UNIVERSITÉ DE LORRAINE | BIBLIOTHÈQUES UNIVERSITAIRES

AVERTISSEMENT
DROIT D'AUTEUR

Ce document est le fruit d'un long travail approuvé par le jury de soutenance et mis à disposition de l'ensemble de la communauté universitaire élargie.

Il est soumis à la propriété intellectuelle de l'auteur. Ceci implique une obligation de citation et de référencement lors de l'utilisation de ce document. Toute contrefaçon, plagiat ou reproduction illicite encourt une poursuite pénale.

Contact bibliothèque : ddoc-theses-contact@univ-lorraine.fr
(Cette adresse ne permet pas de contacter les auteurs)

Code de la Propriété Intellectuelle. articles L 122. 4
Code de la Propriété Intellectuelle. articles L 335.2- L 335.10

UNIVERSITÉ DE LORRAINE
34 COURS LÉOPOLD — CS 25233
54052 NANCY CEDEX
TÉL : 03 54 50 54 00
CONTACT@UNIV-LORRAINE.FR
WWW.UNIV-LORRAINE.FR

UNIVERSITÉ DE LORRAINE

Ecole Doctorale IAEM
Laboratoire Loria

UNIVERSITEIT VAN AMSTERDAM

En cotutelle avec l'Université d'Amsterdam

Thèse

Présentée et soutenue publiquement pour l'obtention du titre de

DOCTEUR DE l'UNIVERSITE DE LORRAINE

Mention : INFORMATIQUE

par **Valentin Dominique Marie RICHARD**

Sous la direction de Philippe de Groote,

Reinhard Muskens et Floris Roelofsen

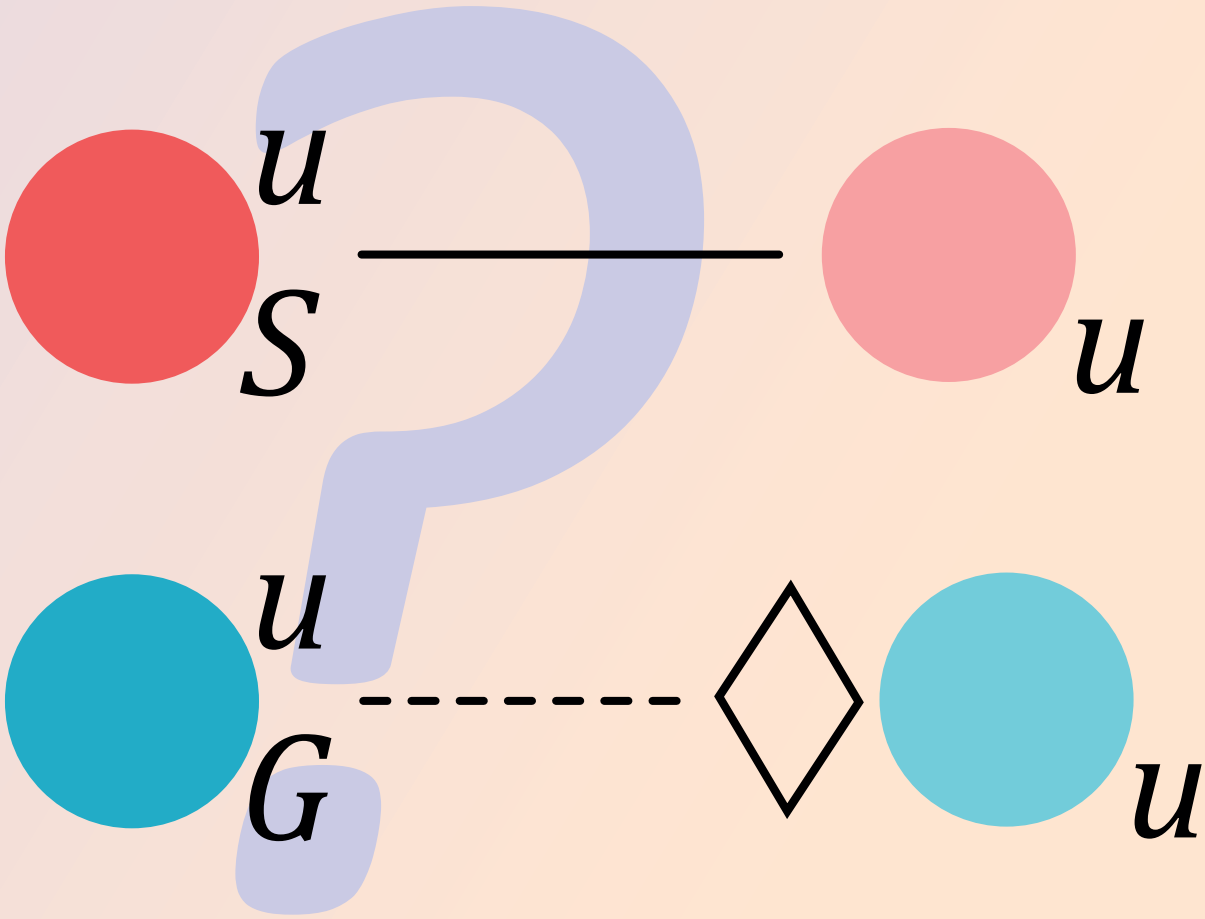
Aspects dynamiques et présuppositionnels des questions

3 février 2026

Membres du jury :

Directeurs de thèse :	**M. Philippe DE GROOTE**	**Directeur de recherche, Centre Inria de l'Université de Lorraine, Villers-lès-Nancy**
	M. Reinhard MUSKENS	**Professeur émérite, Université d'Amsterdam, Amsterdam**
	M. Floris ROELOFSEN	**Professeur, ILLC, Université d'Amsterdam, Amsterdam**
Présidente de jury :	**Mme. Claire GARDENT**	**Directrice de recherche, Loria, CNRS, Vandœuvre-lès-Nancy**
Rapporteurs :	**M. Lucas CHAMPOLLION**	**Professeur associé, New York University, New York**
	M. Jonathan GINZBURG	**Professeur, Laboratoire de Linguistique Formelle, Université Paris Cité, Paris**
Examinateurs :	**M. Maxime AMBLARD**	**Professeur, Loria, Université de Lorraine, Vandœuvre-lès-Nancy**
	Mme. Maria ALONI	**Professeure associée, ILLC, Université d'Amsterdam, Amsterdam**
	M. Robert VAN ROOIJ	**Professeur, ILLC, Université d'Amsterdam, Amsterdam**

Presuppositional and Dynamic Aspects of Questions



Valentin D. Richard

Presuppositional and Dynamic Aspects of Questions

Valentin D. Richard

Presuppositional and Dynamic Aspects of Questions

ILLC Dissertation Series DS-2026-03



Institute for Logic, Language and Computation

For further information about ILLC-publications, please contact

Institute for Logic, Language and Computation
Universiteit van Amsterdam
Science Park 107
1098 XG Amsterdam
phone: +31-20-525 6051
e-mail: illc@uva.nl
homepage: http://www.illc.uva.nl/

This PhD was funded by a Contrat Doctoral Spécifique Normalien (CDSN) provided by the École Normale Supérieure Paris-Saclay for a three-year PhD at Université de Lorraine, and by the Institute for Logic Language and Computation (ILLC).

Copyright © 2026 by Valentin D. Richard

Cover design by Valentin D. Richard.
Printed and bound by Ipskamp Printing.

ISBN: 978-94-6536-026-3

Presuppositional and Dynamic Aspects of Questions

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad van doctor
aan de Universiteit van Amsterdam
op gezag van de Rector Magnificus
prof. dr. ir. P.P.C.C. Verbeek
ten overstaan van een door het College voor Promoties ingestelde commissie,
in het openbaar te verdedigen in de Agnietenkapel
op dinsdag 3 februari 2026, te 16.00 uur

door Valentin Dominique Marie Richard
geboren te Saint-Sébastien-sur-Loire

Promotiecommissie

Promotores:	dr. P. De Groote	Inria
	dr. R.A. Muskens	Universiteit van Amsterdam
Copromotores:	prof. dr. F. Roelofsen	Universiteit van Amsterdam
Overige leden:	dr. L. Champollion	New York University
	dr. M.D. Aloni	Universiteit van Amsterdam
	prof. dr. ing. R.A.M. van Rooij	Universiteit van Amsterdam
	prof. dr. M. Amblard	Université de Lorraine
	prof. dr. C. Gardent	CNRS

Faculteit der Natuurwetenschappen, Wiskunde en Informatica

Dit proefschrift is tot stand gekomen binnen een samenwerkingsverband tussen de Universiteit van Amsterdam en de Université de Lorraine met als doel het behalen van een gezamenlijk doctoraat. Het proefschrift is voorbereid aan de Faculteit der Natuurwetenschappen, Wiskunde en Informatica van de Universiteit van Amsterdam en à L'Ecole Doctorale IAEM-Lorraine de l'Université de Lorraine.

This thesis was prepared within the partnership between the University of Amsterdam and Université de Lorraine with the purpose of obtaining a joint doctorate degree. The thesis was prepared in the Faculty of Science of the University of Amsterdam and in the IAEM - Lorraine Doctoral School of Université de Lorraine.

À mon usine à phrases

Contents

Acknowledgments **xiii**

1 Introduction **1**
1.1 Scope of this thesis 1
1.1.1 Presuppositions 1
1.1.2 Anaphora 1
1.1.3 Presuppositions and anaphora 2
1.1.4 Main questions 2
1.2 Variation in presupposition and reference 3
1.3 Reading ambiguity and local contexts 4
1.4 Modeling referents and contexts 5
1.5 Outline of the thesis 6

Part One: Presuppositions of Questions

2 Existential Inference of Wh-Questions **11**
2.1 Introduction 11
2.2 Properties of the existential inference of simplex wh-words 13
2.2.1 Spontaneous triggering 13
2.2.2 Cancellation and suspension 14
2.2.3 Null answers are full-fledged answers 15
2.2.4 Sensitivity to semantic environments 17
2.2.5 Weak islands 18
2.2.6 Sensitivity to the asker's goal 19
2.2.7 Interim summary 20
2.3 Motivating a specific vs. generic duality 21
2.3.1 Focus marking in Mongolian, Ngamo and French 21
2.3.2 Duality in non-interrogative wh-items 27

2.3.3 Duality of anaphoric properties . . . 29
2.4 The specific vs. generic ambiguity theory . . . 31
2.4.1 Epistemic pronouns and identification . . . 31
2.4.2 Null answers . . . 35
2.4.3 Explaining spontaneous triggering . . . 39
2.5 Applying the specific vs. generic theory . . . 40
2.5.1 Weak NPIs in questions . . . 40
2.5.2 Question types and the reading likelihood gradient . . . 44
2.5.3 Explaining weak islands . . . 47
2.6 An alternative analysis . . . 49
2.6.1 Null individuals are marked . . . 49
2.6.2 Pragmatic referential anchoring . . . 50
2.6.3 Drawbacks of the pragmatic approach . . . 50
2.7 Comparison with other approaches . . . 51
2.7.1 Grammatical derivation . . . 51
2.7.2 Pragmatic derivation without an asymmetry . . . 53
2.7.3 Ambiguity due to a semantic operator . . . 54
2.7.4 The bias hypothesis . . . 55
2.7.5 The type-flexibility hypothesis . . . 56
2.8 Conclusion . . . 59

Part Two: Dynamic Aspects of Questions

3 Dynamic Inquisitive Semantics 63
3.1 Inquisitive semantics . . . 63
3.1.1 Introduction . . . 63
3.1.2 Semantic space . . . 64
3.1.3 Inquisitive logic . . . 65
3.2 Dynamic inquisitive objects . . . 66
3.2.1 Introduction . . . 66
3.2.2 Contexts . . . 68
3.2.3 Relations between contexts . . . 69
3.3 Syntax-semantics interface . . . 70
3.3.1 Logical language . . . 70
3.3.2 A fragment of English . . . 73
3.3.3 The left periphery . . . 73
3.3.4 Example . . . 74

4 Dynamics of Modalized Questions 77
4.1 Modalized questions . . . 77
4.1.1 Possibility weakens questions . . . 77
4.1.2 Modal subordination with modalized questions . . . 78

4.2 Previous works . . . 79
4.2.1 Non-distributive test semantics for modals . . . 79
4.2.2 Stack-based treatment of local contexts . . . 80
4.3 Modal dynamic inquisitive semantics . . . 81
4.3.1 Inquisitive possibility . . . 81
4.3.2 Dynamic inquisitive possibility . . . 82
4.3.3 Pushing and percolating . . . 83
4.3.4 Meeting the requirements . . . 84
4.4 Conclusion . . . 87

5 Wh-Anaphora **89**
5.1 Introduction . . . 89
5.1.1 Referring to a wh-word . . . 89
5.1.2 Anaphora theory . . . 90
5.2 Answers and co-reference to wh-words . . . 91
5.2.1 Issues . . . 92
5.2.2 Corpus study . . . 93
5.2.3 Different anaphora types in answers . . . 95
5.2.4 Theoretical arguments . . . 98
5.2.5 Interim summary . . . 100
5.3 Specific and generic co-reference to wh-words in non-answers . . . 100
5.3.1 Modal subordination with generic questions . . . 101
5.3.2 Identification and coherence . . . 103
5.4 Modeling wh-anaphora . . . 105
5.4.1 Modeling generic questions . . . 105
5.4.2 Modeling specific co-reference . . . 109
5.4.3 Modeling answers . . . 111
5.5 Conclusion . . . 113

6 Dynamic and Compositional Aspects of ADUs **115**
6.1 Introduction . . . 115
6.2 Dependence statements in inquisitive semantics . . . 117
6.3 Limits of previous models . . . 118
6.3.1 Dynamic issues . . . 118
6.3.2 Compositional issues . . . 120
6.3.3 Interim Summary . . . 123
6.4 Inquisitiveness triggers . . . 123
6.4.1 Inquisitiveness in the antecedent . . . 124
6.4.2 Inquisitiveness in the consequent . . . 126
6.5 Modeling assertive dependence utterances . . . 129
6.5.1 Modeling anaphoric properties . . . 129
6.5.2 Modeling exemplification . . . 130
6.5.3 Example . . . 131

6.6 Dependence utterances with an interrogative consequent 132
6.6.1 Dependence questions 132
6.6.2 Embedded interrogatives in the consequent 134
6.6.3 Presuppositions of dependence questions 134
6.7 Conclusion 135

7 Conclusion **137**
7.1 General summary 137
7.1.1 Properties of wh-words 137
7.1.2 Interaction between questions, modals and conditionals . . 138
7.2 Limits and future directions 138
7.2.1 More finely-grained properties for wh-words 139
7.2.2 Scope and instantiation of wh-words 141

A Extracting French Interrogatives from Treebanks **143**
A.1 Introduction 143
A.2 Related works 145
A.2.1 Corpus studies on French interrogatives 145
A.2.2 Question detection 145
A.2.3 Clause type features in UD 146
A.3 Method 147
A.3.1 Development phase 147
A.3.2 Description of FUDIA modules 148
A.3.3 Evaluation phase 151
A.3.4 Extraction 152
A.4 Results 153
A.4.1 Annotation task 153
A.4.2 Evaluation results 154
A.4.3 Other difficulties 156
A.4.4 Statistics on the French Interrogative Bank 157
A.5 Limits 157
A.5.1 Limits of the rule design policy 158
A.5.2 Evaluation limits 158
A.6 Conclusion 159

B Appendix of Chapter A **161**
B.1 Grammaticalized QU-expressions: some linguistic background . . 161
B.2 Coveney's classification of French interrogative forms 162
B.3 Linguistic conceptual limits 164
B.4 Baseline codes 166

C Corpus extracts **169**

Bibliography **177**

List of symbols and abbreviations **209**

Résumé **211**

Samenvatting **217**

Abstract **219**

Acknowledgments

I am very grateful to all the people I met during these past four years. Some of them helped me shape my ideas, my plans and my career. Some of them made me feel good and included in the professional world. And some of them supported me personally.

First and foremost, I would like to thank my supervisors. They accepted and managed to build this joint PhD despite the administrative supplement. Despite the challenges, we got through them. Thank you, Philippe, for believing in my projects, trusting me, and letting me explore what I wanted. Your suggestion to organize a workshop on the semantics of questions in Nancy pushed me to extend my skills and network. I am grateful for that. Thank you, Floris, for your very relevant linguistic examples and your advice about article writing. Thank you, Reinhard, for your daily support. You took the time to listen to my problems, even though you were already retired. You remained optimistic and happy about my ideas, kindly commenting on my mistakes. I would have lost my mind without you.

Je remercie très chaleureusement tout l'équipe Sémagramme, qui a été une fantastique équipe et dans laquelle j'ai pu m'épanouir professionnellement et personnellement. Travailler dans le grand bureau B230 a été source de joie, d'entraide et de complicité. Merci Karën Fort pour toute ton énergie, pour l'atelier d'écriture et pour avoir pris le temps de me former sur la relecture d'article et l'éthique des modèles de langue. Merci pour ton engagement syndical et militant, de nous avoir toujours poussé vers l'avant. Merci Bruno Guillaume pour ton investissement, ta joie de vivre contagieuse et tes conseils sur l'annotation syntaxique. Merci pour les soirées Sémagramme. Merci Jacques Jayez pour les conseils précieux relatifs au monde des linguistes. Merci Maxime Amblard pour les cours et IWCS 2023. Merci à Sylvain Pogodalla. Merci à Hee-Soo Choi pour le soutien mutuel. Merci à Marie Cousin, Amandine Decker, Amandine Lecomte, Fanny Ducel, Siyana Pavlova, Maxime Guillaume, Khenza Amani Daoudi, Betrand Remy et Vincent Tourneur pour l'ambiance incroyable parmi les non permanents, les soirées et les

jeux à la pause du midi. Merci aux doctorant·es que j'ai connu et qui ont quitté l'équipe : Maria Boritchev, Chuyuan Li (Lisa), Priyansh Trivedi. Merci à toutes les autres personnes qui sont passées dans l'équipe.

Merci à toutes les autres personnes que j'ai pu rencontrer au Loria et à l'Université de Lorraine, notamment Amaury et Ambroise. Merci aussi à Gabriel pour toutes les soirées que tu as organisées, pour l'ambiance de l'enclave Carbone et les journées de Twilight Imperium.

I am grateful for the lunches with my colleagues at the ILLC, Aditya Surikuchi, Alina Leidinger, Marco Degano, Louis Ravn, and others. I would like to especially thank Tomasz Klochowicz for his support and mentorship. I am also grateful to some MoL students I got to know, such as Stefano Volpe and Gidon Kaminer.

Je me dois aussi de remercier les gens qui m'ont donné par ci par là quelques indications utiles, comme Marie Candito, Djamé Seddah, Sylvain Kahane et Guy Perrier pour les annotations et corpus de syntaxe, Jan Fliessbach, Hélène le Levier et les personnes avec qui j'ai pu discuter lors de diverse conférences.

I had the opportunity to meet and exchange with several great people, including Maria Aloni, Robert van Rooij, Ivano Ciardelli, Alda Mari, Anastasia Giannakidou, Mora Maldonado, Veneeta Dayal, Gennaro Chierchia, Ivano Caponigro, Anamaria Fălăuş, Orin Percus, Hamida Demirdache, and many more. In particular, I am very grateful for the invitation to the LLing seminar in Nantes. I met a fantastic team of people sharing the same passion for semantics and the syntax-semantics interface. I am also very grateful to all the participants in the InqBnB4 workshop.

Merci aux personnes qui m'ont accueilli à Montpellier, notamment Éléonore Yasri-Labrique, Ksenija et Léo Léonard, Nadia Makouar et Marion Dufour.

Parmi les personnes que j'ai apprécié avoir rencontrées, je remercie Baptiste Loreau Unger pour les discussions qui m'ont bien guidé, Marie Boscaro, Pauline Gillet, et une spéciale dédicace à Adèle Hénot-Mortier, qui a ma plus grande amitié et mon profond respect. Je suis aussi chanceux d'avoir rencontré Jeanne Lecavelier, et j'espère qu'on continuera à collaborer.

Enfin, ce doctorat n'aurait pu se passer sans peine sans l'investissement que j'ai pu concrétiser auprès des différentes associations que j'ai rejointes. Un énorme merci à Christophe Benzitoun et aux Linguistes Atterrées, notamment Anne Abeillé, Laélia Véron, Maria Candea mais aussi Anne-Catherine Simon. Il y a encore plein de grands moments à créer ensemble. Merci à Emmanuel Fruchard, Valentin Gatien-Baron, Danièle Cogis, Mario Chabot, Mario, Périard, Mireille Elchacar et tous les autres membres d'Érofa. C'est en continuant à soutenir un tel projet qu'on maintient l'espoir.

Je termine en remerciant mes amis et toute ma famille, mes parents, mes frères et sœurs, mes neveux et nièces, pour le soutien inconditionnel qu'ils m'ont apporté, psychologique, matériel et scientifique, en essayant par exemple de comprendre à leur manière qui est allé dans la cave hier. Merci de vous être intéressé à ces questions techniques et de reconnaitre mon investissement.

Chapter 1

Introduction

1.1 Scope of this thesis

In a dialogue, interlocutors can bring information using assertions or request information using questions. Asking a question proposes to update the common ground with a set of possible resolutions. However, the pragmatic effects of questions extend beyond their resolution conditions. Like assertions, questions can impose background assumptions on the common ground, called presuppositions, and provide material for subsequent turns, e.g. discourse referents. This thesis aims to study these additional discourse effects in wh-questions, i.e. questions containing a wh-word (*who*, *what*, *which*, etc.).

1.1.1 Presuppositions

A presupposition is a background inference taken for granted. For example, assertion (1-a) implies that there was an exam. This information is not at issue but is necessary for the sentence to be felicitous. This inference constitutes an existential presupposition of (1-a), triggered by the determiner *the*. Similarly, wh-words often trigger existential presuppositions, like the existence of a cheating student in (1-b).

(1) a. Mary cheated at the exam.
$\overset{\text{presup}}{\Longrightarrow}$ There was an exam.
b. Which student cheated?
$\overset{\text{presup}}{\Longrightarrow}$ A student cheated.

1.1.2 Anaphora

Some expressions may refer to individuals introduced earlier in the discourse. For example, in (2-a), *she* refers to the woman introduced in the first sentence.

The interpretation of *she* depends on the referential expression *a woman*, called its antecedent. This connection between entity-denoting phrases is named anaphora. Modern anaphora theories assume that interlocutors keep track of the individuals at stake via discourse referents (or *drefs* for short), symbolized by indexes, e.g. $u, v, ...$ [1] Wh-words can also be antecedents of anaphoric relations. In (2-b), *she* refers to the woman sitting in the indicated seat.

(2) a. A^u woman was sitting here. She$_u$ forgot her bag.
b. SITUATION: *John saw a woman leaving her seat and forgetting her bag. He addresses someone sitting next to that seat.*
JOHN: Excuse me, who^u was sitting here? She$_u$ forgot her bag.

1.1.3 Presuppositions and anaphora

Anaphora and presuppositions are linked. Introducing an individual available for reference usually requires the existence of this individual. An existential presupposition typically ensures this existence. Conversely, an existential presupposition is often sufficient to make an individual salient for subsequent reference.

1.1.4 Main questions

In this thesis, I investigate the presuppositions and anaphoric properties of the wh-words in questions. I bring new data suggesting that the existential presupposition and reference pattern of simplex wh-questions exhibit some variation that has not been accounted for before. In modalized questions, similar effects were observed regarding the uniqueness presupposition of *which+sg* and anaphora. The goal of this dissertation is to connect these two phenomena and provide a unified analysis and model. The research questions are the following:

RQ1 What explains the variation observed in simplex wh-questions regarding their existential inference and anaphora to the wh-word?

RQ2 Can modalized questions help us analyze these observations?

RQ3 How can we model these behaviors?

In this thesis, I focus on English questions, but I also study some French constructions.

Section 1.2 presents the variation in simplex wh-questions. Section 1.3 summarizes the analysis I developed. A glimpse of the model created is presented in section 1.4. Finally, section 1.5 presents the outline of this dissertation.

1. Antecedents are identified by a superscript while a subscript identifies anaphoric expressions. The index is put on the determiner when possible.

1.2 Variation in presupposition and reference

Research in formal semantics has traditionally focused on assertions, for which various linguistic tests have been developed to classify inferences. Presuppositions are known to project through negation and conditional antecedents: (3-a) and (3-b) still imply that there was an exam. However, such constructions are impossible with a question.

(3) a. It is false that Mary cheated at the exam.
$\overset{\text{presup}}{\Longrightarrow}$ There was an exam.
b. If Mary cheated at the exam, then she will be punished.
$\overset{\text{presup}}{\Longrightarrow}$ There was an exam.

The restrictive distribution of interrogatives may explain why question presuppositions have been less intensively studied. In Chapter 2, I show that we can turn to various techniques to circumvent this issue. Controlling the discourse context, the semantic environment, and the presence of certain operators or interveners shed light on differences that have been overlooked.

One major debate I address concerns the existential inference of questions with a simplex wh-word (*who* or *what*). Some authors (Katz and Postal 1964; Dayal 1996 among others) argue that a question like "*Who went into the cellar yesterday?*" presupposes that "*Someone went into the cellar yesterday*". Others, like Groenendijk and Stokhof 1984 and Ginzburg 1992 argue that this inference is a conversational implicature.

The limit of previous approaches is their monolithic analysis, i.e. predicting that simplex wh-questions always have (or never have) an existential presupposition. In Chapter 2, I provide evidence against such an analysis. For example, in (4-b), John's question requires that someone went into the cellar to be felicitous, whereas it is not the case in (4-c).

(4) a. SITUATION: *John went into the cellar two days ago. There were no socks on the floor and no weird smell. This morning, he went back into the cellar. Then, he says to the rest of the family:*

b. (i) JOHN: I saw a sock in the cellar this morning.
(ii) JOHN: Who went in the cellar yesterday?
$\overset{\text{presup}}{\Longrightarrow}$ Someone went in the cellar yesterday.

c. (i) JOHN: There was a weird smell in the cellar this morning. It might be a gas leakage. I wonder when it started.
(ii) JOHN: Who went in the cellar yesterday? And was there any weird smell?
$\not\rightsquigarrow$ Someone went in the cellar yesterday.

The second part of this variation is the type and availability of anaphora with wh-words as antecedents. In (2-b), pronominal co-reference is possible. Conversely, after question (5-a), reference to *who*^{u} must include a modality, like in (5-b).

(5) a. A: Who^u knows how to model free-choice effects in questions?
b. B: I don't know, but [such a person]$_u$ would probably work at the ILLC.
c. B: #I don't know, but [such a person]$_u$ works at the ILLC.

These observations are the starting point of the reflection proposed in this thesis. A comprehensive model of this data should be expressive enough to account for these subtleties. However, it should also provide reasons for these behaviors. In this perspective, the theory I propose is explanatory because it takes ideas proposed for (non-interrogative) indefinite NPs in assertions and extends them to wh-words.

1.3 Reading ambiguity and local contexts

Indefinite noun phrases (NPs) can be construed as specific or non-specific (von Heusinger 2002). In (6-a), the NP *a pretty girl* is ambiguous between these two readings. Continuation (6-b) orients towards the specific reading while (6-c) orients towards the non-specific one.

(6) a. John is looking for a pretty girl.
b. ..., namely for Mary. (specific)
c. ..., whoever he will meet, he will take her to the movies. (non-specific)

Similarly, the anaphora tradition distinguishes between specific mentions and non-specific ones, also called generic mentions. [2] Specific mentions are used when the speaker has some particular individual in mind. Conversely, generic mentions refer to general, unspecified, or prototypical individuals.

Chapter 2 extends the specific vs. generic duality to simplex wh-words. For example, John is looking for a specific person in (4-b) while he is not in (4-c).

Generic mentions have different anaphoric patterns than specific ones. In particular, they are more sensitive to modal operators scoping over them. A typical consequence is modal subordination (Roberts 1989). Under a non-specific reading of *a*^{u} *wolf* in (7-a), the discourse referent u it introduces cannot be picked up by *it*$_u$ in indicative sentence (7-b). However, using a modal in subsequent (7-c)

2. The terms *generic mention* (anaphora theory) and *generic statements* or *generics* (e.g. Chierchia 1995, among others) encompass different ideas. In this dissertation, I use both concepts interchangeably. I explain in Chapter 2 why it makes sense to consider "generic" wh-words as generic in both acceptations.

makes the anaphora felicitous.

(7) a. A^u wolf might walk in. (*might* > *a*)
b. #It$_u$ is grey.
c. It$_u$ would eat you first.

Sentence (7-a) does not imply the existence of a wolf in the common ground. The wolf is mentioned relative to a hypothetical situation called a local context. Modalities can introduce and refer to such local contexts. The operator *would* in (7-c) points towards the hypothetical situation described in (7-a) and thus licenses the anaphora.

The same phenomenon appears in modalized questions. A similar analysis is also relevant for generic questions. The data in (5) suggests that generic questions create local contexts in which the wh-drefs–the discourse referents introduced by the wh-word–are defined. Extending the theory of modal subordination to modalized questions in Chapter 4 and to generic questions in Chapter 5 explains these patterns.

In this dissertation, we adopt a broad definition of modal subordination as an anaphora accessibility pattern mediated by a wide range of semantic modalities, including modal verbs (epistemic or deontic), modal adverbs, conditionals, and teleological modalities (future, goal constructions).

1.4 Modeling referents and contexts

This PhD thesis presents a descriptive theory of presuppositions and discourse referents raised in wh-questions. Additionally, it provides a formal model capturing these phenomena.

The model needs to represent questions and discourse referents in a combined way. A question like (8) is resolved by propositions corresponding to its possible answers. In every resolving proposition, the dref u must be valued with the actual cheater. Dynamic Inquisitive Semantics DInq (Roelofsen and Dotlačil 2023b) offers an architecture to implement this idea.

(8) Which^u student cheated?

Let me illustrate the semantics of (8) in DInq. Take two individuals, the students a and b, and four worlds, w_a, w_{ab}, w_b and $w_\emptyset$. The subscript indicates the cheaters in each world.

The representation of question (8) is displayed in Fig. 1.1. Each dot is a possibility, a pair ⟨world, assignment function⟩. Resolving propositions are symbolized by regions containing possibilities. A context is a set of resolving propositions. [3]

3. For simplicity, I do not take the downward-closedness of resolving propositions intro account in this introduction. I only represent the output of the denotation of question (8) on an

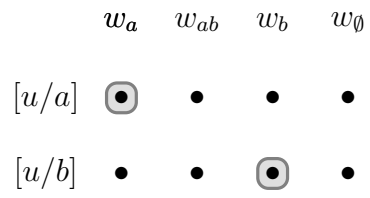


Figure 1.1 – Semantic representation of question (8) in Dynamic Inquisitive Semantics

Question (8) is represented by a context containing two resolving propositions: "*Only a cheated*" $\{\langle w_a, [u/a]\rangle\}$ and "*Only b cheated*" $\{\langle w_b, [u/b]\rangle\}$.

Chapter 4 develops Modal Dynamic Inquisitive Semantics (MDInq). This extension of Dynamic Inquisitive Semantics enables the management of multiple contexts structured in a nonempty stack (S. Kaufmann 2000). At the bottom, the stack contains the global context c_0 corresponding to the common ground. Modalities can push new local contexts on top of c_0. Some discourse referents may be defined only in the upper local contexts. Modal subordination is implemented by restricting the stack. Indicative mood requires the stack to be reduced to c_0, while modal sentences can retrieve the topmost element and refer to the local drefs defined therein.

For example, on an initial stack $\langle c_0\rangle$, sentence (7-a) pushes a new context c_1 corresponding to the proposition "A^u *wolf walks in*", as illustrated in (1.1). The dref u is only defined in c_1. Thus, a modalized sentence like (7-c) is required in order to access u in c_1.

$$\langle c_0\rangle \xrightarrow{\text{A}^u \text{ wolf might walk in.}} \left\langle \begin{array}{c} c_1 \\ c_0' \end{array} \right\rangle \tag{1.1}$$

The stack-based model was developed to treat modalized statements and conditionals uniformly. Adopting this system thus allows us to represent conditional statements and conditional questions. I apply this feature in Chapter 6 to model dependence propositions. In particular, I study French interrogative-based conditionals and model them in Modal Dynamic Inquisitive Semantics.

1.5 Outline of the thesis

The rest of the thesis is organized as follows.

Chapter 2 studies the existential inference of simplex wh-questions. I present the specific vs. generic ambiguity theory of wh-words capturing this inference.

Chapter 3 is a gentle introduction to Dynamic Inquisitive Semantics. This model is the basis of the formal system used in the following chapters.

---

input context where the existential presupposition is satisfied.

In **Chapter 4**, I develop Modal Dynamic Inquisitive Semantics (MDInq). Wide-scoping modal operators in questions weaken the uniqueness presupposition of singular *which*. This phenomenon, as well as modal subordination, orient the construction of MDInq. This chapter is an extended version of the following publication:

> Valentin D. Richard (Dec. 2024b). "Dynamic Effects of Modalized Questions". In: *Proceedings of the 24th Amsterdam Colloquium*. Ed. by Fausto Carcassi et al. Amsterdam, The Netherlands, pp. 289–307. URL: https://hal.science/hal-04866506

The anaphoric properties of simplex wh-words are studied in depth in **Chapter 5**. I propose to consider anaphora to wh-words as structured anaphora. I set out the differences between anaphora to a specific wh-word and to a generic wh-word, and model them in MDInq. This chapter includes a corpus study that led to a subsequent publication:

> Valentin D. Richard (Oct. 2025a). "Évaluation d'un corpus annoté en anaphores : le cas des chaines contenant un mot interrogatif". In: *Lift2-2025 : Journées scientifiques du réseau thématique LIFT2 - linguistique informatique, formelle et de terrain*. Ed. by Timothée Bernard et al. Paris, France, pp. 1–6. URL: https://hal.science/hal-05326668

In **Chapter 6**, I turn to dependence propositions and French interrogative-based conditionals. I use MDInq to model their anaphoric behavior and solve a compositional issue about how alternatives are raised. A shorter version of this chapter was accepted for a presentation at the 16th Syntax and Semantics Conference in Paris (CSSP) and will appear in the following publication:

> Valentin D. Richard (Nov. 2025b). "Raising Alternatives to Express Dependence: A Compositional Issue". In: *16th Syntax and Semantics Conference in Paris (CSSP 2025)*. Ed. by Charlotte Hauser, Céline Pozniak, and Laurent Roussarie. Paris, France, pp. 1–7. URL: https://conf.llf-paris.fr/cssp2025/content/abstracts/CSSP_2025_paper_5.pdf

Finally, **Chapter 7** concludes.

I include a related work in **Appendix A**. This chapter presents a program named FUDIA, which was developed to annotate and extract the interrogative clauses of French corpora annotated in Universal Dependencies (UD). This chapter is an extended version of the following publication:

> Valentin D. Richard (Nov. 2023). "Est-ce que l'extraction des interrogatives du français peut-elle être automatisée ?" In: *5èmes journées du Groupement de Recherche CNRS " Linguistique Informatique, Formelle et de Terrain " (LIFT 2023)*. Ed. by Karën Fort,

Claire Gardent, and Yannick Parmentier. Nancy, France: CNRS, pp. 69–76. URL: https://hal.science/hal-04359947

Appendix C contains the dialogues from the corpus studied in Chapter 5.

Part One

Presuppositions of Questions

Chapter 2

Existential Inference of Wh-Questions

2.1 Introduction

Wh-questions trigger an existential inference (EI) (Belnap 1963), e.g. (1-a) and (1-b).

(1) a. Who cheated?
   ⇝ Someone cheated.
   b. Which student cheated?
   ⇝ Some student cheated.

This inference varies in strength depending on the wh-word. The EI of *which+sg* is consistently analyzed in the literature as a presupposition (Karttunen and Peters 1976; Rullmann and Beck 1998). Conversely, there is debate over the nature of the EI of simplex wh-words like *who* and *what*. Some authors claim that this inference is a presupposition (Katz and Postal 1964; Postal 1971; Horn 1972; Lawler 1971; Keenan and Hull 1973; Karttunen and Peters 1976; Comorovski 1996; Dayal 1996; Gawron 2001; Abusch 2010; Uegaki 2021), [1] while others argue for a conversational implicature (Groenendijk and Stokhof 1984; Ginzburg 1992, 1995b). [2] This chapter addresses this debate by investigating the properties of the existential inference (EI) of simplex wh-words.

The EI of *who*-questions is not systematically triggered. Abusch 2010 remarks that it can be suspended, as illustrated in (2).

(2) I've alienated my colleagues completely. Who will vote for me? Probably nobody.

1. Although they prefer the term *conventional implicature* to *presupposition*, Karttunen and Peters 1976's analysis of the EI seems to correspond to modern definitions of presuppositions.

2. Groenendijk and Stokhof 1984 use the term *existential suggestion*, and assume pragmatic principles to derive it. Therefore, their analysis is close to a conversational implicature.

This instability challenges models that treat the EI as a fixed or invariant inference. Moreover, some questions are compatible with null answers (aka. negative answers). For example, sentence (3) is not contradictory, so the embedded interrogative "*who came*" must contain the null answers "*no one came*" in its answer set. This observation suggests that EI derivations based on the disjunction of positive answers (Keenan and Hull 1973; Abusch 2010) do not paint the entire picture. A more nuanced analysis of EIs is thus required.

(3) Both Bill and Mary think no one came to the party, and so they agree on who came. (Fitzpatrick 2005)

To explain these phenomena, I suggest–like Elliott, Nicolae, and Sauerland 2022–that simplex wh-words are ambiguous between two semantic interpretations. The hypothesis I defend is that simplex wh-words are either specific or generic. The specific reading (who_S) refers to individuals identified by some means. This reading triggers an existential presupposition. In contrast, the generic reading (who_G) refers to hypothetical individuals and is related to epistemic free-choice items. This reading does not trigger any existential inference and is open to null answers. The likelihood of either reading in a given sentence influences the likelihood of triggering the EI.

To support my position, I present new data illustrating how various factors disambiguate or favor one reading over the other, such as semantic environments, the speaker's goal, focus, question type, and NPIs. I demonstrate that the distinction between who_S and who_G correlates with cross-linguistic and language-internal variations in referentiality in non-interrogative wh-items and anaphoric properties of interrogative words.

This chapter is organized as follows. In section 2.2, I present the properties of existential inferences triggered by simplex wh-phrases, supported by new empirical data. Accounting for these observations requires a novel analysis. I suggest a specific vs. generic opposition to explain the variation. In section 2.3, I initiate this theory with independent arguments related to focus, non-interrogative wh-items, and anaphora to wh-words. Using the previous two points as basis, I argue that the source of the specific vs. generic duality is the wh-word; hence there are two readings: who_S and who_G. Section 2.4 provides further details about this theory, the construction of null answers, and how pragmatics determines the appropriate reading. Applications to this theory are shown in 2.5. I also illustrate how the theory provides insightful explanations for the following unrelated issues: NPIs in questions, social pressure to answer, and weak islands. I outline an alternative analysis of these data in section 2.6. In section 2.7, I explore competing theories and offer a critical evaluation. In particular, I examine Elliott, Nicolae, and Sauerland 2022's type-flexible *who* in the light of my data. Finally, section 2.8 concludes the chapter.

2.2 Properties of the existential inference of simplex wh-words

In this section, I review the key properties and present new supporting data regarding the existential inference (EI) of matrix simplex wh-questions. EIs exhibit some variations in their availability and strength. Moreover, contrary to assertion inferences, they are subject to the asymmetry between the speaker (i.e. the asker) and the listener (i.e. the responder). Finally, interrogatives have a limited syntactic distribution, preventing us from using common embedding tests (e.g., projection under negation or conditionals). Because of these issues, EIs have resisted classification into the usual inference types for a long time.

To investigate question inferences, we must search for specific contexts that impact their triggering. These signature phenomena may shed light on their origin and the factors they depend on.

I start with basic observations about the default nature of existential inferences in section 2.2.1. In section 2.2.2, I review data about the cancelability and suspendability of EIs to advocate for a more nuanced analysis compared to previous attempts. I emphasize in section 2.2.3 that the null answer to questions does not only serve a pragmatic purpose but is also needed for their truth conditions. Finally, I present new data showing that EIs are sensitive to some semantic environments (sections 2.2.4 and 2.2.5) and the asker's mental state (section 2.2.6).

2.2.1 Spontaneous triggering

In non-biased, out-of-the-blue *who*-questions, the existential inference (EI) spontaneously arises, as in (1-a). This inference arises intrinsically from the act of questioning and the semantic content of the utterance. No prior evidence or authority of the responder in the asker is sufficient *on its own* to derive the EI. To see why, consider situation (4-a).

(4) a. SITUATION: *Mary supervised and graded an exam. She is certain nobody cheated, even if the average grade is unusually high. She tells John, a colleague who never taught, that the average is unusually high.*
b. JOHN: And who cheated?
⇝ Someone cheated.

