As these examples illustrate, both the answer type $A$ and the question type $B$ in the type definition of wh-questions (see definition 25) may be complex categories. As an

\(^1\)For present purposes, we treat declarative clauses and question-answer sequences as syntactically uniform. Further research is needed to discover how the syntactic distinction between different statements in a discourse can be captured in a type-logical grammar approach.
example, we list a number of possible types of wh-questions with corresponding semantic types.

\[
\begin{align*}
\text{‘Which man . . . ’} & \quad s/np \\ 
\text{‘Who . . . ’} & \quad s/(s/(np\backslash s)) \\ 
\text{‘When . . . ’} & \quad s/(iv\backslash iv) \\ 
\text{‘Who . . . what’} & \quad (s/np)/np
\end{align*}
\]

\[
e \rightarrow t
\]

Sample derivation To illustrate how a wh-question combines with an expression with a matching type, we derive the constituent question ‘Who saw Mary?’. This question can be answered with a higher-order typed answer, for instance the generalized quantifier ‘Nobody’. The constituent question has goal type \(s/(s/(np\backslash s))\). The required answer type is the higher-order type \(s/(np\backslash s)\) (abbreviated to \(gq\)) which maps to semantic type \((e \rightarrow t) \rightarrow t\). ‘Nobody’ can be used as an answer because it matches the required type. The wh-question selects its answer and the combination yields a question-answer sequence of type \(t\) with semantic type \(t\). The derivation step of merging a wh-question ‘Who saw Mary?’ with a possible answer ‘Nobody’ is as follows.

\[
(\text{Who} \circ (\text{see} \circ \text{Mary})) \vdash s/gq \quad \text{Nobody} \vdash gq
\]

\[
\frac{(\text{Who} \circ (\text{see} \circ \text{Mary})) \circ (\text{Nobody}) \vdash s}{[E]}
\]

Now that we know the syntactic and semantic type of wh-questions, let us see how this type determines the meaning assembly of wh-questions.

5.2.2 Meaning assembly of wh-questions

In chapter 3, we introduced the semantic operator \(\omega\) to represent the meaning assembly of any type of wh-question. The function of \(\omega\) was to bind the gap hypothesis in the semantic term of a wh-question. We repeat the inference rule along with the meaning assembly that accounted for merging a wh-type schema with a question body:\(^2\)

\[
\begin{array}{c}
\Gamma \vdash \omega : WH(A, B, B' / \backslash A) \\
\Delta[x : A] \vdash B \\
\Delta[\Gamma] \vdash (\omega \lambda x. t) : B'/\backslash A' \\
\end{array}
\]

\[
[WH]
\]

The inference rule shows that the merge step is both an abstraction step of the gap hypothesis as well as an application step for merging the wh-phrase to the body of the question. After merging the wh-phrase we obtain a wh-question of type \(B'/\backslash A'\), which is the decomposed type of wh-questions. Notice that the gap hypothesis \(A\) and the answer type \(A'\) are related. As we will show, the answer type is derivable from the type of the gap hypothesis \((A \vdash A')\).

\(^2\)We abstract here over the different operations that have been proposed for the different versions of the wh-type schemata. As explained in chapter 2, structural differences are not reflected in the semantic representation.
This relation between the answer type and the gap hypothesis must also be encoded in the meaning assembly of wh-questions. We have used $\omega$ to capture the meaning assembly of a wh-question. Similar to the decomposition of the wh-question type, we will see how the $\omega$-operator can be captured as a lambda term.

The precise meaning representation of a wh-question depends, however, on the kind of wh-phrase that forms a wh-question. We argue that, at least for argument wh-phrases, different wh-type schema each can be derived from a single wh-type schema. The basic case for wh-phrases is a wh-type schema that ranges over higher-order typed answers: $\text{WH}(np, s, s/gq)$. The $\omega$-operator that captures the meaning assembly of this wh-type schema can be regarded as a logical constant. The definition of the $\omega$-operator generalizing over different types of wh-phrases is as follows:

**Definition 26** Semantic term decomposition of wh-operator $\omega$

\[
\omega = \lambda P^{\tau_1 \rightarrow \pi}. \lambda Q^{(\tau_2 \rightarrow \pi) \rightarrow \pi}. (Q P)
\]

From the basic wh-type schema $\text{WH}(np, s, s/gq)$, we can derive instances of the wh-type schema whose meaning assembly can be determined by filling in the meaning assembly of $\omega$. Before we show how instances of the wh-type schema are derived, we first illustrate that the meaning assembly of the $\omega$-operator does indeed yield the right meaning assembly for the wh-question ‘Who saw Mary?’. 

**Sample derivation** In a single constituent question, ‘who’ is merged with a question body of type $s$ and associates with a $np$ typed gap hypothesis. After merging ‘who’ with the question body, the wh-phrase replaces the gap hypothesis in the question body. The sentence becomes of type $s/(s/(np \backslash s))$, a sentence which is incomplete for an answer of type $s/(np \backslash s)$. For ease of exposition, we abbreviate $s/(np \backslash s)$ to $gq$. The syntactic type and the lexical meaning assembly of the wh-phrase ‘who’ is:

\[
\text{who} \vdash \lambda P^{(et)}. \lambda Q^{(et)}. (Q P) : \text{WH}(np, s, s/gq)
\]

The following derivation illustrates that this semantic term assignment to ‘who’ derives the right meaning assembly for a wh-question ‘Who saw Mary?’:

\[
\begin{align*}
\lambda P \cdot \lambda Q (Q P) : & \text{who} \\
\text{WH}(np, s, s/gq) & \xrightarrow{x : np \quad \text{see} : \text{saw} \quad \text{m} : \text{mary}} \ (np \backslash s) / np \\
\text{[/E]} & \xrightarrow{x : np \circ (\text{saw} \circ \text{mary})} \ ((\text{see m}) x) : s \\
\text{[\WH]} & \xrightarrow{\text{who} \circ (\text{saw} \circ \text{mary})} s / gq \\
\lambda Q (Q \lambda x. ((\text{see m}) x)) & \end{align*}
\]
To derive a question-answer sequence the derivation continues as illustrated in section 5.2.1. The answer ‘nobody’ has the semantic term and syntactic type-assignment: \( \lambda P \rightarrow t. \neg \exists \lambda y. (P y) : s/(\eta s) \). We can derive the following meaning assembly for the question-answer sequence ‘Who saw Mary? Nobody’:

\[
\text{Who saw Mary?} \quad \lambda Q. (Q \lambda x. ((\text{see } m) x)) : s/s/\eta s \quad \text{Nobody} \quad \lambda P. \neg \exists y. (P y) : s/(\eta s) \quad [E]
\]

\[
\text{Who saw Mary? Nobody} \vdash s
\]

\[
(\lambda P. \neg \exists y. (P y) \lambda x. ((\text{see } m) x))
\]

\[
\sim \exists y. \lambda y. ((\text{see } m) y)
\]

On the basis of the simplest case for argument wh-phrases, we can derive different instances of the wh-type schema. Using the logical constant \( \omega \), we can determine how the meaning assembly is changed accordingly.

### 5.2.3 Derivability patterns of wh-type schemata

Incorporating the answer type into the wh-type schema enables us to derive different instances of the wh-type schema from a single wh-type schema. For instance, we will show that we can account for multiple wh-questions. The derivation relations between different instances of wh-type schema can be described as a derivability pattern, similar to the derivability pattern of unary feature decorated sentence types \( \Diamond s \vdash s \vdash \Box s \).

The derivability pattern is based on characteristic laws in semantic type language.

#### 5.2.3.1 Semantic derivability

The derivability pattern of wh-type schemata is based on three theorems that are derivable in semantic type language: type-lifting, geach and exchange. We illustrate each rule in semantic type language and present the meaning assembly for each type-shifting rule.

\[\text{[type-lifting]} \quad A \vdash (A \rightarrow B) \rightarrow B \quad x \mapsto \lambda y. (y x)\]

\[\text{[geach]} \quad B \rightarrow A \vdash (C \rightarrow B) \rightarrow (C \rightarrow A) \quad x \mapsto \lambda y \lambda z. (x (y z))\]

\[\text{[exchange]} \quad C \rightarrow (D \rightarrow E) \vdash D \rightarrow (C \rightarrow E) \quad x \mapsto \lambda z \lambda y. ((x y) z)\]
Using these theorems, we can derive two additional laws argument lowering and dependent geach.

**argument lowering** The type-lifting rule shows how an arbitrary type $A$ is lifted to a type $(A \to B) \to B$. In the previous section, we have illustrated that type lifting alters the answer type to fit the answer type requested by the wh-question. From the type-lifting rule, we can derive the rule for argument lowering. Argument lowering applies to the dependent instead of the main type. Each type can be a complex type.

\[
\text{[argument lowering]} \quad ((A \to B) \to B) \to C \quad \vdash \quad A \to C \\
\quad x \mapsto \lambda y.(x \lambda z.(z y))
\]

**dependent geach** The geach rule adds an additional dependent to both the main clause type $A$ and its argument type $B$. Again, each type may be a complex type. The exchange rule captures the reordering of two dependents. From the combination of geach and exchange, we can derive another valid type shifting rule which we will refer to as dependent geach because the geach rule applies to the dependent of a complex type. The geach rule is now applied to a complex type $(D \to E) \to A$. The geach rule alone would change this type into a complex type $(C \to (D \to E)) \to (C \to A)$. Additionally, we apply exchange to the result of the geach type. We obtain the following type-shifting rule.

\[
\text{[dependent geach]} \quad (D \to E) \to (B \to A) \quad \vdash \quad (D \to (C \to E)) \to (B \to (C \to A)) \\
\quad x \mapsto \lambda z.\lambda y.\lambda v.((x \lambda u.((z u) v)) y)
\]

### 5.2.3.2 Syntactic derivability

The theorems in the semantic type language reveal that under certain assumptions a number of type alterations are also derivable in the syntactic formula language. Let us investigate under what assumptions argument lowering and dependent geach are derivable in the grammatical reasoning system. We will illustrate the derivability of each type-shifting rule by presenting the syntactic derivation of a specific wh-type schema. Along with the syntactic derivation, we can show how the meaning assembly of the wh-phrase changes using the $\omega$-operator as a logical constant.
Argument lowering Argument lowering is derivable in the grammatical reasoning system without making any structural assumptions. As an example, we apply argument lowering to the subject wh-phrase ‘who’ and show how the meaning assembly is changed.

‘Who’ is typed in the lexicon as WH(np, s, s/sgq). The wh-type schema associates with an np gap hypothesis in a s-typed question body and yields a wh-question of type s/sgq, which requires an sgq-typed answer. The following derivation illustrates how we derive an argument lowered answer type for the subject wh-phrase ‘who’. We show how argument lowering can be derived for the wh-type schema of ‘who’. For a clearer presentation, we use the decomposed type for the wh-type schema (see chapter 3).

who :: WH(l ∅ (np, s, s/sgq)) = (s/sgq)/(np\s)

For the lexical semantics of ‘who’, we use the logical constant ω. The hypotheses are each decorated with term variables. In the derivation, each syntactic step reveals the meaning assembly.

\[
\begin{align*}
\text{who} & \vdash \omega : (s/sgq)/(np\s) \\
\text{who} \circ \text{np}\s & \vdash (\omega \text{ Q}) : s/sgq \\
\text{who} \circ \text{np}\s & \vdash \lambda x.((\omega \text{ Q}) \lambda P.(P x)) : s/np\s \\
\text{who} & \vdash \lambda Q.\lambda x.((\omega \text{ Q}) \lambda P.(P x)) \ (s/np\s)/(np\s) \\
\end{align*}
\]

Using the definition of the logical constant ω, we can compute the meaning assembly for the argument lowered type. In the semantic term that is computed after argument lowering (step 1), we substitute the semantic term of the ω-operator of definition 26 (step 2). After several β-reductions (step 3) and a single η-reduction (step 4), we obtain a term that is equal to the identity function.

\[
\begin{align*}
\lambda Q.\lambda x.((\omega \text{ Q}) \lambda P.(P x)) & \equiv (1)
\end{align*}
\]

Schematically, the derivability pattern for argument lowering can be captured in the following derivability relation:

\[
\begin{align*}
\text{WH}(np, s, s/sgq) & \vdash \text{WH}(np, s, s/np) \\
\omega & \mapsto \lambda Q.\lambda x.((\omega \text{ Q}) \lambda P.(P x))
\end{align*}
\]
Dependent Geach  Unlike argument lowering, dependent Geach is not freely derivable in the grammatical reasoning system. Moortgat (1997) shows that the Geach law is derivable in the syntactic formula language with the addition of associativity. The restricted set of displacement postulates provides the necessary restructuring mechanisms for deriving Geach.

To use the postulates specifically for the derivation of Geach variants of the wh-type schema, we must allow the postulates to reason over the structural binary operator $\circ$. Figure 5.1 presents the alternations of displacement postulates that the structural module needs in order to reason over answer hypotheses of ex-situ wh-phrases. Similarly, we must add alternations to the postulates that underlie wh-in-situ phrases (see appendix A for an overview of the in-situ postulates).

$$\begin{align*}
\Gamma[\langle \, \Delta_1 \circ \Delta_2 \rangle] \vdash C & \quad \text{[Pl1]} \\
\Gamma[\langle \, \Delta_1 \circ (\Delta_2 \circ \Delta_3) \rangle] \vdash C & \quad \text{[Pr1]} \\
\Gamma[\langle \, \Delta_1 \circ (\Delta_2 \circ \Delta_3) \rangle] \vdash C & \quad \text{[Pl2]} \\
\Gamma[\langle \, \Delta_1 \circ (\Delta_2 \circ \Delta_3) \rangle] \vdash C & \quad \text{[Pr2]}
\end{align*}$$

Figure 5.1: Restricted set of displacement postulates where $i \in \{\emptyset, \?\}$

With this change in the structural module, we can derive dependent Geach variants for wh-ex-situ and wh-in-situ types. Schematically, the geach variant of a wh-type schema can be captured by the following type change:

$$\text{WH}(A, B, B' / \gamma A') \vdash \text{WH}(A, B / \gamma C, (B' / \gamma C) / \gamma A')$$

The question body type $B$ and the goal type $B' / \gamma A'$ have an additional dependent type $C$ in the dependent geach type. We will use this type to capture the dependency of a wh-phrase on the occurrence of another wh-phrase, i.e. multiple wh-questions. We will show how we can derive the dependent geach type syntactically for both the ex-situ and the wh-in-situ type schema.

wh-ex-situ  The wh-ex-situ type schema $\text{WH}_r(\langle \, \diamond \mathit{np}, s, s / \mathit{gq} \rangle)$ is a basic wh-type schema which has the semantic term assignment captured by the logical operator $\omega$. We will show that we can derive a geach variant of this wh-type schema and present the meaning assembly accordingly. The meaning assembly of the derived type is computed substituting the $\omega$-operator in the term that is derived for the geach type.

Again, for a clearer presentation, we decompose the type into the usual logical connectives of the type-logical grammar system. Additionally, we add subscripts $1$ and $2$ to distinguish between the $\mathit{gq}$ answer type ($\mathit{gq}_1$), which is part of the wh-type schema, and the answer type($\mathit{gq}_2$), which is added through the derivation of the geach type. In the derivation, we abbreviate the structural occurrence of the wh-type schema to ‘wh$_{ex}$’.
\[ \text{wh}_{\text{ex}} \vdash \text{WH}'_{\text{ex}}(\diamondsuit np, s, s/\Diamond gq_1) = (s/\Diamond gq_1)/(s/\Diamond np) \]

Figure 5.2 illustrates the derivation of the geach variant of the wh-ex-situ type. Along with the syntactic derivation, we compute the meaning assembly for each derivation step.

\[
\begin{align*}
\frac{\left[ P : (s/\Diamond gq_1)/(\Diamond np) \right]}{(s/\Diamond gq_1)/(\Diamond np) \circ \Diamond np \vdash (P z) : s/\Diamond gq_2} & \quad \frac{\left[ Z : (\Diamond np) \right]}{(s/\Diamond np) \circ \Diamond np \vdash (P z) : s/\Diamond gq_2} \quad \frac{\left[ R : \Diamond np \right]}{\Diamond np \vdash s/\Diamond gq_2} \quad \frac{\left[ \omega \right]}{(s/\Diamond np) \circ \Diamond np \vdash \lambda z.(P z) : s/\Diamond np} \\
\frac{(s/\Diamond gq_1)/(\Diamond np) \circ \Diamond np \vdash (P z) : s/\Diamond gq_2} & \quad \frac{(s/\Diamond np)/(\Diamond np) \circ \Diamond np \vdash (P z) : s/\Diamond gq_2} \quad \frac{(s/\Diamond np)/(\Diamond np) \circ \Diamond np \vdash (P z) : s/\Diamond gq_2} \quad \frac{(s/\Diamond np)/(\Diamond np) \circ \Diamond np \vdash (P z) : s/\Diamond gq_2} \quad \frac{(s/\Diamond np)/(\Diamond np) \circ \Diamond np \vdash (P z) : s/\Diamond gq_2} \\
\frac{(s/\Diamond gq_1)/(\Diamond np) \circ \Diamond np \vdash \lambda ^{1}.(P z) : s/\Diamond np} & \quad \frac{(s/\Diamond np)/(\Diamond np) \circ \Diamond np \vdash \lambda ^{1}.(P z) : s/\Diamond np} \quad \frac{(s/\Diamond np)/(\Diamond np) \circ \Diamond np \vdash \lambda ^{1}.(P z) : s/\Diamond np} \quad \frac{(s/\Diamond np)/(\Diamond np) \circ \Diamond np \vdash \lambda ^{1}.(P z) : s/\Diamond np} \quad \frac{(s/\Diamond np)/(\Diamond np) \circ \Diamond np \vdash \lambda ^{1}.(P z) : s/\Diamond np}
\end{align*}
\]

\[
\frac{\omega : (s/\Diamond gq_1)/(s/\Diamond np)}{(s/\Diamond gq_1)/(\Diamond np) \circ \Diamond np \vdash (\omega \lambda z.(P z) : s/\Diamond gq_2)) = (s/\Diamond np)/(s/\Diamond np)} \quad \frac{Q : (s/\Diamond np)/(s/\Diamond np)}{(s/\Diamond gq_1)/(\Diamond np) \circ \Diamond np \vdash (\omega \lambda z.(P z) : s/\Diamond np) = (s/\Diamond np)/(s/\Diamond np)}
\]

Figure 5.2: Derivation of geach variant of wh-ex-situ type

The result is a dependent geach variant for \( \text{wh}_{\text{ex}} \) which can be rewritten as a wh-type schema:

\[
(s/\Diamond np)/(\Diamond np) = \text{WH}'_{\text{ex}}(\Diamond np, s, s/\Diamond gq_1)/(s/\Diamond np)
\]

Similar to the derivation of the argument lowering type, we determine the meaning assembly of the derived instance of the wh-type schema by substituting the term definition of the \( \omega \)-operator\(^3\) in the obtained term (step 2). After \( \beta \)-reduction, we obtain the term of the geach wh-type schema.

\[
\begin{align*}
\lambda P.\lambda Q.\lambda R.((\omega \lambda z.(P z) : s/\Diamond gq_2)) & \quad \frac{(1)}{\lambda P.\lambda Q.\lambda R.((\lambda P'\lambda Q'.(Q' P') \lambda z.(P z) : s/\Diamond gq_2))} \\
\sim_{\gamma_{\omega}} & \quad \frac{(2)}{\lambda P.\lambda Q.\lambda R.((\gamma_{\omega} P' \lambda Q'.(Q' P') \lambda z.(P z) : s/\Diamond gq_2))} \\
\sim_{\gamma_{\beta}} & \quad \frac{(3)}{\lambda P.\lambda Q.\lambda R.((\gamma_{\beta} \lambda Q'.(Q Q' \lambda z.(P z) : s/\Diamond gq_2)))}
\end{align*}
\]

\(^3\)See definition 26 on page 159
5.2. Question and answer types

The derivation of a geach type also extends to the left-ex-situ variant and to the argument lowered variant of both ex-situ wh-type schemata. For the lowered geach type variant, we compute the semantic term using the term for the lowered wh-type schema ($ω_{id} = λQ.λz.((P y) x)$). We can compute the following term assembly for an ‘argument lowered geach’ wh-ex-situ type.

\[
\text{WH}_{ex}(♦\square np, s/s\triangleleft np_1) ⊢ \text{WH}_{ex}(♦\square np, s/np_2, (s/♦\square np_2)/\triangleleft np_1)
\]
\[
ω_{id} \mapsto λP.λx.λy.((ω_{id} λz.((P y) x)) z)
\]

This term can be further reduced by substituting the semantic term for the lowered wh-ex-situ type for $ω_{id}$.

\[
\begin{align*}
\lambda P, λx.λy.((ω_{id} λz.((P y) x)) z) \\
\sim_{[ω_{id}]} \lambda P, λx.λy.((λQ.λz.((P y) x)) z) \\
\sim_{β'} \lambda P, λx.λy.((P y) x)
\end{align*}
\]

wh-in-situ  In a similar fashion to the wh-ex-situ types, we can derive dependent geach types for the wh-in-situ type schema. The syntactic proof for deriving a geach type for a wh-in-situ wh-phrase proceeds along similar lines as the proof for the wh-ex-situ type schema.