In situation (4-a), Mary puts no authority in John about the existence of a cheater. She has no reason to believe that he could know better than her. Despite this, John's question (4-b) still triggers an EI. Moreover, this inference does not come up as John's belief[3] but is imposed on the common ground. Mary must find question (4-b) infelicitous because this EI contradicts her assumptions about

3. The inference is not "*John believes that someone cheated*".

John and her beliefs.

This example suggests that the EI is derived using some element intrinsic to the utterance content and is not only due to the enunciative context or participants' beliefs.

2.2.2 Cancellation and suspension

As remarked by Karttunen and Peters 1976, the EI cannot be canceled by the speaker (5).

(5) A: #I know that Mary doesn't read anything. What (exactly) does she read?

However, any responder can cancel the EI without background disagreement, e.g. (6).

(6) a. A: Who cheated?
b. B: Nobody.

Karttunen and Peters 1976 use speaker non-cancelability to argue against a conversational implicature. However, this non-cancelability is more probably due to a competing Ignorance Principle (7), as Fitzpatrick 2005 explains (see also Onea and Zimmermann 2019, (85)a. p. 51).

(7) **Ignorance Principle**
A speaker can only ask an information-seeking question if he or she does not know the answer(s). (Fitzpatrick 2005)

As I will argue in section 2.2.3, null answers are full-fledge answers to questions. Therefore, by the Ignorance Principle, the speaker cannot simultaneously ask a question and believe in its null answer. This explains the asymmetry between speaker cancellation (5) and responder cancellation (6).

The EI is not as strong as *bona fide* presuppositions. In particular, contrary to cleft questions (8-b), the existential presupposition of *which+sg* (8-c), and presuppositions projected from the scope of the question (8-d), simplex-wh EIs can be suspended (8-a). Here, *suspended* means that the utterance U remains non-contradictory when conjoined with a suspender clause (Horn 1972), i.e. a clause expressing that the inference χ may be false (viz. $U \wedge \Diamond\neg\chi$ is non-contradictory).

(8) a. A: Who cheated? Maybe nobody cheated.
b. A: Who is it that cheated? #Maybe nobody cheated.
c. A: Which student cheated? #Maybe no student cheated.
d. A: Who cheated at the exam? #Maybe there was no exam.

Suspension is also possible before the question. In a court, judge A might hold

it unlikely that anyone has any objection. However, if it is technically possible that someone has an objection, question (9-b) is acceptable as an information-seeking question, where A sincerely waits for potential objections.

(9) a. A: I know that, most probably, nobody has any objection. But I must still ask because of the procedure.
b. A: Who has any objection?

The pattern of cancelability and suspendability of EIs prevents us from classifying them as regular implicatures or presuppositions. It appears that this phenomenon is essentially neither of them. I argue, however, that a combination of pragmatic principles, a reading ambiguity, and a lexical presupposition trigger can capture these properties.

2.2.3 Null answers are full-fledged answers

The pragmatic status of inter-speaker EI cancellation is debated. Mittwoch 1979, Comorovski 1996, and Hagstrom 1998 assert that responses like (10-b) deny the existential presupposition triggered by the question. Haida 2007 views it as rejecting the question as an unsuitable speech act. [4] Unlike them, Groenendijk and Stokhof 1984, Ginzburg 1992, Han and Siegel 1996, and Fitzpatrick 2005 claim that they truthfully resolve the wh-question.

(10) a. A: Who cheated?
b. B: Nobody cheated.

In some cases, like (4-b) above (see also (23-c) below), Mittwoch, Comorovski, and Hagstrom seem to be right. [5] However, in some contexts (see section 2.2.6 below), it appears that null answers can genuinely bring some information sought by the asker. Negative answers truthfully resolve polar questions; why not also wh-questions? Moreover, null answers can be made as salient as necessary, e.g. by biasing questions using an emphatic weak NPI (see (19)). The Ignorance Principle (7) also requires that the null answer is an answer to account for the speaker's non-cancelability of the EI.

4. Haida 2008 claims that, if simplex wh-questions have no existential presupposition, then they should accept existential responses like (i-b), going against the observation that (i-b) is not a felicitous answer to (i-a). Onea 2016 retorts that, if proposition (i-b) were entertained by A as a possible response, then A should have asked "*Did anyone cheat?*" instead. As A uttered (i-a) and not "*Did anyone cheat?*", response (i-b) is ruled out by this reasoning.

(i) a. A: Who cheated?
b. B: #Someone cheated.

5. In section 2.4.3, I explain the case of (4-b) by showing how, outside of any relevant context, the specific reading of *who* is derived by default, thus triggering an existential presupposition.

The preceding arguments indicate that null answers must be incorporated into the use conditions of questions. Fitzpatrick 2005's example (3) (repeated in (11)) and Groenendijk and Stokhof 1984's inference (12) prove that null answers are even required for truth conditions of embedded interrogatives. [6]

(11) Both Bill and Mary think no one came to the party, and so they agree on who came. (Fitzpatrick 2005)

(12) John knows who left.
No one left.

John knows that no one left. (Groenendijk and Stokhof 1984)

Individuals d agree on Q if and only if there is a strongly exhaustive answer q to Q such that d all believe q (Spector and Egré 2015). [7] Take $d = \text{Bill} \oplus \text{Mary}$ and $Q =$ "*who came (to the party)*". In (11), d's belief state is the proposition p containing the worlds where nobody came to the party. The *and so* coordination in (11) indicates that $\mathbf{believe}_d(p)$ entails $\mathbf{agree}_d(Q)$. Consequently, there exists an answer q to Q such that p entails q. Clearly, p does not entail any answer where at least one person came. Therefore, p ("*nobody came*") is an answer to Q ("*who came*"). [8]

(13) a. They agree on who came.
b. $\rightsquigarrow$ Someone came.
AND/OR $\rightsquigarrow$ They believe that someone came.

In isolation, sentence (13-a) spontaneously triggers the inference (13-b). However, it is erroneous to analyze this inference as a presupposition and the first clause in (11) as canceling this presupposition. The clause "*Both Bill and Mary think no one came to the party*" cannot cancel presuppositions. To see why, compare (11) with (14), where existential presupposition "*There was an exam*" projects through *agree* into Bill and Mary's epistemic state. This projection contradicts

6. Unfortunately, Fitzpatrick 2005's example doesn't seem to extend to other predicates. For example, Dayal 2016 points out that (11) is at odds with the infelicity of (i). It is unclear whether the felicity of (11) is due to a special property of *agree*. Despite this, I assume here that sufficient evidence has been shown in favor of considering null answers as necessary for the resolution conditions of questions.

(i) #No one left so John knows who left.

7. Here and in the rest of the text, I assume that a question interpreted by as a set of resolving propositions, like in Inquisitive Semantics (Ciardelli, Groenendijk, and Roelofsen 2018). Nevertheless, my reasoning does not hinge on this formalism. It would also work within alternative semantics (Hamblin 1973; Karttunen 1977).

8. Technically, I only proved that some answer q to Q is entailed by p. But as p is minimally informative with respect to the question, it is reasonable to assume that this answer q is equal to p and even that p is an alternative of Q.

the first embedded clause "*there was no exam*", making (14) infelicitous. If the EI of "*who came*" was a presupposition, it should project and make (11) infelicitous. But this is not the case. [9]

(14) #Both Bill and Mary think that there was no exam, and so they agree on who cheated at the exam.

I argue in section 2.4.3 that the embedded interrogative of (11) triggers no EI in this particular case because of the contextual information.

2.2.4 Sensitivity to semantic environments

Some constructions inside the question weaken the EI's strength and leave more options to answer using negation. The data I provide suggests that environments licensing Free Choice Items (FCIs) tend to license null answers and inhibit the EI.

Free Choice Items are known to be ungrammatical in episodic sentences, i.e. sentences relating exactly one specific event that happened at a specific time (15-a). On the contrary, FCIs are licensed by habitual aspect, future tense, and generics. [10]

(15) a. *Mary read any book yesterday. (episodic)
b. When she was young, Mary used to read any comic book. (habitual)
c. Tomorrow, we'll take advantage of any situation. (future)
d. Any student knows that. (generic)

Existential inferences appear weaker in questions occurring within environments licensing FCIs. Compare the following pairs.

(16) a. What do you usually read when you relax? (habitual)
b. What did you read when you were relaxing yesterday?

(17) a. What will you read during that meditation retreat? (future)
b. What did you read during that meditation retreat?

9. The same opposition appears with singular *which*:

(i) #Both Bill and Mary think no student came at the party, and so they agree on which student came.

10. Other constructions licensing FCIs include conditional antecedents, comparatives, imperatives, permissive attitudes, and restrictors of universals. However, interrogative words are not possible in their scope. Regarding possibility epistemic modals, the judgments are not conclusive. Although Krifka 2011 claims that question (i) has no existential import, it might still have a very weak EI: "*It might be that somebody is able to solve this problem*".

(i) Who can solve this problem?

(18) a. What do Dutch people eat for dessert? (generic)
b. What did Guido eat for dessert?

Questions (16-a), (17-a) and (18-a) trigger a weaker or no EI and are more open to a null answer compared to episodic questions (16-b), (17-b) and (18-b) respectively.

Moreover, weak Negative Polarity Items (weak NPIs) favor a reading that does not trigger any EI (19-a). Emphatic weak NPIs even bias questions towards their null answer (Jeong and Roelofsen 2023). For example, (19-b) and (19-c) each imply that it is more probable, according to the asker, that nobody did anything to help when Mary was ill.

(19) a. What did Robin ever do to help Mary?
b. Who did ANYTHING to help when Mary was ill?
c. Who lifted a finger to help when Mary was ill?

These examples reveal that various semantic constructions impede the EI. A theory of question EIs needs to justify why *these environments specifically* produce this interference.

2.2.5 Weak islands

Embedded interrogatives are weak islands for wh-word extraction. Comorovski 1996 and Cinque 1996 observe that this weak island constraint can be violated by *which+sg* phrases (20-b), but not by simplex wh-word (20-a).

(20) a. ?What are you wondering how to fix?
b. Which car are you wondering how to fix?

Kroch 1998 retorts that this difference is not imputable to D-linking (Pesetsky 1987) and brings a counterexample (21).

(21) Who were you wondering whether to visit on your vacation?

In Kroch 1998's argumentation, escaping a weak island is licensed by an existential presupposition provided by the wh-word and the uniqueness of the referent introduced by this presupposition.

The semantic environments in section 2.2.4 affect weak island escapability. Sentence (21), which pertains to a specific state of affairs that happened in the current world, is acceptable. Conversely, escaping is infelicitous with a habitual aspect (22-a), future mood (22-b) or a generic term (22-c). However, these sentences can be accepted if *what* refers to a specific individual, which is rigid over regular events / possible futures / Dutch people, respectively.

(22) a. ?What do you usually wonder whether to visit? (habitual)

b. ?What will you wonder whether to visit? (future)
c. ?What do Dutch people wonder whether to eat? (generic)

The weak island constraint confirms that there is an inherent question reading ambiguity related to some semantic environments.

2.2.6 Sensitivity to the asker's goal

Unsurprisingly, the EI's availability is sensitive to contextual information. In assertions, many inferences (e.g. scalar implicatures) are sensitive to the Question Under Discussion (QUD). An analogous tendency is likely in our case. I follow Ginzburg 1995a and Asher and Lascarides 1998's idea that the asker's goal is a key factor influencing how a question is construed.

Consider the situation in (23-a). Depending on John's expectations or assumptions, the question "*Who went into the cellar yesterday?*" does not have the same interpretation. In SITUATION A (23-b), John has somebody in mind, the person who left the sock. Here, *who* refers to someone specific. John wants to know the identity of that person. Conversely, in SITUATION B (23-c), John has no idea whether anyone went into the cellar yesterday. However, he can ask whoever did so whether there already was a weird smell. Question (23-c) is thus generic (i.e. non-specific).

(23) a. CONTEXT: *John went into the cellar two days ago. There were no socks on the floor and no weird smell. This morning, he went back into the cellar. Then, he asks to the rest of the family:*
b. SITUATION A: *A sock was on the floor in the cellar this morning.*
JOHN: Who went into the cellar yesterday? (specific)
c. SITUATION B: *This morning, the cellar smelled weird. John suspects that it might be a gas leakage. He wants to know when it started.*
JOHN: Who went into the cellar yesterday? (generic: MA or MS)

The following test helps distinguish between the two readings. Sometimes, a question "*Who/What P?*" can be rephrased as "*Someone/Something P, and I would like to know the identity of that (plural) individual*". If such a rephrasing is possible and preserves the communicative effects, then the question is specific. If not, the question is generic.

In SITUATION A, "*Who went into the cellar yesterday?*" can be rephrased as (24). This is impossible in SITUATION B because John does not know whether anybody went into the cellar yesterday.

(24) Someone went into the cellar yesterday, and I would like to know that person's identity.

The pragmatic status of the null answer (25) with respect to the question

is another divergence feature. In SITUATION A, John would take (25) as Mary lying or trying to hide the truth that someone went into the cellar. To accept this response, John would need to revise his question. Conversely, (25) is helpful to John in SITUATION B. He would then know that he must search for other information sources to know when the gas leakage started. In other words, a negative response is a non-cooperative move after a specific question, while it is a cooperative move after a generic question.

(25) MARY: Nobody went into the cellar yesterday.

Another context favoring null answers is small talk. A naïve question like (26-a)–often called a *generic question* or *general question*–has a wide domain to avoid restricting the responder's answer too much. Rephrasing (26-b) feels too strong or restrictive. [11]

(26) a. A: What did you do during your weekend?
b. A: #You did things during your weekend. I would like to know those things.
c. B: Nothing. I stayed locked up because I was sick.

Modeling how goals select one interpretation over the other is another challenge for a theory of question EIs.

2.2.7 Interim summary

Sections 2.2.1 through 2.2.3 demonstrate that EIs are neither presuppositions nor implicatures in a straightforward sense; rather, they represent a distinct and complex type of inference that demands a more nuanced analytical framework than those previously proposed. Furthermore, the need to account for null answers as part of the question's resolution conditions undermines many theories. A more detailed critique of these other theories can be found in section 2.7.

Sections 2.2.4, 2.2.5, and 2.2.6 exhibit the sensitivity of EIs to semantic and pragmatic contexts. These phenomena suggest two readings of questions which can be teased apart using targeted diagnostic tests. In particular, section 2.2.6 introduces the idea that this ambiguity can be characterized in terms of specificity. This idea will be further developed and defended in the following sections.

Simplex wh-questions exhibit ambiguity between a specific and a generic reading. A specific question has an existential presupposition, while a generic question has no existential inference and includes null answers in its resolution conditions.

The existential inference of *which+sg* questions has been attributed to the semantics of the wh-word (Rullmann and Beck 1998; Uegaki 2021). Similarly, I argue that the specific/generic ambiguity in simplex wh-questions arises from a

11. I explore further the discrepancy in answer expectation in section 2.5.2.

semantic ambiguity [12] in wh-words: specifically, between a specific interpretation $who_S/what_S$ and a generic interpretation $who_G/what_G$.

I provide independent arguments supporting my theory in section 2.3. In section 2.4, I flesh out the characteristics of both readings and describe how they explain the data presented above. Section 2.5 provides evidence for the specific vs. generic duality by showing that it solves unrelated unaccounted phenomena.

2.3 Motivating a specific vs. generic duality

This section presents independent evidence supporting a duality in the interpretation of wh-questions with respect to existential inferences (EIs). Section 2.3.1 demonstrates that certain languages can use grammatical devices (e.g. focus marking) to create unambiguously specific or generic questions. Sections 2.3.2 and 2.3.3 further motivate the claim that the source of the specificity/genericity distinction lies in the semantics of the wh-word itself. Specifically, section 2.3.2 explores the closely related referential-attributive opposition in non-interrogative wh-items, while section 2.3.3 summarizes key insights from Chapter 5 regarding the divergent anaphoric properties of interrogative expressions.

2.3.1 Focus marking in Mongolian, Ngamo and French

Some languages can distinguish between two kinds of questions: those in which the asker has a specific individual in mind and those where this is not the case. Mongolian, Ngamo (a West Chadic language) and French–though unrelated–each exhibit mechanisms for expressing this distinction through information focus. This section demonstrates that non-ambiguous presuppositional readings of simplex wh-questions consistently correlate with a specific interpretation and are marked by focus on the wh-word.

Mongolian focus movement

In Khalkha-Mongolian (the main Mongolian dialect), the canonical word order is SOV. Furthermore, it allows *in situ* questions, like (27-a). A focused object may move before the subject. Wh-words can also move, creating *ex situ* questions, e.g. (28-a).

(27) a. Tuya hen-tei gerle-j bai-san be?
Tuya who-COM marry-CVB be-PST Q
'*Whom did Tuya marry?*' (Mongolian, O&G)

12. In section 2.6, I consider an alternative hypothesis where the ambiguity is a pragmatic ambiguity.

b. Tuya hen-tei ch gerle-j bai-gaa-güi.
Tuya who-COM FOC marry-CVB be-NPST-NEG
'*Tuya married nobody.*'

(28) a. Hen-tei Tuya gerle-j bai-san be?
who-COM Tuya marry-CVB be-PST Q
'*Whom did Tuya marry?*' (Mongolian, O&G)

b. #Hen-tei ch Tuya gerle-j bai-gaa-güi.
who-COM FOC Tuya marry-CVB be-NPST-NEG
'*Tuya married nobody.*' (The speaker doesn't want to tell who Tuya married)

Onea and Guntsetseg 2011 (O&G) observe that *in situ* question (27-a) can be answered with the negative response (27-b). By contrast, *ex situ* question (28-a) is "strongly presuppositional": "native speakers get the impression that the person uttering [(28-b)] is lying, i.e. he does not want to divulge the secret, who Tuya married?", (O&G).

These two question forms are not felicitous in the same types of contexts. The *ex situ* question in (29-c) is infelicitous when asked out of the blue if the asker has no particular individual in mind. However, it becomes acceptable when referring to a known event or a set of events already under discussion, such as in contexts (29-a) or (29-b). Conversely, *in situ* question (30-b) is not appropriate in context (30-a) but it is acceptable as a general question with no salient event. [13]

(29) a. CONTEXT: I just saw that Peter kissed a girl, but I didn't recognize her.

b. CONTEXT: I'm sure, Peter kissed some girl or another in his life.

c. Hen-ig Peter üns-sen be?
who-ACC Peter kiss-PST Q
'*Whom did Peter kiss?*' (Mongolian, O&G)

(30) a. I just saw that a woman kissed someone, but I didn't recognize the kissee.

b. #Ter hen-ig üns-sen be?
she who-ACC kiss-PST Q
'*Whom did she kiss in the event you saw?*' (Mongolian, O&G)

Ngamo background particle

Ngamo is a West Chadic (Afro-Asiatic) language spoken in Nigeria. [14] Its canonical word order is SVO. However, when the subject is focused or questioned, it undergoes inversion to the right edge of vP. Focused subjects require the clitic

13. Question (30-b) is redeemed in context (30-a) with a demonstrative NP subject because it constitutes a salient sentence topic (Onea and Guntsetseg 2011).

14. Data presented here is from the Gudi dialect.

=i on the previous word (Grubic and Zimmermann 2011). This particle marks elements on its left as backgrounded, thus indicating the phrase on its right is focused (Grubic 2015).

Subject questions in Ngamo can appear either unmarked, as in (31-a), or marked with *=i*, as in (31-b). Grubic 2015 reports that (31-b) has an existential presupposition contrary to (31-a). In (31-b), "the speaker is specific about the person that Dimza called, and the speaker knows that he has called somebody", (Grubic).

(31) a. Dimza esha lo?
Dimza call.PFV who
'*Who did Dimza call?*' (Ngamo, Grubic)

b. Dimza esha=i lo?
Dimza call.PFV=BM who
'*Who did Dimza call?*' (Ngamo, Grubic)

Clefted wh-words in French

In many languages, clefts (or cleft-like constructions) serve to mark a narrow, non-contrastive focus on a particular element–the clefted phrase–while assigning background status to the remainder of the sentence. In French, the construction *c'est X que* ('*it is X that*') exemplifies this type of specificational construction (Lambrecht 1988). [15]

Like example (23), the question in (32-a) is ambiguous between a specific and a generic reading. By contrast, the clefted question (32-b) presupposes that someone went into the cellar yesterday. For this utterance to be felicitous, the speaker must have some reason to assume that the event occurred–for example, she may have a witness in mind, even if she is unable to identify that individual precisely.

(32) a. Qui est allé dans la cave hier ?
who is gone into the cellar yesterday ?
'*Who went into the cellar yesterday?*'

b. C' est qui qui est allé dans la cave hier ?
it is who.INT that.SBJ is gone into the cellar yesterday ?
same translation, lit. '*Who was it that went into the cellar yesterday?*'

c. Il y a qui qui est allé dans la cave hier
EXPL there has who.INT that.SBJ is gone into the cellar yesterday
?
?

15. Note that questions with *c'est* and *il y a* clefts are non-standard in French and more common in spoken French. An alternative transcription of (32-c) taking the actual pronunciation into account would be "*Y'a qui qu'est allé dans la cave hier ?*".

same translation, lit. '*Who was there that went into the cellar yesterday*'

By contrast, *il y a X que* ('*there is X that*') clefts do not place the clefted element in focus; instead, they assign focus to the sentence as a whole–a pattern known as sentential focus (Lambrecht 2001).[16] Destruel 2017 notes that *il y a* assertions typically respond to broad or mention-some (MS) questions (e.g. "*What's happening?*"), unlike *c'est* clefts, which tend to answer more specific queries.

Question (32-c) is infelicitous if the asker is requesting the identity of someone particular. It is a mention-some question requesting examples of people that went into the cellar yesterday, e.g. in SITUATION B of (23). It is compatible with a context where no one did so. Although less common, *il y a* clefts seem possible with a mention-all reading, e.g. the last question of dialogue (33).

(33) a. MARY: Je vends tous mes livres pour préparer le CAPES. Je n'en ai plus besoin.
'*I am selling all my CAPES preparation books. I don't need them anymore.*'
b. ALICE: Ça m'intéresse. Mais j'en ai déjà quelques uns.
'*I'm interested. But I already have a few.*'
c. ALICE: Y' a quoi exactement que tu vends ?
there has what exactly that you sell ?
'*What exactly are you selling?*'

According to Baunaz 2016, *il y a NP* is incompatible with definite NPs (except if they express a member of a list) and specific quantifiers. This claim supports the observation that *il y a*-clefted questions block specific readings.

I assume that *il y a* clefts are infelicitous with *quel(le)* '*which.SG*'.[17] For

16. Even if narrow (aka. argument) focus is possible with *il y a* clefts (Destruel 2017), I argue that the sentences presented here do not exhibit narrow focus in their most intuitive interpretations.

17. Few occurrences of *il y a quel(le)* $\bar{N}$ *que/qui* were found on the large web corpus `frTenTen23`. Most of them use subjunctive mood (i-a), conditional tense (i-b) or are rhetorical questions (i-c). I analyze the first two cases as a modality quantifying over the wh-word, weakening the inferences triggered by *quel(le)*, making it compatible with the *il y a* cleft. I analyze *il y a quel journal que* with an individual-level predicates (i-d) similarly, using the Gen operator (Chierchia 1995). More generally, we could assume that a D-linked wh-word leads to a generic reading if a modality scopes over it.

(i) a. il y a quelle probabilité que ce tir entre ?
EXPL there has which.SG probability that this shot enters.SJV ?
'*What's the probability of this shot going in?*' (`frTenTen23`)
b. il y a quel sentier ou portion que tu coterais T5 ou T4 ?
EXPL there has which.SG trail or section that you rate.COND T5 or T4 ?

example, sentence (34-a) is awkward due to the stage-level predicate *you read yesterday* and improves only when followed by a phrase like *par exemple* ('*for example*') eliciting partial answers. Moreover, copular *il y a*-clefted questions are incompatible with definite predicates, as illustrated in (34-b), because the mention-some (MS) reading conflicts with the exhaustivity inference associated with definiteness. [18] [19]

(34) a. ??Il y a quel journal que tu as lu hier ?
EXPL there has which.SG newspaper that you have read yesterday ?
Intended: MS '*Which newspaper did you read yesterday?*'

b. *Il y a qui qui est le tricheur ?
EXPL there has who.INT that.SBJ is the cheater ?
Intended: MS '*Who is the cheater?*'

Analysis

Onea and Guntsetseg 2011 account for the strong existential presupposition of Mongolian *ex situ* questions by positing that they are referential with respect to a salient event. However, because they assume that all wh-questions carry an existential presupposition, their analysis cannot explain why *in situ* questions accept null answers. Similarly, Grubic and Zimmermann 2011 analyze $=i$-marked vPs as being referential to a salient event.

I argue that the data instead demonstrates that Mongolian, Ngamo and French

'*Which trail or section would you rate T5 or T4?*' (frTenTen23)

c. y a quelle équipe dans le monde qui peut se targuer d' une
there has which.SG team in the world that.SBJ can REFL boast of a
telle affluence en rugby ??
such crowd in rugby ??
'*Which team in the world can boast such a high attendance in rugby?*'
(frTenTen23)

d. Y' a quel journal qui a une rubrique "économie" ?
there has which newspaper that.SBJ has a section economy ?
MS '*Which newspaper has an economy section?*'

For more discussion on *which+sg* and modality, see fn. 34 and Chapter 4. For rhetorical questions, see section 2.5.2.

18. *Il y a*-clefted questions–and more generally generic questions–can have an EI if the context provides it. For example, as sport competition selections necessarily contain at least one player, (i) is read with an existential presupposition.

(i) Y' a qui que l' entraineur a sélectionnné ?
there is who that the trainer has selected ?
'*Who did the trainer select?*'

19. Additionally, *il y a* clefts may induce a more-than-one inference on the witness set, conflicting with the uniqueness presupposition of the definite NP.

are capable of syntactically distinguishing between specific and generic questions through focus marking on the wh-word. My hypothesis is that the specific variant (who_S) can be selected by placing the wh-word in narrow focus, which yields a clear existential presupposition. In contrast, the generic variant (who_G) cannot be focused. Focus marking is only available when the asker has a particular individual in mind; conversely, non-focused wh-words tend to yield non-specific interpretations and thus allow for null answers.

In French, the data further suggests that *il y a* clefts are only compatible with non-specific contexts. Accordingly, *c'est* clefts select who_S while *il y a* clefts select who_G. The existence of two different readings for French simplex wh-words is confirmed by prosody. Baunaz 2016 remarks that non-presuppositional *qui* ('*who*') has no special accent while specific *qui* is always prosodically prominent, both *in situ* and fronted. [20]

This analysis remains compatible with the event-referential view proposed by Onea and Guntsetseg 2011 and Grubic and Zimmermann 2011, insofar as the existence of a salient event entails the existence of at least one participant–such as agent, theme or experiencer–associated with that event. [21]

---

20. Baunaz 2016 also argue for a third reading, called the partitive reading. In a partitive question, the responder must select one of the alternatives that are presupposed by the asker, as in (i). Partitive *in situ qui* ('*who*') is produced with a different pitch contour. According to Baunaz 2016, partitive quantifiers are compatible with *il y a* but not with null answers. Partitive questions can accept weak islands, but do not assume that the asker knows the true witness in some way. This partitive reading could constitute a third case between the specific and generic readings.

(i) a. SITUATION: *Eva organizes a Christmas party in the department. She suggests three menus: vegetarian, turkey, and salmon. She asks one of her colleagues.* (Baunaz 2016)
b. EVA: et toi, tu préfères quoi dans la liste ?
and you you prefer what in the list ?
'*And you, what do you prefer in the list?*'
c. COLL.: Turkey/Salmon/Vegetarian/#Nothing.

21. An alternative analysis more aligned with Onea and Guntsetseg 2011 and Grubic and Zimmermann 2011 would be to model specific questions with a wide-scope specific event. Focus on the wh-word implies background on the question scope P, thus indicating that P's event is a background assumption. For example, (i-a) could be interpreted with a formula like (i-b). On the contrary, a non-specific reading could receive formula (i-c), where the wh-word scopes over the event closure operator. This would also explain the difference between the global uniqueness presupposition of who_S ("definite" reading) and its absence with who_G.

(i) a. Who went into the cellar?
b. $\exists e.\ \mathbf{go_into_cellar}(e) \wedge \textit{who}\ x.\ \textsc{agent}(e, x)$ (specific event)
c. $\textit{who}\ x.\ \exists e.\ \mathbf{go_into_cellar}(e) \wedge \textsc{agent}(e, x)$ (non-specific event)

2.3.2 Duality in non-interrogative wh-items

Referential and attributive uses of wh-items in free relatives

Donnellan 1966 identifies two uses of plural definite descriptions. The referential use presupposes the existence of a referent (35-a), whereas the attributive use does not, as shown in (35-b). Sentences with attributive plural definites are compatible with a context in which no individual actually satisfies the predicate.

(35) a. A golden coin was given to the sailors who captured a pirate.
⇝ Some sailors captured a pirate.
b. A golden coin will be given to the sailors who capture a pirate.
↛ Some sailors will capture a pirate.

Dayal 1997 and Giannakidou and Cheng 2006 observe a similar referential-attributive distinction in free-relatives introduced by a wh-word–or more broadly, by wh-items (i.e. a word morphologically related to an interrogative word). In the referential reading, as in (36-a), the free relative denotes a salient individual satisfying the predicate associated with *what*. In contrast, in the attributive reading, the truth of the assertion does not depend on a particular individual. In English, the distinction is partially reflected in the contrast between *what* vs. *whatever*.[22]: free relatives headed by *whatever* typically receive a free-choice interpretation, as in (36-b).

(36) a. I ordered what John ordered for dessert.
⇝ John ordered something for dessert.
b. John will read whatever Bill assigns.
↛ Bill will assign something.

This interpretative ambiguity also appears in French free relatives introduced by bare wh-words such as *qui* ('*who*') and *où* ('*where*') (Godar 2021, §XIII-5.3 p.1529).[23] For example, sentence (37-a) presupposes "*You told me to see someone*" and entails "*I saw someone*". Conversely, (37-b) is compatible with

22. Actually, there are two uses of *wh-ever* in free relatives (Dayal 1995). The *identity* reading is about a unique individual whose identity is unknown or irrelevant (i-b). In particular, it presupposes the existence of such an individual. Conversely, the free-choice reading is quantifying over broader situations and is suspendable with respect to the EI (i-a). Both readings are labeled as attributive by Dayal 1997.

(i) a. John will read whatever Bill assigns.
↛ Bill will assign something to read.
b. Everyone who went to whatever movie the Avon is now showing said it was boring.
⇝ The Avon is now showing a movie.

This opposition shows that the specific vs. generic duality differs a bit from the referential vs. attributive duality. Identity readings are attributive but specific.

23. The *identity* reading is not available in French free relatives.

contexts where no such individual exists, such as "*You want to go with nobody*" and "*You (will) go with nobody*".

(37) a. J' ai vu [qui tu m' as dit].
I have seen [who you to_me have said].
'*I saw the person you told me (to see).*' (definite use, French)

b. Pars avec [qui tu veux] !
Go with [who you want] !
'*Go with whoever you want (to go)!*' (free choice use, French)

As with wh-questions in Mongolian and Ngamo, wh-free relatives in several languages exhibit a semantic duality. One reading presupposes the global existence of an individual, while the other one–typically associated with a free-choice interpretation–is compatible with contexts in which no such individual is assumed to exist.

Existential inference variation in indefinite wh-items

In some languages, some interrogative words can also function as indefinites. These are referred to as quexistentials (a portmanteau of *question* and *existentials*) by Hengeveld, Iatridou, and Roelofsen 2023. Cross-linguistically, the existential component of quexistentials behave either like a regular existential or like a negative polarity item (NPI), depending on the language and context.

Dutch *wat* (38-a) and German *was* can be construed either as '*what*' or as '*something*', depending on their syntactic position and whether they are focused. When they are interpreted as '*something*', they carry an existential inference. [24]

(38) a. Miranda heeft wat gegeten.
Miranda has what eaten.
'*Miranda has eaten something.*' (Dutch, Hengeveld, Iatridou, and Roelofsen)

Some quexistentials are restricted to environments that are compatible with the absence of an individual satisfying their descriptive content. Examples include *opjosdhipote* (39-a) in Greek, derived from the wh-word *pjos* '*who*' (Giannakidou and Cheng 2006), and *nǎ*-CL (39-b) in Chinese, also meaning '*which*' (Lin 1998).

(39) a. An kimithis me opjondhipote tha se skotoso.
if sleep.2sg with FCI-person FUT you kill.1sg
'*If you sleep with anybody, I'll kill you.*' (Greek, Giannakidou and Cheng)

24. Under a modal, like in (i), the Dutch pronoun *wat* can also have a non-specific reading, resembling the *identity* reading of English free-relative *wh-ever*.

(i) Johnny wil wat eten.
Johnny wants what eat.
'*Johnny wants to eat something, it could be anything.*'

b. Tā bù xiǎng mǎi nǎ-běn shū.
he not want buy which-CL book
'*He doesn't want to buy any book (in particular).*'

(Mandarin, Giannakidou and Cheng)

Like universal or definite wh-items, indefinite wh-items appear to come in (at least) two varieties: some of them induce an existential inference, while others do not.

Analysis

Von Heusinger 2002 argues that specificity is a referential property of noun phrases (NPs). Specific NPs are anchored to another referent in the discourse, typically the speaker. The denotation of a specific expression is determined by this anchor, rather than by its semantic environment. Both definite and indefinite NPs may be either specific or non-specific (i.e., generic) in their interpretation.

A parallel distinction can be observed in the behavior of wh-items. The referential use of certain wh-items is typically specific, while their free-choice and NPI uses are typically non-specific. The existence of a specific vs. non-specific contrast in non-interrogative wh-expressions provides strong motivation for positing a similar distinction in interrogative expressions.

2.3.3 Duality of anaphoric properties

The study of the anaphoric properties[25] of wh-words is the primary topic of Chapter 5. In the following section, I provide a summary that supports my hypothesis.

Modal subordination

Wh-words always introduce a discourse referent, making them inherently referential (H. Li 2020). However, the accessibility of that referent depends on several factors.

In questions where the asker is uncertain whether a true answer exists, modal subordination become necessary to license subsequent reference to the wh-word. For example, sentence (40-b) or (40-c) can felicitously respond to (40-a), whereas (40-d) cannot.

(40) a. A: Who^{u} knows how to model free choice effects in questions?
b. B: I don't know, but if [such a person]$_{u}$ exists, they$_{u}$ would work at the ILLC.

25. By *anaphoric properties*, I mean the faculty of NPs to raise discourse referents or be anaphoric. An NP is called referential (in the sense of anaphora theory) if it can introduce a discourse referent.

c. B: I don't know, but [such a person]$_u$ would probably work at the ILLC.
d. B: #I don't know, but [such a person]$_u$ works at the ILLC.

Co-reference to the wh-word is also possible in follow-up questions (41).

(41) A: Who^u went to the party? And what did he$_u$/they$_u$ bring as a present? (van Rooij 1998)

Indicative co-reference and deictic use

Despite this behavior, co-reference with wh-words is possible in subsequent indicative assertions. In example (42-a), we infer that A knows someone was sitting here, and she is asking for that person's identity. In (42-b), *what*$_u$ refers to a salient conversation topic that just happened. This use can thus be construed as a deictic use. [26]

(42) a. A: Who^u was sitting here? She$_u$ forgot her$_u$ bag.
b. A: What$_u$ were you talking about? It$_u$ made Mark laugh so much.

Eckardt 2007 reports that, in some sign languages, certain wh-words are morphologically related to deictic expressions. For example, in Langue des Signes Française (LSF, France), the sign for '*when*' involves a pointing gesture that touches the palm of the non-dominant hand, with the palm facing the dominant side. [27] Similarly, Ruys 2023 proposes to consider wh-words as demonstratives.

Analysis

I take the two anaphoric behaviors of wh-referents as evidence for their dual semantic nature. Modal subordination is characteristic of modal environments. However, question (40-a) does not overtly contain a modal operator. The specific/generic opposition offers an potential solution to this issue. Question (40-a) is a generic question. We could assume that generic questions are inherently modal. [28] This would explain why (40-a) requires modal subordination to refer to its wh-word.

By contrast, when the asker has a specific individual in mind–such as a salient discourse referent–the referent is explicitly established, like in (42-a). This allows

26. Pushing the analogy with definite/demonstrative pronouns vs. indefinites further, we could hypothesize that *who*$_S$ can only be anaphoric to familiar referents while *who*$_G$ can only raise new drefs. I leave this hypothesis for subsequent research.

27. This non-dominant hand position is used in other signs related to time, e.g. complementizer '*when*' or '*timetable*'.

28. In section 2.4.1, I argue that mention-some question are necessarily generic. Fox 2018 claims that mention-some questions are all modal in some way, e.g. thanks to an individual-level predicate like *know how to X*. This hypothesis corroborates mine.

for straightforward anaphoric reference in indicative sentences.

In the examples above, the referent u takes the wh-word as its antecedent. This pattern reinforces the claim that the specific/generic ambiguity in wh-questions originates from a semantic ambiguity of the wh-word itself.

2.4 The specific vs. generic ambiguity theory

This section outlines my theory of the specific vs. generic ambiguity of wh-words. Section 2.4.1 introduces characterizations of who_S and who_G. In section 2.4.2, I develop a theory of null answers based on the concept of the null individual. Finally, section 2.4.3 explains how the interpretative duality accounts for the observed weakness and instability of existential inferences (EIs).

2.4.1 Epistemic pronouns and identification

Epistemic pronouns

Some determiners, adjectives, and pronouns are sensitive to the speaker's epistemic stance–that is, to whether the referent is known or identifiable (Haspelmath 1997). For instance, the English determiner or adjective *certain* requires that the individual it introduces is known in some way. Conversely, Italian adjective *qualunque* and the French adjective *quelconque* are only felicitous when the speaker has no description of this individual (Jayez and Tovena 2006).

In (43-a), *certain* signals that the boy is identified–either by his name or by some other salient property–by the speaker or Gianni. Conversely, in (43-b), the boy is not identified, either because no distinguishing information is available or because such information is deemed irrelevant.

(43) a. Gianni walked into the classroom and addressed a certain boy.
b. Gianni è entrato in classe e si è rivolto ad un ragazzo qualunque.
'*Gianni walked into the classroom and addressed a (random) boy*'

Von Heusinger 2002 argues that *certain* prominently marks a specific reading of an indefinite. On the contrary, *qualunque* marks a non-specific one.

Building on this distinction, I propose that wh-words function as epistemic pronouns (or determiners). The determiner *which+sg* is always specific.[29] However, *who*, *what* and *which+pl*[30] are ambiguous: they can yield either a specific

29. Except in rhetorical questions, cf. section 2.5.2.

30. The EI of *which+pl* has the same properties as the EI of simplex wh-words. It is variable and can be suspended (i). This is probably due to plural being semantically unmarked (Sauerland, Anderssen, and Yatsushiro 2008).

(i) Which students cheated? Maybe no student cheated.

interpretation, akin to *certain*, or a non-specific interpretation, akin to *qualunque*.

Specific *who*$_S$

The specific *who*$_S$ is characterized by two main properties: transparency and referential anchoring.

According to von Heusinger 2002, "[t]he interpretation of a specific NP does not depend on the interpretation of the matrix predicate or semantic operators such as modal verbs". In other words, a specific NP is transparent. Likewise, the (plural) individual referred to by *who*$_S$ is not dependent on the interpretation of the predicates that scope over its trace. This transparency account for the different behaviors of *who*$_S$ and *who*$_G$ with respect to weak islands (cf. section 2.5.3).

Referential anchoring is the property whereby the value assigned to a discourse referent u, introduced by a specific NP, is functionally dependent on another discourse referent v. Typically, v may be the speaker, implying that she knows the value d of u.[31] However, v may also be another discourse referent introduced in the same sentence as u, such as Gianni in (43-a).

Under this definition, referential anchoring seems problematic for interrogative words, since the asker does not know the identity of u. However, the asker may still identify u in some way while inquiring about u in another way (Aloni 2001). For example, question (44) identifies u as *the woman who was sitting here* but asks about her name. Individual concepts offer a more suitable tool to grasp this subtlety.

(44) Who^u was sitting here? (She$_u$ forgot her bag.)

An individual concept is a function from possible worlds to individuals $f : s \to e$, e.g. *the president of the US*, *the woman with the weird hat*, *John's dog*, etc. In every world w, f identifies a unique individual $f(w)$, but that individual may vary from world to world.

A wh-word is specific when there exists a description that can be interpreted as an individual concept f such that the value of u is determined by the value of f. Even though the asker does not know the exact identity of u, she might identify u via f. For example, in (44), f = *the woman who was sitting here*.

Referential anchoring for *who*$_S$ can also depend on a discourse referent v introduced within the same sentence. In (45), the value of the specific *what*^{u} is determined by each person v, rather than by the speaker. It shows that specific wh-words are not simply wide-scope pronouns. Questions with a quantifier or modal operator scoping over the wh-word may be specific.