Allowing the restricted set of structural rules underlying the wh-in-situ type schema to apply over the ◦? composition, we derive the following derivability patterns of wh-in-situ type schema. The semantic terms that are computed can be further reduced using term reduction.

\[
\begin{align*}
\text{WH}_{in}(np, s/s\triangleleft gq_1) ⊢ \text{WH}_{in}(np, s/s\triangleleft gq) \\
ω \mapsto λQ.λz.((ω Q) λz.((P x) x))
\end{align*}
\]
\[
\begin{align*}
\text{WH}_{in}(np, s/s\triangleleft np) ⊢ \text{WH}_{in}(np, s/np, (s/np)/\triangleleft np) \\
ω_{id} \mapsto λP, λx.λy.((ω_{id} λz.((P y) x)) z)
\end{align*}
\]
\[
\begin{align*}
\text{WH}_{in}(np, s/s\triangleleft gq_1) ⊢ \text{WH}_{in}(np, s/gq_2, (s/gq_2)/\triangleleft gq_1) \\
ω \mapsto λP.λQ.λR.((ω λz.((P z) R)) Q)
\end{align*}
\]

5.2.3.3 Overview

We have shown that the two theorems of argument lowering and dependent geach are syntactically derivable in the grammatical reasoning system. Applying the two rules on different instances of the wh-type schema gives us the following derivability patterns between instances of wh-type schema. In figure 5.3, the
Figure 5.3: Derivability pattern of wh-ex-situ schema

Figure 5.4: Derivability pattern of wh-in-situ schema

Figure 5.5: Meaning assembly of derivability patterns
5.3 Linguistic application

The syntactic decomposition of wh-question types into types that are part of an question-answer sequence adds polymorphism to the wh-type schemata. The semantic representation of wh-questions reflects the question’s requirement for certain types of answers. We have shown that two theorems of argument lowering and dependent geach leads to a derivability pattern of wh-type schema. In this section, we continue to explore the syntactic and semantic aspects of this polymorphism for wh-question formation. The cross-linguistic data that has been explored in chapter 4 provides support for our hypothesis that the wh-type schema accounts for a uniform interpretation of wh-questions, but also that a polymorphic stance is needed to account for variation in question-answer sequences. We show that the derivational possibilities of the types largely depend on the right encoding of the wh-type schema for wh-phrases.

In section 5.3.1, we focus on the derivation of single constituent questions in English. We discuss the syntactic and semantic consequences of argument lowering for the derivation of question-answer sequences in local and non-local wh-questions. In section 5.3.2, we discuss multiple wh-questions in English, Serbo-Croatian and Japanese. We show that we can account for the derivation of multiple wh-questions on the basis of deriving geach types for both ex-situ and in-situ type schema. And as a result derive the correct meaning assembly of multiple wh-questions. Additionally, in section 5.3.3 and 5.3.4, we discuss the semantic representation of two special instances of wh-type schema that were syntactically explored in chapter 4. In section 5.3.3, we explore pied-piping constructions in English and show that the analysis applies similarly to complex NP constructions in Japanese. In section 5.3.4, we discuss scope marking constructions and show that the decomposition of wh-question types gives a direct question interpretation to such questions.
5. Syntax-semantics interface of wh-questions

5.3.1 Single constituent questions

A single constituent question requires a single constituent answer. We concentrate here on argument wh-phrases to illustrate the relation between a wh-question and possible answers. We first look at direct questions where the associated gap hypothesis appears in the local domain. We will furthermore illustrate the contrast between wh-pronouns and wh-determiners. Subsequently, we discuss the meaning assembly of indirect wh-questions. Lastly, we briefly show how the meaning assembly of non-local wh-questions does not differ from the meaning assembly of local wh-questions.

5.3.1.1 Direct questions

In a direct question in English, a fronted wh-phrase associates with a np gap hypothesis. The expected answer, however, differs on the wh-phrase. Wh-questions with argument wh-phrases ‘what’ or ‘who’, expect either a referential or a quantified noun phrase. Wh-questions with which-determiners only expect a referential noun phrase as an answer. On the basis of the derivability pattern of wh-ex-situ types we can account for the distinction between the two types of wh-phrases. First, we discuss the lexical type-assignment of wh-pronouns. Then, we present the contrast with wh-determiners.

Wh-pronouns A suitable answer to a wh-question such as ‘Who saw Mary?’ might be a referential noun phrase e.g. ‘John’, as well as a generalized quantifier phrase e.g. ‘everyone’. To allow both types of answers, ‘who’ and ‘whom’ are assigned the following wh-type schema in the lexicon.

who :: λP.λQ.(Q P) : WH(np,s,s/(np\$s))
whom :: λP.λQ.(Q P) : WH(ex\$\triangleright np,q,s/s/(np\$s))

We use the following lexical entries to illustrate the syntactic derivation and meaning assembly of direct questions such as: ‘Who saw Mary?’ and ‘Whom did John see?’.

john, mary :: j.m : np
saw :: λx.λy.(see x y) : (np\$s)/np
see :: λy.(see y) : inf/np
did :: λP.(\pi 1 P \pi 1 P) : q/(np • inf)
every man :: λP.∀λy.(man y) \rightarrow (P y) : s/(np\$s)
some woman :: λP.∃λx.(woman x) \wedge (P x) : s/(np\$s)

In section 5.2.1 we showed that a generalized noun phrase matches the required type of the wh-question directly. For a definite noun phrase with category np to be used as an answer, the type needs to be lifted to s/(np\$s). We argued that type-lifting can be derived in the grammatical reasoning system without making any further structural assumptions.
The following derivation illustrates how a higher-order type for the noun phrase *John* can be derived. The derived type matches the required type of the wh-question and this is also reflected in the meaning assembly of the question-answer sequence.

Who saw Mary?  
\[ \text{Who saw Mary?} \vdash \lambda \text{P} \cdot (\text{P} \text{ }} \text{John} ) : s / (n p \backslash s) \]

\[ \text{Who saw Mary?} \vdash \lambda Q^{\text{et}}. (Q \text{ }} \text{John} ) : s / (n p \backslash s) \]  

The sentences in example 5.2 and 5.3 present an overview of the different kinds of question-answer sequences that can be derived using the given type-assignments for wh-pronouns. The type that is derived for subject or non-subject wh-questions is a *s*-typed clause which is incomplete for a higher-order typed *np, (s/(np\backslash s)). A generalized quantifier phrase can be merged directly, while referential noun phrases such as 'John' and 'Mary' in example 5.2b and 5.3b have to be lifted before they can be merged. Along with the syntactic type, lifting alters the semantic type of the answer in such a way that the lifted type matches the semantic type requested by the interrogative clause. The semantic term is computed as usual.

(5.2) Who saw Mary? \vdash \lambda Q^{\text{et}}. (Q \lambda x.((\text{see m}) x)) : s / (s / (n p \backslash s))

a. Answer: 'every man' \vdash \lambda P \cdot \forall y.((\text{man } y) \rightarrow (P y)) : gq  
   Meaning assembly: \( \forall y.((\text{man } y) \rightarrow ((\text{see m}) y)) \)

b. Answer: 'John' \vdash \lambda P \cdot (P \text{ }} \text{John} ) : s / (n p \backslash s)  
   Meaning assembly: \( (\lambda P (P j) \lambda x.((\text{see m}) x)) \)  
   \( \sim_j (\lambda P (P m) \lambda y.((\text{see y}) j)) \)   

(5.3) Who(m) did John see \vdash \lambda Q^{\text{et}}. (Q \lambda y.((\text{see y}) j)) : s / (s / (n p \backslash s))

a. Answer: 'some woman' \vdash \lambda P \cdot \exists x.((\text{woman } x) \land (P x)) : gq  
   Meaning assembly: \( \exists x.((\text{woman } x) \land ((\text{see } x) j)) \)

b. Answer: 'Mary' \vdash \lambda P \cdot (P m) : n p  
   Meaning assembly: \( (\lambda P (P m) \lambda y.((\text{see y}) j)) \)   
   \( \sim_j (\lambda P (P m) \lambda y.((\text{see y}) j)) \)
Wh-determiners Suitable answers to wh-questions that are built with wh-determiners like ‘which’ are restricted to definite noun phrases. The semantic difference between wh-phrases and wh-determiners lies in the specific denotation of the which-phrases. As Pesetsky (1987) notes in his paper, ‘which phrases’ are Discourse linked, while ‘who’ or ‘what’ are generally not D-linked. Answers to D-linked phrases are limited to np-typed phrases. For instance, the wh-question ‘Which man saw Mary?’ can be paraphrased as ‘Who is the man that saw Mary?’. The person who utters the question and the hearer already have the background knowledge that the person who saw Mary is a man. A definite answer is the only possible response. This gives us evidence to assume that a wh-determiner has a minimal type assignment such that it derives a question of type: s/np. On the basis of this assumption, wh-determiners are assigned the following wh-ex-situ types along with its lexical term assignment.

\[ \text{which} \vdash \lambda V. \lambda P. \lambda x.(x = i y.((V y) \land (P y))) : \text{WH}(n p, A, s/np)/n \]

where \( A \in \{s, q\} \)

On the basis of this type-assignment we can derive the following question-answer sequence. After merging the wh-question with its answer, we derive a question-answer statement of type \( s \), which has a corresponding semantic term of a proposition.

\[
\frac{\text{Which man saw Mary}}{\lambda x.(x = i y.((\text{man } y) \land ((\text{see } m) y)) : s/np} \quad \text{John} \vdash n p
\]

\[
\frac{\text{Which man saw Mary} \circ \text{John} \vdash s}{j = i y.((\text{man } y) \land ((\text{see } m) y))}
\]

Quantified noun phrases such as ‘everyone’ and ‘some woman’ cannot be used as answers to D-linked wh-phrases. These answers are ruled out on the basis of the type-assignment for wh-determiners. The question-answer sequences in example 5.4a and 5.4b are derived using the type-assignment of ‘which’ above. A higher-order type answer as given in example 5.4c cannot be derived.

(5.4) a. Which man saw Mary? \vdash \lambda x.(x = i y.((\text{man } y) \land ((\text{see } m) y)) : s/np

Answer: ‘John’ \vdash j : np

Meaning assembly: \( j = i y.((\text{man } y) \land ((\text{see } m) y)) \)

b. Which woman did John see?

\( \vdash \lambda x.(x = i y.((\text{woman } y) \land ((\text{see } y) y)) : s/np \)

Answer: ‘Mary’ \vdash m : np

Meaning assembly: \( m = i y.((\text{woman } y) \land ((\text{see } y) y)) \)

c. Which man saw Mary? \vdash s/np

Answer: *‘every man’ \vdash \lambda P. \forall \lambda y.((\text{man } y) \rightarrow (P y)) : \emptyset q
Further restrictions on possible question-answer combinations for single constituent questions can be finetuned using the derivability pattern of syntactic types: $\Box A \vdash A \Box A$ (Bernardi, 2002). On the basis of this derivability schema, Bernardi accounts for the interaction and the ordering between different types of quantified noun phrases. Along the same lines, the interaction between question-answer pairs can be studied further.

### 5.3.1.2 Indirect wh-questions

The decomposition of main clause wh-question types applies similarly to embedded wh-questions. Main clause types have been typed as $s$, $q$ and $wh$, while embedded clauses are typed as $s'$, $q'$ and $wh'$, respectively (see chapter 3, section 3.3). The decomposition of wh-question types into a complex type $A/\Box B$, where $A$ is the type of the question-answer combination and $B$ is the type of the answer, will also be applied to embedded question types. Schematically, embedded interrogatives are typed as: $A'/\Box B$ where $A'$ can be $s'$, $s'/\Box gq$, etcetera.

Verbs such as ‘know’ and ‘ask’ select for an embedded interrogative. Unlike main clause questions, the intrinsic meaning of an embedded interrogative is that a referring noun phase is expected to fill the argument role in the embedded sentence. A sentence such as ‘John knows who left’ can never generate a meaning where the embedded wh-phrase can be filled by a quantifier phrase. It must be the selectional requirements of the interrogative verbs that imposes constraints on the type of the embedded interrogative. As an example, we present the type-assignment of ‘ask’ and ‘who’. We abbreviate the higher-order type $s/(np\setminus s)$ to $gq$.

\[
\begin{align*}
\text{ask} &:: \lambda P.\lambda x.((\text{ask} P) x) : ((np\setminus s)/(s'/np)) \\
\text{who} &:: \lambda P.\lambda Q.(Q P) : \text{WHl}_s(np.s'.s'/\Box gq)
\end{align*}
\]

As a result of this type-assignment, the embedded interrogative must first undergo argument lowering before being merged with the interrogative verb. We illustrate the analysis of the sentence ‘Mary asked who left’ in figure 5.6. In the lambda term for the complete sentence, the embedded interrogative is an argument to the predicate ‘ask’.

### 5.3.1.3 Non-local wh-questions

For non-local wh-questions where a wh-phrase associates with a gap hypothesis in an embedded clause, the analysis of the meaning assembly is similar to the analysis of local questions (see chapter 3, section 3.3.3 for the syntactic derivation of non-local wh-questions). The syntactic restrictions on long-distance wh-questions do not effect the meaning assembly. We derive the following question-answer sequences using the decomposed types for wh-phrases.
5. Syntax-semantics interface of wh-questions

**Figure 5.6: Natural deduction derivation with meaning assembly of ‘Mary asked who left.’**

wh-phrases :: WH_{ts}(\Box np, s, q ; gq)

(5.5) Who did Sue believe saw Mary? ⊢ s ; gq

a. Answer: John ⊢ np
   Meaning assembly: ((believe ((see m) j)) s)

b. Answer: Every man ⊢ gq
   Meaning assembly: ∀ y((man y) → ((believe ((see m) y)) s))

(5.6) Who did Sue believe John saw? ⊢ s ; gq

a. Answer: Mary ⊢ s
   Meaning assembly: ((believe ((see m) j)) s)

b. Answer: some woman ⊢ gq
   Meaning assembly: ∃ x((woman x) ∧ ((believe ((see x) j)) s))

5.3.2 Multiple wh-questions

With the derivability pattern of wh-type schema using dependent Geach, as presented in section 5.2.3, we can derive multiple wh-questions from a single
type-assignment of the wh-type schema to a wh-phrase in the lexicon. Using the derivability relations between wh-ex-situ and wh-in-situ type schema, we can account for the syntactic differences between multiple wh-questions in English, Serbo-Croatian and Japanese. English and Serbo-Croatian are wh-ex-situ languages. Nevertheless, English only allows one wh-phrase to be fronted, whereas Serbo-Croatian allows multiple wh-fronting. In Serbo-Croatian, the derivation of multiple wh-questions follows from the derivability of a dependent Geach type for wh-ex-situ type schema. Additionally, we discuss how the restrictions on possible answers to multiple wh-questions in English can be accounted for. Lastly, we show that Japanese multiple wh-questions can be derived from the derivability pattern of wh-in-situ types.

5.3.2.1 Multiple wh-fronting

In chapter 4, section 4.3, we presented data on multiple wh-questions in Serbo-Croatian. In Serbo-Croatian, all wh-phrases occur fronted. We illustrate that we can derive multiple fronting in Serbo-Croatian from a single lexical type-assignment to wh-phrases based on the derivability pattern of wh-ex-situ type. The free order of fronted wh-phrases is derived because we may use a geached type for either wh-phrase. The wh-phrase with the geach type variant is the leftmost wh-phrase. We limit the analysis to the derivation of the direct question ‘ko koga vidi’ (= who whom sees).

Wh-phrases ‘ko’ (= who) and ‘koga’ (= whom) are assigned the following category and semantic term in the lexicon:

\[
\begin{align*}
ko &: \lambda P \cdot \lambda x. (P x) : WH_{\text{ex}}(\square{\text{NOM}}, s, s/\text{NOM}) (= 'who') \\
koga &: \lambda Q \cdot \lambda y. (Q y) : WH_{\text{ex}}(\square{\text{ACC}}, s, s/\text{ACC}) (= 'whom')
\end{align*}
\]

For the sake of simplicity, we choose to assign the wh-phrases an argument lowered wh-ex-situ type. For a proper treatment of Serbo-Croatian, we should treat the wh-phrases along similar lines as English where we derived wh-questions allowing a generalized quantified noun phrase as an answer.

Using the above type assignments, we can derive the following multiple wh-questions, where either ‘ko’ precedes ‘koga’ as in example 5.7a, or where ‘koga’ precedes ‘ko’ as in example 5.7b.

(5.7)  
\begin{align*}
a. \quad & Ko \quad koga \quad vidi? \\
& \text{who whom sees} \\
& 'Who sees whom?' \\
\hline
b. \quad & Koga \quad ko \quad vidi? \\
& \text{whom who sees} \\
& 'Whom was seen by who?'
\end{align*}

With the use of the dependent geach type, we can now derive multiple wh-questions in Serbo-Croatian from the above single type-assignments. To
derive the multiple wh-question ‘Ko koga vidi’, the fronted wh-phrase ‘ko’ is first derived as a dependent geach type (see section 5.2.2 for the syntactic derivation). Along with the syntactic type change, the lambda term of ‘ko’ is also changed. The meaning assembly reflects the dependency of ‘ko’ on the occurrence of another wh-phrase:

ko :: \( \lambda R. \lambda x. \lambda y. ((R y) x) : WH'_t(\diamond \Box \text{NOM}, s/\gamma \text{ACC}, (s/\gamma \diamond \text{ACC})/\gamma \text{NOM}) \)

We illustrate the use of the geach type by deriving ‘ko koga vidi?’ (= who whom sees). We divide the derivation in two parts. The first part of the analysis shows the derivation of the question body with two gap hypotheses for both the nominative and the accusative argument. ‘Koga’ merges with the wh-question body and replaces the accusative gap hypothesis. The resulting structure, which still contains a gap hypothesis for the nominative argument phrase, is of type \( s/\gamma \text{ACC} \), the category of single wh-questions.

\[
\begin{align*}
\text{koga} & \vdash WH'_t(\diamond \Box \text{ACC}, s, s/\gamma \text{ACC}) \\
\text{\lambda Q. \lambda y. (Q y)} & \vdash (\text{see} \, v) \, u \\
\text{koga} & \circ (\diamond \Box \text{NOM} \circ \text{vidi}) \vdash s/\gamma \text{ACC} \\
\text{\lambda y. ((see} \, y) \, u) & \vdash \text{WH}'_t
\end{align*}
\]

The second part of the analysis is where the geach wh-type of the nominative wh-phrase ‘ko’ merges with the previously derived question body. The partial structure \( (\text{koga} \circ (\diamond \Box \text{NOM} \circ \text{vidi})) \) contains the right gap hypothesis \( \diamond \Box \text{NOM} \) and has the right type for the geach type variant of ‘ko’.

ko :: WH'_t(\diamond \Box \text{NOM}, s/\gamma \text{ACC}, (s/\gamma \text{NOM})/\gamma \diamond \text{ACC})

‘Ko’ merges with the structure and yields a multiple wh-question type of type \( (s/\gamma \diamond \text{ACC})/\gamma \text{NOM} \). In the derivation, we omit the geach type of ‘ko’ to focus on the semantic representation of the multiple wh-question.

\[
\begin{align*}
\text{\lambda R. \lambda x. \lambda y. ((R y) x)} & \vdash \text{koga} \circ (\diamond \Box \text{NOM} \circ \text{vidi}) \vdash s/\gamma \text{ACC} \\
\text{\lambda y. ((see} \, y) \, u) & \vdash (s/\gamma \diamond \text{ACC})/\gamma \text{NOM} \\
\text{\lambda x. \lambda y. ((see} \, y) \, x) & \vdash \text{WH}'_t
\end{align*}
\]

The syntactic type derived for the wh-question reveals that the question requires two noun phrases, a nominative noun phrase and an accusative noun phrase. As we explained in section 5.2.3, in order to derive the type required for multiple wh-questions one needs access to the displacement postulates. Therefore, the answer type in the geach type for ‘ko’ is decorated with features, \( \diamond \Box \text{ACC} \).
Due to the feature decoration on the added dependent we can apply the displacement rule to this result and derive the following argument switching. This switching is not only structurally derived, but also changes the underlying syntactic ordering, which in turn has semantic consequences. The following derivation illustrates how the ordering of the arguments for a multiple wh-phrase is derived.

\[
\lambda y. \lambda x. ((\text{see } y) \; x)
\]

\[
\begin{align*}
\text{ko} \circ (\text{koga} \circ \text{vidi}) & \vdash (s/\text{NOM})/\gamma \diamond \text{ACC} \quad [\gamma \diamond \text{ACC}] \\
\text{ko} \circ (\text{koga} \circ \text{vidi}) & \circ \gamma \diamond \text{ACC} \vdash s/\text{NOM} \\
\text{ko} \circ (\text{koga} \circ \text{vidi}) & \circ \gamma \diamond \text{ACC} \circ \text{nom} \vdash s \\
\text{ko} \circ (\text{koga} \circ \text{vidi}) & \circ \gamma \diamond \text{ACC} \circ \text{nom} \vdash s/\gamma \diamond \text{ACC} \\
\text{ko} \circ (\text{koga} \circ \text{vidi}) & \vdash (s/\gamma \diamond \text{ACC})/\gamma \text{NOM}
\end{align*}
\]

\[
\lambda x. \lambda y. ((\text{see } y) \; x)
\]

The result of this argument switching is that in possible responses to a multiple wh-question in Serbo-Croatian, the argument order is free. An answer to the question ‘ko koga vidi’ in a context where ‘Boris saw Ivana’, the two answers ‘Boris Ivana’ and ‘Ivana Boris’ are equally acceptable.

This leads to a further speculation on the status of this feature decoration. We could interpret the additional feature information as a focus marker. This would be in line with Bošković (1998) who claims that the wh-phrase that follows the first fronted wh-phrase is moved for focus reasons. We leave it for further research to check whether these constructions indeed give rise to such interpretations.

5.3.2.2 Simple wh-fronting

Multiple wh-questions in English are recognized by a single wh-phrase that appears fronted at the main clause, whereas the other wh-phrases appear in-situ. The ordering between wh-phrases follows a strict pattern. In chapter 4, we discussed how we can account for the strict ordering between wh-phrases using the derivability pattern of feature decorated s-types to distinguish between different wh-phrases. We now want to focus on the use of the derivability patterns between wh-type schema to derive multiple wh-questions in English.

In chapter 3, section 3.3.5, we explored the syntax of multiple wh-phrases. Wh-phrases that occur in-situ have been typed as \(WH_{in}(np, wh, wh)\). This encoded that the phrase may only appear in-situ in a wh-question body of type \(wh\). On the basis of the decomposed type for wh-questions, the type for wh-in-situ phrases changes into the type-assignment below. Notice that this type occurs in the derivability patterns of wh-in-situ type schema and is derived from argument lowering and dependent geach.
The wh-type schema encodes that a wh-phrase merges with a question body of type $s/\gamma np$, which contains a gap hypothesis of type $np$. This instance of the wh-in-situ type schema is derived from $WH_{in}(np, s, s/\gamma gq)$ using argument lowering and dependent geach. By assigning wh-in-situ phrases the above minimal type, we correctly derive that ‘whom’ can never be of type $WH_{in}(np, s, s/\gamma gq)$. In English, a wh-phrase does not occur in-situ in a $s$-typed body. With this minimal type-assignment the wh-in-situ phrase is always dependent on the occurrence of another wh-phrase ($s/\gamma gq$). Nevertheless, we can add another dependent to the question body and derive a wh-in-situ type phrase which can be used in multiple wh-questions with more than two wh-phrases, such as, ‘Who gave what to whom?’.

\[
\text{multiple wh-in-situ} :: WH_{in}(np, (s/\gamma np)/(\gamma np), ((s/\gamma np)/(\gamma np))/(\gamma np))
\]

The semantic term of the wh-in-situ type reflects this dependency on another wh-phrase. In section 5.2.2, we computed the following meaning assembly for a lowered geach type.

\[
\lambda P^{e \rightarrow (c \rightarrow 0)} \lambda x^{e}. \lambda y^{e}. ((P y) x) :: WH_{in}(np, s/\gamma np, (s/\gamma np)/(\gamma np))
\]

The order in which the answer types are expected is encoded both in the syntactic type as well as in the semantic term. Syntactically, the wh-in-situ phrase is dependent on the occurrence of the subject wh-phrase. Semantically, the lambda abstraction binds the type of the subject wh-phrase over the object wh-phrase. On the basis of this type-assignment and the usual wh-type schema assigned to the subject wh-phrase, we derive the multiple wh-question ‘Who saw whom’.

\[
\text{who} :: \lambda R. \lambda Q. (Q R) :: WH^{l}_{gq}(np, s, s/\gamma gq)
\]

\[4\]Echo-questions form an exception to the occurrence of a wh-in-situ in a $s$ typed question body. Echo-questions, question that repeat a previously uttered sentence of which a part is not understood, should get a different analysis than wh-questions. We leave this analysis for future research.
5.3. Linguistic application

Possible answers to a multiple wh-questions with wh-pronouns are limited to referential noun phrases. By assigning wh-phrases that occur in multiple wh-questions the wh-type schema that we have presented here, we can account for these types of answers. The type of generalized quantifier noun phrases do not match with the type required by the multiple wh-question. The following examples illustrate the correct derivation of two referential noun phrases as answers and the type mismatches with generalized quantified noun phrases as answer.

\[
(5.8) \quad \text{Who saw whom} \vdash \lambda x.\lambda y.((\text{see } y) x) : (s/\text{np})/\text{np}
\]

a. John (np) Mary (np) \vdash ((\text{see m}) j) : s
b. * Every man (gq) some woman (gq) \not \vdash s
c. * John (np) some woman (gq) \not \vdash s
d. * Every man (gq) Mary (np) \not \vdash s

We have shown that assigning wh-phrases a wh-in-situ type correctly accounts for the analysis of multiple wh-questions in English. The question which now arises is why does English allows only simple wh-fronting while we can derive multiple wh-fronting in Serbo-Croatian? On the basis of the derivability pattern of wh-ex-situ types, one could assume that similar to Serbo-Croatian it should be possible to derive a dependent geach type for wh-ex-situ types in English. Nothing prevents the derivation of such types, but because of do-support these types are not applicable.

We briefly point out why geach variants of the wh-ex-situ types in English cannot be used to derive multiple wh-fronting. Non-subject argument wh-phrases have been assigned a wh-ex-situ type that is used in a \( q \)-typed question body to form a single constituent questions, e.g., ‘Whom did John see?’. In theory, we can derive a geach variant of this type which would yield the following type:

\[
\text{whom} :: \text{WH}_{\text{ex}}(\diamond \Box \text{np}, q, s/\text{gq}) \vdash \text{WH}_{\text{ex}}(\diamond \Box \text{np}, q, s/\text{gq}), (s/\emptyset \diamond \text{gq})/\text{gq})
\]

\[
\text{Who saw whom} \vdash \lambda x.\lambda y.((\text{see } y) x) : (s/\text{np})/\text{np}
\]
The geach variant is a wh-ex-situ type which can only be merged with a question body of type $q/\sim gq$. Because question-answer combinations have been typed as $s$-typed sentences, the geached variant cannot be merged with any derived wh-question. As Serbo-Croatian does not have do-support and the question body of all wh-phrases is based on a $s$ type sequence, we can derive multiple wh-fronting. Further research should be done to see whether this line of reasoning also extends to languages with verb second phenomena.

### 5.3.2.3 Multiple wh-in-situ

In this section, we show that the geach type for wh-in-situ type schema also accounts for the analysis of multiple wh-questions in a wh-in-situ language such as Japanese (chapter 4, section 4.2). In chapter 4, we noted that the choice for the syntactic types are led by the interpretation that a wh-question may get. We derive the same semantic representation of multiple wh-questions with wh-in-situ types as multiple wh-questions in a wh-ex-situ language (cf. Serbo-Croatian).

In chapter 4, we presented data from Japanese that explored the different kinds of wh-questions. Recall that Japanese wh-questions have the same argument ordering as declarative clauses. The wh-phrase stays in-situ. The clause where the wh-phrase is embedded in is marked by a question marker, i.e. ‘$ka’’. The position of the question marker determines which clause is interpreted as a question. Sentences with just a question marker and no wh-phrase get the interpretation of a polar question. We refer to chapter 4 for an overview of examples of Japanese wh-questions.

Before we treat multiple wh-questions, we illustrate that the meaning of wh-questions on the basis of wh-in-situ type schema is the same as the meaning assembly of wh-ex-situ wh-phrases. In Japanese, the type for single constituent wh-questions is $q/\sim np$. Based on the decomposed type for wh-questions, wh-in-situ wh-phrases in Japanese are assigned wh-type schema $WH_p(np, q/\sim np)$, where $q$ is the type assigned to the body of the question headed by the question marker ‘$ka’’. An expression of type $q$ denotes a yes-no question. We leave the semantic representation of yes-no questions for further research and concentrate on the use of the $q$ type for the derivation of wh-questions. Again we abstract away from the assignment of $gq$-typed answers and use the lower typed $np$-phrases as required answer types. We list the lexical entries that we have used in chapter 4 to account for the derivation of wh-questions along with their lambda term assignments. We assume the wh-phrases to appear fully inflected in the lexicon.

---

1. The wh-phrase does not carry case feature information over to the answer type. In Japanese it is possible to give a single np-typed answer with the sentence-ending particle ‘desu’ where the answer has no case marking. For instance, the question ‘John-ni nani-o tabe-masi-ta ka’ (= What did John eat?) can be answered with ‘ninzin desu’ (= carrots (it is)) (Nishigauchi, 1990, ex.71,p.49).
5.3. Linguistic application

\[ nani, \ dare \quad :: \quad \lambda P \cdot P : WH_{in}(np, q, q/\neg np) \]
\[ \text{dare-ga} \quad :: \quad \lambda P \cdot \lambda x.(P \cdot x) : WH_{in}(\text{NOM}, q, q/\neg \text{NOM}) \]
\[ \text{ka} \quad :: \quad \lambda Q, Q : s\neg q \]
\[ \text{John, Mary, hon} \quad :: \quad j, m, \text{book} : np \rightarrow \neg \text{ni} \]
\[ \text{katta} \quad :: \quad \lambda y, \lambda x.((\text{buy} \quad y \cdot x) : \text{ACC}(\text{NOM}\neg s)) \]

The meaning assembly of a single constituent wh-question is computed on the basis of the following derivation and the above lexical term assignments.

\[
\begin{align*}
\text{mary} \circ \text{ga} \\
\text{nani} \\
\text{WH}_{in}(np, q, q/\neg np) \\
((\text{mary} \circ \text{ga}) \circ ((np \circ o) \circ \text{kaimasita}) \circ \text{ka} \rightarrow q / \neg np)
\end{align*}
\]

\[
\lambda x.((\text{buy} \quad x) \quad m)
\]

The wh-phrase ‘\text{nani}’ is inserted after the question marker ‘\text{ka}’ is merged with the structure. The meaning assembly of merging the wh-in-situ type schema yields a semantic representation which reflects the binding of the object argument variable of the predicate ‘buy’.

Let us now turn to multiple wh-questions and show that the geach type variants of wh-in-situ type schema yields the correct meaning assembly. The following examples illustrate multiple wh-questions in Japanese.

(5.9) a. \text{dare-ga} \quad \text{nani-o} \quad \text{katta} \quad \text{ka} \quad ? \\
\text{Who ate what?}

b. \text{John-wa} \quad \text{[dare-ga} \quad \text{nani-o} \quad \text{katta} \quad \text{ka]} \quad \text{tazuneta.} \\
\text{John-[-top]} \quad \text{who-[-nopr]} \quad \text{what-[-acc]} \quad \text{buy[-past]} \quad \text{Q asked} \\
\text{John asked who bought what.} \\
\text{NOT} \quad \text{‘Who did John ask bought what?’}

c. \text{‘ka’ marks an embedded wh-question clause} \\
\text{John-wa} \quad \text{[dare-ga} \quad \text{nani-o} \quad \text{katta} \quad \text{to]} \quad \text{itta} \quad \text{ka?} \\
\text{John-[-top]} \quad \text{who-[-nopr]} \quad \text{what-[-acc]} \quad \text{buy[-comp]} \quad \text{said} \quad \text{Q} \\
\text{‘Who did John say bought what?’}
Along the same lines as for Serbo-Croatian, the ‘dependent geach’ rule applies to either of the two embedded wh-phrases. The wh-phrase for which the geach wh-in-situ type is derived becomes dependent on the occurrence of the other wh-phrase and merges with the question body after the other wh-phrase is inserted. Using these geach types we derive sentences with multiple embedded wh-phrases. Similar to single constituent questions, the embedded wh-phrases are interpreted at the level where the closest question marker appears.

The geach type is either used for the subject wh-phrase causing the subject wh-phrase to take scope over the object wh-phrase or vice versa.

\[
\text{nani} :: \lambda Q. \lambda y. \lambda x.((Q y) x) : WH_{in}(np,q/np,(q/np)/np)
\]

\[
\text{dare} :: \lambda P. \lambda x.\lambda y.((P y) x) : WH_{in}(np,q/np,(q/np)/np)
\]

With the geach type assigned to ‘dare’ we can derive the multiple wh-question of example 5.9a, where ‘dare’ is merged after ‘nani-o’ is merged with the question body.

\[
\frac{[\text{np}] \quad \text{np} \quad \text{NOM}}{\text{np} \quad \text{ga} \quad \text{NOM}} \quad \frac{[\text{np}] \quad \text{np} \quad \text{ACC}}{\text{np} \quad \text{o} \quad \text{np} \quad \text{ACC}} \quad \frac{[\text{np}] \quad \text{katta} \quad \text{NOM}s}{\text{katta} \quad \text{NOM}s} \quad \frac{[\text{np}] \quad \text{ka} \quad \text{np}}{\text{ka} \quad \text{np}}
\]

\[
\frac{\text{(np o ga) o ((np o o) o katta) o ka} \quad \text{q}}{\text{WH}_{in}} \quad \frac{\text{nani} \quad \text{WH}_{in}(np,q,q/np)}{\text{WH}_{in}(np,q,q/np)}
\]

\[
\frac{\text{(dare o ga) o ((nani o o) o katta) o ka} \quad \text{q}/np}{\text{WH}_{in}(np,q,q/np,q/np)/np}
\]

\[
\lambda x. \lambda y.((\text{buy} y) x)
\]

Similarly, we can derive the wh-question where ‘nani’ is merged after ‘dare’. The order in which the answers are expected reflects this difference. This analysis predicts that the question can generate two readings; one where the subject argument is questioned with respect to the object argument and another where we have the reverse. In Japanese, focus is needed to disambiguate the wh-question.

Additionally, we can compute the following meaning assemblies for the wh-question in example 5.9 (see appendix B.2.7 for on-line derivation).

\[(5.10) \quad \text{a. Dare-ga nani-o katta ka?} \quad \vdash (q/np)/np
\]

\[
\lambda x. \lambda y.((\text{buy} y) x) (= \text{Who bought what?})
\]
b. Dare-ga nani-o katta ka? ⊢ (q/\np)/\np
   λy.λx.((\textbf{buy} y) x) (= What did who buy?)

c. John-wa [dare-ga nani-o katta to] itta ka? ⊢ (q/\np)/\np
   λx.λy.((\textbf{say} ((\textbf{buy} y) x)) y) (= Who did John say bought what?)

d. John-wa [dare-ga nani-o katta to] itta ka? ⊢ (q/\np)/\np
   λy.λx.((\textbf{say} ((\textbf{buy} y) x)) y) (= What did John say who bought?)

With the derivability patterns between ex-situ and wh-in-situ type schemata we have shown that we can account for syntactic and semantic differences in question-answer pairs in single and multiple wh-question formation on the basis of these polymorphism. We have shown that multiple wh-question formation in multiple wh-fronting languages and wh-in-situ languages can be derived from a single type-assignment for wh-phrases. Additionally, we have shown that in English we must identify the lexical type-assignment for wh-phrases. In the coming section, we broaden the idea of polymorphism on wh-type schema further and move to the analysis of pied-piping and scope marking constructions. The proposed syntactic analyses in chapter 4 are refined by incorporating the decomposed type for wh-questions into the suggested wh-type schema. As a result, we can compute the correct meaning assembly for these phenomena.

5.3.3 Pied-piping

In chapter 3, we presented a syntactic analysis of pied-piping for English. Pied-piping is the phenomenon where the wh-phrase drags along additional material. Consider the following examples of pied-piping. Example 5.11a illustrates that the wh-determiner is part of a possessor phrase and pied-pipes the whole noun phrase that is being modified. Example 5.11b illustrates that the prepositional phrase can pied-pipe along with the wh-determiner. Example 5.11c illustrates the pied piping of a complex np.

\[ (5.11) \]

a. Which man’s picture did John see?

b. On which topic did John read a book?

c. The author of which novel did John like?