(45) What^u did everybody^v read? (*every* > *what*)

31. See section 5.4.2 for a formal definition of the specificity condition for *who*$_S$ using Dynamic Inquisitive Semantics.

Existential presupposition of *who*$_S$

Specific *who*$_S$ entails the existence of an individual concept f. This existential inference (EI) is spontaneously triggered in simplex wh-questions. I argue that this inference should be analyzed as a presupposition.

Notably, the EI of embedded interrogatives projects exactly like the existential presupposition of *which+sg*, as shown in (46). Previous research has argued that the existential presupposition carried by *which+sg* question is triggered by *which* (Rullmann and Beck 1998; Uegaki 2021). This similarity between the two cases suggests that a unified mechanism–presupposition projection–underlies both.

(46) **Similarity Observation**

a. John doesn't know { who / which student } cheated at the exam.
   $\rightsquigarrow$ { Someone / some student } cheated at the exam.

b. John wonders { who / which student } cheated at the exam.
   $\rightsquigarrow$ John believes that { someone / some student } cheated at the exam.

c. John doesn't agree with Mary on { who / which student } cheated at the exam.
   $\rightsquigarrow$ { Someone / some student } cheated at the exam.
   OR $\rightsquigarrow$ John believes that { someone / some student } cheated at the exam.

In particular, the EI projects through the negation of the attitude verbs like the presupposition of a definite phrase (47-a) and unlike the at-issue existential inference of an indefinite (47-b).

(47) a. John doesn't agree with Mary that the student cheated.
   $\rightsquigarrow$ There was a student.
   OR $\rightsquigarrow$ John believes that there was a student.

b. John doesn't agree with Mary that a student cheated.
   $\not\rightsquigarrow$ A student cheated.
   AND $\not\rightsquigarrow$ John believes that a student cheated.

This opposition suggests that *who*$_S$'s inference is a presupposition, and it not part of its at-issue content. That justifies why so many authors claim that the EI is a presupposition.

Generic *who*$_G$

The generic reading *who*$_G$ functions similarly to an epistemic free-choice item (e.g. French *quelconque*). It is felicitous only when the discourse referent introduced remains unidentified. Questions including generics, future tense, or habitual aspect favor *who*$_G$ because these environments license free-choice items.

By calling who_G *generic*, I do not mean that who_G refers to kinds. This terminology is motivated by some examples where who_G typically triggers modal quantification similar to the *Gen* operator (Chierchia 1995). The term *generic* also refers to generic questions like (48-b), i.e. questions asking for a general overview or opinion, not delving into a specific detail or situation. Finally, *generic* is a feature of anaphora annotation referring indiscriminately to non-specific NPs, i.e. NPs interpreted under the scope of another operator, e.g. in ARRAU (Poesio et al. 2024). [32]

The individuals quantified over by who_G may be absent from the common ground and only exist in local contexts. They are hypothetical individuals. As a result, reference to them requires modal subordination.

Questions with who_G support both mention-some and mention-all readings. [33] By contrast, who_S supports only mention-all readings, as the speaker seeks to identify a particular (possibly plural) individual. [34] Cross-linguistic evidence corroborates this distinction. In Mongolian, *ex situ* questions are answered with sentences using *ex situ* focused constructions, thus being exhaustive answers (Onea

32. A potential common feature to all the questions involving who_G presented in this thesis is Unselective Binding (Lewis 1975). In this theory, indefinites are interpreted as free variables that some overt or covert operator must bind. Unselective Binding has been developed for generics, habitual, and future. It has also been advocated for to interpret *in situ* wh-words (Cresti 1998) and model Quantificational Variability Effects (Lahiri 1991). In this setting, I could argue that who_G (but not who_S) requires an operator quantifying over situations to bind its referent. Examples involving quantificational variability, like (i), would then be analyzed as selecting who_G.

(i) John knows, for the most part, who cheated.
'*John knows the identity of most of the students who cheated.*'

33. A possible idea to bring my theory and the exhaustification paradigm (Chierchia 2013, a. o.) closer is the following. FCIs contribute domain alternatives that can be pre-exhaustified to derive the FC inference. Like an FCI, who_G contributes domain alternatives. Non-pre-exhaustified questions are mention-some readings, while pre-exhaustified ones are mention-all readings.

34. The interaction of who_S, focus, and modality would require further investigation. A non-modalized specific question cannot be mention-some. For example, in the cleft question (i-a), the asker refers to a particular person d that she wants to be reminded of. Answering with a strict sub-individual $d' < d$ would be considered non-cooperative. Wide-scope modality can make a *which+sg* question mention-some (see Chapter 4). Therefore, a wide-scope modality can make a specific question mention-some. However, the interaction between focus devices (selecting who_S) and modality is not entirely elucidated. Questions (i-b) and (i-c) are specific. Although (i-c) is clearly mention-all (with respect to what was uttered before by the responder), it is unclear to me what the most natural reading of (i-b) is. If (i-b) is interpreted as mention-some, I would analyze it as *can* scoping over the cleft.

(i) a. A: Who was it that knew Python?
b. A: Where is it that I can find an Italian newspaper? (attr. to Robert van Rooij)
c. A: Where is it again that I can find an Italian newspaper?

and Guntsetseg 2011). Similarly, French *c'est* clefts trigger an exhaustivity inference (Lambrecht 2001).

Differences in the domain

Finally, the two readings differ in the size and structure of their quantification domains.

Schwarzschild 2002 argues that specific readings of indefinites arise when the domain is a singleton. While this analysis cannot be directly applied to interrogative words, I have shown above that who_S can be construed as referring to a single individual concept in a comparable way.

By contrast, free-choice readings are argued to widen the (saliently given) domain of indefinites (Kratzer and Shimoyama 2002). For example, in (48-a), *what* ranges over books that the interlocutor might not yet be aware of. A similar question with past tense would restrict the domain to books the interlocutor has already read.

I advocate for modeling null answers as instantiations with the null individual $\star$. The who_G reading licenses null answers, so its quantification domain contains $\star$. This inclusion is consistent with domain widening.

Moreover, the domain of who_G may be less well structured. It might support fewer semantic operations, such as when quantifying over events rather than individuals, as in (48-b). In some cases, the domain may also depend on external parameters. For example, in in (48-c), the possible desserts may vary depending on which Dutch are being referred to, or their region of origin.

(48) a. What will you read during that meditation retreat?
b. What did you do during your weekend?
c. What do Dutch people eat for dessert?

2.4.2 Null answers

Null answers–also referred to as negative answers–are sometimes characterized as the Strawson (i.e. narrow-scope) negation of positive answers (Ginzburg 1995a; Ginzburg and Sag 2001; Duží and Číhalová 2015). However, other scholars have proposed an alternative characterization, treating null answers as negative instantiations of the question itself (Fitzpatrick 2005; Fox 2018; Roelofsen and Dotlačil 2023a).

In this section, I examine cases where the two competing hypotheses make divergent predictions. I argue that the Null Individual hypothesis better fits the observed patterns in question resolution.

The null individual

The null individual, symbolized by $\star$, is defined as the bottom element in the mereological lattice (Landman 2004). It corresponds to the empty set of individuals and has cardinality zero: $\#\star = 0$. Therefore, every predicate holds on $\star$.[35] In what follows, I call D the set containing atomic and strictly plural non-null individuals.

Bylinina and Nouwen 2018 propose using $\star$ to model the meaning of the numeral *zero*.[36] Intuitively, $\star$ can be interpreted as "at least nobody". To derive the intended "nobody" interpretation, exhaustification (EXH) is required: EXH negates the non-null domain alternatives. For example, sentence (49-a), interpreted as (49-b), is trivially true because $\mathbf{student}(\star)$ and $\mathbf{pass_test}(\star)$ are true in every world. However, when EXH is applied, as in (49-c), the sentence receives the interpretation "*At least no student passed the test and no other plural individual did*".

(49) a. φ = Zero students passed the test.
b. $[\![\varphi]\!] = \exists x \in D \cup \{\star\}.\ \#x = 0 \wedge \mathbf{student}(x) \wedge \mathbf{pass_test}(x)$
c. $[\![\text{EXH}\ \varphi]\!] = [\![\varphi]\!] \wedge \neg\exists y \in D \cup \{\star\}.\ (\#y > 0 \wedge \mathbf{student}(y) \wedge \mathbf{pass_test}(y))$

Theories of null answers

I define a *null answer* to a question Q as any answer to Q that is incompatible with Q's existential inference (EI). I use Fitzpatrick 2005's sentence $\mathbf{believe}_d(p) \rightarrow \mathbf{agree}_d(Q)$ to test whether p is an answer to Q (see section 2.2.3).

The theory I defend here–Null Individual–holds that a null answer to Q is computed by instantiating the wh-referent u with the null individual $\star$ and applying exhaustification $\text{EXH}_\star$. This operator only targets occurrences of $\star$ and conjoins the proposition φ with $\forall y \neq \star.\ \neg\varphi[\star/y]$. In cases where Q allows multiple instantiations of the wh-word (e.g., in pair-list readings), a null answer includes at least one instantiation by $\star$. By default, the scope of $\text{EXH}_\star$ aligns with the scope of the wh-word relative to other operators in the question.[37]

The competing theory–Negation of Positives–computes a null answer as the conjunction of the negations of all positive instantiations of Q, viz. $\bigwedge_{d \in D} \neg Q[u/d]$.

While both theories make equivalent predictions for simple questions, their predictions diverge when an operator scopes over the wh-word. In such cases, Negation of Positives yields a global negation, whereas Null Individual predicts a local negation.

35. See also the discussion on the pragmatics of $\star$ in section 2.6.1.
36. See also discussions in Buccola and Spector 2016, § 8.3.
37. See section 5.4.1 for a formal model using Dynamic Inquisitive Semantics.

Evidence from pair-list readings

Universally quantified wh-questions, like (50-a) give rise to a pair-list reading: a pair enumeration is expected (Chierchia 1993; Ciardelli, Roelofsen, and Theiler 2017), e.g. (50-b) (suppose here that the students are Anne, Bob and Charlie). The existential inference (50-c) conveys the existence of a witness for every student.

(50) a. What did every student read? (*every > what*)
b. Anne read Ulysses, Bob read Moby Dick, and Charlie read Madame Bovary.
c. EI: *Every student read something.*
d. Anne read Ulysses, Bob read Moby Dick, and Charlie read nothing.
e. Both Bill and Mary believe that Anne read Ulysses, Bob read Moby Dick, and Charlie read nothing, and so they agree on what every student read.

Any response like (50-d)–in which at least one student did not read anything–is incompatible with the existential inference (EI) of (50-a). In addition, (50-d) constitutes a resolving proposition for (50-a), as evidenced by the non-contradictory nature of (50-e). According to the definition introduced earlier, (50-d) therefore qualifies as a null answer to (50-a). More generally, any enumeration where at least one student fails to read anything is considered a null answer. [38]

Here, I adopt the standard assumption that pair-list questions quantify over functions from individuals to individuals (Engdahl 1980; Chierchia 1993; Roelofsen and Dotlačil 2023b).

The Negation of Positives theory predicts that (50-a) has exactly one null answer–namely (51-a), which is the negation of (50-c): "*Some student read nothing*". However, this prediction is wrong because (51-a) does not even constitute a valid answer to (50-a). As demonstrated by (51-b), such a response is infelicitous.

(51) a. $\bigwedge_{f:\mathbf{student}\to D} \neg(\forall x.\ \mathbf{student}(x) \to \mathbf{read}(x, f(x)))$
b. #Both Bill and Mary believe that some student read nothing and so they agree on what every student read.

On the contrary, Null Individual predicts that there are as many null answers as functions $f : \mathbf{student} \to D \cup \{\star\}$ such that $f(d) = \star$ for at least one $d \in \mathbf{student}$. Each null answer is of the form (52-a). An example with $f(a) = \mathsf{U}$, $f(b) = \mathsf{MD}$ and $f(c) = \star$ is given in (52-b).

(52) a. $p_f = \forall x \in D.\ \mathbf{student}(x) \to \text{EXH}_\star\, \mathbf{read}(x, f(x))$
b. $\mathbf{read}(a, \mathsf{U}) \wedge \mathbf{read}(b, \mathsf{MD}) \wedge (\mathbf{read}(c, \star) \wedge \forall d \in D.\ \neg\mathbf{read}(c, d))$

38. This contradicts Agüero Bautista 2001, who claims that pair-list reading is only possible with presuppositional wh-words.

Thus, Null Individual better fits the behavior of pair-list questions.[39]

Evidence from modalized questions

Questions containing an epistemic possibility verb can be interpreted with the modal operator scoping over the wh-word (Hirsch and Schwarz 2019; Richard 2024b). For example, (53-b) as a short answer to (53-a) is construed as the free-choice conjunction "*Brad might have killed Doug, and Crystale might have killed Doug*".

(53) a. Who might have killed Doug? (*might* > *who*)
b. Brad or Crystale.
c. It might be that somebody killed Doug.
d. It must be that somebody killed Doug.

It is unclear what the EI of (53-a) is. Hirsch and Schwarz 2019 predict the existential presupposition projection from under the possibility modal (53-c). In contrast, Richard 2024b predicts the stronger version (53-d).

Regardless, we observe that any response that includes a null individual is a resolving proposition of (53-a). For example, (54-b) being non-contradictory shows that (54-a) resolves (53-a). The stronger version, (54-c), assumed to be strongly exhaustive, also resolves question (53-a).

(54) a. It might be that nobody killed Doug. (*might* > *nobody*)
b. Both Criquette and Ashley believe that it might be that nobody killed Doug, and so they agree on who might have killed him.
c. Nobody might have killed Doug. (*nobody* > *might*)
d. Both Criquette and Ashley believe that nobody might have killed Doug, and so they agree on who might have killed him.

Negation of Positives predicts that (53-a)'s null answer is the conjunction of the negations of "*d might have killed Doug*", for $d \in D$. This amounts to the strong null answer (54-c).

In contrast, Null Individual predicts the weak null answer (54-a) by default (55-a). However, Null Individual can also predict the strong null answer (54-c) by letting $\text{EXH}_\star$ scope above the possibility modal ((55-b)).

39. This result extends to multiple-wh questions. The highest wh-word in LF is presuppositional, but the lowest can instantiate null individuals: (i).

(i) a. Who read what?
b. Anne read Ulysses, Bob read Moby Dick, and Chalie read Madame Bovary.
c. EI: *Everybody read something.*
d. Anne read Ulysses, Bob read Moby Dick, and Charlie read nothing.
e. Both Bill and Mary believe that Anne read Ulysses, Bob read Moby Dick, and Charlie read nothing, and so they agree on who read what.

(55) a. $\Diamond\, \text{EXH}_\star\, \mathbf{kill}(\star, Doug) = \Diamond(\mathbf{kill}(\star, Doug) \wedge \forall y \in D.\, \neg\mathbf{kill}(y, Doug))$
b. $\text{EXH}_\star\, \Diamond\mathbf{kill}(\star, Doug) = \Diamond\mathbf{kill}(\star, Doug) \wedge (\forall y \in D.\, \neg\Diamond\mathbf{kill}(y, Doug))$

To summarize, the Null Individual theory can account for both (54-a) and (54-c), whereas the Negation of Positives theory predicts only (54-c). The evidence from quantified and modalized questions therefore supports Null Individual as the more empirically adequate theory.

2.4.3 Explaining spontaneous triggering

In section 2.2.1, I noted that existential inferences (EIs) arise spontaneously–that is, by default–in non-biased, out-of-the-blue questions. This section accounts for that phenomenon, as well as the other properties discussed in section 2.2.2.

Given the ambiguity between who_S and who_G, discourse participants adopt strategies to infer which reading is intended. Although the speaker need not commit to either interpretation at the time of utterance, the responder is inclined to attribute a particular intention in the asker's utterance.

When the context clearly indicates that the asker is open to a null answer, the responder will interpret who_G. Fitzpatrick 2005's example (11) illustrates this well. Similarly, speaker suspension, as in example (8-a), is possible only under the generic reading.

In ambiguous contexts, I argue that responders rely on a communicative principle, namely Maximize Presupposition (Heim 1991; Schlenker 2012).[40] I assume that who_S and who_G share the same at-issue content: both introduce (or bind) a discourse referent (Dotlačil and Roelofsen 2021), quantifying over a domain including the null individual. However, who_S additionally carries an existential presupposition, eliminating the null individual. These two variants thus form a scalar relationship, with who_G as the weaker item and the who_S as the stronger. Because both readings are homophonous, the responder–following Maximize Presupposition–assumes that, whenever possible, the speaker intended the stronger form.[41] As a result, who_S is interpreted by default, giving rise to the spontaneous EI.

Nevertheless, the responder can revise this interpretation. If subsequent context warrants it, she may backtrack and reinterpret the question as involving who_G. This flexibility explains how null answers can be given without background disagreement, even if an EI was initially inferred.

40. I refer to Maldonado 2020, § 2 for details on the application of Maximize Presupposition in questions.

41. Baunaz 2016 assumes that the non-presuppositional and the specific *who* are not homophonous readings but one single lexical entry produced by different nanosyntactic trees, who_S having an extra *specific phrase* layer. This analysis is compatible with my proposal. Although Maximize Presupposition usually assumes that the expressions must have the same complexity–that is, the same number of syntactic nodes–to compete, here we can assume that this requirement is blind to sublexical/feature decompositions.

2.5 Applying the specific vs. generic theory

2.5.1 Weak NPIs in questions

Weak negative polarity items (weak NPIs) appear to disambiguate questions towards an interpretation that lacks an existential inference (EI), as illustrated by example (19-a), repeated here as (56). Roelofsen 2015 argues that weak NPIs are licensed in a question if and only if that question has no presupposition.

(56) What did Robin ever do to help Mary?

In this section, I show that Roelofsen 2018's characterization–along with other existing models–fails to account for certain data. Specifically, I present unexpected occurrences of weak NPIs under *be surprised*, which challenge previous assumptions and call for a more refined theoretical account. I argue that incorporating the specific vs. generic duality, along with its association with focus, provides the necessary framework to explain these cases.

Weak NPIs in weakly exhaustive readings

Guerzoni and Sharvit 2007 observe that a strongly exhaustive (SE) interpretation is sufficient to license weak NPIs in both matrix and embedded questions, as illustrated in (57).

(57) Andy wonders who has ever been to Paris.

The predicate *be surprised* is known to reject weak NPIs in interrogative complements, as seen in (58-b). Guerzoni and Sharvit 2007 argue that *be surprised* only accepts weak exhaustive (WE) complements, and therefore conclude that WE blocks weak NPIs. However, experimental results by Cremers and Chemla 2017 show that *be surprised* is compatible with both WE and SE interpretations. Yet, sentence (58-b) remains unacceptable under either reading (as illustrated by context (58-c) for SE and (58-d) for WE). This suggests that the unacceptability of (58-b) must be explained by something other than the WE/SE distinction.

(58) a. SITUATION: *John, Bob and Marc went to Angela's birthday. John brought her two gifts, Bob brought her one gift, and Marc didn't bring anything.*
b. *It surprised Angela, who brought her any gifts.
c. SE (only): *Angela was expecting everybody to bring her a gift.*
d. WE (and SE): *Angela was expecting only John to bring her a gift.*

I propose that weak NPIs can be licensed under *be surprised* in certain configurations. When the wh-word's trace is embedded and a different sentence element

(not the wh-word) takes focus, NPIs such as *any* and *ever* are acceptable within the scope of that trace. [42]

(59) a. SITUATION: *Each student said which movies they thought merited their success. Mary is an outstanding student and knows a lot about movies. However, this time, she mentioned very bad or controversial movies. Therefore, the teacher John is surprised that Mary, in particular, mentioned these movies.*
b. It surprises John, [which movies]$_i$ Mary$_F$ thinks [t_i have any merit].

(60) a. SITUATION: *Each student said which books they thought had already been censored. Mary is an outstanding student and knows a lot about publication history. However, this time, she mentioned uncontroversial books. Therefore, the teacher John is surprised that Mary, in particular, mentioned these books.*
b. It surprised John, [which books]$_i$ Mary$_F$ thought [t_i had ever been censored].

This novel data reveals that scope and focus configurations need more attention to account for weak NPIs in interrogatives. [43] Before presenting my proposal, I will first discuss an alternative theory that helps clarify the relevant scope effects.

Scope effects in NPI licensing

Han and Siegel 1996 observe that weak NPIs in questions are subject to scope effects. When the NPI scopes under the trace of the wh-word, the question can be information-seeking or rhetorical (61-a) (61-b). However, when it scopes over the trace, only the rhetorical reading is available (61-c) (61-d).

(61) a. Who said anything at the semantics seminar?
b. Who$_i$ did Jeff introduce t_i to anyone at the party?
c. What$_i$ did anybody say t_i at the semantics seminar?
d. Who(m)$_i$ did Jeff introduce anyone to t_i at the party?

In addition to the scope requirement, weak NPIs are not licensed in islands, e.g. in a relative clause (62). [44]

42. Thanks to Gidon Kaminer for help on these sentences.

43. The NPI environment in these examples is not downward-entailing. For example, in situation (59-a) augmented with (i-a), (i-b) is true but (i-c) is false.

(i) a. SITUATION: *Marie said that movies a and b won an Oscar and c and d won a César. Movies a and b indeed won an Oscar. However, c and d are very bad and won no prize.*
b. It surprises John, [which movies]$_i$ Mary$_F$ thinks [t_i have won a prize].
c. It surprises John, [which movies]$_i$ Mary$_F$ thinks [t_i have won an Oscar].

44. This scope effect invalidates Roelofsen 2018's model. Some non-presuppositional ques-

(62) *Who read the book which has any missing pages?

According to Han and Siegel 1996, a question licenses an NPI whenever some answer licenses it.

(63) a. {Mary said something, Lucas said something,..., Nobody said anything}
b. {Someone said that it was sunny, Someone said that the first talk was good,..., Someone said nothing}

For example, the alternatives of (61-a), represented in (63-a), contain the null answer "*Nobody said anything*", so (61-a) is felicitous with *anything*. Conversely, the alternatives of (61-c), represented in (63-b), do not contain the proposition "**Anybody said nothing*", so (61-c) disallows *anybody* under an information-seeking reading.

Contrary to non-reductive approaches (van Rooij 2003; Schwarz 2017a,b; Roelofsen 2018; Jeong and Roelofsen 2023), Han and Siegel (1996) correctly predict that the felicity of weak NPIs is highly sensitive to their position in the sentence. According to Fitzpatrick 2005, it also explains the divergence in EI strength between different wh-words.

A final argument in favor of their model concerns mention-some readings. Weak NPIs are licensed in mention-some readings, with a modal verb (64-a) or without (64-b). Mention-some readings are typically open to null answers (Dayal 2016, § 3.3.3.). [45]

(64) a. Which university might host any semantics conference?
b. Who has any programming skills?

Like Han and Siegel 1996, I recognize that the wh-trace plays a crucial role in weak NPI licensing. However, the licensing process cannot be explained using null answers under *be surprised*. Emotive factives have an existential presupposition (Roelofsen, Herbstritt, and Aloni 2019) and thus block null answers, e.g. (65) is degraded.

(65) #It surprises John which books, if any, Mary thinks have any merit.

Therefore, an alternative explanation is required.

---

tions do not license weak NPIs. For example, (i-a) can be answered by *Nobody* and has no presupposition, but replacing the indefinite NP by a weak NPIs is ungrammatical (i-b).

(i) a. Who read exactly one book that has a missing page / missing pages?
b. *Who read exactly one book that has any missing pages?

45. Mayr 2013 rejects that NPIs are licensed in mention-some readings based on (i). However, making a teleological conversational goal explicit improves their acceptability (Dayal 2016).

(i) *Where can I buy any newspaper?

Unifying theories

Taking stock, we must identify a licenser associated with the wh-trace–typically the locus of null answer licensing. Generic who_G emerges as a promising candidate.

Within the specific vs. generic distinction, Roelofsen's model could be reinterpreted as stating that only generic questions license weak NPIs. Contextual influence supports this claim: question (66-a) is infelicitous in situation (66-b), where John has somebody in mind, but become acceptable as a generic question, e.g. in situation (66-c).

(66) a. JOHN: Who ate any cookies?
b. #SITUATION: *John baked 20 cookies. Now, one cookie is missing.*
c. SITUATION: *John baked a large number of cookies. He wants to bake more cookies, but only for those who haven't had any cookies.*

Section 2.4.1 suggests that who_G allows both mention-some and mention-all readings, whereas who_S yields only a "definite" reading. The hypothesis developed below can thus be seen as integrating Han and Siegel 1996, Roelofsen 2018 and my theory in a unified framework, while also accounting for the empirical observations made by Guerzoni and Sharvit 2007.

Addressing the *be surprised + weak NPI* problem

Section 2.3.1 introduced the idea that the existential inference (EI) is correlated with focus on the wh-word. It was proposed that who_S can be focused while who_G cannot. Emotive factives like *be surprised* are focus-sensitive. By default, focus falls on the embedded wh-word. When the wh-word is focused, only who_S is available. In contrast, when focus targets another constituent, the wh-word can receive a generic interpretation (who_G).

I propose the following hypothesis: the trace t_G left by who_G licenses weak NPIs within its scope. Combined with the focus-sensitivity observation, this predicts that weak NPIs are licensed under *be surprised* only when (i.) they scope below the wh-trace, and (ii.) the wh-word itself is not focused. This prediction explains why NPIs are permitted in sentences (59-b) and (60-b), but not otherwise under *be surprised*.

Further support for the idea that t_G licenses NPIs comes from intervention effects. In declaratives, the focus-sensitive element *even*, when positioned between negation and a weak NPI, blocks licensing (67-b) (Chierchia 2013). The same effect appears in questions (67-c).

(67) a. A: Mary and Bill read something interesting.
b. A: *But, I don't think that even John$_F$ read anything interesting.
c. A: *Who thinks even John$_F$ read anything interesting?

However, when the wh-trace scopes below *even*, the NPI is rescued, as shown in (68-b).

(68) a. A: Mary said that Alice, Bob and Charlie read something interesting. Miriam said that Anne, Bob and Charlie read something interesting.
b. A: Who$_i$ did even John$_F$ say [[t_G]$_i$ had read anything interesting]?

More generally, any focused element between the trace and the NPI is an intervener (69). This phenomenon explains why *Mary*$_F$ must scope over the wh-trace in (59-b) and (60-b).

(69) *It surprises John, who thinks Mary$_F$ invited anyone.

To summarize, the specific/generic duality offers a coherent and elegant [46] solution to the puzzle of weak NPIs in question. Moreover, the interaction between *be surprised*, weak NPIs, and focus supports the proposed link between focus structure and the *who*$_S$/*who*$_G$ distinction.

2.5.2 Question types and the reading likelihood gradient

So far, I have focused on information-seeking questions. In this section, I examine other types of questions and argue that they, too, conform to the specific vs. generic duality. Due to their lack of ambiguity, these question types offer additional evidence in support of the theory.

Non-information-seeking question types

Rhetorical questions are typically biased towards the null answer (70-a). Moreover, Schwarz 2017b and Jeong and Roelofsen 2023 observed that *which+sg* questions–which usually block NPIs–can exceptionally license them in rhetorical questions (70-b). [47] [48]

46. In particular, it does not require the complex machinery proposed by Nicolae 2015.

47. Note that this phenomenon cannot be attributed to a hypothetical presupposition weakening triggered by rhetorical questions. Scope presuppositions project from rhetorical questions as usual, e.g. (i).

(i) Which sane human being would ever vote for the captain?
$\overset{\text{presup}}{\Longrightarrow}$ There is a captain.

48. Some informants of Schwarz 2017b rate (i) as acceptable (see also Krifka 1995, p. 251). He hypothesizes that they accept it if and only if they take (i) as compatible with a potential situation where no student ever went to China. Perhaps these informants re-interpret *which student* as *which students* to make sense of the NPI. However, it is also possible that they let *which+sg* be construed as generic.

(i) Which student has ever been to China?

(70) a. And who cared for me when I was working so hard?
b. Which sane human being would ever vote for Prof. Jones?

By contrast, exam questions[49]–that is, questions where the asker already know the answer and tests the responder's knowledge–disallow null answers. Consider dialogue (71): if PROF. A poses a question she knows has no positive answer, the question will be perceived by STUDENT B as non-cooperative.

(71) a. PROF. A: When did the Belgian monarchy end?
b. STUDENT B: I don't know.
c. PROF. A: Never. Belgium is still a monarchy.
d. STUDENT B: Wait a minute, that was a trick question!

Similarly, echo questions (72) refer to a contextually salient content whose existence is therefore presupposed. In English, they are formed by stressing the wh-word, which can appear in either canonical or *in situ* position. Following Erteschik-Shir 1986, I treat echo questions as involving explicit focus on the wh-word.

(72) I didn't hear well. You bought WHAT?

Exam questions are mention-all by default. For instance, if it is known that two people independently invented dynamics semantics, (73-a) is interpreted as requesting the identities of both individuals. To allow non-exhaustive answers, exams often include an explicit qualifier, such as (73-b). Echo questions are also mention-all.

(73) a. Who invented dynamic semantics?
b. Give at least one name.

Specific vs. generic likelihood gradient

These three question types align with the specific vs. generic ambiguity theory. Rhetorical questions are necessarily generic. Even singular *which* is interpreted generically in rhetorical questions, thereby licensing weak NPIs in the scope of its trace (cf. section 2.5.1). In contrast, exam and echo questions are specific (Starke 2001), with echo questions being linked to focus.[50] In both cases, the asker has a individual in mind, which makes any null answer appear non-cooperative.

The lack of ambiguity in these question types confirms the existence of purely specific and purely generic questions, reinforcing the hypothesis that information-seeking questions are inherently ambiguous between these two poles. More broadly,

49. Exam questions are also called quiz master questions (Higginbotham 1996).

50. More precisely, echo questions are not specific about an individual but about an expression the interlocutor said. The existential presupposition of (72) is "*There is an expression x such that you said "I bought x"*".

Reading	Semantic-pragmatic contexts
who_S only	exam questions, echo questions
who_S more likely	episodic
who_G more likely	habitual, generic, future
who_G only	(emphatic) NPI questions, rhetorical questions

Table 2.1 – Gradient of who_S vs. who_G likelihood and associated semantic/communicational contexts

the likelihood of each reading can be seen as lying on a continuum. Episodic questions are more often interpreted with who_S, while questions containing habitual aspect, generics, or future tense tend to favor who_G.This gradient is summarized in Table 2.1.

Social pressure to answer

Section 2.2.6 proposes a test to determine whether a question is specific: attempt to rephrased it as a conjunction of an existential statement and a request, e.g. (74-b). This rephrasing works with specific questions but fails with generic ones, such as (75-b), because generic questions do not presuppose the existence of a particular individual.

Additionally, rephrasing (75-a) with a strong request, like the second sentence of (75-b), feels unnatural. Question (75-a) is more akin to a suggestion or invitation, where the responder is not obliged to answer. A more accurate rephrasing would be (75-c), which better captures the open-ended nature of the question. For example, in the case of (75-a), the responder can politely decline to answer if she does not want to talk about it. This refusal has few social consequences. In comparison, failing to answer (74-a) may have wider social consequences–for instance, the asker might hold it against the responder, supposing the asker believes that the responder knows the answer.

(74) a. Who was sitting here? (She forgot her bag.)
b. Someone was sitting here. I would like to know her name.
c. #If someone was sitting here, I propose you share her name with me.

(75) a. What did you do during your weekend?
b. #You did things during your weekend. I would like to know those things.
c. If you did things during your weekend, I propose you share them with me.

These examples suggest that specific who_S is associated with a higher degree of social pressure to answer than who_G.

Other question types support this hypothesis. Exam questions, by definition,

require an answer. Echo questions also place strong pressure on the hearer to provide information. Conversely, rhetorical questions are not expected to be answered.[51] Between these extremes, the level of social pressure varies.[52]

This communicative gradient highlights another dimension in which the specific/generic distinction clarifies the dialogical effects of questions.

2.5.3 Explaining weak islands

In light of my theory, I address the puzzle about weak islands with *who* and *what* presented in section 2.2.5.

Specificity and long distance

Rizzi 1990 and Cinque 1990 argue that being D-linked (Pesetsky 1987), i.e. "refer[ring] to specific members of a pre-established set", is a condition to escape weak islands. In their account, referential wh-phrases can be long-distance linked to their traces, whereas non-referential wh-phrases are prone to several intervention effects. More generally, Kratzer 1998 and Baunaz 2016 argue that specific readings tend to violate scope islands more easily.

Along these lines, I argue that specificity is a redeeming factor for weak islands with simplex wh-words. Questions like (76-a) are weird because the generics *Dutch people* favors who_G and who_G is sensitive to weak islands, contrary to who_S. However, a cleft (76-b), interpretation as an exam question (76-c) or an echo question (76-d) salvage this utterance because these constructions force a specific reading.

51. The observations about rhetorical questions probably extend to reflective questions. Giannakidou and Mari 2019 and Mari and Giannakidou 2021 remark that some epistemically existentially modalized questions are open to not being answered, called *reflective questions*. For example, question (i-a) is typically uttered by someone who wonders why the door is opened and is open to non-stereotypical individuals. For example, the asker might entertain that the door might have opened because of the wind. Reflective questions also license NPIs (i-b). Therefore, they should probably be construed with who_G.

(i) a. Qui a bien pu ouvrir la porte ?
who has well can.PPRT open the door ?
'*Who may have opened the door? (I really wonder.)*' (French)

b. Qui a bien pu faire quoi que ce soit d' aussi stupide ?
who has well can.PPRT do something.NPI of so stupid ?
'*Who could have done anything so stupid?*'

In reflective questions, the asker does not pressure other participants to provide information but instead aims to raise inquisitiveness. Contrary to rhetorical questions, the asker of reflective questions does not know the answer(s).

52. The gradient in Table 2.1 is not correlated with authority. In exam and rhetorical questions, the asker has authority and knows the answer(s). Conversely, in echo questions, the responder has the authority. In reflective questions, no one is expected to have authority (Giannakidou and Mari 2019).

(76) a. ?What do Dutch people wonder whether to eat for breakfast? (generic)
b. What is it that Dutch people wonder whether to eat for breakfast? (cleft)
c. According to a Dutch study conducted in 2025, what do Dutch people wonder whether to eat for breakfast? (exam question)
d. Sorry, WHAT do Dutch people wonder whether to eat for breakfast? (echo question)

The ability of who_S to escape weak islands boils down to transparency: its trace is construed independently from its surrounded semantic context (von Heusinger 2002).[53] This characterization should not be confused with wide-scope taking. As mentioned in section 2.4.1, who_S does not necessarily take wide scope with respect to operators in the question nucleus. For example, a pair-list specific reading of (76-c), with the generics scoping over *who*, is possible because it accepts pair-list answers like (77). In this case, the individuals *dairy products* and *gluten* are specific with respect to the respective halves of Dutch people.

(77) Half of them wonder whether to eat dairy products for breakfast, and the other half wonder whether to eat gluten.

Non-specific wh-words are poorer extractees because their interpretation depends on the semantic environment around their trace. This property is most visible in rhetorical questions, which are sensitive to more intervention effects than just weak islands (Han and Siegel 1996). For example, weak NPI *anyone* scoping over t_i forces a rhetorical reading and thus fails extraction in (78-b).

(78) a. Who does John thinks [Mary has kissed t_i]?
b. *Who$_i$ does John think [anyone has ever kissed t_i]? (Han and Siegel 1996)

Domain effect

Szabolcsi and Zwarts 1993 advocate that bad weak island extractees have a domain lacking common semantic operations provided by the structures producing weak islands (e.g. set complementation for negation). Typically, non-individuated domains (i.e. with overlapping elements) are poorer wide-scope takers than domains of distinct individuals.

Section 2.4.1 claims that wide and not well-behaved domains are characteristics of who_G. Therefore, Szabolcsi and Zwarts 1993's analysis supports that who_G is sensitive to weak islands and not who_S.[54]

53. Both echo questions and exam questions license *in situ* wh-words in English (Cole 1987). This exceptional syntactic behavior is probably due to their specificity.

54. Following Pesetsky 1987, we could take *the hell* as a non-specific marker. Indeed, it opens the question to non-stereotypical individuals. However, it is acceptable when the asker

2.6 An alternative analysis

In this section, I explore the alternative hypothesis where the specific vs. generic opposition is not a semantic ambiguity but is due to pragmatics.

2.6.1 Null individuals are marked

Landman 2011 argues that the null individual $\star$ is needed in the domain of plural predicates, e.g. *dogs*. Determiners with a downward-entailing restrictor often trigger an existential inference (79). However, this inference is not a presupposition. These determiners can be used *stricto sensu* with an empty domain (80-a), e.g. in mathematics (80-b). This is not possible with *the+sg* (80-c) and definite numerals (80-d).

(79) Every dog is barking.
$\rightsquigarrow$ There are dogs.

(80) a. The students that studied for the test got a good grade, but nobody did. (Landman 2011)
b. Every negative prime number is even.
c. #The student that studied for the test got a good grade, but nobody did. (Landman 2011)
d. #The two negative prime numbers are even.

Landman 2011 argues that the existential inference of NPs with a downward-entailing restrictor is a conversational implicature. He attributes this inference to Grice's Maxims of Quality and Quantity. A downward-entailing predicate P is true on the empty individual $\star$. Nonetheless, as $P(\star)$ is a triviality, the speaker should avoid uttering it. [55]

Extending Landman 2011's analysis to simplex wh-words amounts to assuming

---

has evidence over an individual but wonders about its identity (i-b) (cf. *identity* readings of *wh-ever* in fn. 22).

(i) a. ??Who the hell are you wondering whether to invite? (Pesetsky 1987)
b. SITUATION: *B has been madly searching through the dictionary for twenty minutes.* A TO B: What the hell do you still not know how to spell? (Szabolcsi and Zwarts 1993)

Therefore, although *the hell* may favor a generic reading, it does not impede a specific one. Like *-ever* for English in wh-free relatives and reflective questions (see fn. 51), domain widening and attributivity appear to favor but not select who_G.

55. Aloni 2022 suggests that human beings tend to reject sentences that are satisfied by virtue of the empty set: the neglect-zero effect. We could assume that such a constraint also applies to the individual domain: a neglect-zero-individual effect. In other words, human beings tend to reject sentences that are satisfied by virtue of the null individual because it has a higher cognitive cost. Thus, cognition could be another reason why $\star$ is marked.

that the domain of wh-words includes $\star$ by default–corresponding to the denotation of who_G. Due to a communicative principle, an existential conversational implicature would be triggered in simplex wh-questions, but it could be waived. This would explain the EI spontaneous triggering, responder cancellation without background disagreement, and speaker suspension.

However, modeling the EI as a conversational implicature is not sufficient. The EI can be strong, can project, and null answers can be non-cooperative in some contexts (e.g. section 2.2.6, exam and echo questions). Therefore, an analysis of the EI of simplex wh-questions *à la* Landman 2011 cannot account for the observed data alone.

2.6.2 Pragmatic referential anchoring

Dekker 1998 argues that referential anchoring (cf. section 2.4.1) is a pragmatic phenomenon. By default, NPs receive an attributive reading. However, their domain can be narrowed to a singleton by anchoring the (plural) individual the speaker has in mind.

Building on this idea, the default denotation of simplex wh-words can, in certain contexts, be referentially anchored–yielding a perceived interpretation similar to who_S, with an existential presupposition.

This pragmatic ambiguity theory thus predicts that the existential inference (EI) may arise in two distinct ways: either as a conversational implicature via rejection of the null individual $\star$, or as a presupposition triggered by referential anchoring.

2.6.3 Drawbacks of the pragmatic approach

Although a pragmatic analysis seems more parsimonious, it faces several issues confronted with the data presented in this chapter.

First, the pragmatic approach fails to connect interrogative wh-words with their non-interrogative counterparts. As discussed in section 2.3.2, the NPI and free choice readings of certain non-interrogative wh-items appear related to the epistemic pronoun nature of who_G. Referential anchoring does not offer an adequate explanation for this connection. Nor does it account for why generics, habitual aspect, and future tense disfavor referential anchoring.

The most significant problem concerns NPIs. The data in section 2.5.1 are best explained by the hypothesis that weak NPIs within questions are licensed in the scope of who_G's trace–only in those cases. This theory is supported by previous works emphasizing the role of scope and the correlation between null answers and weak NPI licensing. For a pragmatic approach to replicate these predictions, it would need to posit that NPI licensing by the wh-trace is the default, but that referential anchoring somehow removes this property. However, this is implausible, since specific NPs neither block nor interfere with NPIs.

In summary, treating the specific vs. non-specific opposition as a matter of pragmatic ambiguity offers an alternative explanation for the EI of simplex wh-questions. Yet, this approach fail to account for the full range of empirical observations. The semantic ambiguity hypothesis, by contrast, provides a more explanatory and robust framework.