To account for these questions, we have presented an analysis where the wh-determiner is instantiated with the following wh-type schema.

\[ \text{The pied-piping of a complex noun phrase is judged as marginally acceptable. Most speakers only accept these sentences under an echo interpretation with stress intonation on ‘which’. Nevertheless, we incorporate this sentence into the analysis of pied-piping constructions as there are no theoretical grounds on which this sentence can be ruled out in comparison to other pied-piping analysis. Furthermore, there seems an additional contrast with clearly unacceptable direct questions such as ‘s The mother of whom did John meet?’. This suggests a further distinction between wh-pronouns and wh-determiners.} \]
which :: WH_{in}(np, A, WH_{ex}(\square A, q, wh))/n
where A \in \{np, pp\}

Before we continue with the meaning assembly of this wh-type schema, we first paraphrase and recapitulate the use of this wh-type schema. The wh-phrase embedded in a noun phrase or prepositional phrase is a wh-in-situ type which fills the position of an np-argument. After the wh-phrase merges with the np or pp and replaces the np or pp gap hypothesis, the whole construction becomes a wh-ex-situ type schema. Thus, the whole construction becomes like a wh-phrase that merges with a q-typed question body and replaces the np or pp gap hypothesis in the question body. In this analysis, additional material is not dragged along with the wh-phrase, but the whole constituent functions as a wh-phrase.

In chapter 3, we limited the presentation of pied-piping construction to a syntactic analysis. Here, we will show how the meaning assembly of such constructions is computed. For the meaning assembly we follow Morrill’s (1994) analysis of pied-piping in relative clause constructions. To illustrate the close resemblance between pied-piping constructions in relative clauses and wh-questions, we first present the meaning assembly for pied-piping in relative clauses, before presenting the meaning assembly for pied-piping in wh-questions.

### 5.3.3.1 Pied-piping in relative clauses

With a polymorphic type-assignment to wh-phrases, Morrill (1994) accounts for the pied-piping of additional material along with the wh-phrase. The type needs to be polymorphic, because the pied-piped material may form a prepositional phrase as well as a noun phrase. The following examples illustrate a pied-piping construction of a noun phrase and a prepositional phrase in a relative clause.\(^7\)

\[(5.12)\]
\[
a. \text{I like the author [whose book] John read.}

b. \text{I like the topic [on which] John read a book.}

c. \text{I read the novel [the author of which ] John liked.}
\]

We illustrate the meaning assembly of pied-piping constructions by analyzing the complex np construction of example 5.12. To indicate the different semantic relations in the complex np, we first decompose the construction and give a brief description of the different components.

\[(5.13)\] (I read) the novel\(_1\) [[the author of which\(_1\) ]_2 John liked \_2]
5.3. Linguistic application

The indices in this example indicate the semantic relations between three different parts of the relative clause:

- the relative clause [[the author of which] John liked] modifies the noun ‘novel’;
- the relative pronoun ‘which’ occupies the gap position in the direct object [the author of t1] to which the noun phrase ‘the novel’ is associated;
- the gap t2 in the relative clause relates to the whole direct object NP, [the author of which]2.

The interaction between the three parts is reflected in the syntactic type that Morrill (1994) proposes for the analysis of pied-piping constructions.8 We use the q-operator (Moortgat, 1991) to incorporate the three parts involved in the construction of a relative clause. The meaning assembly reveals the relations between the different components of the relative clause: the relative clause (P), whose argument position is filled by the clause modifying the noun phrase, the relative clause (N), the common noun (V) and the gap hypothesis (x), which associates the common noun (V x) to the gap in the relative clause (N x). The syntactic type of the relative pronoun along with its lexical semantics can be represented as follows:

\[
\text{which} :: q(np, np, (n\backslash n)/(s/\diamond np)) \\
\lambda N.\lambda P.\lambda V.\lambda y.((V y) \wedge (P N x))
\]

The meaning assembly of the relative clause follows the syntactic derivation steps. For a detailed presentation of the syntactic analysis of relative clause constructions we refer to Morrill (1994); Carpenter (1997). We limit the presentation here by showing the meaning that is computed for the whole construction:

\[
\lambda x. (\text{novel } x) \wedge ((\text{like } y. ((\text{author } y) \wedge ((\text{of } x) y)) y))
\]

5.3.3.2 Pied-piping in wh-questions

The meaning assembly of the relative pronoun is very close to the meaning assembly of the wh-determiner for left-branch constructions. The wh-question that we are going to derive is presented below. The analysis to incorporate the semantics of the pied-piped material in the meaning representation of a wh-question is similar to the analysis of pied-piping constructions in a relative clause.

\[(5.14) \text{ [the author of which novel] does John like?} \]

8Morrill (1994) uses a different mechanism for deriving scope than we have discussed throughout this thesis.
The indices in this example indicate the semantic relations within the complex np are the same as the semantic relations in the relative clause constructions. Let us now step-by-step decompose the syntactic use of the wh-phrase to determine the semantic representation of pied-piping constructions in wh-questions. The different syntactic parts of the derivation are mapped to a semantic representation. To formulate a lexical semantics for the pied-piping case of the which-phrases, we look at the distinct components and see how each component contributes to the meaning assembly of the whole construction:

**which novel**: the wh-phrase 'which novel' embedded inside the construction is the actual argument phrase that is being questioned over. Isolating the meaning assembly of the wh-phrase from the other components, the semantic representation of the wh-phrase becomes:

\[
\lambda V\cdot \lambda x. (x = \iota z. (V z)) \lambda y. (\text{novel } y))
\]

\[
\sim^*_j \lambda x. (x = \iota z. (\text{novel } z))
\]

**the author of which novel**: the np, with the wh-phrase embedded in it, restricts the semantic denotation of the wh-phrase. Similar to other wh-in-situ phrases, the wh-phrase is semantically represented as a lambda term that binds a term variable in the question body. For 'which', the question body is the complex np. The meaning assembly of the complex np construction with the embedded wh-phrase can therefore be represented as:

\[
\lambda N\cdot \lambda x. (N x) \lambda z. \iota y. [(\text{author } y) \land ((\text{of } z) y)]
\]

\[
\sim^*_j \lambda x. \iota y. [(\text{author } y) \land ((\text{of } x) y)]
\]

**wh does John like**: The meaning assembly of the question body which the complex wh-phrase will be merged with is similar to other direct questions. The argument variable of the predicate 'like' is bound by the λ operator.

\[
\lambda P\cdot \lambda x. (P x) \lambda u. ((\text{like } u) j))
\]

\[
\sim^*_j \lambda x. ((\text{like } x) j)
\]

Combining the different components in the order that matches the wh-type schema, we obtain the following semantic representation for 'which':

\[
\text{which } = \text{WH}_{wh}(\text{np }, \text{np }, \text{WH}_{ex}(\Diamond \Diamond \text{np }, q, s/ \text{np })))/n
\]

\[
\lambda V. \lambda N. \lambda P. \lambda x. (x = \iota z. [(V z) \land (P (N z))])
\]
### Sample derivation

We present the analysis of *The author of which novel does John like?* in three parts, the same parts as we used to explain the meaning assembly of the wh-determiner. Firstly, we merge the wh-determiner with the noun *novel*. Secondly, we present part of the derivation where the wh-phrase combines with its pied-piping material. Thirdly, we merge the result type of the complex wh-phrase to the body of the question. Step-by-step, we resolve the meaning assembly of the wh-question along with the syntactic derivation.

\[
\begin{align*}
\text{which novel} & \\
\frac{\text{WH}_{\text{lin}}(np, np, WH_{\text{ex}}(\Diamond \Box np, s/np))/n}{\text{WH}_{\text{lin}}(np, np, WH_{\text{ex}}(\Diamond \Box np, q, s/np))} & \quad \frac{n}{n} & \quad [E] \\
\end{align*}
\]

\[
\lambda N. \lambda P. \lambda x. (x = \iota z. [(\text{novel } z) \land (P (N z) y)])
\]

\[
\text{the author of which novel} \quad \frac{\text{WH}_{\text{lin}}(np, np, WH_{\text{ex}}(\Diamond \Box np, q, s/np))}{\text{of \ author \ (of \ (which \ novel))} \quad \frac{\text{of \ np} \vdash n \Box n}{\text{\author{np}}}} \\
\end{align*}
\]

\[
\lambda P. (\lambda x. (x = \iota z. [(\text{novel } z) \land (P (N y) [(\text{author } y) \land (of \ z) y])]))
\]

\[
\text{wh does John like} \quad \frac{\text{WH}_{\text{ex}}(\Diamond \Box np, s/np)}{\text{the \ author \ of \ which \ novel} \quad \frac{\text{the \ author \ of \ which \ novel}}{\text{\author{np} \ (\text{of \ np})} \quad \frac{\text{\author{np}} \vdash n \Box n}{\text{\author{np}}}} \\
\end{align*}
\]

\[
\lambda x. (x = \iota z. [(\text{novel } z) \land (\text{like } \iota y. [(\text{author } y) \land (of \ z) y])])
\]

The analysis of *the author of which novel* illustrates that we can derive a complex wh-phrase of type \(WH_{\text{ex}}(\Diamond \Box np, q, s/np)\). The meaning assembly of the complex \(np\) shows that we need to combine with a question body to fill in...
the predicate variable. The last part of the derivation shows that the derived
wh-ex-situ phrase as a whole merges with the question body and replaces
the associated gap hypothesis, ‘does John like np’. The meaning assembly of
the whole sentence captures the meaning of the pied-piping construction as a
direct question.

5.3.3.3 Pied-piping in Japanese

In chapter 4, section 4.2.5, we presented wh-questions where the wh-phrase
is embedded in a complex noun phrase. The syntactic analysis of these wh-
questions shows many similarities to pied-piping constructions. We reanalyze
the syntactic account and show that a pied-piping analysis yields a semantics
that matches the interpretation of direct questions.

Example 5.15 illustrates a complex noun phrase where the wh-phrase is
embedded in a relative clause construction (Nishigauchi, 1990, ex.57). Al-
though embedded in a complex NP, the wh-phrase is interpreted at a main
clause level. As a result, the whole sentence is interpreted as a direct question.

(5.15) Mary-wa [[John-ni nani-o ageta] hito-ni] atta-ka?
‘What did Mary meet the man who gave to John?’

Both Pesetsky (1987) and Nishigauchi (1990) notice that a possible re-
sponse to such wh-questions is to repeat the complex noun phrase where the
wh-phrase is embedded in. The answer given in example 5.16A only pro-
vides a minimal response to the embedded wh-phrase which Pesetsky (1987)
indicates as ungrammatical. Nishigauchi (1990) presents the short answer as
grammatical, but claims that the short answer is a truncated form of 5.16B. Ni-
shigauchi provides further evidence that short answers to other complex NP
wh-expressions are completely unacceptable. The most salient answer to wh-
questions where the wh-expression is embedded in a complex NP must repeat
at least some part of the complex NP. Example 5.16B shows such a possible
response.

(5.16) A. */* Konpyuutaa desu
      computer   [Cop]
‘It’s a computer’

      B. [[Konpyuutaa-o ageta] hito] desu
      computer[Acc] gave man [Cop]
‘It’s the man who gave a computer to him’

A proper analysis of Japanese complex NP’s must account for the semantic
embedding of the wh-phrase in the relative clause. In chapter 4, section 4.2, we
showed that complex NP’s are correctly analyzed by assigning the embedded
interrogative the following wh-type schema:
Syntactically and semantically the type is comparable to the type assigned to wh-phrases that occur in pied-piping constructions in English. The wh-phrase ‘nani’ (= what) occurs in an np-type construction which as a whole functions as a wh-in-situ. For instance, assume that ‘nani’ occurs in the complex noun phrase expression ‘John-ni nani-o ageta hito’ (= man who gave what to John) which, for the sake of simplicity, we abbreviate to ‘whcomplex’. The complex wh-expression has the same syntactic type and semantic representation as other wh-phrases in Japanese:

\[ \text{whcomplex} :: \lambda P. \lambda x. (P x) : \text{WH}_{\text{in}}(np, q, q/\text{np}) \]

The semantic interpretation of the complex noun-phrase needs to be incorporated into the semantic representation of the embedded wh-phrase. The semantic representation of the wh-in-situ type becomes:

\[ \text{nani} :: \lambda N. \lambda P. \lambda x. (P (N x)) : \text{WH}_{\text{in}}(np, np, \text{WH}_{\text{in}}(np, q, q/\text{np})) \]

We illustrate the meaning assembly of wh-questions with a complex np. First, we derive the semantic representation of the complex np with the embedded wh-phrase. Secondly, the whole complex np is merged with the main clause and replaces the associated argument position which is reflected in the meaning assembly.

**Complex np:** The syntactic analysis of complex noun phrases in Japanese is in many respects equal to the analysis of English relative clause constructions. As we noted in chapter 4, Japanese lacks an overt complementizer such as ‘that’. We therefore proposed the following type for noun phrases that are restricted by a relative clause:

\[ \text{hito} :: \lambda Q^v. \lambda y. ((\text{man } y) \land (Q y)) : (\text{NOM} \setminus \text{s}) \setminus \text{np} \]

The derivation of the complex noun phrase ‘John-ni nani-o ageta hito’ (= man who gave what to John) with ‘nani’ as embedded wh-phrase is the following.
The complex noun phrase yields the same type as a wh-in-situ phrase. The meaning assembly corresponds to the meaning assembly of a normal wh-in-situ. Only the variable that refers to the gap hypothesis is embedded in the predicate that is part of the complex np. The next step is that the complex np as a wh-in-situ is merged with the body which will provide the predicate $P$ that is requested by the wh-in-situ type.

We use the following abbreviation to refer to the complex noun phrase 'john-ni nani-o ageta hito' (= man who gave what to John).

\[
\text{whcomplex} \vdash \lambda P. \lambda x. \left( (\text{man} \land \left( ((\text{give} \ x) \ j) \ y \right)) : \text{WHin}(np, q, q/np) \right)
\]

The following derivation illustrates the steps of merging the complex noun phrase to the main clause. After the question marker is merged with the structure and forms the body of the question of type $q$, the complex noun phrase is merged.

The meaning assembly of the wh-question shows that the argument inside the complex noun phrase is questioned at the main clause level. The lambda abstraction ($\lambda x$) abstracts over the argument position embedded inside the complex noun phrase. The meaning assembly of the relative clause construction modifying 'hito' restricts the interpretation of the wh-phrase. Possible answers to such a wh-questions therefore have to match the restrictive meaning representation of the argument.
5.3.4 Scope marking constructions

In chapter 4, we showed how scope marking constructions in German and Hindi can be analyzed syntactically using the wh-type schema. The decomposition of the wh-type as a sentence which is incomplete for an answer also affects the type assigned to the scope marker. With the decomposition, the meaning assembly computed for scope marker constructions is the same meaning assembly for long-distance wh-questions.

In section 5.3.4.1, we repeat the wh-type schema for German and show how the type unfolds with the decomposition of the wh-type. Step-by-step, we build up a meaning assembly which equals that given for wh-questions without a scope marker. In section 5.3.4.2, we follow a similar line of reasoning inspecting scope marking constructions in Hindi. In section 5.3.4.3, we discuss the analyses of scope markers of Hindi and German in type-logical grammar in comparison to analyses for scope marking constructions that have been suggested in generative syntactic research (Dayal, 2000; van Riemsdijk, 1983; McDaniel, 1989).

5.3.4.1 German scope marking

Scope marking constructions in German are sentences with an embedded interrogative clause and a left peripheral scope marker. The matrix verb is a bridge verb such as ‘glauben’ (= believe) that normally does not select for an embedded interrogative. The scope marker ‘was’ (= what) appears clause-initial at the main clause, while the actual wh-phrase appears clause-initial in the embedded clause. The overall interpretation of a scope marking construction is similar to wh-questions where the wh-phrase appears fronted. The following examples illustrate a direct question and the same question as a scope marking construction.

(5.17) Welches Bild, glaubt Miro dass Picasso t, gemalt hatte
     with whom believe Miro that Picasso painted had
     “Which picture do Miro believe that Picasso had painted?”

(5.18) Was, glaubt Miro welches Bild, Picasso t, gemalt hatte?
     what believes Miro which picture Picasso painted had
     “Which picture does Miro believe that Picasso had painted?”

In chapter 4, we assigned ‘was’ the following wh-type schema:

\[ \text{was} :: WH_{2}^{\text{I}}(\Diamond \square(s'/wh'), s, wh) = wh/(\Diamond \square(s'/wh')\langle s\rangle) \]

On the basis of this wh-type schema, the scope marking construction is derived as follows. We only display the last steps in the derivation, where ‘wbpgh’ is an abbreviation for an embedded interrogative clause, ‘welches Bild Picasso gemalt hatte’ (= wh’).
The derivation of the scope marker construction can be paraphrased in prose as follows. After the embedded interrogative is built up as usual, the gap hypothesis of the scope marker (\( \diamond \Box (s'/wh') \)) merges with the interrogative clause at the position where, normally, a complementizer such as 'dass' would be expected. The scope marker functions as a 'lever' and changes the category of the embedded interrogative ("wh") into the category for embedded declarative clauses ("s"), which in turn can be selected by the bridge verb 'glauben'. After the matrix clause is merged and the gap hypothesis is displaced to the left edge of the structure (P[2]), the scope marker merges with the question body replacing its hypothesis. The whole sentence becomes of type "wh".

For the syntactic analysis, we used wh-types in an abbreviated form. For the semantic analysis, we will unfold the "wh"-type inside the wh-type schema scope marker to the question-answer type, "s/\(\rightarrow\)np" for direct questions and "s'/\(\rightarrow\)np" for embedded questions. After unfolding, we are left with the following wh-type schema for the scope marker 'was'.

\[
\frac{\text{wh}/(\diamond \Box (s'/wh'))/s}{(s/\rightarrow np)/(\diamond \Box (s'/\rightarrow np))/s} \text{ [unfold]}
\]

**Wh-phrases and scope markers**  So far, we have concentrated on the similarity between scope markers and complementizers. However, we believe and argue that the wh-scope marker is more similar to the object wh-phrase. The difference is that the object wh-phrase associates with np gap hypotheses and the scope marker associates with a s'/\(\rightarrow\)np gap. Mapping the syntactic type to a semantic type reveals that the object wh-phrase reasons over gap hypothesis of type e, while the scope marker is reasoning over "lifted" types: (e \(\rightarrow\) t) \(\rightarrow\) t.

object wh-phrase 'was' :: WH\(e\rightarrow\)(\(\diamond\)(np,s,s/np))

scope marker 'was' :: WH\(e\rightarrow\)(\(\diamond\)(s'/\(\rightarrow\)np),s,s/\(\rightarrow\)np)

Also semantically, the similarity between object wh-phrases and scope markers is visible. Object wh-phrases and scope markers have the following lexical semantics. Because the scope marker is reasoning over a lifted e type, instead of applying a predicate P to the argument variable x, the predicate P is applied to the lifted argument variable: \(\lambda Q.(Q \; x)\).
object wh-phrase 'was' :: $\lambda Q x. (Q x)$
scope marker 'was' :: $\lambda P y y. \lambda x. (P (\lambda Q x. y))$

We illustrate the meaning assembly of scope marker constructions with a step-by-step analysis and show that meaning assembly is similar to long-distance wh-questions. For the analysis of scope marking constructions, we focus on the meaning assembly at the point where the scope marker is merged with the question body. We repeat the meaning assembly for the inference rule for merging a wh-type schema to a question body. $\omega$ is the term variable that is instantiated by the semantic term assigned to the wh-phrase.