2.7 Comparison with other approaches

Now that my theory is laid out, I turn to potential counterarguments. I first explore hypotheses of EI variation located in the grammar (section 2.7.1) or based on pragmatic principles without any ambiguity (section 2.7.2). I show that they cannot account for the data *on their own*, for an asymmetry is required somewhere in the model.

Then, I turn to other potential asymmetries. In section 2.7.3, I discuss the potential contribution of a semantic operator. I confront the idea that the existential inference (EI) is due to bias in 2.7.4. Finally, I point out the similarities and differences with Elliott, Nicolae, and Sauerland 2022's type-flexible *who* in section 2.7.5.

2.7.1 Grammatical derivation

Soft presupposition

Abusch 2010 proposes that a general grammatical principle derives the EI. Constructions invoking alternatives generate a soft presupposition computed from the disjunction of their alternatives. As Abusch 2010 only considers positive instantiations, disjoining them produces the EI, as illustrated in (81).

(81) a. Who cheated?
b. $\mathcal{A} = \{$Anne Cheated, Bob cheated,... $\}$
c. $\bigvee \mathcal{A} =$ "*Someone cheated.*"

I proved in section 2.2.3 that null answers are full-fledged answers (and alternatives) of questions, both with respect to felicity and resolution conditions. If nothing else is posited, then there is *a priori* no reason to reject null answers or to rank them differently from positive answers. This constitutes the Symmetry Problem (82).

(82) **Symmetry Problem**
All positive and null answers are, theoretically, equally available and equally likely.

If null answers are part of the alternative set representing a question, e.g. (83-a), then taking the disjunction of this set does not entail the EI. Thus, Abusch

2010's derivation fails.

(83) a. $\mathcal{A} = \{$Anne Cheated, Bob cheated,..., Nobody cheated$\}$
b. $\bigvee \mathcal{A} = \top$

Structured questions

Dayal 2016 and Onea and Zimmermann 2019 mention that the EI can be accounted for while keeping null answers as native if we assume that questions are structured. More precisely, they posit that a simplex wh-question like "*Who cheated?*" contains a conditional like (84-a). Null answer (84-b) negates the condition.

(84) a. A: Who, if anyone, cheated?
b. B: Nobody.

This proposal cannot explain the variation in EI strength. It fails to model that some simplex wh-questions have a clear-cut existential presupposition. Moreover, it incorrectly predicts that a null answer does not constitute a resolution condition for a question: Fitzpatrick 2005's example is odd under this structured interpretation (85).

(85) #John and Mary both believe that nobody cheated and so they agree on who, if anyone, cheated.

As a consequence, I reject the structured question hypothesis. [56]

Exhaustification

The anti-singular inference (ASI) of *which+pl* question (e.g. (86-a)) is another question inference that is not a presupposition. Like the EI, it can be suspended (86-b).

(86) a. A: Which students cheated?
$\rightsquigarrow$ Several students cheated.
b. A: Maybe only one student cheated.

Elliott, Nicolae, and Sauerland 2022 claim that the plural marking of the wh-complement has singular marking as an alternative (here *alternative* in the sense of scalar implicature theory). They derive the ASI by using Maximize Presupposition, encoded in the exhaustification operator EXH (Chierchia, Fox, and Spector 2012).

56. The rephrasing suggested in section 2.5.2, involving a similar conditional, is a test to distinguish specific from generic questions. I do not maintain that it is semantically equivalent to the respective question.

Such derivation crucially builds on the singular vs. plural opposition. In order to proceed similarly with the EI, an inherent asymmetry in simplex wh-words must be posited.[57] Operator EXH cannot provide such an asymmetry on its own, but my theory does.

2.7.2 Pragmatic derivation without an asymmetry

Maximal informativity principle

Dayal 1996 introduces a pragmatic derivation of the EI using the Maximal Informativity Principle (87). Like Abusch 2010, Dayal 1996 assumes that the answer set does not contain null answers. Again, this is an incorrect assumption.

(87) **Maximal Informativity Principle**
A question Q presupposes the existence of a unique, maximally-informative true answer to Q.

Assuming null answers belong to the answer set, principle (87) fails to predict EI triggering.

Weak presuppositions

Chierchia and Caponigro 2013 claim that the EI of simplex wh-questions is a "weak" presupposition that only holds in the speaker's epistemic state. When asking a question, the asker proposes to project her epistemic state to the common ground. Contrary to question's scope presuppositions, "weak" presuppositions of questions do not project to the common ground (Gennaro Chierchia, p.c.). This trigger distinction explains why responders may retort *Nobody* without background disagreement.

Contra Chierchia and Caponigro 2013, I argue that speaker suspension is possible in some contexts (see section 2.2.2), namely in generic questions. This phenomena contradicts that all simplex wh-questions project a presupposition to the asker's epistemic state.

Chierchia and Caponigro 2013 predict that this "weak" presupposition only projects to the attitude holder's mental state for embedded interrogatives. This is not always true. For example, the EI (88-b) of the embedded interrogative in (88-a) can project to the matrix level. Therefore, the EI of (88-a) is a regular presupposition.

57. I do not claim that exhaustification-based systems cannot produce null answers. They can, by two means. First, by allowing null answers in the alternative set Q. Second, by allowing the answer operator to select (a conjunction of) propositions that do not belong to Q. What I claim is that this strategy cannot predict that some environments *disfavor* the EI (as opposed to: *prohibit* it). On the other hand, a reading ambiguity is suited to explanations based on preferences as a gradient of interpretation likelihood.

(88) a. John doesn't agree with Mary on who cheated at the last exam.
b. $\overset{\text{presup}}{\Longrightarrow}$ Someone cheated at the exam.

Finally, raising a new class of "weak" presuppositions only for simplex wh-questions may seem *ad hoc*. My theory is more explanatory as it extends independently motivated properties of indefinites to interrogative words.

Suppositions

Onea and Zimmermann 2019 argue for a similar approach where EI is weaker than a presupposition and originates from the asker's expectations. They call this inference a supposition, referring to the EI of focus constructions (Büring 2004).

The arguments against soft presuppositions and weak presuppositions for simplex wh-questions also hold against this reasoning. See also section 2.7.4 for comments on a closely related analysis.

2.7.3 Ambiguity due to a semantic operator

In this section, I explore a alternative hypothesis where the asymmetry needed to explain existential inferences (EIs) arises from a semantic operator. I examine whether the exhaustivity operator, the question operator, and the **atom** operator could predict the observed properties.

As discussed in sections 2.4.1 and 2.5.1, EIs appear to only be triggered in exhaustive readings, whereas mention-some readings are necessarily compatible with null answers. However, some exhaustive readings also allow null answers, e.g. (89). This shows that EIs cannot be attributed to exhaustivity *per se*, nor to any operator encoding question exhaustivity.

(89) a. SITUATION: *Laura is surveying Anna's start-up.*
b. LAURA: Who, among the employees, is coming to the office by car?
c. ANNA: Nobody.

Similarly, question operators (e.g. Heim 1994) do not predict variation in EIs. While we could technically posit two question operators–one triggering an existential presupposition and one not–this move would lack explanatory power with respect to the broader empirical data presented in this thesis, including the behavior of non-interrogative wh-items, anaphoric properties of wh-words, and scope effects of NPIs.

Finally, consider the **atom** operator contributed by *which+sg*. This operator restricts the domain of the referent u introduced by *which*^{u} to atomic individuals, thereby excluding the null individual $\star$. If the EI were due to an overt **atom** operator, we would predict that EIs always co-occur with a uniqueness inference. However, no such correlation is observed. For example, presuppositional question (90) triggers an inference that several individuals were (distributively) sitting on

those chairs. Even *which+pl* can trigger a spontaneous EI alongside an anti-singular inference (86-a). Thus, **atom** is not the source of EIs in questions.

(90) Who was sitting on those chairs?

In conclusion, the who_S/who_G ambiguity remains the most plausible hypothesis, assuming the relevant asymmetry arises from a semantic component of the question content.

2.7.4 The bias hypothesis

My theory assumes that the inherent asymmetry in the meaning of questions is located at the lexical level. [58] In contrast, Onea 2016 argues that this asymmetry is a general property of questions, linked to the notion of bias.

Bias refers to the asker's preference or prior belief favoring one answer over another. For example, positive polar questions like (91-a) are typically neutral or only weakly biased toward the positive answer (AnderBois 2011). Conversely, English low-negation questions, such as (91-b), tend to be biased toward the negative answer.

(91) a. A: Is John vegetarian?
possibly: $\rightsquigarrow$ A expects John to be vegetarian.
b. A: Is John not vegetarian?
$\rightsquigarrow$ A expects John not to be vegetarian.

AnderBois 2011 models bias using an additional representation tier. Similarly, Roelofsen and van Gool 2010 and Roelofsen and Farkas 2015 introduce a *highlight* tier to account for the behavior of polarity particle responses to alternative and polar questions.

Building on this, Onea 2016 proposes that the existential inference (EI) of simplex wh-questions is derived through default highlighting of positive answers. [59] He introduces a pragmatic principle called The soundness Rule (92) ensuring that the disjunction of positive alternatives is established in the context of utterance.

(92) **The soundness Rule**
The union of highlighted alternatives of a question q should be known to be true in a context c. (Onea 2016)

Crucially, violating the Soundness Rule does not result in a presupposition failure but is instead interpreted as an uncooperative move. [60] However, as dis-

58. Though it is not a lexical ambiguity, but a semantic ambiguity.

59. In a similar fashion, Aloni 2003 estimates that positive answers are under discussion (cf. QUD (Roberts 1996) a. o.) while the negative answer is not.

60. AnderBois 2011 analyzes the inference triggered by highlight as an implicature. This falls into the same pitfalls I exhibited in section 2.6.1.

cussed in section 2.2.6 and elsewhere, null answers can be fully cooperative. This indicates that Onea's model is too restrictive.

While linking question asymmetry to bias is conceptually appealing–especially given bias's sensitivity to various aspectual features–I am unaware of connections between bias and environments that license free-choice items. Moreover, as with other competing theories, Onea's model does not explain the predisposition of wh-referents to undergo modal subordination in non-specific questions. Finally, the use of highlighting offers little insight into the grammaticality contrasts seen in weak islands. Although Herbstritt 2014 and Roelofsen, Herbstritt, and Aloni 2019 apply the highlight framework to embedded interrogatives–arguing that emotive factives like *be surprised* are sensitive to highlight–there is no independent motivation for using this approach to account for weak island escapability in particular.

2.7.5 The type-flexibility hypothesis

According to Spector 2007, wh-words can sometimes quantify over generalized quantifiers. Elliott, Nicolae, and Sauerland 2022 posit that simplex wh-words are type-flexible. They either range over individuals (type e) and trigger an existential presupposition, or they range over generalized quantifiers (type $\langle et, t\rangle$) and trigger no EI. In this section, I list the similarities and differences between their proposal and mine.

The variant who_e ranges over atomic and non-atomic non-null individuals: $[\![who_e]\!] = \lambda P^{et}.\ \exists x^e.\ P(x)$. Therefore, it triggers an existential presupposition due to the Maximal Informativity Principle (87) (Dayal 1996). On the contrary, the variant $who_{\langle et,t\rangle}$ ranges over a set $\mathcal{G}$ of generalized quantifiers, including the constant true and constant false quantifiers ($\lambda p^{et}.\ \top$ and $\lambda p^{et}.\ \bot$ respectively): $[\![who_{\langle et,t\rangle}]\!] = \lambda P^{\langle\langle et,t\rangle,t\rangle}.\ \exists G^{\langle et,t\rangle} \in \mathcal{G}.\ P(G)$. As a consequence, $who_{\langle et,t\rangle}$ does not trigger an existential presupposition. Similarly to my proposal, Elliott, Nicolae, and Sauerland 2022 use Maximize Presupposition to select the most salient variant.

Elliott, Nicolae, and Sauerland's proposal supports the idea of inherent ambiguity in wh-words I also defend. We could imagine that $who_S = who_e$ and $who_G = who_{\langle et,t\rangle}$. However, the two theories make different predictions.

Arguments in favor of a higher-order quantification

Let me begin with locations where Elliott, Nicolae, and Sauerland's proposal fits in the landscape of my data.

Elliott, Nicolae, and Sauerland 2022 aim to model Spanish simplex wh-words (Maldonado 2020). Getting things into a broader variational perspective, as I endeavored to do, would greatly benefit the type-flexible hypothesis while linking it to mine. Let me present some bridging directions.

American Sign Language (ASL) has a syntactic construction made of a question-answer pair. Caponigro and Davidson 2011 mention that the answer constituent

of such a pair can either be referential or non-referential. Non-referential uses include sign NOTHING and sign EVERYWHERE. This supports a connection between higher-order quantification, the possibility for null answers and anaphoric referentiality. [61]

Fox 2018 provides evidence that mention-some questions involve higher-order quantification. Mention-some questions easily allow null answers and weak NPIs. This supports a $\langle et\rangle t$ type for non-specific wh-words.

Iatridou and Varlokosta 1996 argue that *whatever* forms, together with its restrictor, a phrase of type $\langle et, t\rangle$ (e.g. (93-a)), while free-relative *what* forms a phrase of type e (e.g. (93-b)). [62]

(93) a. [Whatever Mary bought]$^{\langle et\rangle t}$ was expensive.
b. [What Mary bought]^{e} was Barriers.

As argued in section 2.3.2, this free-choice vs. definite reading boils down to an attributive vs. referential ambiguity. Exporting the type duality of attributive vs. referential readings to interrogative wh-words amounts to Elliott, Nicolae, and Sauerland 2022's type flexibility. [63]

The last connection concerns the similarity regarding weak islands. In sections 2.5.3, I argued that who_S can escape weak islands while who_G is impeded. Spector 2007 observes similar constraints. The *individual* reading who_e is not restricted by weak islands, while the higher-order reading $who_{\langle et,t\rangle}$ cannot be instantiated in a narrow scope with respect to a weak island. As Spector 2007 says, "The fact that this "higher-order" reading is sensitive to such grammatical constraints strongly suggests that it corresponds to a genuinely distinct reading of wh-questions (and is not simply a pragmatic effect)".

Differences and limits

Despite the apparent closeness between Elliott, Nicolae, and Sauerland's suggestion and mine, I would like to highlight a few points that need to be solved to make their proposal fit the data I raise.

First, as the authors recognize, using constant generalized quantifiers (i.e. $\lambda p^{et}.\ \top$ and $\lambda p^{et}.\ \bot$) is not a common assumption. A way to alleviate that

61. See also section 5.2.4 for similar data in a spoken corpus.

62. However, note that Dayal 1997 challenges this assertion and captures the opposition in (93) in terms of quantification over identity-alternatives.

63. Fox 2018 remarks that free-relative *what* can have a higher-order reading. For example, (i-b) is true if War & Peace and Brothers Karamazov can each be read in one week, but not both. Therefore, "*What Sue is required to read*" has to be of type $\langle et\rangle t$ and not e. Consequently, attributivity reading may not be solely imputed to *-ever*, contra Dayal 1997.

(i) a. SITUATION: *Sue has to read either War & Peace or Brothers Karamazov.*
b. [$_Q$ What Sue is required to read]$^{\langle et\rangle t}$ can be read in one week. ($can > Q$)

non-independently motivated step is to let the empty generalized quantifier (i.e. $\lambda p^{et}.\ p = \{\}$) be in the domain of $who_{\langle et\rangle t}$, what the authors originally proposed in an earlier version of their paper.[64] Admittedly, this goes against Spector 2007's range restriction to upward-entailing quantifiers. However, this exception is clearly motivated by null answers, contrary to constant quantifiers.

Second, higher-order quantification overgenerates. If we assume that $who_{\langle et,t\rangle}$ is always available, responses like (94-b) and (94-c) are incorrectly predicted to be answers to (94-a). They are rather partial answers because they do not bring enough information to fully resolve the question. For example, after (94-c), A can legitimately ask back (94-d). Besides, Fitzpatrick's example is infelicitous with a disjunction (94-e). It could only work with embedded "*who might/must have cheated*". However, assuming that $who_{\langle et,t\rangle}$ is only licensed by modals might be *ad hoc*.

(94) a. A: Who cheated?
b. B: Anne or Charlie cheated.
c. B: Two students cheated.
d. A: Ok, but who?
e. #John and Mary believe that Anne or Charlie cheated, and so they agree on who cheated.

Another problem concerns anaphoric relations. Non-presuppositional questions require modal subordination to refer to their wh-words. However, the type-flexibility theory does not predict that $who_{\langle et,t\rangle}$ requires modal subordination. We would expect that $who_{\langle et,t\rangle}$ can only be co-indexed with generalized quantifiers. But section 5.3.1 shows this is not the case.

Finally, the two approaches differ in how they derive the existential presupposition. My proposal triggers it at the lexical level and aligns with the Presuppositional Answer theory established by Uegaki 2021. On the contrary, their proposal triggers it at the clause level. Hirsch and Schwarz 2019 show that global triggering fails to produce correct predictions when a modal operator scopes over the wh-word.

To sum it up, the type-shift hypothesis is an alternative theory of EIs in simplex wh-questions that can be unified with mine or a promising competitor. Elliott, Nicolae, and Sauerland's proposal is less explanatory in view of the data presented here. Nevertheless, further investigation is required to settle this issue and overcome the flaws of both theories.

Property	**Specific *who*$_S$**	**Generic *who*$_G$**
Referent Identification (§2.4.1)	specific	non-specific
Presuppositional (§2.2.6,2.4.1)	yes (existential)	no
Open to null answers (§2.2.2,2.4.1)	no	yes
Exhaustivity (§2.4.1)	MA	MA or MS
Disambiguating constructions (§2.3.1,2.2.4)	narrow focus, Fr. *c'est* cleft	weak NPIs, Fr. *il y a* cleft
Assoc. epistemic items (§2.4.1)	*certain*	*quelconque, qualunque*
Example related (§2.3.2) non-interrogative wh-items	definite Fr. FR *qui*, D. indefinite *wat*	free-choice Fr. FR *qui*, Ch. NPI *nǎ*-CL
Domain (§2.4.1)	narrow	wide (contains $\star$)
Anaphoric relation (§2.3.3)	indicative	modal subordination
Favoring environments (§2.2.4)	episodic	habitual, future, generics
Sensitive to interveners (§2.2.5,2.5.3)	no	yes (weak islands at least)
Can be focused (§2.3.1,2.5.1)	yes	no
Trace licensing weak NPIs (§2.5.1)	no	yes
Associated question types (§2.5.2)	exam, echo q.	rhetorical q.
Pressure to answer (§2.5.2)	high	low

Table 2.2 – Summary list of the properties of the two readings of *who*.

2.8 Conclusion

The existential inference (EI) of simplex wh-questions (*who*, *what*) varies in both strength and availability–a variation that has not been satisfactorily explained so far. In this chapter, I proposed a theory that accounts for this phenomenon. Table 2.2 summarizes the core claims.

The EI is generally understood as the disjunction of all positive answers of a wh-question, with the null (or negative) answer construed as its negation. I argued that this characterization is inadequate. First, null answers should be treated as valid answers, both on pragmatic and truth-conditional grounds. Second, data from quantified and modalized questions support analyzing null answers as instantiations of the null individual–that is, the empty set of individuals.

Unlike the existential presupposition of *which+sg*, the EI of *who* can be suspended or entirely absent. I provided new data suggesting that the EI is weaker or even excluded from semantic environments that license free-choice items. I presented evidence that negative polarity items (NPIs) disambiguate a question towards a reading compatible with null answers. In addition, weak islands show

64. The same remark holds for Alonso-Ovalle and Rouillard 2023. Their domain of generalized quantifiers $\mathcal{G}^{\wedge\vee}(P)$ does not contain the empty generalized quantifier.

sensitivity to this same variation, suggesting that a purely pragmatics account is insufficient.

Through controlled discourse contexts, I demonstrated that simplex wh-questions allow two distinct readings: specific and generic. The specific reading carries an existential presupposition; the generic does not. This ambiguity accounts for the observed instability and variability in the strength of EIs.

I further supported this theory by showing that certain languages overtly distinguish between these readings via focus. I also explored similar distinctions in non-interrogative wh-items and examined the anaphoric behavior of interrogative expressions. This cross-linguistic and cross-categorial evidence suggests that the ambiguity originates with the wh-word itself: there is a specific *who* (who_S) and a generic *who* (who_G).

The specific who_S functions like the epistemic determiner *certain*: it introduces a transparent, identifiable referent, triggers an existential presupposition, and can take part in indicative anaphora. It can be focused, is immune to intervention effects, and reflects a strong expectation for an answer.

In contrast, who_G resembles epistemic free-choice items such as *quelconque* (French) or *qualunque* (Italian): it introduces a hypothetical, non-identified referent, requires modal subordination, does not trigger an EI, and permits null answers. This reading cannot be focused, is sensitive to interveners, but crucially, licenses weak NPIs. It also carries less social pressure to respond.

By extending the independently motivated specific/generic distinction in noun phrases (von Heusinger 2002) to interrogative expressions, this theory offers an explanatory account of EI variation in simplex wh-questions. Moreover, it opens the door to new insights into phenomena related to questions and scope effects, including weak islands and NPI licensing.

Part Two

Dynamic Aspects of Questions

Chapter 3

Dynamic Inquisitive Semantics

The goal of this section is to introduce Dynamic Inquisitive Semantics (DInq). This formal model constitutes the basis of the model developed in Chapter 4 and used in Chapters 5 and 6. Section 3.1 briefly summarizes the core notions of Inquisitive Semantics (Inq). The two other sections present DInq, its semantic space (section 3.2) and the syntax-semantic interface (section 3.3).

3.1 Inquisitive semantics

3.1.1 Introduction

Inquisitive Semantics (Inq) (Ciardelli, Groenendijk, and Roelofsen 2018) is a formal theory of natural language semantics. The main idea is to represent the denotation of declaratives and interrogatives by the same semantic object: an issue. Instead of truth values, issues have resolution conditions.

There are many reasons to consider an integrated notion of semantic content. [1] Two important assets of Inq are crucial in this dissertation. The first asset is the ability to define a single semantic entry for operators acting on both declaratives and interrogatives. For example, Chapter 4 extends the Kratzerian semantics of modals to any issue. As a result, the denotation of the possibility operator $\Diamond$ works on both assertions and questions.

The second asset is the central role of disjunction in Inq. In this theory, disjunction is one of the sources of inquisitiveness. In a declarative clause, sentence-level disjunction may trigger alternatives. In Chapter 6, this prediction is observed in French embedded complementizer phrases. For example, in (1), the clause A is declarative, but analyzing its denotation as a proposition fails to capture the semantic behavior of this SELON ('*depending on*') construction. Inq

1. I refer to Ciardelli, Groenendijk, and Roelofsen 2018, § 1.1 for a list of these motivations.

allows us to represent the denotation of A identically to the denotation of an alternative question.

(1) Cet effort n' est pas le même selon [$_A$ qu' on est héritier ou que
this effort NEG is not the same SELON [$_A$ that one is heir or that
l'on a que sa force de travail].
one has only one's force of work].
'*This effort is not the same depending on whether you are an heir or only have your labor power.*'

Section 3.1.2 presents issues and basic operations on these objects. A syntax-semantic interface is given in section 3.1.3.

3.1.2 Semantic space

Inquisitive Semantics (InqB) is based on intensional semantics. We assume a non-empty set W of possible worlds (type s) and a non-empty set D of plural (non-null) individuals (type e). Truth values have type t. In InqB, an information state is a set of possible worlds (type $s \to t$). Information states (or just *state* henceforth) are ordered by set inclusion. The trivial state is W–the top element of the lattice–and the inconsistent state is $\emptyset$–the bottom element.

An issue $\mathcal{I}$ is a non-empty downward-closed set of information states. The inconsistent issue is thus $\{\emptyset\}$. A state s resolves an issue $\mathcal{I}$ just in case $s \in \mathcal{I}$.

Issues encapsulate more content than propositions. They have an informative part and an inquisitive part. The informative content of an issue $\mathcal{I}$ is the state $\bigcup\mathcal{I}$. Assertions only contain informative content. Therefore, an assertive issue is defined as an issue $\mathcal{I}$ such that $\bigcup\mathcal{I} \in \mathcal{I}$. Otherwise, the issue is referred to as inquisitive.

Inquisitiveness is the property of raising alternatives. An alternative s of an issue $\mathcal{I}$ is a state $s \in \mathcal{I}$ such that s is maximal in $\mathcal{I}$. The set of alternatives of $\mathcal{I}$ is written $\text{ALT}(\mathcal{I})$. An assertive issue $\mathcal{I}$ contains only one alternative: $\bigcup\mathcal{I}$. An issue containing finitely many elements is inquisitive if and only if $\text{ALT}(\mathcal{I})$ contains two alternatives or more. These alternatives correspond to the minimally informative answers to the question represented by $\mathcal{I}$. In this dissertation, we only illustrate cases with a finite set of worlds.

Visual representation

An issue can be represented by a diagram, like in Fig. 3.1. Here, we assume two individuals, Mary m and John j, and a universe W containing four worlds: w_x is the world where the individual(s) x cheated. A region of space surrounding worlds represents the information state composed of these worlds. An issue is illustrated by its alternatives. For example, assertion "*Mary cheated*" only has



Mary cheated. *Did Mary cheat?* *Who cheated? (mention-some)* *Who cheated? (mention-all)*

Figure 3.1 – Example sentences and their denotation in inquisitive semantics.

one alternative: $\{w_{mj}, w_m\}$. Polar question "*Did Mary cheat?*" has two alternatives, corresponding to "*Mary cheated*" and "*Mary didn't cheat*". Alternatives can overlap, like the ones representing "*Mary cheated*" and "*John cheated*" in the mention-some reading of "*Who cheated?*". The alternatives of a mention-all reading form a partition of W.

3.1.3 Inquisitive logic

Inquisitive semantics (Inq) offers a uniform definition of entailment across assertive and inquisitive issues. An issue $\mathcal{I}$ entails an issue $\mathcal{J}$, written $\mathcal{I} \models \mathcal{J}$, if $\mathcal{I} \subseteq \mathcal{J}$. In terms of resolution, this can be paraphrased as follows: $\mathcal{I}$ entails $\mathcal{J}$ if and only if every state resolving $\mathcal{I}$ also resolves $\mathcal{J}$.

This definition extends the usual entailment on propositions in a conservative way (de Groote and Richard 2021). We can construct an embedding from intensional semantics to inquisitive semantics, mapping propositions to assertive issues while preserving entailment. Furthermore, Inq entailment (partially) captures the answerhood relationship. An assertive issue $\mathcal{I}$ is an answer to an inquisitive issue $\mathcal{J}$ just in case $\mathcal{I} \models \mathcal{J}$.[2] For example, "*Mary cheated.*" $\models$ "*Did Mary cheat?*".

In intensional semantics (Int), a model $\mathcal{M}$ contains an interpretation function I mapping each lexical predicate to a relation. The inquisitivation embedding defined in de Groote and Richard 2021 transforms an intensional model $\mathcal{M}$ into an Inq-model, keeping the interpretation function I. An example is given in (3.1), using the powerset operator $\wp$.

$$\begin{aligned} [\![\mathbf{cheat}]\!]_{\mathsf{Int}} &:= \lambda x^e.\ \{w \in W \mid I(\mathbf{cheat})(w)(x)\} \\ [\![\mathbf{cheat}]\!]_{\mathsf{Inq}} &:= \lambda x.\ \wp(\{w \in W \mid I(\mathbf{cheat})(w)(x)\}) \end{aligned} \tag{3.1}$$

The lexical entry for proper names in Inq is the same as in intensional semantics.

2. The answerhood relationship includes many other important aspects (e.g. granularity, identification method, subject matter,...) that will be ignored in this dissertation.

$$[\![\mathbf{mary}]\!]_{\mathsf{Inq}} := \ I(\mathbf{mary}) \in D \tag{3.2}$$

Like in intensional semantics, the conjunction and the disjunction are interpreted by set intersection and set union, respectively.

$$\begin{array}{rl} [\![\wedge]\!]_{\mathsf{Inq}} := & \lambda\mathcal{I},\mathcal{J}.\ \mathcal{I}\cap\mathcal{J} \\ [\![\vee]\!]_{\mathsf{Inq}} := & \lambda\mathcal{I},\mathcal{J}.\ \mathcal{I}\cup\mathcal{J} \\ [\![\bot]\!]_{\mathsf{Inq}} := & \{\emptyset\} \\ [\![\top]\!]_{\mathsf{Inq}} := & \wp(W) \end{array} \tag{3.3}$$

However, Inq disjunction has a very different behavior: it can generate alternatives. For example: [3]

$$[\![\text{Mary cheated}]\!]_{\mathsf{Inq}} \cup [\![\text{John cheated}]\!]_{\mathsf{Inq}} = [\![\text{Did Mary cheat or did John cheat?}]\!]_{\mathsf{Inq}} \tag{3.4}$$

Alternatives can be erased by using the ! operator defined in (3.5). This operator projects an issue onto the subspace of assertive issues. For instance, formula $!(\varphi \vee \psi)$ recovers the intensional disjunction between formulas φ and ψ. Negation also erases alternatives. The conditional $\rightarrow$ computes the pseudo-complementation of $\mathcal{I}$ with respect to $\mathcal{J}$.

$$\begin{array}{rl} [\![\neg]\!]_{\mathsf{Inq}} := & \lambda\mathcal{I}.\ \wp(W \smallsetminus \bigcup\mathcal{I}) \\ [\![\rightarrow]\!]_{\mathsf{Inq}} := & \lambda\mathcal{I},\mathcal{J}.\ \{s \subseteq W \mid \forall t \subseteq s.\ t \in \mathcal{I} \rightarrow t \in \mathcal{J}\} \\ [\![!]\!]_{\mathsf{Inq}} := & \lambda\mathcal{I}.\ \wp(\bigcup\mathcal{I}) \end{array} \tag{3.5}$$

Polar questions are derived using the ? operator, disjoining an issue with its negation, viz. $?\varphi := \varphi \vee \neg\varphi$. In Inq, wh-words are regarded as indefinites. Indefinites contribute domain alternatives, generating as many alternatives as individuals in their domains, akin to a big disjunction. We do not expand on wh-words because their treatment in Inq is different from their treatment in Dynamic Inquisitive Semantics, and only the latter will be relevant in this thesis.

3.2 Dynamic inquisitive objects

3.2.1 Introduction

Dynamic Inquisitive Semantics (DInq) is an extension of Inquisitive semantics aimed to capture the discourse effects of both assertions and questions uniformly. Utterances are interpreted by update functions, i.e. functions from context to context. DInq models the raising and resolution of issues. This theory describes

3. The implicature of question "*Did Mary cheat or did John cheat?*" that only one of them cheated is ignored here.

how uttering multiple questions adds them up and how assertions may resolve one or several of them. It can also model presuppositions as definedness conditions.

DInq also models anaphora. Discourse referents (or *dref* in the following) can be introduced both in assertions and questions and be referred to in an utterance of either type. An integrated semantics can thus better account for these interpretative dependencies.

Van Rooij 1997 and Groenendijk 1998 proposed that wh-words raise discourse referents. DInq pursues this analysis and pushes it further, taking number marking and exhaustivity into account. Questions require the identification of an individual, i.e. providing the actual value taken by the dref raised by the wh-word. [4] The availability of this wh-dref helps compute the reading of questions by restricting its possible values. In *which+sg* questions, only atomic individuals are in the wh-domain, contrary to questions with *who*, *what* or *which+pl*. In mention-some question "*Who cheated?*", the alternative "*d cheated*" contains worlds where not only *d* cheated. Conversely, in its mention-all reading, the alternatives corresponds to the proposition "*Only d cheated*". Applying operators restricting the values taken by the wh-dref allows us to easily produce these readings.

One primary characteristic of DInq is the decompositional approach used to capture these properties. They are derived by the interaction between different modular operators ($\mathbf{atom}, \mathbf{max}, ?u, \dagger, ...$) contributed by different components and motivated by independent reasons.

DInq combines Inq and the dynamic semantics GSV developed by Groenendijk, Stokhof, and Veltman 1996. It was first presented by Dotlačil and Roelofsen 2019 with atomic individuals only and within a formal language in the style of Groenendijk, Stokhof, and Veltman 1996. Dotlačil and Roelofsen 2021 extended DInq to plural individuals and plural information states, based on Brasoveanu 2007a. A complete description of DInq can be found in Roelofsen and Dotlačil 2023b. [5] In this dissertation, I use a variant of the version by Roelofsen and Dotlačil 2023b: it includes plural individuals but singular information states. [6] The rest of this chapter presents the key elements of DInq needed to understand the model I develop in Chapter 4.

4. Questions about a proposition or a cause (e.g. "*What did you do?*", "*Why did you do that?*") are left aside here.

5. See also Roelofsen and Dotlačil 2023a for a discussion on challenges faced by DInq and potential improvements.

6. Plural information states are necessary to model multiple wh-questions and the difference between weak and strong donkey anaphora. Given that these phenomena are not the focus of this thesis, I opt for a simpler system. This choice benefits ease of comprehension and definition statements.

3.2.2 Contexts

Dynamic inquisitive semantics (DInq) refines Inq by including assignment functions. We assume a non-empty (countably infinite) set of discourse referents Ref. The elements of Ref–typically $u, v, ...$–are constants of a special type r. An assignment function g is a partial function from Ref to D. We write $\mathbf{rdom}\, g$ the definition domain of g.

A possibility is a pair $\langle w, g\rangle$ of a world and an assignment function. An information state is a set of possibilities. The referent domain of a state s is $\mathbf{rdom}\, s = \bigcup_{\langle w,g\rangle \in s} \mathbf{rdom}\, g$. The world component of a state s is $\mathsf{WC}(s) = \{w \mid \exists \langle w, g\rangle \in s\}$. Being in a state s containing several worlds can be understood as ignorance: the worlds in s are undistinguishable, and all equal candidates to be the actual world. Similarly, a state s containing several possibilities $\langle w, g\rangle$ with the same world w can intuitively be understood as ignorance: the values taken by a dref $u \in \mathbf{rdom}\, s$ are all equal candidates to be the actual value of u.[7]

A context is a non-empty set of states which is downward-closed for set inclusion. It has type $k = ((s \times (r \to e)) \to t) \to t$. Similarly to issues, we can define the informative content of a context c as $\bigcup c$ and its alternatives $\textsc{alt}(c)$ as its maximal states for set inclusion. A context is assertive if $\bigcup c \in c$ and inquisitive otherwise.

Visual representations

Like Inq diagrams, we represent contexts by their alternatives. An information state is represented by a region surrounding possibilities. Possibilities are organized as a matrix: worlds are in abscissa and assignment functions in ordinates. We write $[u/d]$ the assignment function mapping dref u to individual d. Except if stated otherwise, when illustrating the denotation of a sentence, we usually represent the output of its denotation on the initial context $c_0 = \wp(\{\langle w, []\rangle \mid w \in W\})$.

Some sentences are illustrated in Fig. 3.2 with the same model as in Fig. 3.1. Note that, in the compositional theory presented in Roelofsen and Dotlačil 2023b, wh-questions have an existential presupposition by default.

In the question "*Did Mary cheat?*", no dref is introduced, so the alternatives contain possibilities with the empty assignment function []. The indefinite *someone*^{u} introduces the dref u. The information content of "*Someone*^{u} *cheated*" contains the possibilities $\langle w, g\rangle$ where $g(u)$ cheated in w. The mention-all reading of "*Who*^{u} *cheated?*" contains three alternatives. Each of them contains only the possibilities $\langle w, g\rangle$ where $g(u)$ is the maximal individual that cheated in w. Explanations for the illustration of the mention-some reading are given in section 3.3.1.

7. In this dissertation, we only encounter states s in which all assignment functions share the same definition domain. Cases where this does not hold might arise if negation is involved.

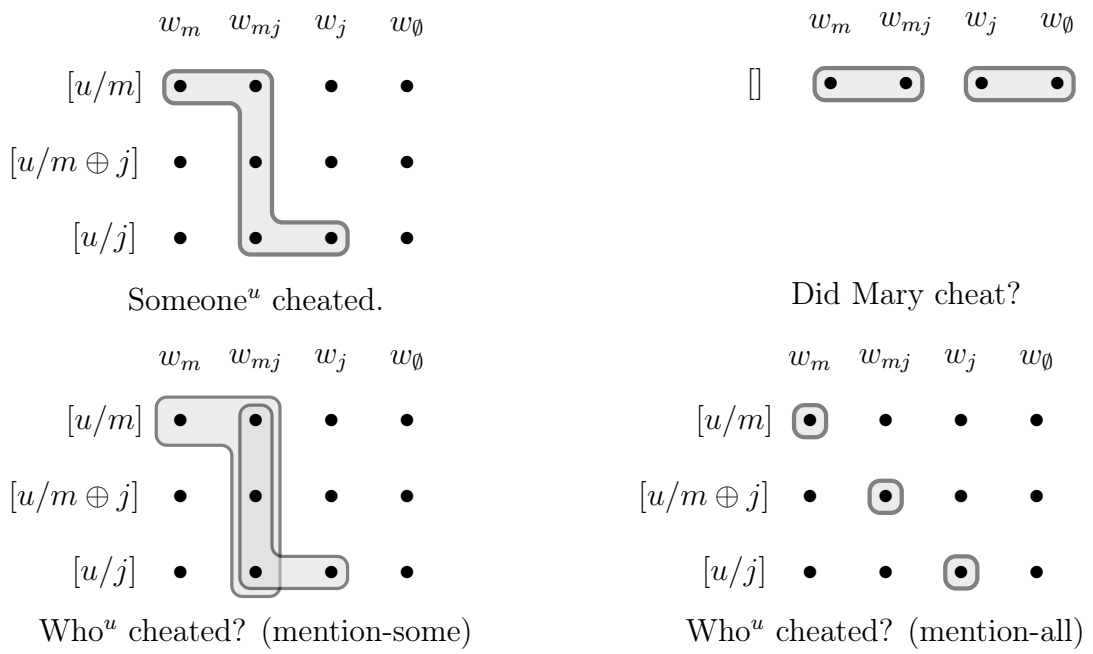


Figure 3.2 – Example sentences and their denotation in DInq.

3.2.3 Relations between contexts

In Inq, a state s is more informative than a state t if $s \subseteq t$. In DInq, set inclusion does not capture informativeness ordering because we also have to take drefs into account. This section presents the relations between contexts used to express the intuitive notions of "more informative than" and "equivalent to" including drefs.

Extension Extension models the relation "is more informative than". An assignment function g' is more informative than an assignment function g if g' is like g but potentially contains more drefs. In this case, we say that g' extends g, written $g \leqslant g'$. A possibility $\langle w', g' \rangle$ extends a possibility $\langle w, g \rangle$ if $w = w'$ and $g \leqslant g'$. For states, the idea is the following: a state s' extends a state s if s' was obtained by an update on s, potentially eliminating worlds and potentially adding drefs. We can define context extension similarly. The formal definitions are given in (3.6).

$$\begin{array}{rcll} g & \leqslant & g' & \text{if } \mathbf{rdom}\, g \subseteq \mathbf{rdom}\, g' \wedge (\forall z \in \mathbf{rdom}\, g.\ g(z) = g'(z)) \\ \langle w, g \rangle & \leqslant & \langle w', g' \rangle & \text{if } w = w' \wedge g \leqslant g' \\ s & \leqslant & s' & \text{if } \forall j \in s', \exists i \in s.\ i \leqslant j \\ c & \leqslant & c' & \text{if } \forall s' \in c', \exists s \in c.\ s \leqslant s' \end{array} \tag{3.6}$$

Subsistence Subsistence aims at extending the notion of "belonging to" or "being equivalent to" while taking dref introduction into account.

A possibility i subsists in a state s', written $i \Subset s'$, if a possibility j in s' extends i. Intuitively, this means that i is a possibility that survived an update with output s'. A state s subsists in a state s', written $s \sqsubseteq s'$, if s' extends s and all the possibilities of s subsist in s'. Intuitively, this means that s' is the result of an update on s where no world component was lost, but drefs might have been added. The formal definitions and the generalization to contexts are given in (3.7).

$$\begin{array}{lcll} i & \Subset & s' & \text{if} \quad \exists j \in s'.\ i \leqslant j \\ s & \sqsubseteq & s' & \text{if} \quad s \leqslant s' \wedge (\forall i \in s.\ i \Subset s') \\ s & \Subset & c' & \text{if} \quad \exists s' \in c'.\ s \sqsubseteq s' \\ c & \sqsubseteq & c' & \text{if} \quad c \leqslant c' \wedge (\forall s \in c.\ s \Subset c') \end{array} \tag{3.7}$$

We adopt a left-to-right notation of extension (e.g. $s \leqslant s'$) to emphasize that the right element of the relation might be the result of an update on the left element. In particular, the right element may contain more drefs. If $\Subset$ is viewed as a kind of "element" relation, then $\sqsubseteq$ could be seen as a kind of "inclusion" relation. However,contrary to the original definition of subsistence by Groenendijk, Stokhof, and Veltman 1996, $\sqsubseteq$-subsistence in DInq entails extension. Therefore, $\sqsubseteq$ should rather be seen as an adaptation of an equivalence relation between states. Though, $\sqsubseteq$ is not an equivalence relation because it is not anti-symmetric: if $s \sqsubseteq s'$, both states have the same world component but s' might contain more drefs. That's why we keep the $\Subset, \sqsubseteq$ notation designed by Dekker 1992 for Groenendijk, Stokhof, and Veltman 1996.

3.3 Syntax-semantics interface

We present here the logical language of DInq and interpretation rules for a fragment of English. The logical language is presented in section 3.3.1. Section 3.3.2 provides the denotation of some basic words. Section 3.3.3 broaches on the left periphery. Finally, a detailed example is given in section 3.3.4.

3.3.1 Logical language

Utterances are detonated by partial functions from context to context: type $T = k \to k$. The definition domain of a function $\mathcal{U}$ of type T is noted $\mathbf{ddom}\mathcal{U}$. If $c \notin \mathbf{ddom}\mathcal{U}$, the utterance interpreted by $\mathcal{U}$ is infelicitous in context c, e.g. due to a presupposition failure.

Following Roelofsen and Dotlačil 2023b, DInq formulas are presented directly as denotational λ-terms: the symbols ;, $\sqcup$, etc., are syntactic sugar.

Lexical predicates Lexical predicates take elements of type r (discourse referents) as arguments. Their entry can be derived from their Inq entry. To simplify matters, we mostly work with distributive lexical predicates. We assume these predicates are closed under mereological sum $\oplus$. The pluralization operator $*$ (Krifka 1989) implements this closure. Here is an example with *cheat*, where I is an interpretation function from a first-order model:

$$\textbf{cheat}\{u\} := \lambda c, s.\ s \in c \wedge \forall \langle w, g\rangle \in s.\ {}^*I(\textbf{cheat})(w)(g(u)) \tag{3.8}$$

The beginning sequence $\lambda c, s.\ s \in c \wedge \forall \langle w, g\rangle \in s$ is very common in DInq. It signifies that the function (i) extends the input context without adding new drefs ($s \in c$), and (ii) does not create inquisitiveness ($\forall \langle w, g\rangle \in s$, scoping over the rest of the formula).

Conjunction and disjunction As expected, conjunction (;) is interpreted as function composition. Similarly to Inq, disjunction ($\sqcup$) is taken as context union and the ? operator for polar questions can be defined as $?\mathcal{U} := \mathcal{U} \sqcup \neg\mathcal{U}$.

$$\begin{aligned} \mathcal{U};\mathcal{V} &:= \lambda c.\ \mathcal{V}(\mathcal{U}(c)) \\ \mathcal{U} \sqcup \mathcal{V} &:= \lambda c.\ \mathcal{U}(c) \cup \mathcal{V}(c) \\ \top &:= \lambda c.\ c \\ \bot &:= \lambda c.\{\emptyset\} \end{aligned} \tag{3.9}$$

Negation, conditional and removing inquisitiveness In DInq, negation ($\neg$) does not only remove inquisitiveness but also blocks discourse referents introduced in its scope. The inquisitiveness-elimination operator ! only removes inquisitiveness. However, both operators propagate inquisitiveness located in the input context. In other words, they do not remove previously asked questions. The conditional $\mathcal{U} \twoheadrightarrow \mathcal{V}$ is similar to the Inq conditional, but considering that updates might introduce drefs. Formula $(\mathcal{U} \twoheadrightarrow \mathcal{V})(c)$ outputs the states $s \in c$ such that all possible descendants t' of any substate $t \subseteq s$ after the $\mathcal{U}$-update subsist after the $\mathcal{V}$-update.

$$\begin{aligned} \neg\mathcal{U} &:= \lambda c, s.\ s \in c \wedge \forall t \subseteq s.\ t \sqsubseteq \mathcal{U}(c) \rightarrow t = \emptyset \\ \mathcal{U} \twoheadrightarrow \mathcal{V} &:= \lambda c, s.\ s \in c \wedge \forall t \subseteq s.\ \forall t' \in \mathcal{U}(c).\ t \sqsubseteq t' \rightarrow t' \sqsubseteq \mathcal{V}(\mathcal{U}(c)) \\ !\mathcal{U} &:= \lambda c, s.\ s \subseteq \bigcup \mathcal{U}(c) \wedge \exists t \in c.\ t \leqslant s \end{aligned} \tag{3.10}$$

Raising and restricting discourse referents Introducing a discourse referent u involves adding u to the reference domain of all assignment functions in the input and assigning it all possible values. The definition (3.11) is close to the definition of subsistence, except that the input and the output assignment functions (resp. states) only differ with respect to u.

$$\begin{array}{rl} g[u]g' \text{ if } & \mathbf{rdom}\, g' = \mathbf{rdom}\, g \cup \{u\} \wedge (\forall z \neq u.\ z \in \mathbf{rdom}\, g \rightarrow g(z) = g'(z)) \\ s[u]s' \text{ if } & (\forall \langle w, g\rangle \in s.\ \exists \langle w', g'\rangle \in s'.\ w = w' \wedge g[u]g') \wedge \\ & \quad (\forall \langle w', g'\rangle \in s'.\ \exists \langle w, g\rangle \in s.\ w = w' \wedge g[u]g') \\ [u] := & \lambda c, s.\ \exists t \in c.\ t[u]s \end{array} \tag{3.11}$$

The $\mathbf{atom}\{u\}$ operator restricts the values taken by u to atoms. The $\mathbf{max}\{u\}$ operator is used to derive mention-all readings. This operator eliminates the possibilities $\langle w, g\rangle$ where $g(u)$ is not maximal compared to the other values taken by u among possibilities with world w. Mereological parthood is written $\leq$, and strict parthood is written $<$.

$$\begin{array}{rl} \mathbf{atom}\{u\} := & \lambda c, s.\ s \in c \wedge \forall \langle w, g\rangle \in s.\ \neg(\exists y.\ y < g(u)) \\ \mathbf{max}\{u\} := & \lambda c, s.\ s \in c \wedge \forall \langle w, g\rangle \in s.\ \forall \langle w', g'\rangle \in \bigcup c.\ w = w' \rightarrow g'(u) \leq g(u) \end{array} \tag{3.12}$$

Requesting a witness The witness request operator $?u$ divides the input context into states that somehow agree on a value taken by u. A state s belongs to $?u(c)$ if, for a given individual d and all possibilities $\langle w, g\rangle$ in s, the world w is such that $g'(u) = d$ for some $\langle w, g'\rangle \in \bigcup c$.

$$?u := \lambda c, s.\ s \in c \wedge \exists x.\ \forall \langle w, g\rangle \in s.\ \exists \langle w', g'\rangle \in \bigcup c.\ w = w' \wedge g'(u) = x \tag{3.13}$$

Note that this definition looks more complex than expected. For a given $s \in c$ and individual d, s does not simply belong to $?u(c)$ when all possibilities $\langle w, g\rangle$ agree on $g(u) = x$. Instead, we require something weaker. For s to belong to $?u(c)$, it is sufficient that $g'(u) = x$ for some possibility $\langle w, g'\rangle$ in $\bigcup c$. In other words, whenever an output state s of $?u$ includes a possibility $\langle w, g\rangle$, then it also includes all possibilities with the same world component. The output on $?u$ does not discriminate on the value given to u but on the world component.

To understand this definition choice, remark that $?u$ derives a mention-some reading by default. In mention-some answers, the instantiations are not in a co-reference chain with the wh-word. Therefore, we have to rely on world information to resolve the question, and not on the dref. Let me illustrate this issue with the question "*Who cheated?*". When a mention-some answer like "*Mary cheated*" is given in the world w_{mj}–where John also cheated–the issue is considered resolved. However we cannot co-index the word *Mary* with u because Mary is not the maximal plural individual who cheated in that world.[8] Therefore, "*Who cheated?*" must be resolved using world information only. This leads to the definition in equation (3.13).

8. In that case, *Mary* is co-indexed with a different discourse referent, e.g. v.

Entailment Like in Groenendijk, Stokhof, and Veltman 1996, entailment is defined as Strawson entailment: the relation $\mathcal{U} \models \mathcal{V}$ is only checked in contexts where $\mathcal{U}$ and $\mathcal{U};\mathcal{V}$ are felicitous. In these cases, entailment consists of checking that uttering $\mathcal{V}$ after $\mathcal{U}$ does not eliminate worlds–but it may add new drefs.

$$\mathcal{U} \models \mathcal{V} \text{ if } \forall c \in \mathbf{ddom}(\mathcal{U};\mathcal{V}).\ \mathcal{U}(c) \sqsubseteq \mathcal{V}(\mathcal{U}(c)) \tag{3.14}$$

3.3.2 A fragment of English

Verbs and nouns are interpreted as lexical predicates from the logical language, e.g. **student** for *student(s)* and **cheat** for *cheat(ed)*. Singular nouns also contribute the **atom** operator. Proper nouns are assumed to be co-indexed by a discourse referent.

$$\begin{aligned} [\![\textit{cheat(ed)}]\!] &:= \lambda z.\ \mathbf{cheat}\{z\} \\ [\![\textit{student}]\!] &:= \lambda z.\ \mathbf{atom}\{z\};\mathbf{student}\{z\} \\ [\![\textit{Mary}_u]\!]_{:=} &:= \mathbf{name}\{u, \textit{“Mary”}\} \end{aligned} \tag{3.15}$$

Conjunction, disjunction, and negation are interpreted by logical operations.

$$\begin{aligned} [\![\textit{and}]\!] &:= \lambda\mathcal{U},\mathcal{V}.\ \mathcal{U};\mathcal{V} \\ [\![\textit{or}]\!] &:= \lambda\mathcal{U},\mathcal{V}.\ \mathcal{U} \sqcup \mathcal{V} \\ [\![\textit{no/not}]\!] &:= \lambda\mathcal{U}.\ \neg\mathcal{U} \end{aligned} \tag{3.16}$$

Indefinites introduce a discourse referent u. Wh-words behave the same. The only difference lies in the movement undergone by wh-words in the left periphery. To derive a mention-some reading, the mention-some version (MS) of a simplex wh-word is used, while to derive a mention-all reading, the mention-all version (MA) is used, requesting maximal witnesses. These readings were proposed by Roelofsen and Dotlačil 2023b and are independent from the specific vs. generic readings presented in Chapter 2, although a specific mention-some reading was claimed to be impossible. Singular *which* is assumed to have a mention-all version exclusively.

$$\begin{aligned} [\![a^u]\!] &:= \lambda R, S.\ [u]; R(u); S(u) \\ [\![\textit{who}^u_{\text{MS}}]\!] &:= \lambda S.\ [u]; \mathbf{person}\{u\}; S(u) \\ [\![\textit{who}^u_{\text{MA}}]\!] &:= \lambda S.\ [u]; \mathbf{person}\{u\}; S(u); \mathbf{max}\{u\} \\ [\![\textit{which}^u_{\text{sg}}]\!] &:= \lambda R, S.\ [u]; \mathbf{atom}\{u\}; R(u); S(u); \mathbf{max}\{u\} \end{aligned} \tag{3.17}$$

3.3.3 The left periphery

We adopt a syntactic theory based on the transformational tradition (Chomsky 1993). The syntactic representation of a matrix sentence is divided into a Finite Phrase (FinP) and a left periphery. According to Rizzi 1997 (and subsequent works), the left periphery includes several functional heads, which have

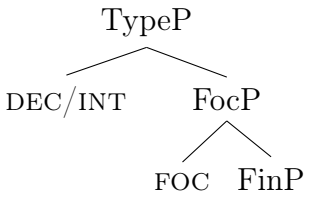


Figure 3.3 – Deep syntactic layers and functional heads in the left periphery

distinct syntactic and semantic effects. In this dissertation, we assume a Focus Phrase (FocP) and a Type Phrase (TypeP). A summary of the left periphery is given in Fig. 3.3.

The FOC_u head of the Focus Phrase contributes the witness request operator. Moreover, any inquisitiveness triggered under the Focus Phrase is erased. In questions, the wh-phrase is generated *in situ* and moves to the specifier of the Focus Phrase–its surface position in English–to co-index FOC with its discourse referent u. Before interpretation, the wh-phrase is reconstructed through an Across-the-Board movement. This way, the wh-phrase can be interpreted *in situ*. Reconstruction is required to explain the ban on weak crossover (Nishigauchi 1990). Despite the fronted surface position of wh-words in English, their dref cannot co-index reflexive pronouns that are not in the scope of their trace, as illustrated by (2).

(2) *Who^u [does his$_u$ mother [admire t_u]]?

$$\begin{aligned} [\![\text{FOC}_u]\!] &:= \lambda\mathcal{U}.\ !\mathcal{U}; ?u \\ [\![\text{DEC}]\!] &:= \lambda\mathcal{U}.\ !\mathcal{U} \\ [\![\text{INT}]\!] &:= \lambda\mathcal{U}.\ \dagger\langle ? \rangle\mathcal{U} \end{aligned} \tag{3.18}$$

The head of the Type Phrase is a syntactic type marker: INT for interrogatives and DEC for declaratives. DEC ensures that the utterance is not inquisitive thanks to the inquisitiveness-elimination operator !. INT ensures that the utterance is inquisitive thanks to $\langle ? \rangle$ and non-informative thanks to $\dagger$. The presuppositional closure operator $\dagger$ maps informativeness to presupposition: $\dagger\mathcal{U}$ is undefined on contexts c where applying $\mathcal{U}$ eliminates worlds.

$$\begin{aligned} \dagger\mathcal{U} &:= \lambda c.\ \begin{cases} \mathcal{U}(c) & \text{if } \bigcup c \sqsubseteq \bigcup \mathcal{U}(c) \\ \text{undefined} & \text{otherwise} \end{cases} \\ \langle ? \rangle\mathcal{U} &:= \begin{cases} ?\mathcal{U} & \text{if } \mathcal{U} \text{ is not inquisitive yet} \\ \mathcal{U} & \text{if } \mathcal{U} \text{ is already inquisitive} \end{cases} \end{aligned} \tag{3.19}$$

3.3.4 Example

To exemplify DInq, consider sentence (3-a). From its logical form (before Reconstruction) in Fig. 3.4, we derive its denotation (3-b). The operators $\langle ? \rangle$

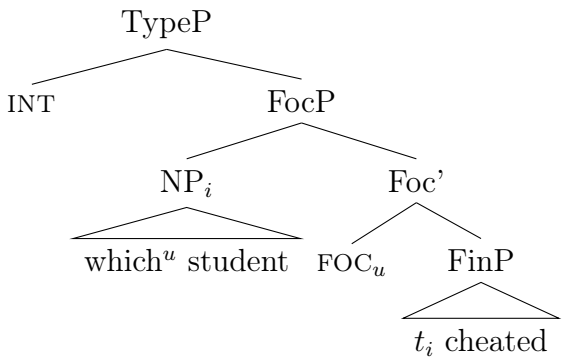


Figure 3.4 – Logical form of (6-a).

and ! are here vacuous. A step-by-step computation of formula $\mathcal{U}; ?u$ on the initial context c_0 is given in Fig. 3.5.

(3) a. Which^u student cheated?
b. $\dagger\langle ?\rangle(!\mathcal{U}; ?u)$, with $\mathcal{U} = [u]; \mathbf{atom}\{u\}; \mathbf{student}\{u\}; \mathbf{cheat}\{u\}; \mathbf{max}\{u\}$

Update function $\mathcal{U}; ?u$ is informative on c_0: worlds w_{ab} and $w_\emptyset$ are not part of the output. Consequently, the denotation (3-b) of question (3-a) is undefined on c_0 because of the $\dagger$ operator. This models the uniqueness and existential presuppositions of question (3-a). However, this denotation is defined on context $c_1 = \wp(\{\langle w_a, []\rangle, \langle w_b, []\rangle\})$ and produces the same context depicted in the bottom right-hand corner of Fig. 3.5, viz. $\{\{\langle w_a, [u/a]\rangle\}, \{\langle w_b, [u/b]\rangle\}, \emptyset\}$.

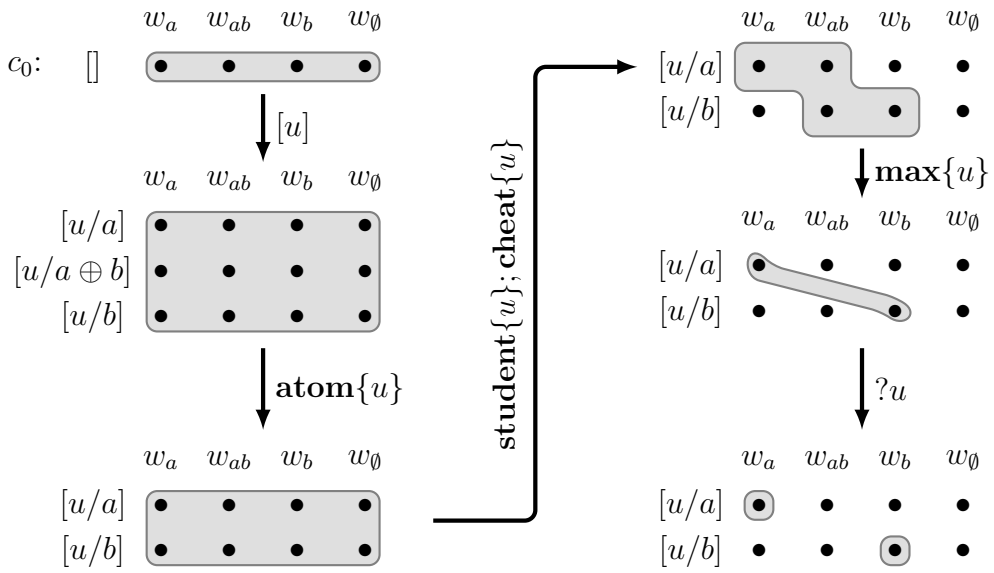


Figure 3.5 – Step-by-step semantic actions of the FocP of "*Which*^{u} *student cheated?*" on c_0.

Chapter 4

Dynamics of Modalized Questions

4.1 Modalized questions

4.1.1 Possibility weakens questions

Singular *which* questions have a uniqueness presupposition (Higginbotham and May 1981; Dayal 1996). For example, (1-b) carries the presupposition that a single letter is hidden in the actual world. Either the true answer is *R* (for FORM), or it is *A* (for FOAM).

(1) a. SITUATION: *Alice and Bob are playing a game. Alice wrote a mystery English word of 4 or 5 letters on the board, but her hand hides a part of it. The skeleton is FO__M.*
b. ALICE: Which letter is hidden here?
c. ALICE: Which letter could be hidden here?

Hirsch and Schwarz 2019 observed that this presupposition is weakened with an existential modality operator. Under a *which* $>$ $\Diamond$ reading, modalized question (1-c) could be interpreted as "*Which is the letter that could be hidden*". Such a specific reading presupposes that there exists a unique letter that could be hidden. However, (1-c) is intuitively not as strong as that. Alice does not have a specific letter in mind that she wants Bob to guess. She rather wants to test him about his knowledge of English.

Nevertheless, (1-c) still presupposes something, which could be paraphrased as: "*It might be that a single letter is hidden*".[1] Moreover, for each possible answer, one and only one letter is hidden. Roelofsen and Dotlačil 2023b call this a local presupposition.

1. Kobayashi and Rouillard 2021 claim that this presupposition is too weak. They suggest a presupposition akin to "*Each possible hidden letter group is a one-letter group*". I agree with them that mention-all existentially modalized questions have this presupposition. However, I maintain that mention-some ones, like (1-c), do not have this stronger presupposition.

Furthermore, as opposed to (1-b), where only one true answer exists in every world, several true answers are possible for (1-c) in the actual world. We could truly answer (1-c) by saying *A*, *R*, or *A and/or R* with a free-choice inference (i.e. *A* could be hidden and *R* could be hidden).[2] Ultimately, question (1-c) requires some information about the content of the English lexicon, not about the actual mystery word. In particular, the hidden letters might actually be *RU* (for FORUM). But (1-c) presupposes that the English lexicon does not *only* have *RU* as possible completion.

Plural *which* questions are strongly exhaustive (aka. mention-all) (Groenendijk and Stokhof 1984). For example, in (2-b), Mary requires the whole set of eczema patients.

(2) a. SITUATION: *Mary would like to experiment with two or three patients with eczema. She asks Ann, who knows the medical files of all patients with skin conditions in the hospital, a question.*
b. Which patients have eczema?
c. Which patients could I invite for my experiment?

However, in modalized question (2-c), Mary is open to a non-exhaustive set of eczema patients. Modalized singular and plural *which* questions allow mention-some readings. Xiang 2016 also observed this phenomenon, and Xiang and Cremers 2017 confirmed it through quantitative data.

We follow the analysis of Hirsch and Schwarz 2019 claiming that these phenomena are explained by the modality operator scoping over the *which* phrase and the question operator.

4.1.2 Modal subordination with modalized questions

In declarative sentences, a modal scoping over an indefinite blocks its anaphoric potential (Karttunen 1969). However, modally subordinated sentences have access to the embedded antecedent (Roberts 1989). We observe the same with modalized questions.

In question (3-a), *which*^{u} raises a discourse referent u under the scope of *might*. This referent cannot be accessed in the follow-up indicative sentence (3-b). However, in sentence (3-c), u can be referred to by it_u because of modal subordination.

(3) a. A: Which^u university might want to host the next ACL conference?
b. B: ?I don't know, but it$_u$ needs a lot of support.
c. B: I don't know, but it$_u$ would need a lot of support.

2. See Alonso-Ovalle and Rouillard 2023 for a derivation of the free-choice effect of disjunctive answers to modalized questions using higher-order quantification.

4.2 Previous works

Our goal is to define an existential modal operator in dynamic semantics that captures all these properties:

(i.) The modal operator must be externally static.

(ii.) But it must also allow modal subordination.

(iii.) The modal operator must obviate the global uniqueness presupposition while still providing a local one.

(iv.) Finally, it must be able to turn a mention-all question into a mention-some one.

Previous dynamic semantic models do not account for all four desired properties. The dynamic model of questions and modalities of Groenendijk 1998 and Groenendijk, Stokhof, and Veltman 1996 satisfies (i.) without allowing (ii.). Van Rooij 1998 manages to encode modal subordination (ii.), but the DPPL substrate (from van den Berg 1996) does not structurally provide external staticity (i.). Similarly, Brasoveanu 2010's account of modal subordination fails to model worlds in which individual discourse referents are *not* accessible.

4.2.1 Non-distributive test semantics for modals

An additional issue of previous dynamic semantics of modals concerns their non-distributive test interpretation. The definition advocated by Veltman 1996, and used by Dekker 1992, Groenendijk, Stokhof, and Veltman 1996, van Rooij 1998, S. Kaufmann 2000, Beaver 2001, and Asher and McCready 2007, considers that epistemic possibility $\Diamond\varphi$ is a test: if φ is true at some world $w \in s$, then $[\![\Diamond\varphi]\!](s) = s$, otherwise $[\![\Diamond\varphi]\!](s) = \emptyset$. This has several drawbacks.

First, it fails to model other modal flavors, like circumstantial or deontic modalities. Second, it fails to model knowledge variations because it assumes that all worlds are accessible from any world. Finally, it implies that a modalized declarative either brings no at-issue information or leads to a contradiction. As a consequence, Veltman 1996's existential modality cannot account for weakenings (iii.) and (iv.). Worse, extending this definition to inquisitive semantics would prevent an interrogative formula ψ from raising alternatives because $[\![\psi]\!]$ would either output the input context or the contradiction state.

We argue against modeling modals as tests, for both declaratives and interrogatives. Modalities can safely remove possible worlds from the global context. For example, sentence S in (4) does provide some new (non-contradictory) at-issue information about the content of the English lexicon. Suppose we are in an initial information state s which contains the actual world plus a world w_R^R (see Tab. 4.1) where, in English, the only word of the form FO__M is FORM. Then, updating s with $[\![S]\!]$ should remove w_R^R from s. With a test, on the contrary, updating s with $[\![S]\!]$ would not remove w_R^R. This is unwanted.

(4) A vowel can be inserted in FO__M to make an English word.

We can solve these problems by adopting a Kripkean accessibility semantics for modals (Kripke 1959), and more generally, Kratzerian theory of modals (Kratzer 1977, 1991). Our account is both eliminative and distributive. [3]

4.2.2 Stack-based treatment of local contexts

Implementing externally static modal operators prompts us to be able to retrieve modal local contexts somehow (Stalnaker 1981). Many authors resolve local context as anaphora, using world referents (Brasoveanu 2010; Hofmann 2019) or context referents (Frank 1997; Geurts 1999). Kibble 1998 and Asher and McCready 2007 claim that, unlike individual anaphora, modal local contexts are more restricted and have no real resolution ambiguity. Here, we follow their arguments and opt for a different system.

We adopt a stack-based semantics (S. Kaufmann 1997, 2000). Sentences are interpreted by functions updating stacks as macro-contexts. A stack τ contains local contexts. The bottom element c_0 corresponds to the common ground. Modal operators push on top of τ a new local context, which can contain more discourse referents than c_0. A subsequent modal sentence is evaluated in that topmost local context. Indicative sentences presuppose that the topmost element is c_0. We assume here that popping elements out of τ is performed by pragmatics (e.g. via discourse relations (Asher and Lascarides 2003)).

Isaacs and Rawlins 2008 use S. Kaufmann 2000's system to model modal subordination of conditional questions within partition semantics (Groenendijk and Stokhof 1984). Hara and Sano 2017 revise their model to provide better predictions for conditional questions, using inquisitive semantics (Ciardelli, Roelofsen, and Theiler 2017; Ciardelli, Groenendijk, and Roelofsen 2018). We extend their approach by (i.) adding referents in the model, by upgrading to dynamic inquisitive semantics DInq (Dotlačil and Roelofsen 2019, 2021); (ii.) defining the denotation of modals in DInq based on Kratzerian theory; and (iii.) specifying their presuppositions and actions on stacks.

The choice of DInq is motivated by another factor. Socolof, Hirsch, and Schwarz 2020 showed that exhaustivity and presupposition weakenings also happen with disjunction. DInq was proven to be able to model desiderata (iii.) and (iv.) for disjunction (Roelofsen and Dotlačil 2023b).

3. The epistemic behavior which motivated Veltman 1996 can be retrieved by requiring the accessibility relation to be an equivalence relation.

4.3 Modal dynamic inquisitive semantics

4.3.1 Inquisitive possibility

In Kratzerian theory, every world is equipped with a modal base (a set of classical propositions) and an ordering source. We only focus on the modal base f here. For each world $w \in W$, call $\mu(w) = \bigcap f(w)$ the modal set at w. The intensional denotation of $\Diamond$ is (4.1), with R_μ being the accessibility relation associated with μ.

$$\begin{array}{rl} [\![\Diamond\varphi]\!]_{\mathsf{Int}} = & \{w \in W \mid \mu(w) \cap [\![\varphi]\!]_{\mathsf{Int}} \neq \emptyset\} \quad = \mathsf{R}_\mu^{-1}([\![\varphi]\!]_{\mathsf{Int}}) \\ \text{where } \mathsf{R}_\mu^{-1}(s) := & \{w \in W \mid \mu(w) \cap s \neq \emptyset\} \end{array} \tag{4.1}$$

The inquisitive existential modality (4.3) proposed by Ciardelli 2016, § 7.1 is sensitive to the inquisitiveness of its complement but produces non-inquisitive issues. Yet, a modalized question is questioning. Therefore, we propose another definition (4.4). By inverting quantifiers $\forall$ and $\exists$, we let inquisitiveness project from φ to $\Diamond\varphi$. In this different modal inquisitive semantics (MInqB), an information state s resolves $\Diamond\varphi$ iff s resolves some $\Diamond\psi$, where ψ is an intensional formula interpreted by a proposition resolving φ.[4]

$$[\![\Diamond\varphi]\!]_{\mathsf{InqBK}} = \{s \subseteq W \mid \forall w \in s.\ \exists s' \in [\![\varphi]\!]_{\mathsf{InqBK}}.\ \mu(w) \cap s' \neq \emptyset\} \tag{4.3}$$

$$[\![\Diamond\varphi]\!]_{\mathsf{MInqB}} := \{s \subseteq W \mid \exists s' \in [\![\varphi]\!]_{\mathsf{MInqB}}.\ \forall w \in s.\ \mu(w) \cap s' \neq \emptyset\} \tag{4.4}$$

$$= \{s \subseteq W \mid \exists s' \in [\![\varphi]\!]_{\mathsf{MInqB}}.\ s \subseteq \mathsf{R}_\mu^{-1}(s')\} \tag{4.5}$$

To illustrate this definition consider example (5).

(5) a. SITUATION: *Mary supervised an exam for students a and b. Mary and John know that exactly one student cheated. John knows that Mary saw one or both of them suspiciously looking at their pencil case during the exam. This counts as a cheat suspicion. He asks her about that.*

b. JOHN: Which student may have cheated?

Let's model this situation with four worlds. The subscript x of world w_x^z is the real cheater: a or b. The superscript z is the set of students Mary saw suspiciously looking at their pencil case: a, b or ab (i.e. a and b). In Mary's epistemic state, world w_x^z is accessible from $w_{x'}^{z'}$ if and only if $z = z'$. The semantics of (5-b) in this model under MInqB is the issue on the right in Fig. 4.1.

4. With a non-epistemic modal base, formula 4.5 needs to be replaced by the definition of $\Diamond$ in modal dependence logic (Väänänen 2008; Kontinen et al. 2016) (4.2) to restrict accessibility correctly. Thanks to Maria Aloni for drawing my attention on this.

$$[\![\Diamond\varphi]\!]_{\mathsf{MDL}} = \{s \subseteq W \mid \exists s' \in [\![\varphi]\!]_{\mathsf{MDL}}.\ s \subseteq \mathsf{R}_\mu^{-1}(s') \wedge s' \subseteq \mathsf{R}_\mu(s)\} \tag{4.2}$$



Figure 4.1 – Illustration of the semantics of (5-b) in MInqB. Arrows represent the accessibility relation R_μ.

The alternatives of the prejacent φ are projected to $\Diamond\varphi$ by mediation of R_μ^{-1}. If the accessibility relation R_μ is an equivalence relation, $\mathsf{R}_\mu^{-1}(s)$ contains s. Hence, the output resolving proposition s is wider than the prejacent resolving proposition s'. This widening is what weakens the question. [5]

The MInqB denotation of $\Diamond$ is more general than Ciardelli 2016's one. We can retrieve InqBK's behavior by appending the ! operator on top of $\Diamond$ in MInqB, as shown by the following lemma.

4.3.1. Lemma. $[\![\Diamond]\!]_{\mathsf{InqBK}} = [\![!\Diamond]\!]_{\mathsf{MInqB}}$

Proof:
Let $\mathcal{I}$ be an issue, i.e. a downward-closed non-empty set of sets of possible worlds.

$\boxed{\subseteq}$ Take $s \in [\![\Diamond]\!]_{\mathsf{InqBK}}(\mathcal{I})$. For each $w \in s$, call $s'_w \in \mathcal{I}$ a set of worlds such that s'_w satisfies $\mu(w) \cap s'_w \neq \emptyset$. Such a set exists thanks to definition 4.3. By definition 4.4, the existence of s'_w entails that $\{w\} \in [\![\Diamond]\!]_{\mathsf{MInqB}}(\mathcal{I})$, so $\{w\} \subseteq \bigcup [\![\Diamond]\!]_{\mathsf{MInqB}}(\mathcal{I})$. Therefore, we have $s = \bigcup_{w \in s}\{w\} \subseteq \bigcup [\![\Diamond]\!]_{\mathsf{MInqB}}(\mathcal{I})$, hence, $s \in [\![!\Diamond]\!]_{\mathsf{MInqB}}(\mathcal{I})$.

$\boxed{\supseteq}$ Now take $s \in [\![!\Diamond]\!]_{\mathsf{MInqB}}(\mathcal{I})$ and $w \in s$. There exists $s' \in [\![\Diamond]\!]_{\mathsf{MInqB}}(\mathcal{I})$ such that $w \in s'$. Call s'' a set of worlds such that $\mu(w') \cap s'' \neq \emptyset$ for all $w' \in s'$. This set exists thanks to definition 4.4. As $w \in s'$, we have $\mu(w) \cap s'' \neq \emptyset$. Therefore, by definition 4.3, $s \in [\![\Diamond]\!]_{\mathsf{InqBK}}(\mathcal{I})$. □

4.3.2 Dynamic inquisitive possibility

In GSV, the dynamic semantics of Groenendijk, Stokhof, and Veltman 1996, information states s are sets of pairs $\langle w, g\rangle$ of a world w and an assignment function g. A Kratzerian externally static denotation of $\Diamond$ in GSV would be (4.6). [6] It restricts the input information states s to the possibilities whose world

5. Information state widening was also advocated by Giannakidou and Mari 2019 for questions with an overt possibility modal.

6. In formula (4.6), quantified assignment g' is independent from g. This allows us to solve the modal identity problem (Beaver 2001, § 8.3.1). For example, in the following discourse, interpreting both it_u in (i-b) and (i-c) with the same assignment g leads to a contradiction because $g(u)$ cannot equal A and R. The modality operator must thus make it possible to

component is in relation R_μ with the world component of some possibility in $[\![\varphi]\!]_{\mathsf{MGSV}}$.

$$[\![\Diamond\varphi]\!]_{\mathsf{MGSV}} := s \mapsto \{\langle w, g\rangle \in s \mid \exists\langle w', g'\rangle \in [\![\varphi]\!]_{\mathsf{MGSV}}(s).\ w' \in \mu(w)\} \tag{4.6}$$
$$= s \mapsto \{\langle w, g\rangle \in s \mid w \in \mathsf{R}_\mu^{-1}(\mathsf{WC}([\![\varphi]\!]_{\mathsf{MGSV}}(s)))\} \tag{4.7}$$

where the world component of an information state is $\mathsf{WC}(s) := \{w \mid \exists\langle w, g\rangle \in s\}$.

Putting all together, we obtain a denotation of $\Diamond$ in Modal Dynamic Inquisitive Semantics MDInq as (4.8), where a context c is a downward-closed nonempty set of information states.

$$\Diamond\mathcal{U} := \lambda c, s.\ s \in c \wedge \exists s' \in \mathcal{U}(c).\ \forall\langle w, g\rangle \in s.\ \exists\langle w', g'\rangle \in s'.\ w' \in \mu(w) \tag{4.8}$$
$$= \lambda c, s.\ s \in c \wedge \exists s' \in \mathcal{U}(c).\ \mathsf{WC}(s) \subseteq \mathsf{R}_\mu^{-1}(\mathsf{WC}(s')) \tag{4.9}$$

Formula (4.9) has the same form as the definition of $\Diamond$ in MInqB (4.5). The only difference is the presence of the world component projection WC, so that the inclusion is an inclusion between sets of worlds. For all context c, $(\Diamond\mathcal{U})(c) \subseteq c$, so $\Diamond$ is externally static.

4.3.3 Pushing and percolating

The syntactic language of our model MDInq has two formula sorts. The first formula sort is type $T = k \to k$, where k is the type of contexts. Formulas of dynamic inquisitive semantics DInq have this type. By taking ℓ the type of stack of contexts $\tau = \langle c_0, ..., c_n\rangle$, the second sort of formula is of type $\ell \to \ell$, i.e. macro-context update functions. Like S. Kaufmann 2000, we take modals to be operators of type $(k \to k) \to (\ell \to \ell)$.[7] After pushing a new local context, modals percolate the information (but not referents) to lower contexts, viz. (4.10). Indicative utterances simply perform an update on the topmost element of the stack.

$$\begin{aligned} [\![\mathsf{if}]\!]_{\mathsf{MDInq}} &:= \lambda\mathcal{U}.\ \textsc{push}\ \mathcal{U} \\ [\![\mathsf{then}]\!]_{\mathsf{MDInq}} = [\![\mathsf{would}]\!]_{\mathsf{MDInq}} &:= \lambda\mathcal{U}.\ \textsc{push}\ \mathcal{U}; \textsc{perc}\ \mathcal{U} \\ [\![\mathsf{might}]\!]_{\mathsf{MDInq}} = [\![\mathsf{could}]\!]_{\mathsf{MDInq}} &:= \lambda\mathcal{U}.\ \textsc{push}\ \mathcal{U}; \textsc{perc}\ \Diamond\mathcal{U} \\ [\![\textsc{indic}]\!]_{\mathsf{MDInq}} &:= \lambda\mathcal{U}.\ \textsc{update}\ \mathcal{U} \end{aligned} \tag{4.10}$$

evaluate it_u in (i-b) with another (accessible) assignment. This is unnecessary for universal modality (4.13).

(i) a. A^u letter is hidden.
b. It$_u$ might be R.
c. But it$_u$ is A.

7. Note that this type lift creates compositional issues for clauses with multiple modalities. A proper treatment of multiple modalities in this framework would require interlacing push and percolation to compute the right common ground updates. A formal analysis of this case would go beyond the scope of this PhD.

where $\ell \to \ell$ conjunction is defined as $\mathcal{T};\mathcal{S} := \tau \mapsto \mathcal{S}(\mathcal{T}(\tau))$.

The analyses of Roberts 1989 and Gillies 2004 for conditionals and modal subordination are given in (4.11).

$$\begin{array}{rcl} \textit{if } \varphi\textit{, then } \psi & \rightsquigarrow & \Box(\varphi \to \psi) \\ \textit{if } \varphi\textit{, then might } \psi & \rightsquigarrow & \Box(\varphi \to \Diamond\psi) \\ \textit{might } \varphi\textit{, would } \psi & \rightsquigarrow & \Diamond\varphi \wedge \Box(\varphi \to \psi) \end{array} \tag{4.11}$$

To obtain these predictions, we define the percolation of a context update function $\mathcal{U}$ as follows. PERC $\mathcal{U}$ applies $\mathcal{U}$ on the penultimate context c_{n-1} and modal conditionalization with respect to c_{n-1} and $\mathcal{U}$ on lower ones (4.12). Intuitively, $c[c' \vdash \mathcal{U}]$ means: context c after learning that *if* c' *then* $\mathcal{U}$.

$$\begin{array}{rcl} \text{UPDATE } \mathcal{U} := & \lambda\langle c_0, ..., c_{n-1}, c_n\rangle.\ \langle c_0, ..., c_{n-1}, \mathcal{U}(c_n)\rangle \\ \text{PUSH } \mathcal{U} := & \lambda\langle c_0, ..., c_n\rangle.\ \langle c_0, ..., c_n, \mathcal{U}(c_n)\rangle \\ \text{PERC } \mathcal{U} := & \lambda\langle c_0, ..., c_{n-2}, c_{n-1}, c_n\rangle.\ \langle c_0[c_{n-1} \vdash \mathcal{U}], ..., c_{n-2}[c_{n-1} \vdash \mathcal{U}], \mathcal{U}(c_{n-1}), c_n\rangle \\ \text{POP} := & \lambda\langle c_0, ..., c_{n-1}, c_n\rangle.\ \begin{cases} \langle c_0, ..., c_{n-1}\rangle & \text{if } n \geq 1 \\ \langle c_0\rangle & \text{if } n = 0 \end{cases} \end{array} \tag{4.12}$$

with $\vdash$ defined as $(\Box(c' \mapsto \mathcal{U}))(c)$. Universal modality and implication (4.13) are consistent with DInq and (4.8).

$$\begin{array}{rcl} c' \mapsto \mathcal{U} := & \lambda c, s.\ s \in c \wedge \forall t \subseteq s. \forall t' \in c'.\ t \sqsubseteq t' \to t' \sqsubseteq \mathcal{U}(c') \\ \Box\mathcal{U} := & \lambda c, s.\ s \in c \wedge \exists s' \in \mathcal{U}(c).\ \forall\langle w, g\rangle \in s.\ \forall w'.\ w' \in \mu(w) \to \langle w', g\rangle \sqsubseteq s' \end{array} \tag{4.13}$$

4.3.4 Meeting the requirements

As an illustration, let's interpret sentence (6-a) in our model. The logical form of (6-a) is displayed in Fig. 4.2 and its interpretation is given in (6-b).[8] After Across-the-Board movement, *which*^{u} *letter* is reconstructed at its trace t_i.

(6) a. Which^u letter could be hidden in FO__M?
b. $\mathcal{T} = \dagger(\text{PUSH } \mathcal{V}; \text{PERC } \Diamond\mathcal{V})$, with
$\mathcal{V} = [u]; \mathbf{atom}\{u\}; \mathbf{letter}\{u\}; \mathbf{hidden}\{u\}; \mathbf{max}\{u\}; ?u$

with $\dagger^{\ell\to\ell}$ defined as

$$\dagger\mathcal{S} := \lambda\tau.\ \begin{cases} \mathcal{S}(\tau) & \text{if } \forall i < |\tau|.\ \bigsqcup \tau_i \sqsubseteq \bigsqcup \mathcal{S}(\tau)_i \\ \text{undefined} & \text{otherwise} \end{cases} \tag{4.14}$$

To weaken the uniqueness presupposition, we require the modality to raise above FOC$_u$ but below Int. If *could* were interpreted at its surface place below

8. As I do not focus on yes-no questions, I do not define the variant of $\langle ? \rangle$ of type $\ell \to \ell$. In this chapter and in Chapter 5, I will disregard this operator.