$\Gamma \vdash \omega : WH(A, B, C)$
$\Delta[x:A] \vdash BODY : B$

$\Delta[\Gamma] \vdash (\omega \lambda x. BODY) : C$

For the derivation of the scope marking construction 'was glaubt Miro welches Bild Picasso gemalt hatte', we used the following lexical entries for the syntactic derivation of these constructions in chapter 4, section 4.4. Notice that to avoid the complexity of a verb movement analysis, we have chosen to capture the different orderings of arguments in main and embedded sentences by fixing the argument ordering in the types assigned to the verb. The lexical semantics of each word is given along with the syntactic type-assignment.

glaubt :: $\lambda P . \lambda x. ((\text{believe } P) x) : IV/s'$
hatte :: $\lambda Q . Q : \text{prt}\_\text{\_AUX}$
gemalt :: $\lambda y. ((\text{paint } y) : np\_\text{\_prt})$
Bild :: $\text{picture} : n$
Picasso, Miro :: $\text{p,m} : \text{nom}$
welches :: $\lambda Q. \lambda P . \lambda y. ((Q y) \land (P y)) : WH_1(\land np, s, wh')/n$

To illustrate the meaning assembly of a scope marker construction, we again build up the semantic term in a stepwise fashion. For the sake of simplicity, we abbreviate the type for embedded wh-questions back to $wh' (= s'/np)$

1. The embedded wh-question is built up standardly:

```
\lambda P . \lambda y. ((\text{picture } y) \land (P y))
```

```
\lambda y. ((\text{picture } y) \land (\text{paint } y) \land (P y))
```

```
\lambda y. ((\text{picture } y) \land ((\text{paint } y) \land (P y)))
```

```
\lambda y. ((\text{picture } y) \land ((\text{paint } y) \land (P y)))
```
In the following derivation, we abbreviate the semantics for the embedded interrogative clause to $\lambda y. (\text{ppp } y)$ and the structural occurrence of ‘welches Bild Picasso gemalt hatte’ to $\text{wbpgh}$.

2. The gap hypothesis of the scope marker is applied to the type of the embedded wh-question and turns the embedded wh-question type into an embedded sentence type. Semantically, the scope marker hypothesis is assigned a predicate variable ($R^{\text{wh}}$), which applies to the embedded interrogative clause yielding a proposition of type $t$.

$$
\begin{array}{c}
R \\
\Diamond \Box (s'/\text{wh}') \vdash s'/\text{wh}' \end{array} \quad \begin{array}{c}
\lambda y. (\text{ppp } y) \\
\text{wbpgh} \vdash \text{wh}' \\
\end{array} \\
\Diamond \Box (s'/\text{wh}') \circ \text{wbpgh} \vdash s' \\
\\(R \lambda y. (\text{ppp } y))
\end{array}
$$

3. The matrix verb selects the embedded sentence with the hypothesized scope marker and yields a main clause structure with the scope marker hypothesis embedded.

$$
\begin{array}{c}
\lambda P. ((\text{believe } P) \ m) \\
glaubt \circ \text{miro} \vdash s/s' \\
\Diamond \Box (s'/\text{wh}') \circ \text{wbpgh} \vdash s' \\
\end{array} \\
\begin{array}{c}
(R \lambda y. (\text{ppp } y)) \\
\vdash (\text{glaubt } \circ \text{miro}) \circ (\Diamond \Box (s'/\text{wh}') \circ \text{wbpgh}) \vdash s \\
\end{array}
$$

4. We have now reached the point where we can merge the scope marker to the question body. The meaning assembly of the wh-type merge rule creates a $\lambda$-binder for redrawing the gap hypothesis. Additionally, the lexical semantics assigned to the scope marker applies to the semantic term of the question body. The result, after applying $\beta$-reduction to the semantic term, is a semantic term where the gap variable inside the embedded clause is bound to $\lambda$-operator, which takes scope over the main clause.

$$
\begin{array}{c}
\lambda P. \lambda x. (P \lambda Q. (Q \ x)) \\
\vdash (\text{was } (\text{believe } (R \lambda y. (\text{ppp } y))) m) \\
\vdash (\text{wH}^t_{\text{wm}} (\Diamond \Box s'/\text{wh}', s, s'/\text{np})) \\
\vdash (\text{glaubt } \circ \text{miro}) \circ (\Diamond \Box (s'/\text{wh}') \circ \text{wbpgh}) \vdash s \\
\vdash (\lambda P. \lambda x. (P \lambda Q. (Q \ x))) \lambda R. ((\text{believe } (R \lambda y. (\text{ppp } y))) m) \\
\vdash \lambda x. ((\text{believe } (R \lambda y. (\text{ppp } y))) m) \\
\vdash \lambda x. ((\text{believe } (\lambda Q. (Q \ x))) m) \\
\vdash \lambda x. ((\text{believe } (\text{ppp } x))) m \\
\end{array}
$$
The semantics of the whole scope marking construction indicates that the scope marker associates with the gap hypothesis in the embedded interrogative. Due to the occurrence of another wh-phrase, the scope marker cannot associate with the hypothesized argument gap directly. With our proposed type for wh-scope markers, the link between the gap hypothesis and the scope marker can be established. The scope marker abstracts over the embedded interrogative and associates with the argument via the answer type that is requested by the embedded interrogative, thereby interpreting the whole sentence as a direct question.

**Multiple scope marker construction** In scope marking constructions where the wh-phrase occurs embedded in a clause intervened by other embedded clauses, each clause that intervenes between the embedded interrogative and the main clause must contain another scope marker. These scope markers pass the semantic representation of the embedded argument position on to the main clause. An example of a multiple scope marker construction is the sentence, ‘Was glaubte Miro was Hans meint welches Bild Picasso gemalt hatte?’ (= Which picture does Miro believe that Hans claims that Picasso has painted?).

In chapter 4, we showed that such constructions are derived by recursively binding the embedded question. We derived the construction adding the following type-assignment to scope markers that occur in the embedded clauses. The type alternates with the type assigned to the main clause scope marker. After merging with the embedded clause, the embedded scope marker returns an embedded question type. The semantic term assigned to embedded scope markers is the same as the semantic term for matrix clause scope markers.

\[
\text{was} : \lambda P. \lambda x. (P \lambda Q. (Q x)) : \text{WH} \text{lex}(\Diamond \Box (s'/\text{osp})), s, s'/\text{osp})
\]

The derivation of a subordinate clause with an embedded scope marker is similar to the derivation of the main clause with a scope marker. Semantically, the scope markers in the intervening clauses pass the lambda abstraction over the term variable of the embedded interrogative on to the main clause. The following meaning assembly is computed for deriving a multiple scope-marker construction (see appendix B.5.5 for an on-line derivation):

\[(5.19) \text{Was glaubt Miro was Hans meint welches Bild Picasso gemalt hatte?}
\[
\vdash \lambda y.((\text{believe} ((\text{claim} (\text{ppp} y)) \text{hans})) \text{miro}) : s/\text{osp}
\]

The alternative construction, where the embedded wh-phrase moves one embedded clause higher up, has the same meaning assembly. In those cases, the scope marker only appears in the clauses preceding the partially moved wh-phrase. The subordinate clauses that follow are headed by the complementizer ‘dass’.

\[(5.20) \text{Was glaubt Miro welches Bild Hans meint dass Picasso gemalt hatte?}
\[
\vdash \lambda y.((\text{believe} ((\text{claim} (\text{ppp} y)) \text{hans})) \text{miro}) : s/\text{osp}
\]
5. Syntax-semantics interface of wh-questions

5.3.4.2 Hindi scope marking

The meaning assembly that accounts for the scope marker construction in German is similarly applicable to Hindi. The difference between German and Hindi is the structural realization of the scope marker. In Hindi, the scope marker occurs cliticized to the matrix verb, as illustrated in the following example.

**Scope marking construction** (Dayal, 2000, ex.5, p.160)

(5.21) Jaun *kyaa* soota hai ki merii *kis-se* baat karegii?

> John what think-[pres] that Mary who-with talk do-[Fut]

> “Who does John think Mary will talk to?”

**'kyaa' as scope marker** We proposed a similar analysis for Hindi scope marking as for German scope marking. We will show that the proposed analysis for Hindi yields the same semantics for scope marking constructions as for German.

Similar to German, the scope marker in Hindi is syntactically similar to the type assigned to wh-phrase. In Hindi, wh-phrases may occur fronted and cliticized to the verb. Because the scope marker only occurs cliticized to the verb, we compare the type for scope markers with the type assigned to wh-phrase that occurs cliticized. Along with the syntactic types, we present lexical semantics on the basis of which we compute the meaning assembly of scope marking constructions in Hindi.

<table>
<thead>
<tr>
<th>scope marker</th>
<th>wh-phrase</th>
</tr>
</thead>
<tbody>
<tr>
<td><em>kyaa</em></td>
<td><em>WHl</em> = (\lambda P. \lambda x. \lambda y. ((P \lambda Q. (Q y)) x))</td>
</tr>
<tr>
<td></td>
<td><em>WHl</em> = (\lambda P. \lambda x. \lambda y. ((P \lambda Q. (Q y)) x))</td>
</tr>
</tbody>
</table>

The scope marker modifies the type of the embedded interrogative and changes the type of the embedded interrogative such that the matrix verb combines with it. Semantically, it binds the embedded wh-phrase such that it is interpreted outside the embedded clause at the main clause level. The meaning assembly of the scope marker reflects the intricate grammatical function that the scope marker has.

As an illustration, we present an analysis of the scope marking construction where the embedded wh-phrase stays in-situ. We use the following lexical entries along with lexical semantics assigned to the types to derive a scope marking construction. In order to improve readability of the derivation, we abbreviate \(s/np\) and \(s'/np\) to \(wh\) and \(wh'\). The type abbreviations \(IV_s\) and \(IV_{wh}\) stand for \(NOM\backslash s\) and \(NOM\backslash wh\), respectively.
5.3. Linguistic application

ki :: \( \lambda Q : A'/A_s \) (where \( A \in \{s, wh\} \))

soctaahai :: \( \lambda P, \lambda x.(\text{think } P) \times : IV_s/s' \)

dekhaa :: \( \lambda y, \lambda x.(\text{see } y) \times : ACC\setminus IV \)

raviine, anu :: r,a : NOM

kisko :: \( \lambda P, \lambda x.((\text{think } P)x : IV_s/s') \)

We present the syntactic analysis in two steps. First, the derivation of the embedded questions. Secondly, the merging of the scope marker. Finally, we present the meaning assembly of the whole derivation.

1. The embedded wh-phrase 'ki raviine kisko dekhaahai' is built up as follows:

\[
\begin{align*}
\text{ki} & \quad \text{:: } \lambda Q : A'/A_s \quad (\text{where } A \in \{s, wh\}) \\
nom & \quad \text{:: } \lambda P, \lambda x.(\text{think } P) \times : IV_s/s' \\
\text{dekhaa} & \quad \text{:: } \lambda y, \lambda x.(\text{see } y) \times : ACC\setminus IV \\
raviine, anu & \quad :: r,a : NOM \\
kisko & \quad :: \lambda P, \lambda x.((\text{think } P)x : IV_s/s') \\
\end{align*}
\]

We present the syntactic analysis in two steps. First, the derivation of the embedded questions. Secondly, the merging of the scope marker. Finally, we present the meaning assembly of the whole derivation.

1. The embedded wh-phrase 'ki raviine kisko dekhaahai' is built up as follows:

\[
\begin{align*}
\text{ki} & \quad \text{:: } \lambda Q : A'/A_s \quad (\text{where } A \in \{s, wh\}) \\
nom & \quad \text{:: } \lambda P, \lambda x.(\text{think } P) \times : IV_s/s' \\
\text{dekhaa} & \quad \text{:: } \lambda y, \lambda x.(\text{see } y) \times : ACC\setminus IV \\
raviine, anu & \quad :: r,a : NOM \\
kisko & \quad :: \lambda P, \lambda x.((\text{think } P)x : IV_s/s') \\
\end{align*}
\]

We present the syntactic analysis in two steps. First, the derivation of the embedded questions. Secondly, the merging of the scope marker. Finally, we present the meaning assembly of the whole derivation.

2. The main clause with the scope marker 'Anu kyaa soctaahai' abstracts over the embedded wh-phrase (wh' = 'ki raviine kisko dekhaahai') as follows:

\[
\begin{align*}
\text{soctaahai} & \quad :: \lambda P, \lambda x.(\text{think } P) \times : IV_s/s' \\
\text{kyaa} & \quad :: \lambda y, \lambda x.(\text{see } y) \times : ACC\setminus IV \\
\text{anu} & \quad :: r,a : NOM \\
\end{align*}
\]

We present the syntactic analysis in two steps. First, the derivation of the embedded questions. Secondly, the merging of the scope marker. Finally, we present the meaning assembly of the whole derivation.

3. The semantic representation of the scope marking construction is built up as follows:

\[
\begin{align*}
\text{ki} & \quad \text{:: } \lambda Q : A'/A_s \quad (\text{where } A \in \{s, wh\}) \\
nom & \quad \text{:: } \lambda P, \lambda x.(\text{think } P) \times : IV_s/s' \\
\text{dekhaa} & \quad \text{:: } \lambda y, \lambda x.(\text{see } y) \times : ACC\setminus IV \\
raviine, anu & \quad :: r,a : NOM \\
kisko & \quad :: \lambda P, \lambda x.((\text{think } P)x : IV_s/s') \\
\end{align*}
\]

We present the syntactic analysis in two steps. First, the derivation of the embedded questions. Secondly, the merging of the scope marker. Finally, we present the meaning assembly of the whole derivation.

1. The embedded wh-phrase 'ki raviine kisko dekhaahai' is built up as follows:

\[
\begin{align*}
\text{ki} & \quad \text{:: } \lambda Q : A'/A_s \quad (\text{where } A \in \{s, wh\}) \\
nom & \quad \text{:: } \lambda P, \lambda x.(\text{think } P) \times : IV_s/s' \\
\text{dekhaa} & \quad \text{:: } \lambda y, \lambda x.(\text{see } y) \times : ACC\setminus IV \\
raviine, anu & \quad :: r,a : NOM \\
kisko & \quad :: \lambda P, \lambda x.((\text{think } P)x : IV_s/s') \\
\end{align*}
\]

We present the syntactic analysis in two steps. First, the derivation of the embedded questions. Secondly, the merging of the scope marker. Finally, we present the meaning assembly of the whole derivation.

2. The main clause with the scope marker 'Anu kyaa soctaahai' abstracts over the embedded wh-phrase (wh' = 'ki raviine kisko dekhaahai') as follows:

\[
\begin{align*}
\text{soctaahai} & \quad :: \lambda P, \lambda x.(\text{think } P) \times : IV_s/s' \\
\text{kyaa} & \quad :: \lambda y, \lambda x.(\text{see } y) \times : ACC\setminus IV \\
\text{anu} & \quad :: r,a : NOM \\
\end{align*}
\]

We present the syntactic analysis in two steps. First, the derivation of the embedded questions. Secondly, the merging of the scope marker. Finally, we present the meaning assembly of the whole derivation.

3. The semantic representation of the scope marking construction is built up as follows:

\[
\begin{align*}
\text{ki} & \quad \text{:: } \lambda Q : A'/A_s \quad (\text{where } A \in \{s, wh\}) \\
nom & \quad \text{:: } \lambda P, \lambda x.(\text{think } P) \times : IV_s/s' \\
\text{dekhaa} & \quad \text{:: } \lambda y, \lambda x.(\text{see } y) \times : ACC\setminus IV \\
raviine, anu & \quad :: r,a : NOM \\
kisko & \quad :: \lambda P, \lambda x.((\text{think } P)x : IV_s/s') \\
\end{align*}
\]
Multiple scope markers In the same way as in German, constructions with multiple scope markers are derived by recursively binding the embedded question. Each embedded clause that intervenes between the embedded interrogative and the main clause must contain another scope marker.

The syntactic type of the scope marker in the embedded clause varies from the main clause scope marker only in the clause type the scope marker appears in: $s'$. Overall, the syntactic and semantic representation of the two scope markers are the same. Both scope markers appear preverbally and both select an embedded interrogative while yielding a wh-question type.

```
intervening ‘kyaa’ :: $\lambda P.\lambda x.\lambda y.((P\lambda Q.(Q\ y))\ x) : WH_2(\emptyset(s'/wh'), IV_2^t, IV_2^s)$
main clause ‘kyaa’ :: $\lambda P.\lambda x.\lambda y.((P\lambda Q.(Q\ y))\ x) : WH_2(\emptyset(s'/wh'), IV_1^t, IV_1^s)$
```

Using the above type for ‘kyaa’ in intervening clauses and the other type for ‘kyaa’ in main clauses, we derive multiple scope marking constructions. Below, we give an example of such a construction.

(5.22) Jaun kyaa soctaahai anu kyaa kahegii ki raviine kisko dekhaa?
$\vdash$ wh = ‘Whom does John think that Anu says that Ravi saw?’

Each subpart of the clause is presented separately. The two scope markers recursively bind the argument variable of the embedded interrogative up to the main clause.

1. An embedded interrogative is selected by the hypothesized modifier type of the scope marker ‘kyaa’. The type modifier changes the type of
the embedded clause from an embedded wh-question type \((wh')\) into an embedded declarative clause type \((s')\)

\[
\begin{align*}
\left[\diamond \sqcap (s'/wh')\right] \\
\frac{\begin{array}{c}
s'/wh' \\
\diamond \sqcap (s'/wh') \circ (ki \circ (ravine \circ (kisko \circ dekhaa))) \vdash wh' \\
\end{array}}{\vdash \left[ wh', \left[ s'/wh' \right] \right]} [/ E] \\
\end{align*}
\]

\((\kappa \lambda y.((\text{see } y) \ r))\)

2. The embedded clause of the previous derivation is selected by the verb in the intervening clause. Then the scope marker is inserted and withdraws the hypothesized modifier. Semantically the embedded argument phrase is now bound by the scope marker. We abbreviate the embedded interrogative ‘ki ravine kisko dekhaa’ to \(whemb\).

\[
\begin{align*}
\text{kahegii} & \vdash IV_s/s' \\
\diamond \sqcap (s'/wh') \circ \text{whemb} & \vdash s' \\
\frac{\text{kahegii} \circ (\diamond \sqcap (s'/wh') \circ \text{whemb})}{IV_{whemb}'} \vdash s' \ \\
\frac{\vdash WH_{\text{whemb}}}{\ \ \vdash \left[ wh', IV_{\text{whemb}}' \right]} [/ E] \\
\end{align*}
\]

\((\kappa \lambda y.((\text{say } (\text{see } y) \ r)) \ a)\)

3. The intervening phrase with ‘kyaa’ in composition with the embedded interrogative forms as a whole an embedded interrogative. The other scope marker projects another sentence modifier which changes the type to a embedded declarative clause \((s')\).

\[
\begin{align*}
\left[\diamond \sqcap (s'/wh')\right] \\
\frac{\begin{array}{c}
s'/wh' \\
\diamond \sqcap (s'/wh') \circ (anu \circ (\text{kyaa} \circ (\text{kahegii} \circ \text{whemb}))) \vdash wh' \\
\end{array}}{\vdash \left[ s'/wh' \right]} [/ E] \\
\end{align*}
\]

\((\kappa \lambda y.((\text{say } (\text{see } y) \ r)) \ a)\)

4. The embedded sentence is selected by the matrix verb ‘soctaahai’ and subsequently, the scope marker withdraws the hypothesis. The whole sentence is typed as a wh-question and is interpreted as such. Once again, ‘kyaa’ binds the lambda abstraction of the embedded interrogative ‘anu kyaa kahegii whemb’.
The proof term of the whole construction reveals that the argument which is questioned over by the embedded wh-phrase is bound in the main clause. This is the result of recursively applying and abstracting over the embedded interrogatives. ‘Kyaa’, like ‘was’ in German serves as a binder operator which reanalyzes the lambda abstractor of the clause where the scope marker is combined with.