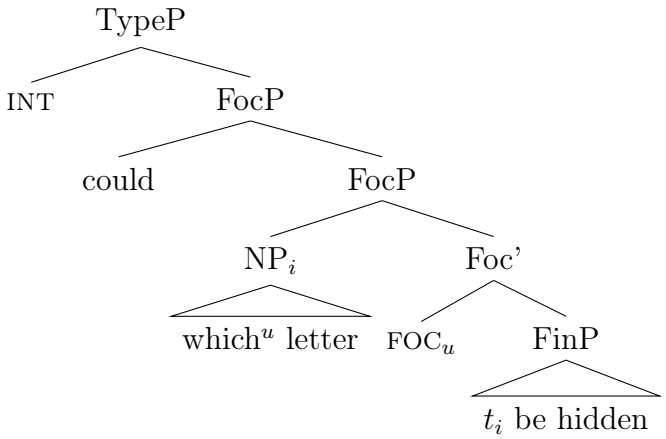


Figure 4.2 – Logical form of (6-a).

World w	w_A^A	w_A^*	w_{RU}^*	w_R^*	w_R^R	w_{RU}^{RU}
Words with skeleton FO__M in English	FOAM	FOAM, FORM, FORUM	FOAM, FORM, FORUM	FOAM, FORM, FORUM	FORM	FORUM
Actual word on the board	FOAM	FOAM	FORUM	FORM	FORM	FORUM
Modal set $\mu(w)$	w_A^A	w_A^*, w_R^*, w_{RU}^*	w_A^*, w_R^*, w_{RU}^*	w_A^*, w_R^*, w_{RU}^*	w_R^R	w_{RU}^{RU}

Table 4.1 – Illustration model $\mathcal{M}$ made of the set of worlds $W = \{w_A^A, w_A^*, w_{RU}^*, w_R^*, w_R^R, w_{RU}^{RU}\}$.

t_i, we would obtain the specific (i.e. *de re*) reading *which* $>$ $\Diamond$. The lexical entry of *which*^{u} introduces the discourse referent u. Given that $\Diamond$ is externally static, u could not be accessed by $?u$ if $\Diamond$ were below FOC$_u$ but above reconstructed *which*. Finally, if $\Diamond$ rose above INT, the global uniqueness presupposition would be computed. Therefore, *could* must raise between FOC$_u$ and Int

Larrivée and Mari 2019, Giannakidou and Mari 2019 and Mari and Giannakidou 2021 advocate for the same relative locations of the modality and the question operator for independent reasons. We take the convergence of their analysis and the technical requirements presented here as an additional piece of evidence that modals can scope over questions.

The illustration model in Table. 4.1 is used to compute the denotation of formula (6-b) displayed in Fig. 4.3.

Dynamic properties Referent u introduced by $\mathcal{U}$ is not projected to the common ground c_0, ensuring external staticity (i.). A subsequent modal operator has access to the topmost local context c_1, containing assignments defined on u, thus

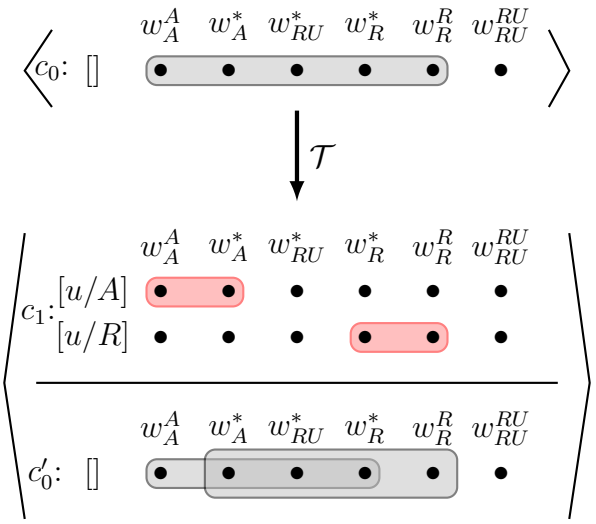


Figure 4.3 – Diagram of the actions of the semantics of question (6-a) in model $\mathcal{M}$ on the initial stack $\tau_0 = \langle c_0 \rangle$.

ensuring modal subordination (ii.).

Mention-some reading In DInq, mention-all reading is triggered by the interaction between the exhaustivity operator **max**, provided by *which*, and the witness request operator $?u$, creating the alternatives. Context $c_1 = \mathcal{V}(c_0)$ contains the alternatives "*d is hidden*" for every single letter *d*. Crucially, $\Diamond$ projects this inquisitiveness onto c_0'. In c_0', the alternatives "*A can be hidden*" and "*R can be hidden*" overlap on $\{w_A^*, w_{RU}^*, w_R^*\}$, creating a mention-some reading (iv.).

Obviated global uniqueness presupposition World w_{RU}^*, where the hidden part RU contains two letters, is included in the alternatives of c_0'. Thus, $\mathcal{T}$ has no global uniqueness presupposition.

Local uniqueness presupposition To prove that $\mathcal{T}$ still has some presupposition, consider context $c_W = \wp(\{\langle w, [] \rangle \mid w \in W\})$, i.e. like c_0 but including world w_{RU}^{RU}. In w_{RU}^{RU}, the English lexicon only has FORUM as a possible completion of FO__M. The condition "*It might be that a single letter could be hidden*" is thus false at w_{RU}^{RU}.

Applying PUSH $\mathcal{V}$; PERC $\Diamond\mathcal{V}$ on stack $\langle c_W \rangle$ results in the same stack as the one at the bottom of Fig. 4.3. In particular, the common ground c_0' of this stack is such that $\bigcup c_W$ does not subsist in $\bigcup c_0'$ (more precisely, we have $\bigcup c_0' \subsetneq \bigcup c_W$). As a consequence, $\mathcal{T}$ is undefined on stack $\langle c_W \rangle$, because of the $\dagger$ operator and world w_{RU}^{RU}. This ensures local uniqueness presupposition (iii.).

Uniform treatment of modalized and conditional questions Finally, our model allows us to compare (6-a) with (7). The conditional antecedent introduces a context where w_{RU}^{RU} is excluded. Thus, the operator † provided in the consequent acts vacuously. MDInq correctly predicts that both (6-a) and (7) have the same effect, except for one point: (7) has no presupposition.

(7) If a^{v} single letter is hidden, which^{u} letter is it$_{v}$?

4.4 Conclusion

We designed a modal dynamic inquisitive semantics which accounts for exhaustivity and presupposition weakening of modalized questions. It also captures modal subordination with an externally static existential modal. To do so, we used stack-based semantics, allowing us to treat conditional and modalized questions uniformly.

Existentially modalized questions are not specific because they are mention-some. Modalized questions with *which+sg* show that specific wh-words can lead to non-specific questions if they scope below a modality. This observation suggests that a factor favoring genericity is wide-scope modality. It supports the claim, made in the next chapter, that generic questions are inherently modal, and so, can be modeled with MDInq.

Chapter 5

Wh-Anaphora

5.1 Introduction

5.1.1 Referring to a wh-word

Interrogative wh-words are referential. They raise a discourse reference that can be co-indexed by an anaphoric expressions (Nishigauchi 1986; Groenendijk 1998; van Rooij 1998), e.g. (1-a) and (1-b). However, the properties of this kind of anaphora are not well understood.

(1) a. No matter who^j comes in, I will meet him$_j$. (Nishigauchi 1990)
b. A: Who^j went to the party? And what did he$_j$/they$_j$ bring as a present? (van Rooij 1998)

In this chapter, I focus on wh-anaphora: cross-sentential anaphora whose antecedent is a wh-word. Nishigauchi 1990 considers wh-anaphora unacceptable when the anaphoric pronoun is in an assertion (2-a). However, van Rooij 1997 finds felicitous examples, like (2-b).

(2) a. A: Who^j is going to visit Kyoto? *I'm sure he$_j$ will enjoy it.
(Nishigauchi 1990)
b. A: Who^j can fully understand Albert's papers on dynamics?
B: I don't know, but he$_j$ will be a trained mathematician.
(van Rooij 1998)

H. Li 2020, § 5.2.1 remarks that "conditions that influence the accessibility of drefs introduced by wh-expressions have been under-explored". This chapter aims to bridge this gap. I mainly focus on cross-sentential pronominal co-reference where the antecedent is a wh-word inside a matrix question.

To investigate wh-anaphora, I gather linguistic examples and corpus occurrences. This data suggests three main ways to refer to wh-words from a ques-

tion. First, answering the question makes the wh-referent fully accessible. In non-answering follow-up utterances, two patterns emerge. Generic wh-words (cf. Chapter 2) can only be referred to via modal subordination (including follow-up questions). In contrast, specific wh-words make indicative co-reference possible, but only in discourse-coherent follow-up sentences.

The rest of the introduction reminds some basic terminology about anaphora. Section 5.2 presents a corpus study and arguments supporting co-reference in answers. Section 5.3 is dedicated to the felicity conditions of wh-referents in non-answers. Section 5.4 presents a model for each of these phenomena in Modal Dynamic Inquisitive Semantics. Finally, section 5.5 concludes.

5.1.2 Anaphora theory

Entity-denoting linguistic expressions may refer to an individual (or a set of individuals) kept track of in discourse. They are called discourse referents (Karttunen 1976) (or *dref* for short). I use here the definition of anaphora by Partee 2014.

5.1.1. Definition (Partee 2014). Anaphora is a relationship between a referentially dependent expression (the **anaphoric expression**, e.g. a pronoun) and a referentially independent expression (the **referential expression**, e.g. a proper name, a definite or indefinite description) that serves as its **antecedent** and from which the anaphoric expression gets its reference (or other semantic value). [1]

Although anaphora with a propositional referent exists, we focus on individual referents here. Following van den Berg 1996, Nouwen 2003 and Schlenker 2011, I consider that all quantifiers can introduce discourse referents, e.g. universal quantifiers.

In the following, I use $u, v, z, ...i, j, k, ...$ as (a countably infinite set of) constants to symbolize drefs. Letters $i, j, k, ...$ are used to report the annotations given by the source. Letters $u, v, z, ...$ represent my analysis. I index anaphoric expressions with a subscript and referential expressions with a superscript. I put the dref on the determiner when possible. For example, in (3), a^u *girl* is the antecedent, and she_u and her_u are anaphoric to a^u *girl*.

(3) A^u girl was sitting here. She$_u$ forgot her bag. Do you know her$_u$?

When one or more expressions refer to the same antecedent, each expression (including the antecedent) is called a *mention*, and their set is a *chain*.

There exist two types of chains: identity chains and bridging chains. Identity chains (aka. *co-reference*) are chains where all mentions are interpreted as the same individual, e.g. in (3). They are further subdivided into *direct references*

1. The boldface is mine.

when the anaphoric NP shares the same noun as the antecedent NP, *indirect references* when the anaphoric NP has a different noun and *pronominal references* when the anaphoric expression is a pronoun.

Bridging anaphora (aka. associative anaphora) concerns anaphoric expressions whose interpretation is not identical to the antecedent but can still be determined by it. Various ontological relations between the two elements are possible, for example, meronymy (whole/part) (4-a), participant (4-b), and so on. I use primes to indicate bridging anaphora (e.g. j and j' are in an associative chain).

(4) a. The thief approached the^{j} house. The$_{j'}$ door was locked.
(ANCOR guidelines, translated)

b. The^{j} concert was great. I loved the$_{j'}$ singer.

5.2 Answers and co-reference to wh-words

Answers offer a favored context for wh-referents. When a responder B answers a question raising wh-dref u, she knows the existence and true identity of u. Thus, she can felicitously talk about u.

The goal of this section is to provide a theory of wh-anaphora between a question and a mention-all answer. I base this theory on theoretical works as well as annotation works. I compare the glosses given by formal semanticists to occurrences of chains containing a wh-word in an anaphora-annotated corpus named ANCOR. First, I highlight the inconsistencies in the anaphora-annotated corpus. Then, I propose a theory based on some previous theoretical works but using more recent concepts. This proposal critiques ANCOR's annotations and pushes for a change in the way anaphora with a wh-word are annotated.

The proposal I defend is made up of the following three points:

1. Inside and after a mention-all answer, a demonstrative NP can refer to the wh-word
2. Mention-all instantiations of wh-words must be annotated as co-reference to the wh-word (i.e. in an identity chain)
3. Wh-anaphora is a structured anaphora in the sense of Brasoveanu 2007b, i.e. expressions can be anaphoric to the range of the wh-word as well as any potential instantiation

Section 5.2.1 briefly introduces the notions at stake in this theory: instantiation and structured anaphora. Section 5.2.2 presents the corpus study on ANCOR. The empirical data is laid out in section 5.2.3. Finally, section 5.2.4 provides the theoretical evidence in favor of the proposal.

ANCOR contains speech disfluencies and transcription spelling mistakes. The original text is reproduced *sic* as in the corpus, but translations include fixes. Transcription mistakes and disfluencies are labeled by ‡.

5.2.1 Issues

Instantiation

In an answer to a wh-question, the phrase that could stand alone as a short answer is called the *instantiation* of the wh-word (or wh-phrase) of that question. For example, in (5-b), the instantiation of *which student* is *Mary* because B could have also answered *Mary* alone. Intuitively, this instantiation is the value (provided by the answer) of the variable abstracted over by the question.

(5) a. A: Which^u student cheated?
b. B: It$_z$ was Mary$_v$.

A mention-all answer to a questions provides the maximal individual satisfying the question. Therefore, it is reasonable to assume that the discourse referent u introduced by the wh-word and the discourse referent v introduced by the instantiation are in an interpretative relation. The problem is: should *which student* and *Mary* be in a co-reference ($v = u$) or should they be in a bridging anaphora ($v = u'$)? There is no agreement in the literature on the answer to this question. The same question between *it* and *which student* arises.

Anaphora between the wh-word and its instantiation

Bridging anaphora includes part-whole relations where the antecedent is a plural entity. For example, *John*$_{u'}$ is a member of the set u of children in (6). They are labeled BE-INST (for *instantiation*) by Nand and Yeap 2013.

(6) The$_j$ children are in the yard. John$_{j'}$ plays with a ball.

(ANCOR guidelines, translated)

Similarly, the denotation of an instantiation is an individual who is part of the wh-domain (Poesio et al. 2024). According to this view, u and v constitute a bridging anaphora in dialogue (5) ($v = u'$). I call this view the Mention-all Instantiation is Bridging approach.

In the following of section 5.2, I refute this analysis and defend that wh-word instantiations in mention-all answers are in an identity chain with the wh-word, e.g. $v = u$ in (5).

Structured anaphora for wh-words

The third claim concerns the nature of wh-referents in general. Discourse referents introduced by generalized quantifiers exhibit behaviors that specific indefinite or definite NPs do not exhibit. They can be referred to by a plural pronoun (7-a) and it is possible to zoom in on each individual and refer to them with a singular pronoun (7-b). This phenomenon is called telescoping (Roberts

1989). Brasoveanu 2007b regroups telescoping and other phenomena akin to quantifier scope extending over sentences (7-c) under the term of quantificational subordination.

(7) a. Each^j degree candidate walked to the stage. They$_j$ were all smiling.
b. Each^j degree candidate walked to the stage. He$_j$ took his$_j$ diploma from the Dean and returned to his$_j$ seat.
c. Harvey courts a^j girl at every convention. She$_j$ always comes to the banquet with him. The$_j$ girl is usually also very pretty. (Karttunen 1976)

The existence of plural reference and quantificational subordination prompts Brasoveanu 2007b to consider anaphora with quantifiers as structured anaphora. The referent introduced by a quantifier is not valued by an individual but by a more complex structure, giving access to the scope set (for plural reference) and each of its values (for quantificational subordination). The following sections provide evidence that specific and generic wh-words involve structured anaphora.

5.2.2 Corpus study

English corpora

There are several English corpora annotated in anaphora. The most known are OntoNotes (Hovy et al. 2006), GUM (Zeldes 2017) and WikiCoref (Ghaddar and Langlais 2016), which is based on OntoNotes markup. However, none of them considers that wh-phrases are referential. Thus, wh-anaphora is not annotated in these corpora. The guidelines of these corpora justify this choice by claiming that quantified NPs (and, in particular, wh-phrases) do not refer to specific individuals but have a non-specific reading.

Corpus ARRAU (Poesio et al. 2024) annotates some "generic" quantified NPs, including wh-phrases. However, its guidelines only mention intra-sentential examples and plural reference to *most N*. No cross-sentential reference to a wh-phrase was found in the free part of ARRAU. [2]

2. Here and throughout this paper, I exclude propositional anaphora to the whole interrogative, e.g. (i) and (ii).

(i) [When are we going to the beach?]^j It$_j$ depends on the weather. (Hui 2021)

(ii) a. SPK1: [depuis quand est-ce que vous l'avez eu]^u ? [à quelle occasion]^v [...] ? (ANCOR:019_C-2)
'*[How long have you had it?]^u [On what occasion?]^v*'
b. SPK2: oh ma foi non je m'en$_{u/v}$ souviens pas
'*Oh my goodness, no, I don't remember that$_{u/v}$.*'

ANCOR

The absence of relevant data from English corpora prompts me to consider French corpora as an alternative. The corpus ANCOR (Muzerelle et al. 2014) is a large French spontaneous speech corpus annotated in co-reference (488,000 words, 30.5 hours recorded). Some interrogative pronouns and determiners are annotated as mentions. [3] In this section, I use ANCOR for two purposes. First, it provides expert annotations to support my proposal. Second, I briefly evaluate these annotations and critique them.

A Python script was used to extract the non-trivial chains containing a mention satisfying the following: the mention has the feature NEW=YES (i.e. annotated as an antecedent) and contains a form among *qui* '*who*', *que/qu'* '*what*', *quel(le)(s)* '*which*', *lequel(el)(s)* '*which one*', *où* '*where*', *quand* '*when*', *comment* '*how*', *pourquoi* '*why*' or *combien* '*how much/many*'. Chains strictly included in another chain were removed. [4]

The script found 400 bridging and 101 identity chains with a wh-word. In 98% of them all, the wh-word was *quel* ('*which*'), or another flexion of this word (in gender or number). [5] Let me first focus on identity reference chains. Bridging chains are explored in the second part of section 5.2.3.

Two false positives (relative pronouns) were removed manually. The other chains were annotated with some syntactic features. We first focus on anaphora to an instantiation. To constitute a corpus of study, we excluded:

— 6 chains where the anaphoric expression could rather be analyzed as referring to the whole interrogative proposition (see fn. 2)

3. The written French corpus DEMOCRAT (Glikman et al. 2022) also contains some anaphora chains including interrogative words, although much fewer, e.g. intra-sentential co-references (i-a).

(i) a. Qui^{j} est-ce qui voudroit gouverner ainsi ses$_{j}$ affaires publiques ?
(DEMOCRAT:montesquieulois)
'*Who^{j} would want to govern their$_{j}$ public affairs in this way?*'

4. The scripts and annotations are available at https://github.com/Valentin-D-Richard/ANCOR_eval

5. A brief look at the corpus revealed that some anaphora with interrogative pronouns are missing. For example, *qui* '*who*' and *moi* in (i) should be in a joint chain because the latter is the instantiation of the former, but no chain contains them both. Anaphora with interrogative words was intended to be included in ANCOR (Jean-Yves Antoine, p.c.). Therefore, this might be a systematic annotation mistake. See Richard 2025a for a more detailed evaluation of the annotations of chains containing an interrogative word in ANCOR.

(i) a. SPK1: et qui^{u} est-ce qui remplissait ‡les les f- les papiers administratifs ah bon
(ANCOR:021_C-6)
'*Who filled out the administrative paperwork?*'
b. SPK2: moi$_{u}$
'*Me.*'

Pronoun	Count
ce (‘*that*’)	17
ça (‘*that*’)	1
tout (‘*everything*’)	1
n’importe quoi (‘*anything*’$_{FC}$)	1
n’importe lequel (‘*any*$_{FC}$ *of these*’)	1
quelque chose (‘*something*’)	2

Table 5.1 – Anaphoric pronouns and their number of occurrences in the corpus of study.

— 24 chains where the referring expressions are repetitions of the anaphoric expression

— 5 chains where the referent is the context domain of the wh-word. [6]

— 4 additional chains where the wh-phrase is in an embedded interrogative

Thus, the corpus of study is reduced to 60 co-reference relations, all of them with *quel(le)(e)*. Out of these relations, 23 have a pronoun as the first co-referring mention following the wh-phrase. In view of arguing on the proposal, we focus on these chains in the following. The list of the pronouns in these chains and their counts is given in Tab. 5.1. The complete list of the dialogues these chains are in is available in Appendix C.

5.2.3 Different anaphora types in answers

In this section, I present the data collected in ANCOR and show how it supports my proposal.

6. Referents representing the context domain (aka. range, or restrictor) of a wh-phrase are necessary to annotate anaphoric negative quantifiers or pronouns (van Rooij 1998), like in (i-b), where *no-one* quantifies over second-year students. ANCOR contains examples like (ii), where the dref j introduced by *votre*^{j} *famille* ‘*your*^{j} *family*’ co-indexes *qui*$_j$ ‘*who*$_j$’ because the value of j determines the context domain of this interrogative word. This kind of anaphora is irrelevant for the study of mention-all instantiations.

(i) a. A: Which^u second-year students cheated?
b. B: No-one$_u$ cheated.

(ii) et en général dans votre^j famille qui$_j$ est-ce qui se sert le plus souvent du dictionnaire ou des dictionnaires ? (ANCOR:CO2_ESLO_003_C)
‘*In general, in your*^{j} *family, who*$_j$ *uses the dictionary or dictionaries most often?*’

Demonstrative pronouns in answers

A total of 19 anaphoric pronouns are demonstratives: *ce* or *ça* ('*that/this*'). All of them (except one: (8) in Appendix C) are subjects of a copular predicative construction whose object is an answer to the question. For example, SPK2 answers question (8-a) with (8-b). The pronoun *ce* (here elided to *c'*) is in an identity chain with *quelle*. By answering the question, SPK2 makes common ground that there was a subject she did best in and what this subject was. Therefore, it_j is felicitous.

(8) a. SPK1: madame lorsque vous étiez encore à l'école dans quelle^j matière étiez vous le plus fort ? (ANCOR:021_C-6)
'*Madam, when you were still at school, which^j subject did you do best in?*'

b. SPK2: ah c$_j$'était littérature$_{j'}$.
'*It$_j$ was literature$_{j'}$*'

After answering a mention-all question, the wh-referent stays accessible in the subsequent discourse and is valued with the individual provided in the answer, as illustrated by *ce$_j$ nom-là* ('*this$_j$ name*') in the third speech turn of (9).

(9) a. SPK3: ah oui ce serait quel^j nom euh ? (ANCOR:023_C)
'*What^j name would it be?*'

b. SPK2: c$_j$'était Ligneaux$_j$ [...]
'*It$_j$ was Ligneaux$_j$.*'

c. SPK2: et ce$_j$ nom-là euh figurent‡ dans les registres de‡ des paroisses
'*This$_j$ name appears in parish registers.*'

Thus, ANCOR's annotations support item 1 of my proposal.

The instantiation annotation

ANCOR agrees with Poesio et al. 2024 regarding the annotation of instantiations in answers. They use the Mention-all Instantiation is Bridging approach. In (8-a), the instantiation *littérature$_{j'}$* is annotated as a bridging anaphora with *quelle^j matière* because literature belongs to the set of school subjects.

In the corpus of study, the 17 mention-all answers of the form *ce + être + NP$_{inst}$* ('*it + be + NP$_{inst}$*') are inconsistently annotated with respect to the relation between the wh-phrase and *NP$_{inst}$*. Twelve have a bridging relation like (8-a), one has a co-reference relation (9), and the three others have no relation at all.

Other parts of ANCOR contain cases where the instantiation is in a co-reference with the wh-phrase. Dialogues (8) and (10) both contain a mention-all question and a mention-all answer. However, the instantiated NP *la$_j$ forme que ça a prit‡ au mois de mai* is annotated as a direct identity chain.

(10) a. SPK1: et d'après vous ça va prendre quelle^{j} forme ce‡ ? (ANCOR:542_C-3)
'*What^{j} shape do you think it will take?*'
b. SPK4: oh ça prendra la$_{j}$ forme que ça a prit‡ au mois de mai
'*It'll take the$_{j}$ shape it took in May.*'

Another example is dialogue (11).[7] In this conversation, *quoi^{j}* and *licence$_{j}$* are annotated as an indirect identity chain and not a bridging anaphora.

(11) a. SPK1: elle souhaite s'inscrire chez nous madame (ANCOR:098_00000063)
'*She would like to register with us.*'
b. SPK2: oui
'*Yes.*'
c. SPK1: en quoi^{j}
'*What$_{j}$ for?*'
d. SPK2: e‡ licence$_{j}$ un AES
'*[First year of AES bachelor]$_{j}$.*'

To sum up, ANCOR seems to apply the Mention-all Instantiation is Bridging approach, although some of its annotations favor the opposite approach: mention-ll instantiations are in a co-reference with the wh-phrase. The existence of the latter approach in ANCOR supports item 2 of my proposal. Additional theoretical evidence to this item is brought in section 5.2.4. I propose that the bridging annotations for mention-all instantiations should be changed in ANCOR.

Anaphoric quantificational pronouns

In the corpus of study, we observe that an anaphoric mention referring to a wh-prase can be a quantifier or a quantificational pronoun. For example, in (12-b), SPK2 says that all individuals of the domain of *quels^{j} gens* '*which.PL people.PL*' are true witnesses.

(12) a. SPK1: pour bien apprendre quels^{j} gens est-ce qu'il faudrait fréquenter ? [...] (ANCOR:015_C-2)
'*To learn properly, which^{j} people should you get to know?*'
b. SPK2: tout$_{j}$ [8]
'*Everybody$_{j}$*'

Pronoun *n'importe lequel* (lit. '*no matter the-which*') is a free choice item. In (13-f), SPK1 either expresses indifference about the question or that any newspa-

7. Dialogue (11) was not extracted by the Python script because the wh-word *quoi* was annotated as not new in the discourse. This choice is not justified. No mention of any training track appears in the preceding dialogue.

8. Here, *tout* ('*everything*') is unexpected. The expression *tout le monde* or *tous* ('*everybody*') would be necessary. I assume that *tout* is nonetheless the right transcription and that this word is acceptable in the French variety spoken by SPK2.

per is a newspaper in which SPK2 do crosswords.

(13) a. SPK2: est-ce que vous faites des mots-croisés ? (ANCOR:012_C-3)
'*Do you do crosswords?*'
b. SPK1: oui bien sûr
'*Yes of course.*'
c. SPK2: oui lesquels ?
'*Which ones?*'
d. SPK1: lesquels dans quel^j journal ‡dans quel^j journal ?
'*Do you mean "In which^j newspaper"?*'
e. SPK2: oui
'*Yes.*'
f. SPK1: oh [n'importe lequel]$_j$ et de préférence dans le$_{j'}$ Canard Enchaîné
'*Any$_j$ (of them) and preferably in the$_{j'}$ Canard Enchaîné*'

The availability of anaphoric quantificational pronouns suggests that expressions referring to a wh-word can refer to their domain as a distributive set (type et). In the Generalized Quantifier theory (Barwise and Cooper 1981), determiners have the type $\langle et, \langle et, t\rangle\rangle$. When combined with a restrictor, of type et, they produce a quantifier of type $\langle et, t\rangle$, like *everybody* or *any of them*. When a referent j indexes a quantifier, this referent determines the quantifier's restrictor. However, if the wh-dref is valued by an individual through a simple assignment function g (like in the first version of DInq by Dotlačil and Roelofsen 2019), there is a type mismatch: $g(j)$ is typed e but is expected to be a restrictor of type et. This issue is solved by suggesting that wh-drefs have to be valued as structured objects, using a set G of assignment functions. Consequently, ANCOR's annotations support item 3 of the proposal. More details are given in section 5.2.4.

5.2.4 Theoretical arguments

Plural reference to the wh-scope and quantificational subordination

Van Rooij 1998 gives an example where a wh-referent can be construed as the maximal plural entity satisfying the question scope. In (14-b), *them$_j$* refers to the plural entity containing all persons that went to Mary's birthday.

(14) a. A: Who^j went to Mary's party? (van Rooij 1998)
b. B: At least John was one of them$_j$.

However, van Rooij 1998 also provides an example where wh-anaphora can zoom in on any potential instantiation of the wh-word. In (15-b) *he$_j$* evaluates to any person able to understand Albert's paper on dynamics, independently of the number of people concerned (maybe nobody). As suggested by van Rooij 1998,

(15-b) indeed parallels telescoping.

(15) a. A: Who^j can fully understand Albert's papers on dynamics?
(van Rooij 1998)

b. B: I don't know, but he$_j$ will be a trained mathematician.

These observations support item 3 of the proposal.

Model for structured wh-anaphora and instantiation wh-anaphora

Brasoveanu 2007b models structured anaphora with assignment matrices, i.e. non-empty sets of assignment functions g. Given an assignment matrix G, the value of a referent u can vary in G. An anaphoric expression can refer to the set $G(u) = \{g(u) \mid g \in G\}$, to its mereological sum $\bigoplus G(u)$, or any $g(u)$ given some quantification over assignments g.

The data I provide suggests that wh-anaphora should also be considered structured anaphora. Dynamic Inquisitive Semantics DInq as presented in (Roelofsen and Dotlačil 2023b) integrates Brasoveanu's assignment matrices G and can thus model structured anaphora for wh-drefs.

As explained in section Chapter 3, Roelofsen and Dotlačil include Brasoveanu's assignment matrices in MDInq in order to model multiple wh-questions and the difference between weak and strong donkey anaphora. These phenomena are not the focus of this thesis. Quantificational subordination and co-indexation of a wh-dref with a quantifier occupy a minor part of this dissertation. Therefore, I opted for a simpler system with singular assignments g because I estimate that it provides more ease of comprehension in the other parts of this manuscript. Consequently, a formal model for quantification subordination and co-indexation of a wh-dref with a quantifier is postponed to a subsequent work. The version of MDInq I use is sufficient to model the other phenomena studied this thesis.

Maximal witness

The Mention-all Instantiation is Bridging approach is based on the following argument: Instantiations of wh-phrases are elements of the wh-phrase's domain. Let me explain why this argument is irrelevant.

Quantifiers raise two discourse referents, one for their restrictor and one for their nuclear scope (Dekker 1992). For example, *they* in (16-b) refers to the set of students that cheated, whereas *they* in (16-c) typically refers to the students.

(16) a. Most$^{restr:v,scope:u}$ students cheated.
b. They$_u$ will have 0 points.
c. Very few of them$_v$ came to the last class.

Scope reference is more typical and frequent than restrictor (or domain) reference. Referents raised by indefinites are typically scope referents. As generalized

quantifiers, wh-phrases function similarly. Restrictor reference with a wh-phrase is possible, like in (17-c) (see also fn. 6), but is less common.

(17) a. A: Which$^{restr:v,scope:u}$ students cheated ?
b. A: Whoever they$_u$ are, they$_u$ will have 0 points.
c. B: None of them$_v$.
d. B: [Anne and Bea]$_u$

Instead of considering that instantiations are elements of the restrictor dref ($Anne, Bea \in g(v)$), we can simply consider that instantiations are the mereological sum of the nuclear scope of the wh-phrase (i.e. the maximal witness): $Anne \oplus Bea = \bigoplus g(u)$.

In this thesis, we consider that the dref raised by a wh-word is its scope referent. Therefore, a mention-all instantiation is in an identity chain with this dref. This ends the proof for item 2 of my proposal.

5.2.5 Interim summary

In this section, I argued for a theory of wh-anaphora repeated below:

1. Inside and after a mention-all answer, a demonstrative NP can refer to the wh-word
2. Mention-all instantiations of wh-words must be annotated as co-reference to the wh-word
3. Wh-anaphora is a structured anaphora in the sense of Brasoveanu 2007b

Crucially, item 2 only applies to complete mention-all answers. Instantiations in partial answers to mention-all questions and answers to mention-some questions do not co-refer to the wh-word because they are partial witnesses. More information about the true witnesses may be added subsequently, as illustrated by (18-d). Therefore, it makes sense to annotated these wh-word instantiations as bridging anaphoric expressions. For example, in conversation (18), u' and u'' constitute a subpart of the sum of the true witnesses represented by u.

(18) a. A: Who^u knows how to code in python?
b. B: Mary$_{u'}$.
c. A: Only Mary?
d. B: No, also John$_{u''}$.

5.3 Specific and generic co-reference to wh-words in non-answers

This section investigates the properties of wh-anaphora in subsequent sentences that are not answers. Parallel to the specific vs. generic duality of wh-

words advocated for in Chapter 2, I argue here that there are two distinct wh-anaphora behaviors. By default, wh-referents are generic and require modal subordination (section 5.3.1). However, under a specific reading, the wh-dref can be referred to by a pronoun in an indicative subsequent assertion, provided that this assertion is coherent with the question in the discourse (section 5.3.2).

5.3.1 Modal subordination with generic questions

In this section, I argue that generic questions behave similarly to modalized questions in terms of anaphoric accessibility. A non-specific discourse referent introduced in a generic question cannot be referred to in a subsequent non-answering indicative assertion. However, reference is redeemed by modal subordination. This observation constitutes evidence I will use in section 5.4.1 to advocate for modeling generic questions with a covert wide-scope modal.

Reference to hypothetical individuals

Indefinite noun phrases introduced by *such* may be anaphoric (Webber et al. 2003). In (19-b), *such people* refers to the witnesses of sentence (19-a), i.e. the farmers who beat donkeys.

(19) a. Some^u farmers beat donkeys.
b. Such$_u$ people are mean.

Contrary to other anaphoric NPs, *such* NPs can refer to hypothetical (atomic) individuals. For example, *such a person* in (20-c) denotes any hypothetical farmer who has beaten a donkey. This denotation is given by co-indexation with the discourse referent u raised in the negatively quantified sentence (20-a).

(20) a. No^u farmer has ever beaten a donkey.
b. #This$_u$ person would be jailed.
c. [Such a person]$_u$ would be jailed.

Hypothetical wh-referents

Van Rooij 1997, 1998 argues that sequences like (15) above is quantificational subordination in wh-anaphora. He argues that this behavior can be modeled similarly to modal subordination: with dref dependence on world referents. I defend the idea that (15) is generic and generic wh-anaphora *is* modal subordination.

Consider the mention-some question (21-a). Suppose responder B does not know whether any person can model free choice effects in questions. B can nevertheless refer to that hypothetical individual in a conditional antecedent with *such* (21-b) or under a modal (21-c). Responding in an indicative assertion would be

interpreted as giving a (partial) answer (21-d). Indeed, when the responder explicitly expresses her lack of knowledge, an indicative co-reference is infelicitous (21-e).

(21) a. A: Who^u knows how to model free choice effects in questions?
b. B: I don't know, but if [such a person]$_u$ exists, they$_u$ would work at the ILLC.
c. B: I don't know, but [such a person]$_u$ would probably work at the ILLC.
d. B: #[Such a person]$_u$ works at ILLC.
e. B: #I don't know, but [such a person]$_u$ works at the ILLC.

The judgments in (21) also hold when interpreting (21-a) as mention-all if *such a person* is replaced by *such persons*. The phenomenon also appears for other drefs raised inside a generic question. For example, non-specific indefinite *a^v car* can only be referred to in (22-b) and (22-c), not in indicative (22-d). [9]

(22) a. A: Who has a^v car I could borrow?
b. A: It$_v$ must be an automatic car. I cannot drive manual cars.
c. A: It$_v$ may be an automatic car. I don't care.
d. A: #It$_v$ is an automatic car.

Example (21) shows that generic mention-some or mention-all questions require modal subordination, even if the question contains no overt modal operator. I advocate for considering generic questions as inherently modal contexts.

Understanding generic wh-word reference

An epistemic modal statement *might p* expresses a relation between the common ground and the hypothetical proposition *p*. The proposition *p* is not considered as a fact, but is hypothetical. Similarly, a question raises alternatives, i.e. resolution propositions that are hypothetical with respect to the common ground.

In Chapter 4, I modeled a hypothetical proposition by a local context. Extending this analogy, each possible answer to a question can be modeled as a local context in discussion. Follow-up questions refine the given question, and can thus access these local contexts. That is why wh-drefs are acceptable in follow-up questions, as seen above (e.g. (1-b)). [10]

9. However, an indicative main verb is possible if the subsequent sentence is in the scope of a generic operator provided by the anaphoric expression. For example, in (i), A expresses her appreciation of any local situation where she gets an automatic car. The car is only defined in those local situations.

(i) A: An$_u$ automatic one is fine for me.

10. I refer to the discussion on exemplification in section 6.3.1 of Chapter 6 for examples of

Non-interrogative wh-words can also behave like modal operators. Typically, the wh-word introducing free relatives can be referred to in the following sentence, but only via modal subordination. For example, modalized assertion (23-b) can refer to any problem Mary faces today, while indicative (23-a) is unacceptable.

(23) a. A: Whatever^u problem Mary is facing today will not block her.
b. A: She would/will solve it$_u$ immediately.
c. A: #She solves it$_u$ immediately.

In out-of-the-blue question-assertion sequences uttered by a single person, H. Li observes that reference to the wh-word is infelicitous. For example, (24-a) feels off when the second sentence is not an answer to the first one. (24-a). Yet, when the pronoun is in the scope of a modal operator (24-b) or an attitude verb (24-c), the sentence is more acceptable.

(24) a. Who^j is Bob dating? #She$_j$ is a doctor. (H. Li 2020, attr. to Anna Szabolcsi)
b. Who^j is Bob dating? She$_j$ might be a doctor.
c. Who^j is Bob dating? I heard she$_j$ is a doctor.

This phenomenon might constitutes additional evidence in favor of the modal behavior of generic questions with respect to anaphora because, outside of any context, (24-b) and (24-c) could be analyzed as generic questions.

5.3.2 Identification and coherence

Specific wh-anaphora

Despite examples like (21) requiring modal subordination, several cases of wh-anaphora in a subsequent indicative assertion have been brought up in the literature, e.g. (25), (26) and (27-b).

(25) a. SITUATION: *John saw Bob talking with a woman who was wearing a hat.*
b. JOHN TO A FRIEND: Who^j is Bob talking with? She$_j$ is wearing a weird hat. (H. Li 2020, attr. to Anna Szabolcsi)

(26) Which^j writer won the Nobel Prize in Literature in 1969? To give you a hint, he$_j$ is Irish. (Haida 2007)

(27) a. A: Who^j went to Mary's party? (van Rooij 1998)
b. B: I don't know, but Mary liked him$_j$ a lot.

The common point to these cases is that they (or their context) establish or presuppose that at least one true witness exists. Let me show that all of these examples are specific anaphora.

---

access to alternative local contexts.

In (25) (taken from H. Li 2020), BOB has somebody specific in mind. I argued in Chapter 2 that this kind of epistemic state is associated with the specific reading who_S. I also claimed that exam questions, like (26), are only felicitous with a specific interpretation (and *which + sg* is specific anyway). In (27-b), the singular pronoun him_j forces its domain to be a singleton. This domain narrowing favors specific readings (cf. Schwarzschild 2002).

In Chapter 2, I explained that who_S comes equipped with an extra identification method. In (25), who^j is identified as the woman with a weird hat Bob was talking to. This property explains why specific wh-anaphora in subsequent assertions by the asker are possible.

However, example (27-b) does not quite work with this approach. A can utter (27-a) as generic, and, in particular, without knowing if anyone went to Mary's party. However, B's answer stipulates that someone (exactly one person) attended her party. To solve this apparent problem, remember that the asker and the speaker may disagree on the wh-word reading because who_S and who_G are homophonous. Therefore, B can interpret (27-a) as specific because he knows that exactly one person went to Mary's party, independently from A's interpretation (provided that the context allows that).

The coherence constraint

There is one additional ingredient for a felicitous specific wh-anaphora: discourse coherence. Take again H. Li 2020's examples in (24). The question "*Who^j is Bob dating?*" is specific by default. But then, why is "*She$_j$ is a doctor*" unacceptable?

Anaphora is sensitive to discourse relations (Asher and Lascarides 2003). I argue that a felicitous specific wh-anaphora requires the subsequent sentence to be coherent in the discourse. [11]

With "*She$_j$ is a doctor*", it is unclear how this sentence relates to the question about j. However, we can redeem this incongruence by setting a situation where this assertion is helpful in the conversation.

In (28-d), A clarifies the goal of her question. This move helps Mark understand why he might be an authority with respect to this question. It indicates to him where to search (here: among his female colleagues) to try to provide an

11. H. Li 2020 takes the discourses in (i-a) and (i-b) as evidence that anaphoric expressions in arguments of psych verbs facilitate reference to wh-drefs. I refute this analysis. These cases are simply specific questions. The question in (i-a) uses *which+sg*, and the follow-up assertion is related to the girl winning the game. Discourse (i-b) can only make sense if, for example, it is uttered by a tourist arriving in Rome, seeing that it has been ransacked, and seeing residents complaining about foreign soldiers.

(i) a. Which^j girl won the game? Bob envies her$_j$. (H. Li 2020, attr. to Philippe Schlenker)
b. Which^j troops occupied Rome? The residents hate them$_j$.

answer.

(28) a. SITUATION: *A desperately wants to know who Bob is dating. Therefore, she goes to the hospital where Bob and her girlfriend work and finds Mark, one of their colleagues.*
b. A: Who^u is Bob Roger dating?
c. *A suddenly realizes that Mark might not understand how he could know that.*
d. A: She$_u$ is a doctor. And she$_u$ is working in your department.

5.4 Modeling wh-anaphora

In this section, I model the three phenomena presented in sections 5.2 and 5.3, that is: reference to a wh-word in an answer, reference to a generic wh-word, and reference to a specific wh-word. Section 5.4.1 starts with generic questions and modal subordination in Modal Dynamic Inquisitive Semantics. Specific co-reference is modeled in section 5.4.2. Finally, section 5.4.3 is dedicated to stack management for answers.

5.4.1 Modeling generic questions

Modal subordination and covert G

In Modal Dynamic Inquisitive Semantics MDInq, utterances are interpreted by functions from macro-context to macro-context. A macro-context τ is a stack of local contexts. Modal subordination is modeled by pushing a new local context onto the stack and updating it with the modal prejacent. Follow-up modal assertions can access this local context, but indicative ones cannot.

To model the modal behavior of generic questions, I assume that generic questions contribute a covert modal operator G. Chapter 4 established that overt modal operators leading to a non-specific reading must raise to Spec FocP. Therefore, I will suppose that G is generated at Spec FocP. Operator G is pragmatically performative, as it creates a new local context. However, more investigation would be required to exhibit its precise semantic contribution. For now, I leave its modal base (μ_{G}) and ordering source unspecified.

$$\begin{aligned} [\![\text{G}]\!]_{\mathsf{MDInq}} &:= \lambda\mathcal{U}.\ \text{PUSH}\ \mathcal{U};\text{PERC}\ \Diamond_{\text{G}}\mathcal{U} \\ \Diamond_{\text{G}}\mathcal{U} &= \lambda c, s.\ s \in c \wedge \exists s' \in \mathcal{U}(c).\ \mathsf{WC}(s) \subseteq \mathsf{WC}(s') \end{aligned} \tag{5.1}$$

In Chapter 2, I argued that the specific vs. generic ambiguity is located at the wh-word level. The presence of G high in the logical form thus requires an explanation. Although wh-words are generated and interpreted *in situ*, they move to Spec FocP to co-index FOC. This is also their surface position in English. I

hypothesize that during this movement, a generic wh-word also marks FocP with genericity. A FocP marked with [+generic] generates covert G just above.

When the wh-question involves future tense, habitual aspect, or a generic term, the respective modal operator contributed by these constructions may move to replace covert G. Giannakidou and Mari 2021 advocate for a similar movement of the operator contributed by the epistemic future above the question operator. Finally, as seen in Chapter 4, even if the wh-word is specific, an explicit possibility modal operator can create a generic reading by rising just above Spec FocP.

Including the null individual

In Chapter 2, I propose to model generic who_G as containing the null individual $\star$ in its domain. To implement this idea, we first have to introduce $\star$ into MDInq.

In MDInq, the domain interpreting type e is a complete lattice D' of plural individuals, including a bottom element $\star$. Following Landman 2011, the pluralization operator * includes $\star$. For example, if the dogs are a and b, then $^*\mathbf{dog} = \{\star, a, b, a \oplus b\}$. Because $\star$ is the bottom element of the mereological lattice, we have $\star \leq d$ and $\star \oplus d = d$ for any plural individual d.

In this setting, the definition of some operators must be updated. The **atom** operator only selects non-null atoms (5.2). Moreover, $[u]$ only introduces a non-null discourse referent (5.3).

$$\mathbf{atom}\{u\} := \lambda c, s.\ s \in c \wedge \forall \langle w, g\rangle \in s.\ g(u) \neq \star \wedge \neg(\exists y \neq \star.\ y < g(u)) \tag{5.2}$$

$$\begin{aligned} g[u]g' \text{ if } \quad & \mathbf{rdom}\, g' = \mathbf{rdom}\, g \cup \{u\} \wedge (\forall z \neq u.\ z \in \mathbf{rdom}\, g \rightarrow g(z) = g'(z)) \\ & \wedge g'(u) \neq \star \\ s[u]s' \text{ if } \quad & (\forall \langle w, g\rangle \in s.\ \exists \langle w', g'\rangle \in s'.\ w = w' \wedge g[u]g') \wedge \\ & (\forall \langle w', g'\rangle \in s'.\ \exists \langle w, g\rangle \in s.\ w = w' \wedge g[u]g') \\ [u] := \quad & \lambda c, s.\ \exists t \in c.\ t[u]s \end{aligned} \tag{5.3}$$

Generic who_G

Generic who_G introduces a dref u and requires that the predicate $\mathcal{P}$ who_G is associated with holds on u. With $[u]_G$, the discourse referent u is authorized to take $\star$ as a value. Definition (5.5) is formally exactly the same as the previous definition of $[u]$, but the inclusion of $\star$ in the model changes its denotation.

$$[\![who_G]\!]_{\mathsf{MDInq}} := \lambda \mathcal{P}.\ [u]_G; \mathcal{P}(u) \tag{5.4}$$

$$\begin{array}{rl} g[u]_G g' \text{ if} & \mathbf{rdom}\, g' = \mathbf{rdom}\, g \cup \{u\} \wedge (\forall z \neq u.\ z \in \mathbf{rdom}\, g \rightarrow g(z) = g'(z)) \\ s[u]_G s' \text{ if} & (\forall \langle w, g\rangle \in s.\ \exists \langle w', g'\rangle \in s'.\ w = w' \wedge g[u]_G g') \wedge \\ & \quad (\forall \langle w', g'\rangle \in s'.\ \exists \langle w, g\rangle \in s.\ w = w' \wedge g[u]_G g') \\ [u]_G := & \lambda c, s.\ \exists t \in c.\ t[u]_G s \end{array} \tag{5.5}$$

Null answer exhaustification

As explained in section 2.4.2, null answers undergo mandatory exhaustification, leading to the *nobody* reading. Unfortunately, adapting the exhaustification operator and its respective grammatical theory to (Modal) Dynamic Inquisitive Semantics is not an easy task and would go beyond the scope of this thesis. [12] Therefore, I propose to model a simplified version of null answer exhaustification at a fixed syntactical position in the left periphery. Null answer exhaustification must be performed after processing the FinP but before computing the question presupposition, so before processing INT. Therefore, I include null answer exhaustification in the witness request operator $?u$.

The definition of $?u$ is modified by (i.) ignoring $\star$ in the initial definition of $?u$ (5.7), and (ii.) disjoining it with a special case for $\star$ (5.8). Intuitively, constraint (ii.) means "*For all possibility $\langle w, g\rangle$ in s, $\star$ and only $\star$ satisfies the question's scope in w*".

$$?u := \lambda c, s.\ s \in c \wedge \tag{5.6}$$

$$(\exists x \neq \star.\ \forall \langle w, g\rangle \in s.\ \exists \langle w', g'\rangle \in \bigcup c.\ w = w' \wedge g'(u) = x) \tag{5.7}$$

$$\vee\ (\forall \langle w, g\rangle \in s.\ \forall \langle w', g'\rangle \in \bigcup c.\ w = w' \rightarrow g'(u) = \star) \tag{5.8}$$

Example

To illustrate these new definitions, consider generic question (29). Its logical form (before reconstruction) is given in Fig. 5.1 and its semantic representation in (29-b). Operator ! is vacuous here. [13]

(29) a. Who^u slept well? (generic)

b. $\dagger(\text{PUSH}\ \mathcal{U}; \text{PERC}\ \Diamond_G \mathcal{U})$, with $\mathcal{U} = !([u]_G; \mathbf{sleep_well}\{u\}); ?u$

I consider a simple model with two atomic individuals a and b and four worlds w_a, w_{ab}, w_b and $w_\emptyset$. The subscript indicates the set of individuals who slept well.

12. I refer to Klochowicz 2025 for an attempt to adapt an exhaustification operator to Inquisitive Semantics.

13. For the absence of $\langle ? \rangle$ in formula (29-b), see footnote 8 of Chapter 4.

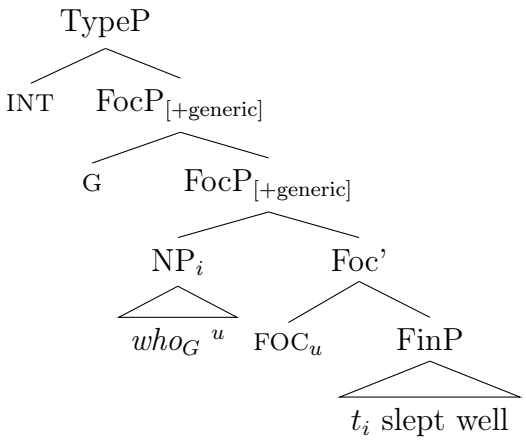


Figure 5.1 – Logical form of (29) before reconstruction.

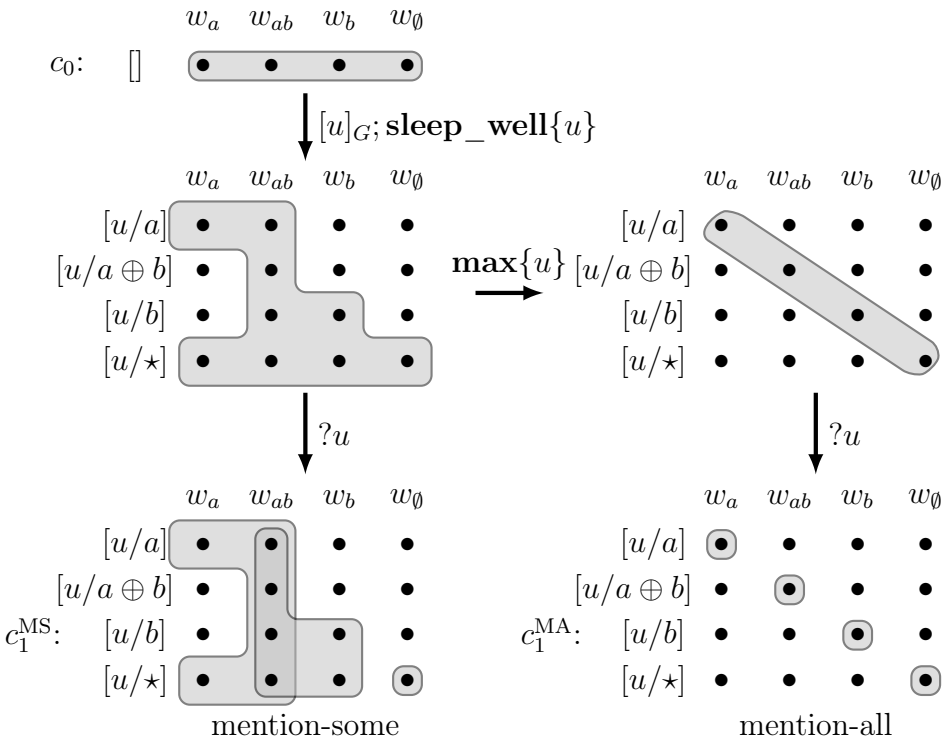


Figure 5.2 – Topmost local context output of the denotation of generic " *Who slept well?*" under mention-some and mention-all reading.



Figure 5.3 – Common ground output of the denotation of generic "*Who slept well?*" under mention-some and mention-all reading.

The denotation of $\mathcal{U}$ from (29-b) is displayed in Fig. 5.2, with the mention-some reading in the bottom left-hand corner and the mention-all reading in the bottom right-hand corner. As expected, the alternative encoding the null answer only contains the possibility $\langle w_\emptyset, [u/\star]\rangle$, both in mention-some and mention-all readings. [14]

On a macro-context $\langle c_0\rangle$, if the update (29-b) is successful, it produces a macro-context as illustrated in equation 5.9 (either mention-some MS or mention-all MA, depending on the exhaustivity interpretation of who_G). The representation of c_0' is displayed in Fig. 5.3, under the respective MS or MA reading. No discourse referent is defined in c_0', but inquisitiveness is projected as expected. We recover the same contexts as in inquisitive semantics for mention-some and mention-all questions.

$$\langle c_0\rangle \xrightarrow{[\![\text{Who}_G \text{ slept well?}]\!]^{\text{MS/MA}}_{\text{MDInq}}} \left\langle \begin{array}{c} c_1^{\text{MS}} \\ c_0'^{\text{MS}} \end{array} \right\rangle / \left\langle \begin{array}{c} c_1^{\text{MA}} \\ c_0'^{\text{MA}} \end{array} \right\rangle \tag{5.9}$$

5.4.2 Modeling specific co-reference

Now, I turn to specific questions.

Specificity condition

Von Heusinger 2002 suggests modeling a specific NP_u by letting the value of u depend on the value of another dref v. In most cases, v is the speaker and u is absolutely specific. As explained in section 2.4.1, von Heusinger's definition cannot work as such in our case because the value of a wh-referent can change depending on the world, i.e., the identity of u is not entirely known. Letting this dependency include the evaluation world as an additional factor–that is, using an individual concept–leads to definition 5.4.1. The set of discourse referents defined in context c is $\mathbf{rdom}\, c := \bigcup_{s\in c} \mathbf{rdom}\, s$.

14. Note that the positive mention-some alternatives now contain a possibility with assignment $[u/\star]$. This situation is not a problem because these possibilities cannot be expressed linguistically. A proposition like $\mathbf{sleep_well}(\star)$ is a tautology. Therefore, it does not answer the question. The question can only be answered by eliminating worlds, not by eliminating assignments. I refer to Roelofsen and Dotlačil 2023b, § 3.3.11 for more discussion on this subtle point.

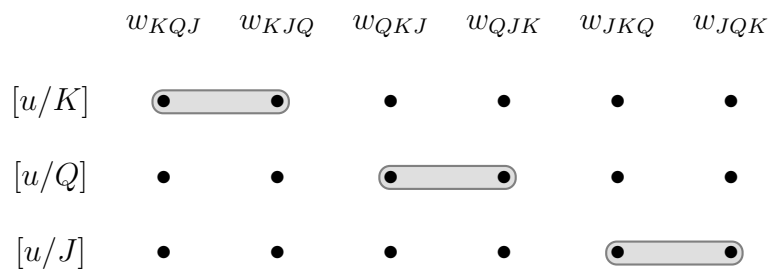
5.4.1. DEFINITION (Specificity condition). An NP_u in a sentence interpreted by a proposition $\mathcal{U}$ with respect to a context c is specific if there is a contextual salient function f and $v \in \mathbf{rdom}\mathcal{U}(c)$, $v \neq u$, such that for all possibility $\langle w, g\rangle$ in $\bigcup\mathcal{U}(c)$, $g(u) = f(w, g(v))$.

5.4.2. EXAMPLE. Consider that there are three cards face down in front of you, one on the left, one in the middle, and one on the right. Among those cards, there is exactly a king K, a queen Q, and a jack J, all of the same suit. The possible worlds are w_{xyz} where xyz is a permutation of KQJ. We assume that the individuals are K, Q, and J. I suppose here that v is the speaker and is fixed.

Question (30) is quantifying over the card figure. The specific identification function is *the card on the left*.

(30) What^u is the card on the left?

The following diagram represents (30) in the example model. We can check that for every possibility $\langle w, g\rangle$, $g(u) = \mathbf{the_card_on_the_left}(w)$.



In most cases, the wh-word is specific with respect to the speaker, i.e. v is the speaker. However, specificity with respect to a dref v introduced in the same utterance $\mathcal{U}$ as u is possible. Examples include universally quantified questions. In (31), for instance, the speaker might not know the identity of the presents that the students bought, but for every student $g(v)$, she can refer to *the* present $g(u) = f(w, g(v))$ that $g(v)$ brought in w.

(31) Which^u present did every^v student buy?

Type-shifting non-modal utterances

As presented in Chapter 4, indicative utterances produce a covert operator INDIC to type-shift their interpretation from type $k \to k$ (context update function) to type $\ell \to \ell$ (stack update function). I assume that when a FocP is not marked with the feature [+generic], it generates a default operator INDIC above Spec FocP. This operator can be replaced by a raising modal operator.

Coherence condition

The identification function f can be used in subsequent indicative assertions to refer back to the wh-dref, but only when the new information is consistent with this individual concept. In particular, if f is determined by the question scope, like in "*Who is Bob dating?*", then a follow-up assertion uttered by the same speaker must be coherent with this question.

To model this sensitivity to discourse coherence, I assume that cross-sentential reference to individual concepts introduced in questions is only possible when there is a defined coherent discourse relation between the two utterances. For example, the second discourse unit in (32-b) elaborates on the established information that Bob is dating someone by judging this romantic relationship. Individual f is a lucky girl because she is dating Bob. Thus, *she* can felicitously refer to f. On the contrary, Bob's girlfriend being a doctor has nothing to do with her dating Bob, so (32-a) is unacceptable out of the blue.

(32) a. A: Who^f is Bob dating? $\overset{?}{\rightarrow}$ #She$_f$ is a doctor.
b. A: Who^f is Bob dating? $\xrightarrow{\textsf{ELAB}}$ She$_f$ is a lucky girl.

5.4.3 Modeling answers

Pragmatics of answers

By responding to a mention-all question with a non-null answer, the responder asserts that the set of true witnesses is nonempty and specifies the maximal true value of the wh-dref. The phrase instantiating a wh-word in this mention-all answer is co-indexed with the wh-word. This co-indexation also allows us to model short answers without ellipsis.

Answers to modalized questions typically agree in modality with the question. Short answers are evaluated in the local context provided by the question. Because of that, the wh-dref and the instantiation dref might stay inaccessible from the common ground even after the question. To demonstrate that, take dialogue (33).

(33) a. A: What^u creature might walk in?
b. B: A$_{u'}$ wolf, a tiger$_{u''}$, or maybe something$_{u'''}$ else. (non-specific)
c. B: #It$_{u/u'/u''/u'''}$ is grey.
d. B: But it$_u$ would eat you first.

The disjunctive short answer (33-b) to (33-a) can have a low scope reading, meaning "*It might be that a wolf, a tiger, or maybe something else walks in*". Thus, the discourse referents u', u'' and u''' need to be evaluated in the local context introduced by *might* in the question. After answer (33-b), this local context is maintained. Responder B cannot refer to any of the creatures that

might walk in in a subsequent indicative sentence (33-c). Modal subordination is still required to talk about the instantiations (e.g. the wolf or the tiger), but also the wh-dref–that is, any creature that might walk in–, as illustrated by (33-d).

As commented in section 5.2.3, answers to specific questions maintain the wh-dref accessible. Therefore, answers to specific or generic questions do not affect the stack. They can be modeled like elaboration.

Short answers are interpreted with a similar modal structure to the questions. If a generic question contributes the modal operator $M : k \to k$, then a short answer whose literal interpretation is $\mathcal{V}$ includes percolation with M. Therefore, their denotation is UPDATE $\mathcal{V}$; PERC $M\,\mathcal{V}$. Short answers to specific questions have no modal structure and are simply type-shifted using UPDATE $\mathcal{V}$, like indicative clauses. [15]

Examples

To illustrate the model for short answers, consider generic question (34-a) and mention-all answer (34-b). Their semantic representations are given in (34-c) and (34-d), respectively. Predicate **KHMFCEQ** means *knows how to model free-choice effects in questions*. Vacuous operators are simplified. The evolution of an initial stack $\langle c \rangle$ on these formulas appears in (5.10).

(34) a. A: Who^u knows how to model free choice effects in questions?
b. B: Tomasz$_u$.
c. $\dagger\mathcal{T}$, with $\mathcal{T} = \text{PUSH}\,\mathcal{U}; \text{PERC}\,\Diamond_{\text{G}}\mathcal{U}$, and $\mathcal{U} = [u]_G; \mathbf{KHMFCEQ}\{u\}; ?u$
d. $\mathcal{S} = \text{UPDATE}\,\mathcal{V}; \text{PERC}\,\Diamond_{\text{G}}\mathcal{V}$, with $\mathcal{V} = \mathbf{name}\{u, \text{“}\textit{Tomasz}\text{”}\}$

As a short answer, (34-b) is modeled with a modal structure copy. Therefore, $\mathcal{V}$ is evaluated in $\mathcal{U}(c)$, where u is defined, and u is still not accessible in the common ground in the output stack. Given that *Tomasz* is a proper name, u should also percolate to the common ground. But implementing this kind of global accommodation is not the focus of this dissertation.

$$\langle c \rangle \xrightarrow{\dagger\mathcal{T}} \left\langle \begin{array}{c} \mathcal{U}(c) \\ \Diamond_{\text{G}}\mathcal{U}(c) \end{array} \right\rangle \xrightarrow{\text{UPDATE}\ \mathcal{V}} \left\langle \begin{array}{c} \mathcal{V}(\mathcal{U}(c)) \\ \Diamond_{\text{G}}\mathcal{U}(c) \end{array} \right\rangle \xrightarrow{\text{PERC}\ \Diamond_{\text{G}}\mathcal{V}} \left\langle \begin{array}{c} \mathcal{V}(\mathcal{U}(c)) \\ \Diamond_{\text{G}}\mathcal{V}(\Diamond_{\text{G}}\mathcal{U}(c)) \end{array} \right\rangle \quad (5.10)$$

Now consider specific question (35-a) and its answer (35-b). Their denotations are shown in (35-c) and (35-d) respectively. Their stack effects are displayed in equation (5.11).

(35) a. A: Who^u was sitting here?
b. B: Constance$_u$.

15. Technically, DEC also contributes ! on top of these formulas. See section 6.6.1 for a discussion on that.

c. $\dagger\mathcal{T}$, with $\mathcal{T} = \text{UPDATE}\ \mathcal{U}$, and $\mathcal{U} = [u]; \mathbf{sitting_here}\{u\}; ?u$
d. $\mathcal{S} = \text{UPDATE}\ \mathcal{V}$, with $\mathcal{V} = \mathbf{name}\{u, \textit{“Constance”}\}$

$$\langle c \rangle \xrightarrow{\dagger\mathcal{T}} \langle \mathcal{U}(c) \rangle \xrightarrow{\mathcal{S}} \langle \mathcal{V}(\mathcal{U}(c)) \rangle \tag{5.11}$$

5.5 Conclusion

In this chapter, I proposed a theory of wh-anaphora–anaphora with wh-words as antecedents in matrix questions. This theory establishes three ways to refer back to an interrogative word. Modal subordination is required when the question is generic. The discourse referents of wh-words in specific questions are accessible in subsequent utterances, provided there is discourse coherence. Finally, pronominal mention is possible in answers. I modeled these phenomena using Modal Dynamic Inquisitive Semantics, extending it to any question type.

I also suggested reconsidering wh-anaphora annotation in two ways. Wh-anaphora should be modeled as a structured anaphora. Moreover, instantiations of the wh-word in mention-all answers should be annotated as co-reference and not bridging anaphora.

Of course, various other pragmatic effects must be considered in the felicity of wh-anaphora. For example, I omitted discussions about gender and number. But I hope this study is a first step towards better comprehending the wide range of anaphoric properties of wh-words.

Chapter 6

Dynamic and Compositional Aspects of Adjunct Dependence Utterances

6.1 Introduction

A dependence statement is an assertion expressing that a question's resolution depends on another question's resolution (Ciardelli 2018). For example, sentence (1) says that if it is sunny, Amy will go to the beach. Otherwise, she will not.

(1) Whether Amy will go to the beach depends on whether it is sunny.
(Hui 2021)

French can express dependence statements using an adjunct clause, called an interrogative-based conditional (de Cornulier 2013; Richard 2024a). This adjunct clause comprises a (complex) preposition and an interrogative. The preposition can either be *selon* (lit. '*in accordance with*'), *suivant* (lit. '*following*') or *en fonction de* (lit. '*in function of*'). Henceforth, I refer to any of these interchangeable variants as SELON. English has a construction with the same meaning using *depending on*.[1]

(2) des jeunes des jeûnes [...] ça peut être paronyme ou homonyme
some youngs some fasts [...] it can be paronymous or homonymous
[suivant comment vous le prononcez]
[SELON how you it pronounce]

1. Note that there exists another use of *selon* and *suivant* that can be translated by *according to*, and that expresses the source of some information, e.g (i). This reading is impossible with *en fonction de*. I thus assume that this use of *selon/suivant* is independent and I do not focus on it here.

(i) Selon les prévisions météo, le mercure va baisser à partir de mardi. (frTenTen23)
'*According to the weather forecast, the mercury will drop from Tuesday.*'

'*French words "jeune" and "jeûne" can be paronymous or homonymous depending on how you pronounce "jeûne".*' (CIENSFO)

In (2), the speaker highlights that *jeûne* has two pronunciation variants, $A_1 = /ʒœn/$ (like for *jeune*) or $A_2 = /ʒøn/$. Words *jeune* and *jeûne* can either be in a homonymy relationship (B_1) or a paronymy relationship (almost the same pronunciation, B_2). Richard 2024a explains that (2) expresses the following: which alternative B_i holds is determined by which alternative A_j holds; in this case: if A_1 then B_1 and if A_2 then B_2.

Interrogative-based conditionals express a complex conditional dependence. Like in a regular conditional *if A, then B*, I call antecedent A the clausal complement of SELON and consequent B the main clause. The adjunct clause [SELON A] can be syntactically before or after the consequent without changing the meaning, which I write $\mathbf{selon}(A, B)$.

I call *adjunct dependence utterance* (ADU) any utterance expressing such a conditional dependence using an adjunct phrase. As I will show below, ADUs include French interrogative-based conditionals and English constructions "*Depending on Q, p*" (M. Kaufmann 2016), but also other variants, like ADUs based on concealed question or inquisitive ADUs. I use the notation $\mathbf{selon}(A, B)$ for the interpretation of any ADU.

The meaning of ADUs seems to be captured by the meaning given to *depend on* by Theiler, Roelofsen, and Aloni 2019 and Hui 2021. If $\mathbf{dep}(B, A)$ represents the denotation of *B depends on A*, then $\mathbf{selon}(A, B) = \mathbf{dep}(B, A)$, as glossed in (2).

However, the definition suggested by Theiler, Roelofsen, and Aloni 2019 and Hui 2021 has some limits regarding the dynamic properties of ADUs and regarding the compositional variation exhibited by SELON structures in French. I aim to improve their definition and provide a model for dependence utterances in Modal Dynamic Inquisitive Semantics. Although I only provide a syntactic analysis for French, a similar structure can be extrapolated for English. The semantic and pragmatic analyses concern both French and English.

Throughout this Chapter, I take occurrences from corpora whenever possible. I use corpus CIENSFO (Richard 2024a), which contains spoken extracts from various online sources. I also use `frTenTen23`, a French corpus of written web (Jakubíček et al. 2013).

Section 6.2 presents the lexical entry for *depend on* given by Theiler, Roelofsen, and Aloni 2019. Then, in section 6.3, I provide data showing that this definition is insufficient and needs to be adapted. I exhibit three problems that I solve in the rest of the chapter. In section 6.4, I present my compositional theory of SELON constructions. I show that by letting the adjunct clause attach at FOC, we can derive the correct alternatives of declarative consequents. Section 6.5 explores the anaphoric properties of wh-words in the antecedent and accessibility in follow-up sentences. In section 6.6, I explain how SELON can relate to an embedded or

matrix interrogative in the consequent by generalizing its lexical entry. Section 6.7 concludes.

6.2 Dependence statements in inquisitive semantics

Saying that question B depends on question A amounts to saying that there exists a functional relation between A's alternatives and B's alternatives.[2] That is, there is a function $f : \textsc{alt}(A) \to \textsc{alt}(B)$ such that for every proposition $t \in \textsc{alt}(A)$, if t holds then $f(t)$ holds.

Conditionals have a modal flavor and can be expressed with the formula $\Box(A \to B)$ (Gillies 2004). Similarly, dependence statements are modal statements (Karttunen 1977; Ciardelli 2018). The relation between the resolution conditions is not evaluated in the actual world but in possible worlds respectively to some modal base. A modifier clause can provide this modal base. For example, (3) holds in the actual world if there exists a function f from age to income tax rate such that, for every world w where the Dutch law applies and some people x is y years old in w, x has income tax rate $f(y)$ in w.

(3) According to Dutch law, one's income tax rate depends on one's age.
(Theiler, Roelofsen, and Aloni 2019)

Let $\mu : s \to (s \to t)$ be a modal base function. Theiler, Roelofsen, and Aloni 2019's lexical entry for *depend on* is given in (4). It is slightly modified to fit the denotation of inquisitive semantics, i.e. issues are evaluated with respect to sets of worlds rather than possible worlds.

(4) a. $[\![\text{depend on}]\!]_{\mathsf{Inq}} := \lambda\mathcal{I}, \mathcal{J}, \lambda s.\ \forall w \in s.$
$\exists f : \textsc{alt}(\mathcal{I}) \to \textsc{alt}(\mathcal{J}).\ \mathbf{fmd}(f, \mu(w))$,

b. with $\mathbf{fmd}(f, M)$ holding if and only if

(i) **conditional dependence:**
$\forall w \in M.\ \forall t \in \mathbf{dom}\, f.\ w \in t \to w \in f(t)$, and

(ii) **non-triviality:**
$\exists t, t' \in \mathbf{dom}\, f.\ t \cap M \neq \emptyset \wedge t' \cap M \neq \emptyset \wedge f(t) \neq f(t')$

As explained above, $[\![\text{depend on}]\!]$ asserts that there is some functional modal dependence $\mathbf{fmd}(f, \mu(w))$ between the antecedent $\mathcal{I}$ and the consequent $\mathcal{J}$. The first condition (4-b-i) requires that, for every evaluation world w in the modal base M, if alternative t of $\mathcal{I}$ is true at w then $f(t) \in \textsc{alt}(\mathcal{J})$ is also true at w. The second condition (4-b-ii) eliminates trivial dependencies (i.e. constant functions

2. To simplify matters, we focus on complete dependence here, excluding cases where a question's resolution might partially depend on another question's resolution. For a discussion, see Karttunen 1977, fn. 6.

f) by requiring that at least some inputs t, t' intersecting with the modal base are mapped to different outputs $f(t) \neq f(t')$.

Condition (4-b-ii) impacts the semantic objects selected by *depend on*. It forbids issues $\mathcal{I}$ and $\mathcal{J}$ to be assertive, i.e. they must be inquisitive. This prediction correctly captures the syntactical restrictions of *depend on*. Verbs of dependency neither accept declarative clausal subjects nor declarative clausal objects (5).

(5) a. *That the light is on depends on whether the switch is up.
b. *Whether the light is on depends on that the switch is up.

6.3 Limits of previous models

Theiler, Roelofsen, and Aloni's model does not consider the anaphoric properties of referents introduced in the antecedent. In section 6.3.1, I show that they are accessible in the consequent but not in a follow-up indicative sentence. However, they can be accessed by modal subordination.

Section 6.3.2 raises compositional issues. Contrary to *B depends on A*, most adjunct dependence utterances (ADUs) SELON *A, B* contain a declarative clause as consequent B. But some items in B can trigger inquisitiveness. I first discuss how this inquisitiveness can be captured in Modal Dynamic Inquisitive Semantics. Then, I turn to cases where the main clause B is or contains an interrogative clause. I show that by positing a different modifier clause attachment site and generalizing the lexical entry for SELON, I can also model these cases within my framework.

6.3.1 Dynamic issues

Conditionals have a particular anaphoric accessibility pattern (Kamp 1981). A discourse referent introduced in the antecedent can be referred to in the consequent but not in a subsequent sentence, except with modal subordination. Here, we show that ADUs have the same anaphoric accessibility pattern.

Wh-donkey anaphora

Nishigauchi 1986, 1990 shows that a pronoun can refer to the discourse referent introduced by a wh-word if the wh-word is in an adjunct clause and the pronoun is after in the main clause. This is possible with *no matter + wh* concessive clauses, e.g. (6-a). The anaphora in (6-a) is a donkey anaphora because neither *who*^{u} nor its trace has the pronoun *him*$_u$ in its scope.

(6) a. No matter who^u comes in, I will kiss him$_u$.

Moreover, Nishigauchi notes that this referent is not accessible higher in the syntactic tree (7-a). It is also not accessible in a follow-up indicative sentence (8-b), but reference is possible with a modal operator (8-c).

(7) a. *Mary's claim that no matter who^u comes in she must kiss him$_u$ will bother him$_u$. (Nishigauchi 1990)

(8) a. No matter who^u comes in, we will be happy.
b. #My wife kiss / kissed him.
c. My wife will probably kiss him$_u$.

We observe the same phenomena in ADUs. A wh-referent introduced in the antecedent is accessible in the consequent, e.g u in (9).

(9) En fonction de [$_A$ qui^u tu veux aider] [$_B$ tu dois être précis dans ta communication pour qu'ils$_u$ sentent que c'est à eux$_u$ et à eux$_u$ seuls que tu parles.] (frTenTen23)
'*Depending on [$_A$ who^u you want to help], [$_B$ you must be precise in your communication to let them$_u$ feel that you speak to them$_u$ and only them$_u$].*'

Moreover, this wh-referent cannot be referred to in a follow-up indicative sentence (10-b). But it can using modal subordination (10-c). These observations also hold for other indefinites in the antecedent, e.g. see (14) below.

(10) a. En fonction de qui^u tu veux aider, tu dois être précis dans ta comminucation.
'*Depending on who^u you want to help, you must be precise in your communication.*'
b. #Ils$_u$ sentent que c'est à eux$_u$ que tu parles.
'*They$_u$ feel that you speak to them$_u$.*'
c. Ils$_u$ doivent sentir que c'est à eux$_u$ que tu parles.
'*They$_u$ must feel that you speak to them$_u$.*'

To model these properties, we need to keep track of the local context introduced by SELON and the referents that are only accessible in that local context.

Exemplification

Modal subordination also includes reference in a conditional antecedent, e.g. (11-b) following ADU (11-a).

(11) a. Suivant [$_A$ où^u se trouve ta fuite] [$_B$ pas la peine de tout vidanger].
SELON [$_A$ where^u REFL find your leak] [$_B$ not the trouble of all drain]

'*Depending on* $[_A$ *where*^{u} *your leak is*$]$, $[_B$ *there's no need to drain everything*$]$.' (frTenTen23)

b. Si $[_C$ c' est un radiateur [à l' étage$]_u]$, $[_D$ vidange seulement
if $[_C$ it is a radiator [at the floor$]_u]$, $[_D$ drain only
jusqu' au niveau de l' étage].
until at-the level of the floor.]
'*If* $[_C$ *it's an upstairs*$_u$ *radiator*$]$, $[_D$ *drain only up to floor level*$]$.'

The phrase *à l'étage* is an instantiation of the wh-word. However, in this case, clause C is not the true answer to the ADU antecedent A in the actual world, but rather a possible answer. Sentence (11-b) provides more details regarding the dependence introduced by (11-a) by describing what the consequent resolution is for the example antecedent resolution C. I call this an *exemplification*.

Exemplification might occur without any overt conditional. For example, (12-b) completely specifies the unspecified dependence asserted by (12-a): what the two relevant locations are, what the two actual pronunciations of *figatellu* (a dry sausage) are and the function between those two variables.

(12) a. $[_B$ La^v prononciation du figatellu est différente] suivant $[_A$ où^u vous vous situez en Corse]. (frTenTen23)
'$[_B$ *The*^{v} *pronunciation of figatellu is different*$]$ *depending on* $[_A$ *where*^{u} *you are in Corsica*$]$.'

b. $[_C$ Au nord$]_u$ $[_D$ c$_v$'est figadellu] et $[_{C'}$ au sud$]_u$ $[_{D'}$ figateddu].
'$[_C$ *In the north*$]_u$, $[_D$ *it*$_v$*'s figadellu*$]$ *and* $[_{C'}$ *in the south*$]_u$ $[_{D'}$ *figateddu*$]$.'

The theory of ADUs I present models anaphora in exemplification sentences.

6.3.2 Compositional issues

Declarative consequents

Condition (4-b-ii) in the lexical entry for *depend on* imposes that its consequent must be inquisitive. However, all ADUs provided by Richard 2024a, and almost all examples found in `frTenTen23` have a declarative consequent. Despite this apparent contradiction, I argue that all attested declarative consequents do raise alternatives. Alternatives can be raised by many elements, including a disjunction, e.g. (2) (whose translation is repeated in (13)) or an indefinite adjective, e.g. *certain* in (14).

(13) a. '$[_B$ *French words jeune and jeûne can be paronymous or homonymous*$]$ *depending on how you pronounce jeûne.*'

b. B_1 = it is homonymous, B_2 = it is paronymous

(14) a. parce que on projette le fait que selon comment une^{v} personne parle [$_{B}$ elle$_{v}$ aurait ‡elle aurait une certaine identité sociale] (CIENSFO)
'*Because we project the idea that, depending on how someone^{v} speaks, [$_{B}$ they$_{v}$ would have a certain social identity].*'

b. e.g. B_1 = they would be a professor, B_2 = they would be an entrepreneur, B_3 = they would be a laborer,...

The consequent's alternatives can also be left implicit. For example, in (11-a), the alternatives of the consequent are B_1 = '*you don't need to drain everything*' (the surface clause) and B_2 = '*you need to drain everything*' = $\neg B_1$.

The model I present in section 6.4 accounts for this variation by attaching SELON adjunct clause at the right place on the left periphery.

The antecedents of ADUs also exhibit some syntactic variation. An antecedent can be an interrogative (polar, alternative or constituent (12-a)), a declarative CP disjunction (a disjunction of complementizer clauses or a complementizer clause with high disjunction (15-a)) [3] or a concealed question (15-b).

(15) a. Selon [$_{A}$ qu' il s' agit d' un permis A2, A1 ou A], le
SELON [$_{A}$ that it REFL is_about of a license A2, A1 or A], the
coût n' est pas le même.
cost NEG is not the same.
'*Depending on whether it's an A2, A1 or A license, the cost is not the same.*' (frTenTen23)

b. Selon [$_{A}$ le type d'investissement], les possibilités de soutiens varient. (frTenTen23)
'*Depending on [$_{A}$ the investment type], support options vary.*'

These variations can also be accounted for in Modal Dynamic Inquisitive Semantics.

Interrogative consequents

We observe that questions can also receive a SELON adjunct phrase. By similarity, we call them inquisitive ADUs. Inquisitive ADUs have an interrogative clause as consequent. For example, (16-a) asks what the set of relevant age periods is and how many eggs it is recommended to eat in those age periods. In other words, the possible answers are the functions f from age periods to diet recommendations. Question (16-a) is used in the title of a section whose content (translated in (16-b)) is the answer to that question.

3. Even if this is not standard French, examples with *en fonction* and CP disjunction were found in frTenTen23:

(i) En fonction que vous êtes expatrié dans un pays de l'Union européenne ou dans un autre pays étranger, la fiscalité sur les plus-values immobilières françaises n'est pas la même.

(16) a. Selon les âges, [$_B$ combien d'œufs peut-on consommer] ? (frTenTen23)
'*At different ages, [$_B$ how many eggs can you eat]?*'
b. '*Before the age of 1: half an egg a day, for example. Up to the age of 10: one egg a day. From age 10 upwards: children and teenagers can eat as many eggs as adults if they wish.*'

Sentence (16-a) also presupposes that such a non-trivial functional relation exists. Therefore, condition (4-b-ii) in the semantics of SELON appears to be a necessary condition for the felicity of inquisitive ADUs.

To model (16-a) we use the concept of inquisitive hole defined by Theiler 2014 while adapting it slightly. The formula **selon**(A, B) is an inquisitive hole with respect to B, that is:

— If the consequent is declarative, then the whole utterance is assertive

— But if the consequent is interrogative, the whole utterance is inquisitive

This property is also needed when the main clause contains an embedded interrogative. There are sentences, like (17-a), where the adjunct clause can somehow attach below the attitude verb and, together with the embedded interrogative, provide an inquisitive proposition feeding the attitude verb.

(17) a. en fonction de [$_A$ combien il te reste d' éléments dans
SELON [$_A$ how_many EXPL to_you remain of elements in
ton objet] tu peux savoir [$_B$ depuis combien de temps il est
your object] you can know [$_B$ since how_much of time it is
en train de se dégrader]
PROG REFL degrade]
'*Depending on [$_A$ how many radioactive elements you have left in your object], you can find out [$_B$ how long it has been decaying].*' (CIENSFO)
b. #'*Whether you can know the decaying duration depends on the element count.*'
c. #'*You can know that the decaying duration depends on the element count.*'
d. '*You can know what the dependence relation f is between the element count and the decay duration.*'

Assertion (17-a) does not mean what we could expect if we took "*tu peux savoir B*" as the whole consequent, that is (17-b). It has to be construed with the SELON clause modifying the embedded interrogative B. However, (17-a) does not just mean that you can know that some functional dependence exists between those two variables, i.e. (17-c). This is too weak. The statement (17-a) means that you can know what this relation is, i.e. (17-d). This is confirmed by the deduction in (18-a).