5.3.4.3 Comparison between German and Hindi

In the generative syntactic framework, different interpretational mechanisms have been proposed for the analysis of scope marking constructions. The two main approaches are the direct dependency approach (van Riemsdijk, 1983; McDaniel, 1989) and the indirect approach (Dayal, 1996, 2000). The crucial difference between these two approaches is the role of the scope marker for the interpretation of the scope marking construction as a direct question. In the direct dependency approach, the scope marker is semantically opaque and is used only to assign matrix scope to the embedded wh-phrases. In the indirect dependency approach, the embedded wh-phrases are not interpreted at main clause level, but play an indirect role. The embedded interrogative forms a semantic restriction to the interpretation of the scope marker which is interpreted as a normal wh-phrase. Both approaches base their analysis on a semantic reconstruction of the phenomenon.

We have argued that the semantics of wh-questions is determined by an interface between syntax and semantics. We have shown that the syntactic derivations of the two scope marking constructions is the same. In both Hindi and German, the scope marker associates with a gap position between the embedded interrogative clause and the main clause. The gap hypothesis merges with the embedded interrogative clause and returns a declarative clause type s which fulfills the selectional requirements of the matrix verb, i.e. a bridge verb that only selects embedded declarative clauses. By identifying the scope marker as a wh-phrase we account for the structural position of the scope marker. Additionally, due to the scope marker’s semantic properties as a wh-phrase, the meaning assembly of scope marking constructions is equal to the meaning assembly of direct questions.
5.4 Concluding remarks

In this chapter, we have discussed the syntactic and semantic consequences of a structured meaning approach to wh-questions. In a structured meaning approach, wh-questions are taken to be incomplete sentences that are part of a question-answer sequence. We have proposed to decompose the \( wh \) type into a type \( A \rightarrow B \) where \( A \) is the type of the question-answer sequence and \( B \) is the type of the answer. Along with the syntactic decomposition of \( wh \)-types, we have been able to express the semantic decomposition of the semantic \( \omega \)-operator as a \( \lambda \)-term.

Additionally, by incorporating the types of possible answers into the type assigned to \( wh \)-questions, we have shown that \( wh \)-type schema can be instantiated in a polymorphic way. Based on characteristic theorems in semantic type language, we have presented a derivability pattern of \( wh \)-type schema. Along with the syntactic derivability, the meaning assembly for each instance of a \( wh \)-type schema could be computed.

With the derivability pattern of \( wh \)-type schemata and the meaning assembly of \( wh \)-questions, we have reanalyzed some phenomena of chapter 3 and 4. In the analyses, we have concentrated on the relation of the polymorphic syntactic use of the \( wh \)-type schema with the meaning assembly of \( wh \)-questions. We have illustrated this polymorphic use of \( wh \)-phrases by providing the syntactic analysis along with the derivational semantic of different constructions: multiple \( wh \)-questions, long-distance binding and embedded \( wh \)-questions in English, complex \( NP \) constructions in Japanese, and scope marking constructions in German and Hindi.

More specifically, we have presented a uniform account of multiple \( wh \)-questions in \( wh \)-ex-situ and \( wh \)-in-situ languages. The semantic difference between \( wh \)-pronouns and \( wh \)-determiners can be accounted for on the basis of the derivability pattern of \( wh \)-type schema. For multiple \( wh \)-questions in \( wh \)-ex-situ languages, \( wh \)-in-situ languages or languages such as English with simple \( wh \)-fronting, we have shown that the same meaning assembly is computed using structurally variant \( wh \)-type schemata. Additionally, we have illustrated that on the basis of the decomposition of \( wh \)-question types \( wh \) into type \( s/np \), the meaning assembly of scope marking constructions is equal to the semantics of direct questions.
Chapter 6

Evaluation and future perspectives

In the introduction of this thesis, we started from the observation that natural language phenomena with a uniform meaning assembly may have different structural realizations. This led to the question, What is the logic of such variation? Let us now discuss whether we have provided an answer to this question.

At the end of chapter 2 we have characterized the logical ‘space’ for variation in terms of a restricted set of structural rule schemata, fixed at the level of UG. With our study of the diversity in wh-question formation, we have shown that three factors play an important role in capturing (a) the structural differences in wh-question formation across natural languages in combination with (b) the uniform meaning assembly of wh-questions.

Firstly, wh-phrases are uniformly typed as higher-order types. With the use of a proposed wh-type schema, WH(A, B, C), for wh-phrases which generalizes over the structural differences between wh-phrases (wh-in-situ, wh-ex-situ), we have accounted for the structural differences of wh-question formation in wh-ex-situ languages, wh-in-situ languages and languages with partial wh-movement. For these different cases, the Curry-Howard interpretation associated with the higher-order types accounts for the uniform meaning assembly of wh-question formation.

Secondly, the possibilities for ‘displacement’ are determined by a restricted set of structural postulates, fixed at the level of UG. The rules apply to those structured expressions that match the structural patterns. A key element of these patterns is the presence of a licensing ♦ feature on the affected substructure. The rules have a built-in symmetry in the sense that affected constituents can be restricted to occur on either left or right branches, where such restric-
tions correlate with OV versus VO typological distinctions.

Thirdly, the sole locus for capturing the cross-linguistic variation in wh-question formation is the lexicon. The lexicon stores syntactic and semantic information of words in type-assignments which the grammatical reasoning system then uses to derive language variation. The binary operators (/, \, •) in the lexical type-assignments encode word-order and constituency which is reflected as predicate-argument relations in the meaning assembly. The unary operators (♦, 2) in the lexical type-assignments encode features licensing or constraining (cf. islands) the application of the structural rules.

Throughout this thesis, we have provided evidence that we can account for wh-question formation within the logic of variation on the basis of these three factors. The question arises whether these three components are necessary and sufficient for the analyses. In order to evaluate our approach, we present a comparison with the grammar formalisms of two other lexicalist approaches: minimalist grammar (MG) and multimodal combinatory categorial grammar (MMCCG). In addition to a brief description of the two frameworks, we discuss their approach to language variation and uniform meaning assembly.

6.1 A comparison with MG and MMCCG

Our goal of formulating a grammar logic to study natural language variation from a formal perspective compares well with the perspectives of minimalist grammars (= MG) (Stabler, 1997, 2001) and multimodal combinatory categorial grammar (= MMCCG) (Baldridge, 2002). Categorial type logics have been compared with minimalist grammars in Vermaat (1999, 2004) and with multimodal CCG in Baldridge and Kruijff (2003).

6.1.1 Minimalist grammars

In a series of papers (Stabler, 1997, 1999, 2001), Stabler explores the possibilities for developing an algebraic grammar system based on the principles of Chomsky’s Minimalist Program (Chomsky, 1995).

A minimalist grammar consists of a lexicon of a finite set of sequences of features and the structure building operations MERGE and MOVE. The features that trigger the application of the operations are the syntactic features. The syntactic features are divided in two groups: category features (N, =N) and control features (+N, −N). The category features are needed for the merging of two constituents, whereas the control features cause the displacement of a constituent in a tree structure.

Each lexical entry consists at least of a base category N, e.g. c, t, vp, d or np, which stand for the more familiar categories of complementizer, tense, verb phrase, determiner, noun phrase. In addition to a base feature, a lexical entry may carry one or more selector features, =N, which encode the selectional requirements of a constituent. Additionally, a feature specification may contain
control features, either licensor features \(+N\) and/or licensee features \(-N\). A licensor feature, \(+N\), is assigned to an element that functions as an ‘attractor’, while a licensee feature, \(-N\), is assigned to the ‘attractee’. Besides these syntactic features, an element may carry non-syntactic features. These non-syntactic features are carried along in the derivation but have no effect on the structure building operations for morphological or semantic reasons. The non-syntactic features may also be empty, e.g. an empty complementizer with no phonological feature information. For simplicity, we only write the headword to indicate that non-syntactic material is present in a derivation.

**Example 6.1 (A minimalist grammar lexicon)** The following lexicon can be built according to the above feature specification.

```
d maria
= d = d v likes
= v + wh c
= d - wh who
```

The feature sequence of each lexical entry obeys a specific order that respects the possibilities given by the structure building operations. Let us now define which are the structure building operations that operate on the lexical entries to build binary tree structures. A tree structure is \(S\) is inductively defined on lexical feature specifications \(F\) as follows:

\[
S ::= F | S < S | S > S
\]

Every node of the tree is labeled with a direction arrow \{<, >\} pointing towards the head of the tree.

The structure building operations are \textsc{Merge}: \(S \times S \rightarrow S\) and \textsc{Move}: \(S \rightarrow S\). \textsc{Merge} combines two tree structures \(S_1\) and \(S_2\) to form a new tree structure. The operation \textsc{Merge} causes the cancellation of the prefixed feature \([=f]\) on the feature sequence of \(S_1\) against \([f]\) of \(S_2\). Technically, \textsc{Merge} can be partitioned into two functions: \textsc{Merge}_\(<\): (\(<, S_1, S_2\)) combines with a tree on the right side, if \(S_1\) is selected from the lexicon. \textsc{Merge}_\(>\): (\(>, S_2, S_1\)) combines with a tree on the left when \(S_1\) is a complex tree structure.

The two merge operations can be represented in a natural deduction style derivation as follows:

\[
\frac{S_1[=f] \quad S_2[f]}{S_1 < S_2} \quad \text{MERGE}_< \quad \frac{S_2[f] \quad S_1[=f]}{S_2 > S_1} \quad \text{MERGE}_>
\]

\textsc{Move} applies to a single tree and operates on a substructure of that tree. In the MG framework, \textsc{Move} is always leftwards. A licensor feature \([+f]\) on the head of tree structure \(S_1[+f]\) attracts the maximal projection\(^1\) \(S_2[-f]\)\(^{-}\)

\(^1\)\(S[-f]\)\(^{-}\) is the maximal projection of \(S[-f]\), i.e. the largest substructure with \([-f]\) as prefixed feature on the head category.
204  6. Evaluation and future perspectives

which has the licensee feature as initial feature in the sequence. After applying \textsc{move}, $S_2$ merges as a specifier to the head of $S_1$, while the part of the tree where $S_2$ is moved from is replaced with the empty string $\lambda$. The \textsc{move} operation can be presented as a natural deduction step. (For a more elaborate introduction to the structure building operations, see Vermaat (1999, 2004); Stabler (1997).)

\[
\frac{S_1\{S_1[-f]\}}{S_2 > S_1\{S_2[-f]/\lambda\}} \textsc{move}
\]

\textbf{Example 6.2 (Derivation of a subject wh-question)} With these operations and the given lexicon, we can derive the subject wh-question, ‘Who likes Maria?’, which belongs to category $c$. To save space, we continue to use the natural deduction style derivation, where every tree is presented as a ‘two-dimensional’ tree. The different levels in the tree are identified by placing brackets around each binary node combining two subtrees.

\[
\begin{align*}
\text{lex} & \quad \text{d-wh who} & \quad \text{d likes} & \quad \text{d maria} \\
=d & \quad v & \quad d & \quad v \quad \text{likes} & \quad \text{maria} \\
\text{MERGE} & \quad \text{MERGE} & \quad \text{MERGE} & \quad \text{MERGE} & \quad \text{MERGE} & \quad \text{MERGE} & \quad \text{MOVE} \\
\text{+wh c} & \quad (-\text{wh who}) & \quad (\text{likes < maria}) & \quad (\text{likes < maria}) & \quad (\text{likes < maria}) \\
\text{who} & \quad (c < (\text{likes < maria})) & \quad (c < (\text{likes < maria})) \\
\end{align*}
\]

\textbf{Generative power} Besides providing an elegant and simple format for analyzing natural language phenomena, minimalist grammars have good parsing and complexity properties. Michaelis (2004) has shown that minimalist grammar languages fall in the mildly context-sensitive languages. Thus, the expressive power of minimalist grammars is such that natural language beyond context-free can be handled while maintaining computational efficiency. The parsing strategies developed by Stabler (1999) provide an insight into the construction and the comparison of analyses for natural language variation along the same lines as the parsing strategies for categorial type logics.\footnote{See appendix B.7 to derive this sentence along with some other sample sentences on-line with the MG parser.}

\textbf{Evaluating MG} Similar to categorial type logics, language variation is controlled through feature checking. The feature specification of lexical entries triggers the application of \textsc{merge} or \textsc{move}. Instead of working with restructuring patterns,\footnote{See http://grail.let.uu.nl/ngcgy for an on-line tool of the minimalist grammar parser.}
6. Evaluation and future perspectives

MG derives structural variation through the application of the \textit{MOVE} operation. Without hypothetical reasoning, \textit{move} operates directly on the elements that occur displaced from their base position. Displaced elements such as wh-phrases have to get to their final position by successive cyclic application of the \textit{MOVE} rule. The application of \textit{move} is controlled by the necessity of feature checking of licensor features assigned to \textit{empty} lexical elements. The use of these empty lexical elements are equal to the use of hypothetical reasoning (Vermaat, 1999).

The question arises whether the use of phonetically empty elements is to be favored over the use of hypothetical reasoning (Retoré and Stabler, 2004). We believe the argument for favoring hypothetical reasoning over the use of empty categories lies in the possibility of having a derivational semantics. The meaning assembly of extraction phenomena where a moved element is associated with a phrasal position further down in the structure is the direct result of using hypothetical reasoning and structural rule schemata.

So far, minimalist grammars have not yet established a direct way of assigning meaning to the syntactic derivations. Other approaches in transformational grammar (Heim and Kratzer, 1988; Szabolcsi, 1997) show that it is possible to compositionally derive a logical form on the basis of functional categories. Such interpretation, however, is of the rule-to-rule type. For the derivational semantics we have pursued in this thesis, the compositional meaning assembly is an automatic effect of the Curry-Howard interpretation that goes with higher-order type-assignments.

6.1.2 Multimodal CCG

Multimodal CCG is an extension of Steedman’s combinatory categorial grammar (Steedman, 2000) with the multimodal resource-sensitive control mechanisms of categorial type logic. To understand how the multimodal approach refines CCG, we present a brief overview of combinatory categorial grammar.

\textbf{CCG} \hspace{1cm} The base logic of CCG is classical categorial grammar with application rules as its basic rule set for combining categories (forward application, $>$; backward application, $<$). The basic rule set is extended with rules based on three combinators: composition ($B$), type-raising ($T$) and substitution ($S$). These combinators give rise to different rule type schema that encode possible type alternations. Certain rule type schemata, however, are excluded from the inventory because they are inconsistent with the directionality of the principal function or they fail to inherit the categorial restrictions that are encoded in the combinatory rule. These constraints are imposed through the principles of “consistency” and “inheritance” (Steedman, 2000).

Let us look at some examples of rule type schemata. For instance, the combinator $B$ gives rise to forward ($>$) and backward ($<$) composition rules which in turn can be further refined into forward and backward crossed ($B_\times$)
Thus, the CCG rule set can be extended with combinations of rules deriving an inventory of language specific rule type schema. With these rule type schemata, along with the basic application rules, words are concatenated into strings.

**Example 6.3 (Derivation of a relative clause in CCG)** To derive a relative clause with an object relative in English, the rule for forward composition is needed. The following example illustrates the use of forward composition to derive the relative NP ‘(the) team that Brazil defeated’ (Baldridge, 2002, ex. 124). (The example also illustrates that CCG uses type-lifting and composition instead of hypothetical reasoning.)

```
\[
\frac{\text{team}}{n} \frac{(n\backslash n)/(s\backslash np)}{s/(np\backslash s)} > T \frac{\text{defeated}}{np\backslash np} s\backslash np \quad > B
\]
```

**Language variation in CCG** The inventory of rule type schema determines which structural generalizations are possible for a grammar. Thus, language variation is captured through the different instantiations of rules which enhance the derivational capacity of the grammar system. The search for rule type schemata is led by empirical research. Steedman (2000) claims that it is exactly this inventory which forms the basis of Universal Grammar. Thus, language variation in CCG is, on the one hand, determined lexically by assigning the correct basic categories to the words and, on the other, by defining the rule type schemata that are needed for the language in question.

As Baldridge (2002) points out certain rule type schema that are needed for one language derive highly ungrammatical orderings in another language. To prevent the derivation of ungrammatical sentences in a language, one either disallows certain rule schemata or one imposes additional constraints to restrict the rules. Baldridge (2002) indicates that only restricting or disallowing certain rule schemata for certain individual grammars or across grammars...
does not provide us with the means to capture structural generalizations in natural language.

**MMCCG**

To reduce the overgeneration of rule type schemata, Baldridge (2002) restricts the set of rule type schemata and adds fine-grained control over the application of rules by adopting the multimodal perspective of type-logical grammar. An inventory of binary composition modes is added to delimit the binary composition modes. Through a hierarchy of modes of composition (e.g. $\ast$, $\triangleleft$, $\triangleright$), Baldridge (2002) ranks the application of rules to a certain level of applicability. The hierarchy ranges from modes that are severely restrictive to complete expressivity.

The distribution of modes on lexical categories in the lexicon imposes control over the application of the rule schemata. To see how the modal control works, we now examine the backward crossed composition rules. These rules are necessary rule type schemata for the derivation of indirect object extraction of English ditransitive verb: e.g. ‘(the) player that the referee gave a yellow card’ (Baldridge, 2002, ex. 216). To restrict the backward composition rule, Baldridge adds constraints such that only categories with a binary modality that matches or that are more expressive than the provided binary modality are subject to the rule application. By assigning the right modalities to the lexical entries in the lexicon, certain the rules may be applied when they match, whereas other rules are blocked when the modalities do not match.

Take, for instance, the following restriction on the backward crossed composition rule which will now only apply to categories with composition mode $\triangleleft$ or composition modes that occur higher in the hierarchy of composition modes.

$$Y /_{\triangleleft} Z \quad Y \setminus X \Rightarrow X \setminus_{\triangleleft} Z \quad [\triangleleft B \_]{\_}$$

In contrast to Steedman (2000), Baldridge (2002) focuses on the universal applicability of the rule type schema. In doing so, he manages to shift the burden of language variation from rule type schemata to the lexicon.

**Generative power**

Baldridge (2002) shows that a multimodal CCG grammar can be mapped to a standard CCG. Standard CCG has been shown to belong to the mildly context-sensitive grammar formalisms (Vijay-Shanker and Weir, 1990). In order to obtain the mild context-sensitivity result, numerical bounds on the combinators used have to be assumed, which means that, from a logical point of view, CCG is incomplete.

**Evaluating MMCCG**

The move of MMCCG to a more lexically controlled grammar system with only a restricted set of rule type schemata imports the technique of mode distinctions that has been pursued in the earlier work in type-logical grammar (cf. Kurtonina and Moortgat (1997); Moortgat and Oehrle (1994); Morrill
(1994)). Our criticism of the multimodal approach in CTL, in other words, also applies to MMCCG: faced with a problem of grammatical analysis, the MMCCG grammar writer is free to introduce a construction-specific mode distinction to control the application of a type-shifting rule. Although Baldridge (2002) proposes a hierarchy of modalities which have a different ranking in accessing the rules, the rules are not universal. Natural language variation in MMCCG is captured through language specific rule type schema which are instantiated in the distribution of binary modalities in the lexical entries of the various languages. An interesting line of research would be to extend the resource-sensitive control mechanism of MMCCG with unary control. The unary features ♦ and ◻ may also offer attractive possibilities for the structural control over type-shifting rules.

6.2 Future research

To conclude, let us briefly mention some lines for future research suggested by the approach taken in this thesis.

Our study of wh-question formation has been mainly concentrated with argument wh-phrases. It will be interesting to see how the results of our research extrapolate to wh-questions with wh-phrases other than argument wh-phrases, e.g. ‘how’, ‘why’ and ‘where’. The logic of variation may also be further tested with cross-linguistic studies of related constructions, such as relative clause constructions, or other displacement phenomena, such as topicalization and scrambling. Our prediction with respect to these phenomena would be that they can be captured with the restricted set of displacement postulates proposed in this thesis. Next to the phenomena one can group under the heading of phrasal displacement, there is a broad set of phenomena involving ‘clustering’ of some kind (such as for instance adjunction, cliticization and clause-union). These phenomena, we would claim, have structural properties distinct from phrasal displacement. We would conjecture that also variation within the domain of these clustering phenomena can be understood in terms of a highly restricted set of structural patterns such as those suggested by the work of cf. Kraak (1998) on French clitics and Moortgat (1999b) on Dutch verb raising.

The derivability pattern of unary features (♦◻A ⊢ A ⊢ ◻♦A) and the derivability schema between wh-type schema as examined in chapter 5 may also be explored further. Both derivability patterns provided explanations for the derivation of word-order and constituency of multiple wh-questions, question-answer sequences and the interaction of wh-phrases with other constituents. The use and possible extensions of these derivability relations have not been exhausted. Further research may lead to the recognition of linear orderings in natural language phenomena on the basis of such derivability patterns (cf. Bernardi (2002) on polarity sensitive expressions).

Research on cross-linguistic variation from a formal perspective may also
result in “linguistic generalizations”. As our cross-linguistic study of wh-question formation shows, the phenomenon is intimately related to other grammatical constructions e.g. topicalization, cliticization, do-support, verb clustering, preposition stranding. We provided an analysis of wh-question formation by isolating the phenomena from other grammatical constructions. However, as our study of the cross-linguistic differences in wh-question formation shows the different wh-type schemata turn out to be intricately related with other grammatical differences. Additional studies on the interaction of grammatical constructions, for instance cliticization with wh-question formation or polarity items in wh-question formation, may lead the way for further generalizations.

As well as offering an explanation for the diversity in natural languages, this thesis raises a number of questions for the formal and computational aspects of the logic of variation. In the evaluation of MG and MMCCG, we noted that both frameworks fall into the group of weakly context-sensitive grammars. The possibility of accommodating our framework among these weakly context-sensitive formalisms is an open question. A start in this direction has been made by Moot (2002), who showed that the complexity of a CTL framework allowing the full class of linear, non-expanding structural rules equals that of context-sensitive grammars. Our restricted set of displacement postulates is a proper subset of the class of postulates that Moot considers, which suggests that the framework we have been using to characterize the logic of variation falls within the class of mildly context-sensitive grammar formalisms.
Appendix A

Overview of rules

A.1  Base logic

Axiom rule

\[ A \vdash A \]

Binary connectives

\[
\begin{align*}
\frac{\Gamma \vdash A/B \quad \Delta \vdash B}{\Gamma \odot \Delta \vdash A} & \quad [/E] \\
\frac{\Delta \vdash B \quad \Gamma \vdash B\setminus A}{\Delta \odot \Gamma \vdash A} & \quad [/E] \\
\frac{\Delta \vdash A \bullet B \quad \Gamma[(A \circ B) \vdash C]}{\Gamma[\Delta] \vdash C} & \quad [\bullet E] \\
\frac{\Delta \vdash \Box A \quad \Gamma \vdash \Box A}{\Delta \vdash \Box A} & \quad [/I] \\
\frac{\Delta \vdash \lozenge A \quad \Gamma[\Diamond (A)] \vdash B}{\Gamma[\Delta] \vdash B} & \quad [\Diamond E]
\end{align*}
\]

Unary connectives

\[
\begin{align*}
\frac{\Gamma \vdash \Box A \quad \Diamond (\Gamma) \vdash A}{\Diamond (\Gamma) \vdash A} & \quad [\Box E] \\
\frac{\Diamond (\Gamma) \vdash A}{\Gamma \vdash \Box A} & \quad [\Box I] \\
\frac{\Gamma \vdash A \quad \Diamond (\Gamma) \vdash \Diamond A}{\Diamond (\Gamma) \vdash \Diamond A} & \quad [\Diamond I] \\
\frac{\Delta \vdash \Diamond A \quad \Gamma[\Diamond (A)] \vdash B}{\Gamma[\Delta] \vdash B} & \quad [\Diamond E]
\end{align*}
\]
A.2 Structural module

left displacement postulates
\[
\Gamma (\Delta_1 \circ (\Delta_2 \circ \Delta_3)) \vdash C \quad [P11] \\
\Gamma (\Delta_2 \circ (\Delta_1 \circ \Delta_3)) \vdash C \\
\Gamma (\Delta_1 \circ (\Delta_2 \circ \Delta_3)) \vdash C \\
\]

right displacement postulates
\[
\Gamma (\Delta_1 \circ (\Delta_2 \circ \Delta_3)) \vdash C \quad [Pr1] \\
\Gamma ((\Delta_1 \circ \Delta_2) \circ \Delta_3) \vdash C \\
\Gamma (\Delta_1 \circ (\Delta_2 \circ \Delta_3)) \vdash C \\
\]

A.3 Wh-type schema rules

wh-ex-situ left
\[
\Gamma \vdash \text{WH}_l(A, B, C) \quad A \circ \Delta \vdash B \\
\Gamma \circ \Delta \vdash C \\
\quad [\text{WH}_l'] \\
\]

wh-ex-situ right
\[
\Gamma \vdash \text{WH}_r(A, B, C) \quad \Delta \circ A \vdash B \\
\Gamma \circ \Delta \vdash C \\
\quad [\text{WH}_r'] \\
\]

wh-in-situ
\[
\Gamma \vdash \text{WH}_i(A, B, C) \quad \Delta [A] \vdash B \\
\Delta [\Gamma] \vdash C \\
\quad [\text{WH}_i] \\
\]

A.4 Syntax-semantic interface

Term labeling of the natural deduction rules
\[
\Delta \vdash u : A \quad \Gamma \vdash v : B \\

\Delta \circ \Gamma \vdash (u \circ v) : A \\ [/E] \\

\Delta \vdash v : B \\
\Delta \circ \Gamma \vdash (u \circ v) : A \\ [/E] \\

\Delta \vdash u : A \quad \Gamma \vdash v : B \\
\Delta \circ \Gamma \vdash (u \circ v) : A \\ [/E] \\

\Delta \vdash t : A \\
\Gamma \vdash u : B \\
\Delta \circ \Gamma \vdash (t, u) : A \circ B \\ [\star] \\
\]
A. Overview of rules

Term labeling of the wh-type schema rules

wh-ex-situ left

\[ \frac{\Gamma \vdash \omega : \text{WH}_{\text{ex}}(A, B, C) \quad x : A \circ \Delta \vdash \text{BODY} : B}{\Gamma \circ \Delta \vdash \omega \lambda x.\text{BODY} : C} \] \[ [\text{WH}_{\text{ex}}^l] \]

wh-ex-situ right

\[ \frac{\Gamma \vdash \omega : \text{WH}_{\text{ex}}^r(A, B, C) \quad \Delta \circ x : A \vdash \text{BODY} : B}{\Gamma \circ \Delta \vdash \omega \lambda x.\text{BODY} : C} \] \[ [\text{WH}_{\text{ex}}^r] \]

wh-in-situ

\[ \frac{\Gamma \vdash \omega : \text{WH}_{\text{in}}(A, B, C) \quad \Delta[x : A] \vdash \text{BODY} : B}{\Delta[\Gamma] \vdash \omega \lambda x.\text{BODY} : C} \] \[ [\text{WH}_{\text{in}}] \]

Reduction rules for lambda term equations

\( \lambda \)-conversion:

\[ (\lambda x^A.t[x]^B)[t/x]^{A\rightarrow B} \]

\( \pi \)-conversion:

\[ (\pi^1 \langle t, u \rangle)^{A\rightarrow B} \stackrel{\lambda \beta}{\sim} t^A \quad (\pi^2 \langle t, u \rangle)^{A\rightarrow B} \stackrel{\lambda \beta}{\sim} u^B \]

\( \eta \)-conversion:

\[ (\lambda x^A.t[x]^B)^{A\rightarrow \eta} \stackrel{\lambda \eta}{\sim} t^A \quad ((\pi^1 \langle t, u \rangle), (\pi^2 \langle t, u \rangle))^A \stackrel{\lambda \eta}{\sim} t^{A\rightarrow B} \]
Appendix B

Sample sentences

B.1 English

B.1.1 Basic sentences

(B.1) Mary saw a bird $\vdash s$
(B.2) did John see Bill? $\vdash q$
(B.3) John said that Bill saw Mary $\vdash s$

B.1.2 Direct questions

(B.4) Who does Sue believe saw a bird? $\vdash wh$
(B.5) *Who does Sue believe that saw a bird? $\n\vdash wh$
(B.6) Who does Bill say that Sue believes saw a bird? $\vdash wh$
(B.7) Who does Bill say Sue believes saw a bird? $\vdash wh$
(B.8) What did Sue believe Mary saw? $\vdash wh$
(B.9) What did Sue believe that Mary saw? $\vdash wh$
(B.10) What did Sue believe that John said that Mary saw? $\vdash wh$
B.1.3 Embedded interrogatives

(B.11) John asked whether Mary saw a bird? ⊢ s

(B.12) John asked what Mary saw? ⊢ s

(B.13) John asked who saw a bird? ⊢ s

B.1.4 Island constructions

Wh-islands

(B.14) *Who did John wonder whether saw a bird? ⊬ wh

(B.15) *Which bird did John wonder whether Mary saw? ⊬ wh

(B.16) *Who did John wonder saw what? ⊬ wh

(B.17) *Which bird did John wonder who saw? ⊬ wh

Adjunct islands

(B.18) *Which topic did John leave when Mary addressed? ⊬ wh

Complex np’s

(B.19) *Which kid did you see [the teacher who punished t]? ⊬ wh

(B.20) *Which man did you hear [the rumor that a kid bit t]? ⊬ wh

Coordinate structures

(B.21) *Which man did you invite Mary and? ⊬ wh

Across-the-board constructions

(B.22) Who wrote a book and reviewed the article? ⊳ wh

(B.23) Which book did Bill write and John read? ⊳ wh

(B.24) Which man did you invite a friend of and a brother of? ⊳ wh

Left branch constructions

(B.25) * Which did you see t picture? ⊬ wh

(B.26) * Which man did you see t ’s picture? ⊬ wh

(B.27) * Which man’s did you see t picture? ⊬ wh
B. Sample sentences

B.1.5  Pied-piping

(B.28) Which man’s picture did you see? ⊬ wh

Meaning assembly of pied-piping

(B.29) To whom did John give a book? ⊬ wh
(B.30) Whom did John give a book to? ⊬ wh

B.1.6  Multiple wh-questions

(B.31) Who saw whom? ⊬ wh
(B.32) What did John give to whom? ⊬ wh
(B.33) Whom did who see? ⊬ wh
(B.34) Which man read which book? ⊬ wh
(B.35) Which book did which man read? ⊬ wh

B.2  Japanese

B.2.1  Basic word order

(B.36) John-ga konpyuuta-o Mary-ni ageta ⊬ s
(= John gave a computer to Mary.)

(B.37) John-wa Mary-ga konpyuuta-o kirateiru to itta ⊬ s
(= John said Mary hates computers.)

B.2.2  Argument scrambling

(B.38) Mary-ni John-ga konpyuuta-o ageta ⊬ s
(= John gave a computer to Mary.)

(B.39) konpyuuta-o Mary-ni John-ga ageta ⊬ s
(= John gave a computer to Mary.)
B.2.3 Questions

Basic questions

(B.40) John-ga hon-o katta ka? ⊢ q
(= Did John buy a book?)

(B.41) Dare-ga hon-o katta ka? ⊢ wh
(= Who bought a book?)

(B.42) John-ga nani-o katta ka? ⊢ wh
(= What did John buy?)

Embedded questions

(B.43) John-wa [Mary-ga hon-o katta ka] tazuneta. ⊢ s
(= John asked if/whether Mary bought a book.)

(B.44) John-wa [dare-ga hon-o katta ka] tazuneta. ⊢ s
(= John asked who bought a book.)

(B.45) John-wa [dare-ga hon-o katta ka] tazuneta ka? ⊢ q
(= Did John ask who bought a book?)

(B.46) * John-wa [dare-ga hon-o katta ka] tazuneta ka? ⊬ wh
(= *Who, did John ask whether t, bought a book?)

Non-local wh-questions

(B.47) John-wa [Mary-ga nani-o katta to] sinzita ka? ⊢ wh
(= What did John believe that Mary bought?)

(B.48) John-wa [dare-ga hon-o katta to] sinzita ka? ⊢ wh
(= Who did John believe t, bought a book?)

(B.49) * John-wa [Mary-ga nani-o katta ka] sinzita ka? ⊬ wh
(= Did John believe what Mary bought?)

(B.50) * John-wa [Mary-ga nani-o katta ka] tazuneta ka? ⊬ wh
(= * What did John ask Mary bought t,?)

B.2.4 Ka versus kadooka

(B.51) John-wa [Mary-ga hon-o katta kadooka] tazuneta ka? ⊢ q
(= Did John ask if/whether Mary bought a book?)

(B.52) * John-wa [Mary-ga hon-o katta to] itta kadooka? ⊬ q
(= * If/whether John said that Mary bought a book?)
B. Sample sentences

(B.53) John-wa [Mary-ga nani-o katta ka] tazuneta ka? \( \vdash q \)

(= Did John ask what Mary bought?)

(B.54) ∗ John-wa [Mary-ga nani-o katta kadooka] tazuneta ka? \( \not\vdash q \)

(= ∗ Did John ask if/whether what Mary bought?)

B.2.5 Relative clauses

(B.55) John-ni hon-o ageta hito-ni \( \vdash \) dat

(= man who gave a book to John)

(B.56) Austen-ga kaita hon-o \( \vdash \) acc

(= book that Austen wrote)

B.2.6 Complex noun phrases

(B.57) John-ni nani-o ageta hito(-ni) \( \vdash \) WH_{\text{in}}(np, q, wh) (=whin)

约翰 what bought man

(= man whom John gave what)

(B.58) Mary-wa [John-ni nani-o ageta hito-ni] atta ka? \( \vdash \) wh

(= Mary met the man who gave what to John?)

(B.59) dare-ga kaita hon(-o) \( \vdash \) WH_{\text{in}}(np, q, wh) (=whin)

who wrote book

(= book that who wrote)

(B.60) Kimi-wa [dare-ga kaita hon-o] yonda ka? \( \vdash \) wh

(= You read the book that who wrote?)

B.2.7 Semantics of questions

(B.61) John-wa [Mary-ga nani-o katta to] itta ka? \( \vdash q / \_ / np \)

\( \lambda y.(\text{say} \ ((\text{buy} \ y \ \text{mary})) \ \text{john}) \)

(= What did John say that Mary bought?)

(B.62) John-wa [dare-ga hon-o katta ka] tazuneta ka? \( \vdash q \)

(\( \lambda x.((\text{buy book} \ x)) \ \text{john} \))

(= Did John ask who bought a book?)

(B.63) ∗ John-wa [dare-ga hon-o katta ka] tazuneta ka? \( \not\vdash q / \_ / np \)

(\( \not\lambda y.((\text{buy} \ y \ \text{mary})) \ \text{john} \))

(= Who did John ask whether t_i bought a book?)
B.2.8 Multiple wh-questions

(B.64) Dare-ga nani-o katta ka? ⊢ (q/np) ⊃ np λx.λy.((buy y) x)
  1. (= Who bought what?)
  2. (= What did who buy?)

(B.65) John-wa [dare-ga nani-o katta to] itta ka? ⊢ (q/np) ⊃ np
  λx.λy.((say ((buy y) x)) y)
  1. (= Who did John say bought what?)
  2. (= What did John say who bought?)

B.3 Serbo-Croatian

B.3.1 Multiple wh-questions

(B.66) Ko koga vidi? ⊢ wh
(B.67) Koga ko vidi? ⊢ wh
(B.68) Ko kome šta kupuje? ⊢ wh
  (Who bought what to whom?)
(B.69) Sta ko kome kupuje? ⊢ wh
  (What did who buy for whom?)
(B.70) Ko kome šta kupuje? ⊢ wh
(B.71) Kome ko šta kupuje? ⊢ wh
(B.72) Sta kome ko kupuje? ⊢ wh

B.3.2 Long-distance extraction

(B.73) Ko Ivan želi da koga vidi? ⊢ wh
   "Who does Ivan want to see whom?"
(B.74) Koga Ivan želi da ko vidi? ⊢ wh
   "Whom does Ivan want who to see?"
(B.75) Ko Ivan tvrdi da koga vidi? ⊢ wh
   "Who does Ivan claim to see whom?"
(B.76) * Koga Ivan tvrdi da ko vidi? ⊬ wh
B. Sample sentences

B.4 Bulgarian

B.4.1 Multiple wh-questions

(B.77) Kogo kakvo Ivan kupuva? ⊢ □♦wh
(To whom did Ivan buy what?)

(B.78) * Kakvo kogo Ivan kupuva? □♦wh

(B.79) Koj kogo kakvo kupuva? ⊢ □♦wh
(Literally: Who did to whom buy what?)

(B.80) Koj kakvo kogo kupuva? □♦wh

B.4.2 Clitic placement

Sample sentences

(B.81) Koj kakvo ti e kazal? ⊢ □♦wh
'Who told you what?'

(B.82) * Koj ti e kakvo kazal? □♦wh

B.4.3 Long-distance extraction

(B.83) Koj kogo misliš če e vidjal? ⊢ □♦wh
(Who do you think saw whom?)

(B.84) * Kogo koj misliš če e vidjal? □♦wh

(B.85) * Koj misliš če e vidjal kogo? □♦wh

(B.86) ?Koj misliš če kogo e vidjal? ⊢ □♦wh

B.5 German

B.5.1 Basic word order

(B.87) Miro sieht ein Bild ⊢ s

(B.88) Picasso hat ein Bild gemalt ⊢ s

(B.89) Hat Picasso ein Bild gemalt? ⊢ q

(B.90) (dass) Picasso ein Bild gemalt hatte ⊢ s,

(B.91) Miro glaubte dass Picasso ein Bild gemalt hatte ⊢ s
B.5.2 Basic wh-questions

(B.92) Wer hat ein Bild gemalt? ⊨ wh
(B.93) Welches Bild hat Picasso gemalt? ⊨ wh
(B.94) Wer hat was gemalt? ⊨ wh
(B.95) Was hat wer gemalt? ⊨ wh
(B.96) Welches Bild glaubte Miro dass Picasso gemalt hatte? ⊨ wh
(B.97) Welches Bild glaubte Miro dass Jacob meint dass Picasso gemalt hatte?
 ⊨ wh
(B.98) * Miro hat was gemalt? ⊭ wh

B.5.3 Scope marking constructions

(B.99) Was glaubte Miro welches Bild Picasso gemalt hatte? ⊨ wh
(B.100) Was glaubte Hans was Jacob meint welches Bild Picasso gemalt hatte?
 ⊨ wh
(B.101) Was glaubte Hans welches Bild Jacob meint dass Picasso gemalt hatte?
 ⊨ wh
(B.102) Was glaubte Miro dass Picasso ein Bild gemalt hatte? ⊭ wh

B.5.4 Constraints on scope marking

(B.103) Welches Bild glaubte Miro dass Picasso gemalt hatte? ⊨ wh
(B.104) Welches Bild möchte Miro dass Picasso gemalt hatte? ⊨ wh
(B.105) Was glaubte Miro welches Bild Picasso gemalt hatte? ⊨ wh
(B.106) * Was möchte Miro welches Bild Picasso gemalt hatte? ⊭ wh
(B.107) Miro glaubte dass Picasso ein Bild gemalt hatte ⊨ s
(B.108) Miro möchte dass Picasso ein Bild gemalt hatte ⊨ s

B.5.5 Questions and answers

(B.109) Was glaubte Miro was Hans meint welches Bild Picasso gemalt hatte?
 ⊨ λy.(believe ((claim ((paint y) picasso)) hans)) miro : s/3 np
(B.110) Was glaubte Miro welches Bild Hans meint dass Picasso gemalt hatte?
 ⊨ λy.(believe ((claim ((paint y) picasso)) hans)) miro : s/3 np
B. Sample sentences

B.6 Hindi

B.6.1 Basic word order

(B.111) Anu kitaab kharidii ⊢ s
        (=Anu bought a book)

B.6.2 Basic direct and indirect questions

(B.112) Anu kyaa karnaa jaantaahai? ⊢ wh
        (=What does Anu know to do?)

(B.113) * Anu kyaa karnaa soctaahai s
        (=Anu thinks what to do.)

(B.114) Anu jaantaahai ki kyaa karnaa hai ⊢ s
        (=Anu know that what is done.)

(B.115) * Anu soctaahai ki kyaa karnaa hai? ⊬ wh
        (=What does Anu think is done?)

B.6.3 Long-distance wh-questions

(B.116) Kisko siitaane jaantaahai ki Raviine dekhaa? ⊢ wh
        (=Who does Sita know that Ravi saw?)

(B.117) Siitaane kisko soctaahai ki Raviine dekhaa? ⊢ wh
        (=Who does Sita think that Ravi saw?)

(B.118) Siitaane jaantaahai ki Raviine kisko dekhaa ⊢ s
        (=Sita knows who Ravi saw?)

(B.119) * Siitaane soctaahai ki Raviine kisko dekhaa ⊬ s

(B.120) * Siitaane soctaahai ki Raviine kisko dekhaa? ⊬ wh

B.6.4 Scope marking constructions

(B.121) Jaun kyaa soctaahai ki merii kisko dekhaa? ⊢ wh
        (=Who does John think Mary saw?)

(B.122) Jaun kyaa soctaahai ki kisko merii dekhaa? ⊢ wh

(B.123) * Jaun soctaahai ki merii kisko dekhaa? ⊬ wh
        (=Who does John think Mary saw? without scope marker)
(B.124) Jaun kyaa soctahai ki kisko merii dekhaa? ⊢ wh  
(=Who does John think Mary saw?)

(B.125) Jaun kyaa soctahai ki anu kyaa kahegii ki kisko merii dekhaa? ⊢ wh  
(=Who does John think that Anu said that Mary saw?)

B.6.5 Questions and answers

(B.126) Sitaane kisko soctahai ki Raviine dekhaa? ⊢ s\_\text{np}  
(=Who does Sita think that Ravi saw?)

(B.127) Jaun kyaa soctahai ki merii kisko dekhaa? ⊢ s\_\text{np}  
(=Who does John think Mary saw?)

(B.128) Jaun kyaa puucha ki kisne kisko dekhaa? ⊢ s\_\text{np}  
(=Whom does John ask that who saw?)

B.7 Minimalist grammar analyses

(B.129) John likes Mary.

(B.130) Does John like Mary?

(B.131) Who likes Mary?

(B.132) Who does John like?
Bibliography


Samenvatting

Talen verschillen in de wijze waarop woorden en zinnen of delen daarvan ge-
combineerd worden om een betekenisvolle expressie te vormen. Bijvoorbeeld,
in de Engelse wh-vraagzin ‘Who does John think left?’ neemt het vraagwoord
(‘Who’) een positie in vooraan de zin, terwijl de eigenlijke positie van het be-
vraagde onderwerp een lege plek achterlaat in de onderschikkende bijzin:
(‘John thinks (Bill) left’).1 In het Japans daarentegen, staat het vraagwoord in
een wh-vraagzin met een gelijke betekenis op dezelfde positie als het bevaar-
de onderwerp, ‘John thinks who left?’ Hoewel de vorm van de wh-vraagzinnen
in deze twee talen verschilt, is de betekenis van de twee zinnen dezelfde: ‘Wie
is de persoon waarvan John denkt dat die persoon wegging?’. Er is dus een con-
trast tussen gelijkheid in betekenisopbouw (betekenisuniformiteit) en variatie
in de structurele opbouw van een zin (structurele variatie). Wat is de logica
achter deze variatie? Deze vraag staat centraal in dit proefschrift.

Taalkundigen streven naar de ontwikkeling van een systeem van Universele
Grammatica dat zowel de structurele variatie tussen talen als de betekenis-
uniformiteit verklaart. Onder invloed van Chomsky’s recente onderzoeks-
programma “The Minimalist Program” (Chomsky, 1995) wordt binnen het
onderzoeksgebied van de generatieve grammatica een theorie van universele
grammatica (UG) ontwikkeld waarin principes worden geformuleerd die de
d mogelijke variatie inperken maar tegelijkertijd ook betekenisuniformiteit ver-
klaren. Taalkundigen in de Chomskyaanse traditie werken vanuit een breed
empirisch perspectief aan het blootleggen van grammaticale verschillen tus-
sen talen. Tegenover deze taalkundigen staan logici die natuurlijke taal vanuit
een wiskundig perspectief beschouwen. Nadat Lambek zijn ideeën uiteen had
gezet in zijn artikel “The mathematics of sentence structure” (Lambek, 1958)
zijn binnen de categoriale grammatica verschillende “grammatica logica’s”
ontwikkeld waarin structurele variatie en betekenisuniformiteit in deductieve
termen worden uitgelegd.

1Vraagzinnen met vraagwoorden zoals, ‘wie’, ‘wat’ of ‘waarom’ worden door taalkundigen
wh-vraagzinnen genoemd, omdat in het Engels dergelijke vraagwoorden met de letters ‘wh’ be-
ginnen.
Zoals gezegd werken taalkundigen in de generatieve traditie en taalkundigen in de traditie van categoriale grammatica vanuit zeer verschillende achtergronden aan een overeenkomstig vraagstuk. Dit proefschrift verbindt deze twee perspectieven en formuleert een antwoord op het vraagstuk van structurele variatie en betekenisuniformiteit.

Dit proefschrift stelt een *logica van variatie* voor die functioneert als een systeem van universele grammatica. Het uitgangspunt voor deze logica is de Categoriale Typenlogica (CTL). Hoofdstuk 2 geeft een overzicht van de werking van CTL. Deze technieken worden later gebruikt om een logica van variatie te formuleren. CTL is een grammaticaal redeneersysteem waarin de vorm- en betekeniscomponenten van talige expressies worden opgebouwd. Het is een sterk gelexicaliseerd systeem, wat betekent dat de afleiding van een uitdrukking volledig bepaald wordt door de typen die worden toegekend aan de afzonderlijke woorden waaruit de uitdrukking is opgebouwd. Het universele grammaticale rekensysteem dat deze lexicaal basiselementen samenvoegt tot zinsdelen en uiteindelijk tot zinnen bestaat uit twee gedeeltes: een basislogica en een structurele module. De basislogica bevat afleidingsregels die nodig zijn om de grammaticale vorm (syntax) en de betekenis (semantiek) van een expressie af te leiden. De structurele module bestaat uit betekenisbehoudende relaties tussen grammaticale vormen van een expressie. De basislogica en de structurele module werken op elkaar in en leiden expressies af die bestaan uit twee componenten: de vorm en de betekenis van een talige expressie. De vormcomponent beschrijft de lineaire volgorde en de constituentstructuur van een expressie. De betekeniscomponent beschrijft de semantische relaties tussen de basiselementen van een uitdrukking. Vorm en betekenis zijn met elkaar verbonden via de basislogica. De structurele module geeft de mogelijkhed om aan één betekenis meerdere structurele realisaties te koppelen.

De structurele module zoals oorspronkelijk geformuleerd in CTL is te expressief om de variatie tussen talen afdoende te verklaren. Aan het einde van hoofdstuk 2, bij de karakterisering van de logica van variatie, wordt de structurele module beperkt tot een aantal structurele regels. De logica van variatie is een universele grammatica waarmee een verklaring kan gegeven worden voor de structurele variatie en betekenisuniformiteit van vraagzinconstructies.

Voor het vergelijken van vraagzinconstructies in verschillende talen door deze te analyseren met behulp van de logica van variatie zijn drie factoren belangrijk: (1) Hogere-orde type toekenning: wh-vraagwoorden krijgen een specifieke type toegekend waarmee de betekenisuniformiteit van wh-vraagzinnen verklaard wordt; (2) Een vaste structurele module: structurele variatie is beperkt door een klein aantal structurele regels dat is vastgelegd in de Universele Grammatica; (3) Sterke lexicaliteit: als consequentie van punten (1) en (2) moet taalvariatie van wh-vraagzinconstructies volledig bepaald worden door lexicale verschillen; er bestaan geen taalspecifieke structurele regels. In dit
proefschrift wordt empirische ondersteuning geleverd voor deze assumpties door een taalvergelijking analyse te geven van wh-vraagzinconstructies.

In hoofdstuk 3 presenteer ik een uniforme aanpak van vraagzinconstructies. Aan de hand van een aantal basiskenmerken van vraagzinconstructies in het Engels, wordt het grammaticale redeneersysteem verfijnd door de introductie van een wh-typeschema $WH(A, B, C)$. Het typeschema is een afgekorte schematische presentatie van afleidingsregels uit de basislogica die nodig zijn om een wh-vraagwoord van een bepaald type $A$ samen te voegen met de basis van een vraagzin van type $B$ om een vraagzin van type $C$ te vormen. Neem bijvoorbeeld de vraagzin ‘Who does John think left?’. Om deze vraagzin af te leiden bouwt het rekensysteem allereerst de basis van de vraagzin, ‘does John think np left’, waarin op de plaats van het onderwerp een hypothese wordt aangenomen van type $np$ (afkorting van noun phrase, zelfstandig naamwoord). Deze hypothese is nodig om de onderschikkende bijzin ‘np left’ te kunnen vormen. De basis van de vraagzin ‘does John think np left’ is van type $q$, een type dat wordt toegekend aan ja/nee vraagzinnen. De volgende stap is het toevoegen van het wh-vraagwoord ‘who’. ‘Who’ is in het lexicon getypeerd als het wh-typeschema $WH(np, q, wh)$ dat betekent dat het wh-vraagwoord toegevoegd kan worden aan een zinsdeel van type $q$ met een hypothese van type $np$. Na het samenvoegen van het vraagwoord met het zinsdeel wordt de hypothese opgeven en wordt het vraagwoord vooraan in de zin geplaatst. De gehele constructie vormt de wh-vraagzin ‘Who does John think left?’ van type $wh$.

Er worden drie varianten van wh-typeschema’s voorgesteld waarmee verschillende voorkomens van wh-vraagwoorden wor den onderscheiden: een tweetal wh-typeschema’s die vooraan in de zin worden geplaatst (ex-situ) maar met een andere positie van de hypothese worden verbonden, en één wh-typeschema dat op dezelfde positie als de hypothese geplaatst wordt (in-situ). Ondanks een onderscheid in het structurele voorkomen van deze wh-typeschema’s is de semantische opbouw van de vraagzinnen dezelfde. Met de verschillende wh-typeschema’s en de bijbehorende syntactische en semantische representatie illustreer ik de opbouw van vraagzinconstructies aan de hand van een analyse van verschillende vraagzinnen in het Engels.

duid met een vraagzinsteken, het woord ‘ka’, in de voorbeeldzin gemarkeerd met Q):

(1) John-wa nani-o tabe-masita ka?  
   John wat at Q  
   ‘Wat had John gegeten?’

Wh-ex-situ talen zijn talen waarvan het vraagwoord vooraan in de zin geplaatst wordt, waar ook het Nederlands onder valt. Een extra onderscheid kan aangebracht worden binnen deze groep, namelijk talen met zogenaamde enkelvoudige wh-ex-situ constructies en talen met meervoudige ex-situ constructies. Het Engels is een voorbeeld van een taal met een enkelvoudige ex-situ constructie, omdat in deze taal als er meerdere vraagwoorden in de zin staan, slechts één vraagwoord voorop geplaatst wordt. In talen met meervoudige ex-situ constructies, zoals in het Servokroatisch en het Bulgaars worden alle vraagwoorden voorop geplaatst. Neem bijvoorbeeld het volgende voorbeeld uit het Bulgaars met twee vraagwoorden:

(2) Koj kakvo vižda?  
    wie wat ziet  
    ‘Wie heeft wat gezien?’

De laatste groep bestaat uit talen waarin naast wh-ex-situ of wh-in-situ constructie wh-vraagzinnen gevormd kunnen worden door het wh-vraagwoord alleen gedeeltelijk te verplaatsen. Een extra kenmerk bij deze constructie is dat de hoofdzin een markering bevatt, veelal ook een wh-vraagwoord, dat de zin aanduidt als een wh-vraagzin. Het volgende voorbeeld uit het Duits illustreert een wh-vraagzin die gevormd is door een gedeeltelijke verplaatsing van het wh-vraagwoord.

(3) Was glaubte Miró welches Bild Picasso gemalt hatte?  
    wat geloofde Miró welk schilderij Picasso geschilderd had?  
    ‘Van welk schilderij geloofde Miró dat Picasso het geschilderd had?’

Op basis van de wh-typeschema’s laat ik zien hoe de structurele variatie in wh-vraagzinconstructies verklaard kunnen worden. De specifieke constructie van een wh-vraagzin in een taal wordt bepaald door de juiste instantie in het lexicon te geven van wh-typeschema voor een wh-vraagwoord. De verschillende instanties van wh-typeschema’s bepalen hoe een vraagzin in een bepaalde taal of in een bepaalde context wordt opgebouwd. De interactie van de wh-vraagwoorden met andere constructies in de zin zorgen voor variatie in woordvolgorde of kunnen in sommige gevallen veroorzaken dat een wh-vraagzin niet afgeleid kan worden en dus ongrammaticaal is.
Hoofdstuk 5 concentreert zich op de koppeling tussen vorm (syntax) en betekenis (semantiek). Op basis van een formele koppeling, het zogenaamde Curry-Howard isomorphisme, tussen syntactische typen en semantische typen zoals vastgelegd in de basislogica, is de uniforme betekenis opbouw van wh-vraagzinnen af te leiden. Allereerst stel ik voor om het type dat aan vraagzinnen wordt toegekend te ontleden door de typen van mogelijke antwoorden te verenigen met het vraagzintype. Een vraagzin zoals bijvoorbeeld ‘Who does John think left?’ krijgt een type dat een zelfstandig naamwoord, bijvoorbeeld ‘Bill’, als antwoord verwacht om een bewering te vormen; de vraag samen met het antwoord vormt een bewering van type $s$. Met deze aanpassing in zowel het syntactische type als de toekenning van semantische termen aan wh-vraagwoorden kan de uniforme betekenisopbouw van wh-vraagzinnen worden afgeleid.

Doordat het antwoordtype is ingebed in het vraagzintype, kun je vanuit één instantie van het wh-typeschema andere instanties afleiden. Op basis van typerende stellingen binnen de taal van semantische typen, presenteer ik een patroon van afleidbaarheidsrelaties tussen wh-typeschema’s. Met dit patroon zijn vanuit één enkele lexicale typetoekenning van een wh-typeschema aan een wh-vraagwoord verschillende fenomenen te verklaren. Ter illustratie geef ik de syntactische analyse en betekenisopbouw van verschillende soorten wh-vraagzinnen, bijvoorbeeld meervoudige wh-vraagzinnen in ex-situ en in-situ talen, ingebedde vraagzinnen, wh-vraagzinnen met complexe zelfstandige naamwoorden en vraagzinnen met gedeeltelijke verplaatsing van het vraagwoord.

In hoofdstuk 6, het afsluitende hoofdstuk van dit proefschrift, keer ik terug naar de centrale vraag van dit proefschrift: Wat is de logica achter taalvariatie?. In dit laatste hoofdstuk evalueer ik mijn bevindingen door een vergelijking te maken met twee andere grammatica formalismen die uitgaan van een lexicaal gestuurd grammatica systeem: minimalistische grammatica’s en multimodale combinatorische categorische grammatica. Naast een korte beschrijving van de twee formalismen, bespreek ik hun verklaring voor het contrast tussen taalvariatie en uniforme betekenisopbouw. Deze vergelijking toont aan dat met de restricties van de structurele regels binnen de logica van variatie en de drie genoemde factoren (hogere orde typen, een vast structurele module eb sterke lexicaliteit) voor de analyse van vraagzinconstructies een brede basis is gelegd voor het verklaren van taalvariatie.

Voor de analyse van vraagzinconstructies is voor iedere taal een grammatica fragment gemaakt. Deze grammatica fragmenten zijn te analyseren met behulp van Grail, de door Richard Moot vervaardigde parser voor categoriale typenlogica’s. Met behulp van de bij dit boek geleverde CD-rom kunnen de fragmenten aan de hand van de zinnen in Appendix B worden geanalyseerd en eventueel verder worden uitgebreid.
Curriculum Vitae

Willemijn Vermaat was born in Rotterdam on the 30th of January, 1974. She finished her pre-university secondary education at S.G. De Krimpenerwaard in Krimpen aan den IJssel in 1992 and obtained an Atheneum Diploma to start university education. From 1993 she studied Cognitive Artificial Intelligence (Cognitive Kunstmatige Intelligentie) at the Department of Philosophy at Utrecht University. In April 1999, she obtained her masters degree with a thesis titled Controlling Movement: Minimalism in a deductive perspective.

In 1999 and 2000, Willemijn worked as junior research fellow at the Utrecht Institute of Linguistics OTS for the development of courseware and web support to enhance the quality of the computational linguistics curriculum. From 2001 until 2005, she held a PhD position (Assistent in Opleiding) at the Utrecht Institute of Linguistics OTS. The present dissertation is the result of research pursued by her. She was awarded a Fulbright AiO/OiO travel grant to carry out research at the Linguistics Department at the University of California, Los Angeles (UCLA) from September until December 2001.