(18) a. - En fonction du nombre d'atomes de carbone 14, Marie sait quel est

SELON A, B	*B* is declarative	*B* is interrogative
A is interpreted as inquisitive	Assertive ADU	Inquisitive ADU
A is interpreted as assertive	*	*

Table 6.1 – Types of Adjunct Dependence Utterances according to the semantic interpretation of the antecedent *A* and the syntactic form of the consequent *B*. * = ungrammatical

l'âge de l'objet.
- Elle sait que l'objet a 8×10^{20} atomes de carbone 14.

Donc elle sait quel est l'âge de l'objet.

- '*Based on the number of carbon-14 atoms, Marie knows how old the object is.*'
- '*She knows that the object has* 8×10^{20} *carbon-14 atoms.*'

'*Therefore she knows how old the object is.*'

Inquisitive ADUs are also possible in English, as illustrated by (19).

(19) Depending on the outcome, what will be the regional and national repercussions of this initiative? (enTenTen21)

To explain this phenomenon, I posit that the SELON adjunct phrase can attach locally above the CP of interrogatives. Like existential modalities above FOC, it propagates inquisitiveness.

6.3.3 Interim Summary

In this section, I showed that ADUs have the same anaphoric accessibility pattern as conditionals. Moreover, the grammaticality of an ADU is determined by the semantic interpretation of its antecedent. Finally, two types of ADUs exists. If *B* is a declarative clause, *SELON A, B* is a statement. However, if *B* is an interrogative clause, *SELON A, B* is a question. Table 6.1 summarizes this distribution.

Sections 6.4, 6.5 and 6.6 present how Modal Dynamic Inquisitive Semantics can model this distribution by adapting Theiler, Roelofsen, and Aloni 2019's definition and keeping a single lexical entry for SELON.

6.4 Inquisitiveness triggers

I start by tackling the compositional issue of antecedents and declarative consequents.

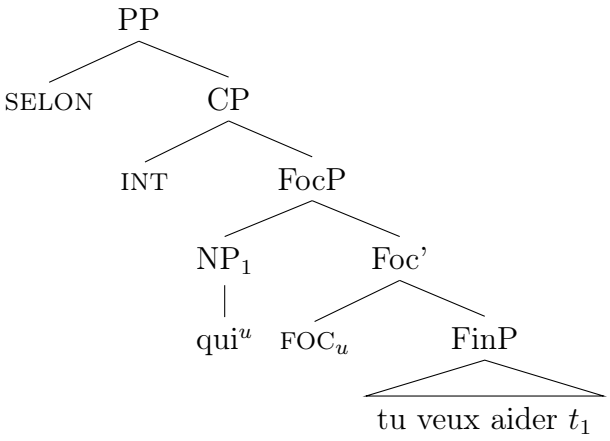


Figure 6.1 – Logical form of a SELON adjunct clause with a wh-interrogative antecedent.

6.4.1 Inquisitiveness in the antecedent

The antecedent A of an adjunct dependence utterance (ADU) $\mathbf{selon}(A, B)$ may be an embedded interrogative. This interrogative may be a polar, alternative, or a wh-interrogative. [4]

According to Theiler, Roelofsen, and Aloni's formula, the non-triviality condition (4-b-ii) requires the complement of SELON to be interpreted by an inquisitive issue, otherwise leading to a contradiction. Alternative questions are derived by letting disjunction $\sqcup$ scope over FOC, thus generating inquisitiveness. The inquisitiveness of polar questions is triggered by $\langle ? \rangle$, contributed by the INT operator in the Type Phrase. Therefore, the complement of SELON must be a Type Phrase. In the following, I prefer using the term Complementizer Phrase (CP) instead fo TypeP whenever I deal with embedded clauses to emphasize that embedded clauses require a complementizer in French. The resulting logical form is displayed in Fig. 6.1.

Declarative sentences are interpreted by assertive issues. In DInq, DEC provides !, thus ruling out any inquisitiveness triggered below, even FocP-level disjunction. Roelofsen and Dotlačil 2023b posit that disjunction over DEC is impossible. However, examples like (15-a) suggest that disjunction inside a declarative CP can raise alternatives for SELON.

Example (20) sheds light on the issue. In example (20), disjunction scopes over the complementizers *que* ('*that*'). ADUs like (15-a) can be rephrased with

4. I only consider single-wh interrogatives here, although ADUs with multiple interrogative words are possible, e.g. (i).

(i) selon qui vend quoi, les prix diffèrent grandement (frTenTen23)
'*Depending on who sells what, prices differ widely.*'

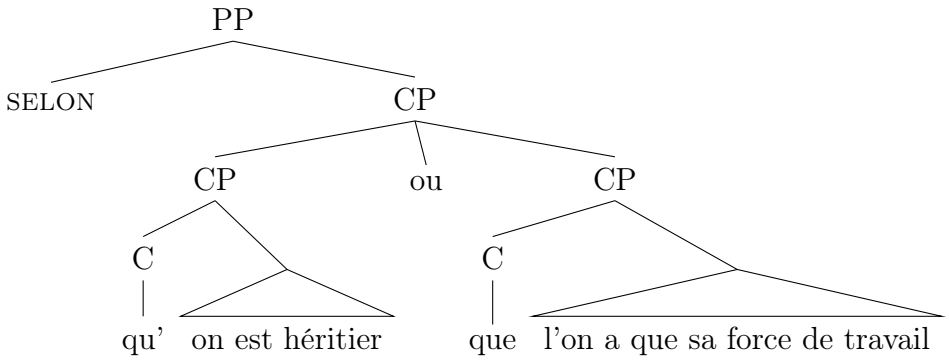


Figure 6.2 – Logical form of a SELON adjunct clause with a declarative CP disjunction.

a similar *que*-clause repetition. I analyze this data as proof that disjunction over DEC is possible in French, and that, as predicted by DInq, it triggers inquisitiveness. [5] This logical form is depicted in Fig. 6.2. Complementizer *que* replaces DEC and has the same denotation.

(20) Cet effort n' est pas le même selon [$_A$ qu' on est héritier ou que
this effort NEG is not the same SELON [$_A$ that one is heir or that
l'on a que sa force de travail].
one has only one's force of work].
'*This effort is not the same for heirs as it is for those who only have their labor power.*' (frTenTen23)

Finally, NP antecedents can be semantically lifted to an inquisitive proposition by analyzing them as concealed questions. Modeling concealed questions in DInq would take me too far. I refer to Aloni and Roelofsen 2011 and Hui 2021 for an analysis in inquisitive semantics.

5. In French, CP disjunction can also be used as a concessive construction (i-a) or with an irrelevance verb (i-b).

(i) a. Que vous accueillez‡ un nouveau salarié ou que l' un d' entre eux
that you welcome.SJV a new employee or that the one of between them
doive vous quitter, nous sommes à vos côtés.
must.SJV you leave, we are at your sides.
'*Whether you're welcoming a new employee or one of your staff is leaving, we're here to help.*' (frTenTen23)

b. La loi se rappellera à vous, car elle se fout que vous
the law REFL call.FUT to you because she REFL not_give_a_damn that you
ayez été payé ou non.
have.SJV been paid or not
'*The law will remember you, because it doesn't care whether you've been paid or not.*'
(frTenTen23)

6.4.2 Inquisitiveness in the consequent

The data presented in section 6.3.2 suggests that SELON has different behaviors depending on the clause type of the consequent. Let me focus here on ADUs **selon**(A, B) that produce assertions. They require that the consequent B is declarative.

Free Association with focus

Free Association with focus (Beaver and Clark 2008) describes how the implicit modal base of some quantificational operators is provided as the union of the alternatives raised via focus. For example, sentence (21-a) means that, in all cases where Kim tells Sandy something, that thing is *to be nice*. The alternatives triggered by focussing *to be nice* are (21-a-i). Their union forms the modal base $M = \bigcup_i B_i$, which restricts the quantificational adverb. Sentence (21-a) is represented by formula **always**(M, B_1). In contrast, the modal base of (21-b) consists of the situations where Kim tells somebody to be nice.

(21) a. Kim always tells Sandy to be nice$_F$.
   (i) e.g. B_1 = Kim tells Sandy to be nice, B_2 = Kim tells Sandy to be strong, B_3 = Kim tells Sandy to be beautiful...
   b. Kim always tells Sandy$_F$ to nice.
   (i) e.g. B_1 = Kim tells Sandy to be nice, B_2 = Kim tells Mary to be nice, B_3 = Kim tells Anne to be nice,...

In the corpus CIENSFO, various inquisitiveness triggers were observed, including disjunction, indefinites, numerals, comparative terms (e.g. (12-a)), scale terms, and variation terms. Disjunction and indefinites are traditional inquisitiveness triggers, but not the others.

M. Kaufmann 2016 proposes that, in the constructions "*Depending on Q, p*", p raises "live issues". Hénot-Mortier 2024b assimilates these "live issues" with the maximal true answers to the question under discussion evoked by p. This question under discussion can be interpreted as a question inquiring about p's focused material (Hénot-Mortier 2024a). I follow this analysis and consider that the inquisitive triggers mentioned are due to focus.

I argue that SELON works as a quantificational adverb using Free Association with focus. Focus in the consequent B generates a set $\mathcal{A}$ of alternatives. When the modal base is not explicitly expressed, the union of these alternatives determines the modal base $\mu(w) = \bigcup \mathcal{A}$ of **selon**$_\mu(\mathcal{I}, \mathcal{J})$ used in definition (4) (and in the subsequent sections). [6] Furthermore, these alternatives constitute an issue $\mathcal{J} = \bigcup_{p \in \mathcal{A}} \wp(p)$ interpreting the consequent of the ADU.

6. To simplify matters, we can assume that the modal function μ is constant so that $\mu(w) = M$ for any w.

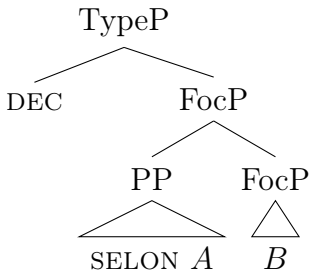


Figure 6.3 – Logical form of an assertive adjunct dependence utterance $\mathbf{selon}(A, B)$.

Many assertive ADUs contain a modal verb or adverb in their consequent, like *might* in (22). Sentence (22) might be read as implying "*If the round goes badly, I might stop (but I might also stay)*". However, it can also be read as implying "*If the round goes badly, I stop, and this situation (where the round goes badly and I stop) is possible*". Under this reading, the modal operator *might* shares its modal base μ with **selon** and indicates the epistemic stance of the surface alternative of the consequent with respect to this modal base. For example, (22) has the at-issue inference that, among the situations wherein the speaker stops or continues, his stopping is possible, i.e. $\mathbf{might}_{\mu}(p)$, with p the intensional denotation of B.

(22) suivant comment se passe le premier tour [$_B$ je vais peut-être arrêter] (CIENSFO)

'*Depending on how the first round goes, [$_B$ I might stop].*'

Structure of assertive ADUs

Modeling focus in (Modal) Dynamic Inquisitive Semantics would take me too far. Therefore, I will restrain illustrations to simple cases (disjunction, indefinites and implicit polar interpretation) and modeled them as if they were triggering "regular" inquisitiveness at FocP. Fig. 6.3 displays this logical form.

Based on the ADU examples collected, it seems that the antecedent and the consequent of ADUs trigger an existential presupposition when they are wh-interrogatives. Therefore, I assume that the interrogative antecedent and the interrogative consequent of ADUs both have a specific reading. I use the notation $\mathbf{selon}(\mathcal{U},\mathcal{U}')$ for the semantic contribution of SELON, which I give in section 6.5. Viewing SELON as a complex modal operator, I assume that **selon** performs stack management, consistently with the model given for modalities in Chapter 4. Therefore, no type-shift function is needed at the FocP of the antecedent or consequent.

Examples

The logical form of sentence (2), repeated in (23-a), is given in (23-b), and its derived semantic representation is written in (23-c). I added context information to clarify the interpretation of pronouns. The consequent is made of a FocP-level disjunction, like in alternative questions. Operator $\langle ? \rangle$ on top of $\mathcal{V}$ is explained below.

(23) a. '*⌊B It$_z$ can be paronymous or homonymous⌋ depending on how^u you pronounce it$_z$.*'
b. DEC ((SELON A) B), with
(i) A = INT (FOC$_u$ (*how^u you pronounce it$_z$*))
(ii) B = *can* (it$_z$ be paronymous or homonymous)
c. $\mathbf{word}\{v, \text{“}jeune\text{”}\}; \mathbf{word}\{z, \text{“}jeûne\text{”}\}; !\mathbf{selon}(\mathcal{U}, \langle ? \rangle \mathcal{V})$, with
(i) $\mathcal{U} = [u]; \mathbf{manner}\{u\}; \mathbf{pronounce}\{z, u\}; ?u$, and
(ii) $\mathcal{V} = \Diamond(\mathbf{paronymous}\{z, v\} \sqcup \mathbf{homonymous}\{z, v\})$

Now, let me turn to sentence (14) (whose translation is repeated in (24-a)). To parallel wh-phrases, I assume that focus raises *a^z certain social identity* to Spec FocP and co-indexes FOC with z, contributing a witness request operator $?z$. Like wh-words, this phrase is then interpreted after reconstruction. The semantic representation of (24-a) is given in (24-b).

(24) a. '*Depending on how^u someone^v speaks, ⌊B they$_v$ would have a^z certain social identity⌋.*'
b. $!\mathbf{selon}(\mathcal{U}, \langle ? \rangle \mathcal{V})$, with
(i) $\mathcal{U} = [u]; \mathbf{manner}\{u\}; [v]; \mathbf{speak}\{v, u\}; ?u$, and
(ii) $\mathcal{V} = \Box([z]; \mathbf{have}(v, z); \mathbf{social_id}\{z\}; \mathbf{certain}\{z\}; ?z)$

If there is no inquisitiveness trigger, the consequent B of an ADU is interpreted as a polar issue *B or not B*. For example, in (11-a) (whose translation is repeated in (25-a)), the implicit alternative is *not B* = "*it is necessary to drain everything*". I take this as the default behavior of ADUs.

I argue that SELON appends a $\langle ? \rangle$ operator on its consequent, viz. (6.1). If the consequent is assertive, it creates a negative alternative. If it is already inquisitive, $\langle ? \rangle$ is vacuous, e.g. in (23-c) and (24-b).

$$\begin{array}{ll} [\![\text{SELON } A, B]\!] = & \mathbf{selon}([\![A]\!], \langle ? \rangle [\![B]\!]) \\ [\![B, \text{SELON } A]\!] = & \mathbf{selon}([\![A]\!], \langle ? \rangle [\![B]\!]) \end{array} \tag{6.1}$$

The semantic representation of (25-a) is given in (25-b).

(25) a. '*Depending on where^u your$_v$ leak is, ⌊B there's no need to drain everything^z⌋.*'
b. $!\mathbf{selon}(\mathcal{U}, \langle ? \rangle \mathcal{V})$, with
(i) $\mathcal{U} = [u]; [v]; \mathbf{leak}\{v\}; \mathbf{loc}\{v, u\}; ?u$, and
(ii) $\mathcal{V} = \neg\Box([z]; \mathbf{max}\{z\}; \mathbf{drain}(z))$

The formulas provided here work with an adaptation of Theiler, Roelofsen, and Aloni 2019's model for dependence statements to DInq. In particular, by updating $\mathcal{V}$ in a context previously updated with $\mathcal{U}$, formula (24-b) correctly lets v introduced in $\mathcal{U}$ be referred to in $\mathcal{V}$. However, it cannot predict that v is accessible in a subsequent utterance under modal subordination. The following section addresses this problem.

6.5 Modeling assertive dependence utterances

In this section, I provide a lexical entry for SELON that takes its anaphoric properties into account. Let me first adapt Theiler, Roelofsen, and Aloni 2019's definition (4) to DInq.

Given a modal base function $\mu : s \to (s \to t)$, $\mathbf{selon}_\mu(\mathcal{U}, \mathcal{V})$ is a context update function. On input context c, it outputs c restricted to the information states s that resolve the conditional dependence $\Rrightarrow_\mu$ between $\mathcal{U}(c)$ and $\mathcal{V}(\mathcal{U}(c))$. This conditional dependence, given in (6.2), is the existence of a function between $\mathcal{U}(c)$'s alternatives and $\mathcal{V}(\mathcal{U}(c))$'s alternatives satisfying the functional modal dependence $\mathbf{fmd}(f, M)$ defined above in (4).

The two definitions in (6.2) will be refined in equations (6.3) and 6.8 respectively.

$$\begin{array}{rcl} \mathbf{selon}_\mu := & \lambda\mathcal{U}, \mathcal{V}.\ \lambda c, s.\ s \in c \wedge (\mathcal{U}(c) \Rrightarrow_\mu \mathcal{V}(\mathcal{U}(c)))(s) \\ c \Rrightarrow_\mu c' := & \lambda s.\ \forall \langle w, g\rangle \in s.\ \exists f : \textsc{alt}(c) \to \textsc{alt}(c').\ \mathbf{fmd}(f, M) \end{array} \tag{6.2}$$

Formula $\mathbf{selon}_\mu(\mathcal{U}, \mathcal{V})$ predicts that the discourse referents introduced in $\mathcal{U}$ are accessible in $\mathcal{V}$, but not in follow-up utterances. However, the data in section 6.3.1 suggests they can be accessed via modal subordination.

Section 6.5.1 addresses this issue by extending this lexical entry to Modal Dynamic Inquisitive Semantics MDInq. Section 6.5.2 is dedicated to the exemplification phenomenon discussed at the end of section 6.3.1. Finally, section 6.5.3 illustrates my model with an example.

6.5.1 Modeling anaphoric properties

To model modal subordination, we must make sure that after $\mathbf{selon}_\mu(\mathcal{U}, \mathcal{V})$ is performed, the stack contains one more local context. This context must contain the content of $\mathcal{U}$ and $\mathcal{V}$ evaluated in the previous topmost element c_n. I treat ADUs as a special kind of conditional: pushing the antecedent, pushing the consequent, and percolating the consequent. The only difference is the semantic relation percolated.

Viewing *SELON* A as a complex modality, pushing its argument (i.e. the consequent B) parallels my model for modal verbs. I assume a final popping

operation to avoid ending with too many local contexts. The new lexical entry for SELON is shown in 6.3.

$$\begin{aligned}
\mathbf{selon}_\mu :=\ & \lambda\mathcal{U},\mathcal{V}.\ \textsc{push}\ \mathcal{U};\textsc{push}\ \mathcal{V};\textsc{perc}^{\mathtt{D}}_\mu\ \ \mathcal{V};\textsc{pop} \\
\textsc{perc}^{\mathtt{D}}_\mu\ \ \mathcal{V} :=\ & \lambda\langle c_0,...,c_{n-2},c_{n-1},c_n\rangle. \\
& \qquad \langle c_0[c_{n-1}\vdash^{\mathtt{D}}_\mu \mathcal{V}],...,c_{n-2}[c_{n-1}\vdash^{\mathtt{D}}_\mu \mathcal{V}],\mathcal{V}(c_{n-1}),c_n\rangle \\
c'[c\vdash^{\mathtt{D}}_\mu \mathcal{V}] :=\ & \lambda s.\ s\in c'\wedge(c\Rrightarrow_\mu \mathcal{V}(c))(s)
\end{aligned} \tag{6.3}$$

D-percolation for dependence statements $\textsc{perc}^{\mathtt{D}}_\mu$ performs similarly to percolation for other modals. It structurally affects the same elements in a similar fashion as equation (4.12) in Chapter 4. The only difference is the turnstile $\vdash$ changed into $\vdash^{\mathtt{D}}_\mu$ to include the characteristics of dependence statements. Intuitively, $c'[c\vdash^{\mathtt{D}}_\mu \mathcal{V}]$ could be read as: "*learning in c' that $\mathcal{V}$ depends on c*". It outputs the restriction of context c' to the states s that resolve the conditional dependence $\Rrightarrow_\mu$ between the previous topmost element c and its updated version $\mathcal{V}(c)$.

To illustrate these definitions, consider again example (26-a). The interaction between ! and stack-managing operations is explained below.

(26) a. '*Depending on where^u your$_v$ leak is, [$_B$ there's no need to drain everything^z].*'
b. $!\mathbf{selon}_\mu(\mathcal{U},\mathcal{V}')$

Given an initial common ground c, $\mathbf{selon}_\mu$ pushes $\mathcal{U}=[\![\text{where}^u\ \text{your}_v\ \text{leak is}]\!]$, pushes $\mathcal{V}'=\langle ?\rangle[\![\text{there's no need to drain everything}^z]\!]$, then D-percolates $\mathcal{V}'$, and finally pops one element. The steps (except the final POP) are displayed in equation (6.4). We obtain a new common ground updated with the information $\mathcal{U}(c)\vdash^{\mathtt{D}}_\mu \mathcal{V}'$, as expected.

$$\langle c\rangle \xrightarrow{\textsc{push}\ \mathcal{U}} \left\langle \begin{array}{c} \mathcal{U}(c) \\ c \end{array} \right\rangle \xrightarrow{\textsc{push}\ \mathcal{V}} \left\langle \begin{array}{c} \mathcal{V}'(\mathcal{U}(c)) \\ \mathcal{U}(c) \\ c \end{array} \right\rangle \xrightarrow{\textsc{perc}^{\mathtt{D}}\ \mathcal{V}} \left\langle \begin{array}{c} \mathcal{V}'(\mathcal{U}(c)) \\ \mathcal{V}'(\mathcal{U}(c)) \\ c[\mathcal{U}(c)\vdash^{\mathtt{D}}_\mu \mathcal{V}'] \end{array} \right\rangle \tag{6.4}$$

6.5.2 Modeling exemplification

Discourse referents introduced in the ADU antecedent can be referred to via modal subordination. This includes conditional antecedents, like in (11-a). When the referential NP is an instantiation of the wh-word, I call it an exemplification.

Some exemplification assertions do not involve any overt modal or conditional operator, e.g. (12-b) repeated in (27). In this case, I posit that a covert conditional operator EXEMPL binds the assertion's two parts C and D. This logical form is displayed in Fig. 6.4.

(27) a. '*The^v pronunciation of figatellu$_z$ is different depending on where^u you are in Corsica.*'

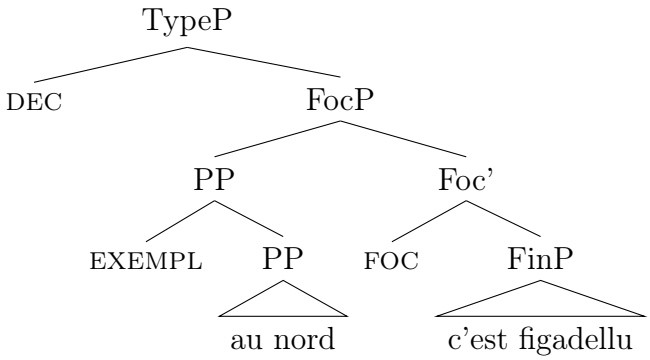


Figure 6.4 – Logical form of an exemplification.

b. '*[$_C$ In the north]$_u$, [$_D$ it$_v$'s figadellu] [...].*'

I analyze EXEMPL(C, D) like a regular conditional, i.e. pushing its antecedent C, then its consequent D, and finally percolating D. Similarly to SELON, I add a final POP to avoid overloading the stack. The formula is given in (6.5).

$$[\![\text{EXEMPL}]\!] := \lambda\mathcal{U}, \mathcal{V}.\ \text{PUSH}\ \mathcal{U}; \text{PUSH}\ \mathcal{V};\ \text{PERC}\ \mathcal{V}; \text{POP} \tag{6.5}$$

6.5.3 Example

Let me illustrate stack management on discourse (27). The semantic representation of (27) is written in (28). A focus phenomenon enables *different* to co-index FOC with v (see section 6.4.2). I also assume that dref z is common ground. For legibility, I omit the modal base function μ.

(28) a. $!(\text{PUSH}\ \mathcal{U}; \text{PUSH}\ \mathcal{V}; \text{PERC}^{\mathtt{D}}\ \mathcal{V}; \text{POP})$, where
   (i) $\mathcal{U} = \mathbf{word}(z, \text{“}\textit{figatellu}\text{”}); [v]; \mathbf{pronunciation}\{z, v\}; ?v$, and
   (ii) $\mathcal{V} = [u]; \mathbf{location}\{u\}; \mathbf{in}\{u, \mathsf{Corsica}\}; ?u$

   b. $\text{PUSH}\ \mathcal{W}; \text{PUSH}\ \mathcal{X}; \text{PERC}\ \mathcal{X}; \text{POP}$, where
   (i) $\mathcal{W} = \mathbf{north}\{u\}$, and
   (ii) $\mathcal{X} = \mathbf{phonological_word}\{v, /\text{figadɛlːu}/\}$

The dynamic effects of (28-a) on an initial stack $\langle c \rangle$ are represented in (6.6).

$$\langle c \rangle \xrightarrow{\text{PUSH}\ \mathcal{U}} \left\langle \begin{array}{c} \mathcal{U}(c) \\ c \end{array} \right\rangle \xrightarrow{\text{PUSH}\ \mathcal{V}} \left\langle \begin{array}{c} \mathcal{V}(\mathcal{U}(c)) \\ \mathcal{U}(c) \\ c \end{array} \right\rangle \xrightarrow{\text{PERC}^{\mathtt{D}}\ \mathcal{V}} \left\langle \begin{array}{c} \mathcal{V}(\mathcal{U}(c)) \\ \mathcal{V}(\mathcal{U}(c)) \\ c[\mathcal{U}(c) \vdash^{\mathtt{D}} \mathcal{V}] \end{array} \right\rangle \tag{6.6}$$

Let us call τ the final stack of (6.6) after an additional POP. It contains the new common ground $c' := c[\mathcal{U}(c) \vdash^{\mathtt{D}} \mathcal{V}]$ and the topmost element $c_1 := \mathcal{V}(\mathcal{U}(c))$. This local context can be rephrased as the conjunction of questions: “ *Where^u are*

you located in Corsica? And how^v is figatellu pronounced?". The effects of (28-b) on τ is represented in (6.7).

$$\tau \xrightarrow{\text{PUSH } \mathcal{W}} \left\langle \begin{array}{c} \mathcal{W}(c_1) \\ c_1 \\ c' \end{array} \right\rangle \xrightarrow{\text{PUSH } \mathcal{X}} \left\langle \begin{array}{c} \mathcal{X}(\mathcal{W}(c_1)) \\ \mathcal{X}(c_1) \\ c_1 \\ c' \end{array} \right\rangle \xrightarrow{\text{PERC } \mathcal{X};\ \text{POP}} \left\langle \begin{array}{c} \mathcal{X}(\mathcal{W}(c_1)) \\ c_1[\mathcal{W}(c_1) \vdash \mathcal{X}] \\ c'[\mathcal{W}(c_1) \vdash \mathcal{X}] \end{array} \right\rangle \tag{6.7}$$

The common ground and the second to last elements are updated with the conditional *if u is northern Corsica, then pronunciation v is figadellu.* Referents u and v are referred to in context c_1 where they are defined.

6.6 Dependence utterances with an interrogative consequent

This section explores the behavior of adjunct dependence utterances (ADUs) with an interrogative in the consequent. If the consequent is a question, then the whole utterance is a question. If the consequent contains an embedded interrogative, the adjunct clause can be interpreted as embedded, taking this interrogative as consequent. These investigations constitute sections 6.6.1 and 6.6.2, respectively. Section 6.6.3 discusses the presupposition of these inquisitive ADUs.

6.6.1 Dependence questions

As explained in section 6.3.2, ADUs whose consequent is a question are read as questions about the dependence relation between the antecedent's alternatives and the consequent's alternatives. For example, (29-a) requests the responder to provide a set of age periods and a function from age periods to egg recommendations.

(29) a. Selon les^u âges, [$_B$ combien^v d^z'œufs peut-on consommer] ?
'*At different^u ages, [$_B$ how^v many eggs^z can you eat]?*'

b. $\dagger(\text{PUSH } \mathcal{U}; \text{PUSH } \mathcal{V}; \text{PERC}^{\text{D}} \mathcal{V}; \text{POP})$, with
(i) $\mathcal{U} = [u]; \mathbf{age_period}\{u\}; ?u$, and
(ii) $\mathcal{V} = [v]; [z]; \mathbf{egg}\{z\}; \mathbf{eating_recom.}\{z, u\}; \mathbf{count}\{z, v\}; ?v$

Syntactic structure

The lexical entry in the previous section cannot model this sentence because SELON attaches below TypeP, which prevents polar question in the consequent. To remedy this issue, I assume that *SELON A* has another possible attachment

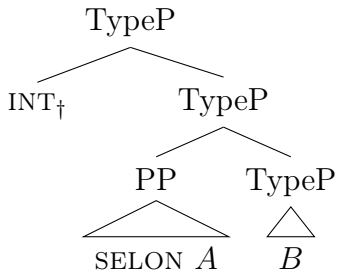


Figure 6.5 – Logical form of an inquisitive adjunct dependence utterance **selon**(A, B), supposing B has $\text{INT}_{\langle ? \rangle}$ in its TypeP.

site. It can attach at TypeP, above the head. This attachment site allows the consequent to be any TypeP phrase headed by INT, including all question forms.

Inquisitive hole

There is a second issue regarding the interpretation of (29-a). The definitions of **selon** in (6.2) and $c'[c \vdash^{\mathtt{D}}_{\mu} \mathcal{V}]$ in (6.3) do not create inquisitiveness. If the input context c' is assertive, so is the output context.

To solve this problem, I use the same method as for $\Diamond$ in Chapter 4: I invert the $\forall w \in s$ and $\exists f$ quantifiers to let each resolving proposition agree on some (potentially different) function f. Formula (6.8) replaces 6.2. This way, the output of **selon** is predicted to be always inquisitive. [7]

$$c \Rrightarrow_{\mu} c' := \quad \lambda s.\ \exists f : \text{ALT}(c) \to \text{ALT}(c').\ \forall w \in s.\ \mathbf{fmd}(f, \mu(w)) \tag{6.8}$$

If this change works for inquisitive ADUs, we might wonder whether it still works for assertive ones. As *SELON* A scopes below DEC in assertive ADUs, assertiveness is kept because DEC contributes the inquisitiveness removing operator !. More precisely, we can assume that ! can type-shift to $(\ell \to \ell) \to (\ell \to \ell)$ similarly to † in Chapter. 4: viz. (6.9). Intuitively, it applies $! : (k \to k) \to (k \to k)$ to every context change $\tau_i \mapsto \tau'_i$ in the stack.

$$\begin{array}{rl} [\![\text{FOC}]\!] := & \lambda \mathcal{T}.\ !\mathcal{T} \\ !\mathcal{T} := & \lambda \tau.\ \tau', \quad \text{where for all } i < |\mathcal{T}(\tau)| : \\ \tau'_i := & \begin{cases} \lambda s.\ s \in !^{kk}(\mathcal{T}(\tau)_i) \wedge (\exists s' \in \tau_i.\ s' \leqslant s) & \text{if } i < |\tau| \\ \mathcal{T}(\tau)_i & \text{if } |\tau| \leq i \end{cases} \end{array} \tag{6.9}$$

where $|\tau|$ is the length of stack τ and τ_i the i-th element of τ.

7. In section 6.4.2, the antecedent and the consequent of an ADU do not generate type-shift operators because stack management is performed by SELON.

6.6.2 Embedded interrogatives in the consequent

Inquisitive ADUs do not only show up as matrix questions. They are also possible as embedded questions. The adjunct phrase can syntactically depend on an interrogative embedded under an attitude verb. However, it can also be dislocated to the left periphery of the whole sentence, like in (30-a). As argued in section 6.3.2, this attitude report expresses a *know-wh* relation and can be paraphrased as (30-b).

(30) a. en fonction de combien^u il te reste d^z'éléments dans ton$_y$ objet tu peux savoir [$_B$ depuis combien^v de temps il$_y$ est en train de se dégrader]
'Depending on how^u many elements$_z$ you have left in your$_y$ object, you can find out [$_B$ how^v long it$_y$ has been decaying].'

b. '*You can know what the dependence relation f is between the element count and the decay duration.*'

c. *can* (*you* (*know* (INT$_\dagger$ ((SELON A) B)))), with
(i) A = INT (FOC$_u$ (*how^u many elements$_z$ are left in your$_y$ object*))
(ii) B = INT$_{\langle ? \rangle}$ (FOC$_v$ (*it$_y$ has been decaying how^v long*))

d. $\Diamond\mathbf{know}(\dagger(\mathbf{selon}(\mathcal{U},\mathcal{V})))$, with
(i) $\mathcal{U} = [u]; [z]; \mathbf{elmt}\{z\}; [y]; \mathbf{object}\{y\}; \mathbf{in}\{z, y\}; \mathbf{count}\{z, u\}; ?u$,
(ii) and $\mathcal{V} = [v]; \mathbf{time_interval}\{v\}; \mathbf{decay}\{y, v\}; ?v$

In this case, I analyze *SELON* A as attached to the CP of the embedded B clause in the logical form, just under the attitude verb. This way, *know* expresses knowledge about the true dependence mapping element count to decay duration.

The logical form of (30-a) is displayed in (30-c) and its semantic representation in (30-d). Modal Dynamic Inquisitive Semantics does not include lexical entries for attitude verbs, like *know*. Providing such analysis would go beyond the scope of this thesis. Therefore, I leave stack management underspecified in (30-d).

6.6.3 Presuppositions of dependence questions

As mentioned in section 6.3.2, question (29-a) requires that there exist age periods and a non-trivial dependence relation from them to nutrition recommendations for a felicitous questioning speech act.

We can implement this requirement as a presupposition by appending the presuppositional operator † on top of $\mathbf{selon}(\mathcal{U},\mathcal{V})$. This makes the question infelicitous whenever there is no dependence relation between the antecedent and the consequent or when this relation is trivial.

INT contributes both the † operator and the $\langle ? \rangle$ operator. This raises a problem. *SELON* A needs to be attached higher than $\langle ? \rangle$ to accept polar interrogatives in its consequent but also needs to be attached lower than INT to be in the scope of †. To solve this issue, I assume that † and $\langle ? \rangle$ are actually contributed at two different levels in TypeP that I call INT$_\dagger$ and INT$_{\langle ? \rangle}$ respectively. The adjunct

clause attaches between these two heads.

Question (29-a) is used on a web page as the title of a section presenting egg recommendations. An inquisitive ADU might not always have such a presupposition in regular information-seeking contexts. For example, it seems that question (31-b) asks, for each possible election winner x, whether we should expect protests against x's win. It might be that the true answer entails that we should expect protests after the win of any candidate. Moreover, ANNE seems open to this eventuality, although we may infer that different people will probably protest depending on the winner.

(31) a. SITUATION: *Before a presidential election, Anne interviews a specialist on current political protest movements.*

b. ANNE: En fonction de qui gagne les élections, doit-on s'attendre à des manifestations ?
'*Should we expect demonstrations depending on who wins the elections?*'

c. '*For each potential winner x, if x wins, should we expect demonstrations?*'

Thus, question (31-b) can be read as the conjunction of conditional statements glossed in (31-c). This reading is weaker than the lexical entry given for SELON in this chapter. More precisely, the non-triviality condition (4-b-ii) does not hold; no such presupposition occurs. However, question (31-b) still raises some presupposition, namely that for any winner x, either we should expect demonstrations or we shouldn't, i.e. the conditional dependence part (4-b-i).

This suggests that non-triviality is maybe not a core feature of SELON. Non-triviality could be derived as a sort of pragmatic strengthening. This would suggest that *bona fide* conditionals and (adjunct) dependence utterances are closer than initially expected. I leave the investigation of the factors favoring non-triviality and the connection between these two kinds of conditionals as future prospects.

6.7 Conclusion

Adjunct Dependence Utterances (ADUs) are a family of adjunct constructions expressing a conditional dependence, including French interrogative-based conditionals introduced by *selon*, *suivant* or *en fonction de* and English "*Depending on Q, p*".

We can view an ADU $\mathbf{selon}(A, B)$ as a complex conditional, calling A the antecedent and B the consequent. This analogy allowed me to model many observations about their dynamic properties. In particular, a wh-referent introduced in A might take part in a donkey anaphora with a pronoun in B if A precedes B. Moreover, this wh-referent can be referred to in a subsequent sentence via modal

subordination. I modeled this phenomenon using Modal Dynamic Inquisitive Semantics.

Both components of ADUs exhibit various syntactic forms. The antecedent can be an interrogative, a declarative CP disjunction, or a concealed question. When the consequent is declarative, the ADU is assertive. ADUs are sensitive to the question under discussion evoked by their consequent, built as the question inquiring about their focused material, e.g. disjunction or an indefinite. It is interpreted as a polar issue by default. When the consequent is a question, the ADU is inquisitive. I provided a lexical entry for SELON that captures all of these observations. It also correctly predicts the presupposition of ADU questions and local attachment to embedded interrogative consequents.

This chapter closes this thesis by showing the complete potential of Modal Dynamic Inquisitive Semantics. In particular, the complex interactions between interrogatives, modal bases, conditionals and discourse referents in ADUs can all be captured in a single lexical entry for **selon** and two attachment sites. This economy is possible thanks to a uniform treatment of declaratives and interrogatives on one side, and modals and conditionals on the other side.

Chapter 7

Conclusion

7.1 General summary

This dissertation investigates the impacts of questions in a conversation. Questions restrain and enrich the discourse context by triggering presuppositions and raising discourse referents. This PhD thesis examines these two related phenomena and demonstrates that they vary according to semantic and pragmatic environments. The supplied empirical data suggests broadening our conception of wh-words and questions. It motivates the extension of analyses designed for indefinites and assertions to wh-words and questions. In particular, this thesis explores the interactions between questions, modals, and conditionals and provides a formal model based on Dynamic Inquisitive Semantics DInq (Roelofsen and Dotlačil 2023b).

7.1.1 Properties of wh-words

Past research has shed light on the various ways wh-words contribute to the meaning of questions, including exhaustivity, D-linking, existential and uniqueness presuppositions, and discourse referent introduction.

This dissertation proposes to add a new dimension to this set of properties: the specific vs. generic duality. This theory stems from a long-standing problem about simplex wh-questions. Chapter 2 brings evidence that the existential inference of simplex wh-questions can be made stronger or weaker in different settings. After observing a similar variation in other determiners and pronouns, I suggest that wh-words are epistemic pronouns or determiners (Haspelmath 1997), like *certain* (specific) and Fr. *quelconque* (non-specific). This hypothesis is the first to posit that the presence of this existential inference depends on the question's reading and that this reading is due to an ambiguity in wh-words. This theory has numerous consequences for the way we describe questions. It impacts exhaustivity, the availability of null answers, weak islands, and NPI licensing.

Chapter 5 presents how differently specific and generic wh-words behave concerning anaphora.

Besides the new research opportunities brought by this proposal, this thesis innovates regarding how question inferences are studied. Distributional restrictions of interrogatives are circumvented by intersecting various unrelated phenomena, testing the effects of different operators, and searching for languages where readings can be split using grammar. Although these techniques have been used extensively in formal semantics, their combination in the study of questions has been under-explored. I hope that this methodology can pave the way to a deeper understanding of questions.

7.1.2 Interaction between questions, modals and conditionals

For a long time, it has been thought that only a few operators can scope over the wh-word in a question's logical form, namely generalized quantifiers (cf. list and choice readings). Recently, Hirsch and Schwarz 2019 showed that existential modalities can also do so. This phenomenon weakens the uniqueness presupposition of *which+sg* questions.

This dissertation dives deeper into the interactions between modals and questions. Chapter 4 models this weakening, as well as modal subordination, for both assertions and questions uniformly. This chapter develops Modal Dynamic Inquisitive Semantics MDInq using a stack to keep track of local contexts. This system captures the dynamic effects of modals and conditionals in an integrated way. Chapter 6 applies this setting to dependence propositions, exhibiting the interplay of alternatives, complex conditionals, and modal bases. This research reveals new insights into the interaction between focus alternatives and inquisitiveness.

One of the main conclusions of this thesis is that generic questions are modal environments. We can view alternatives as epistemic local contexts raised by the asker. They can serve as modal bases. As long as the question remains unresolved, referring to these alternatives constitutes modal subordination. This novel perspective may have many repercussions on our perspective of question meanings. For example, modeling the precise contribution of wide scope generics, future tense, and habitual mood in questions is a promising prospect in that inquiry.

7.2 Limits and future directions

Although I attempted to analyze data and define concepts as precisely as possible, this dissertation contains some weaknesses and unexplained aspects that

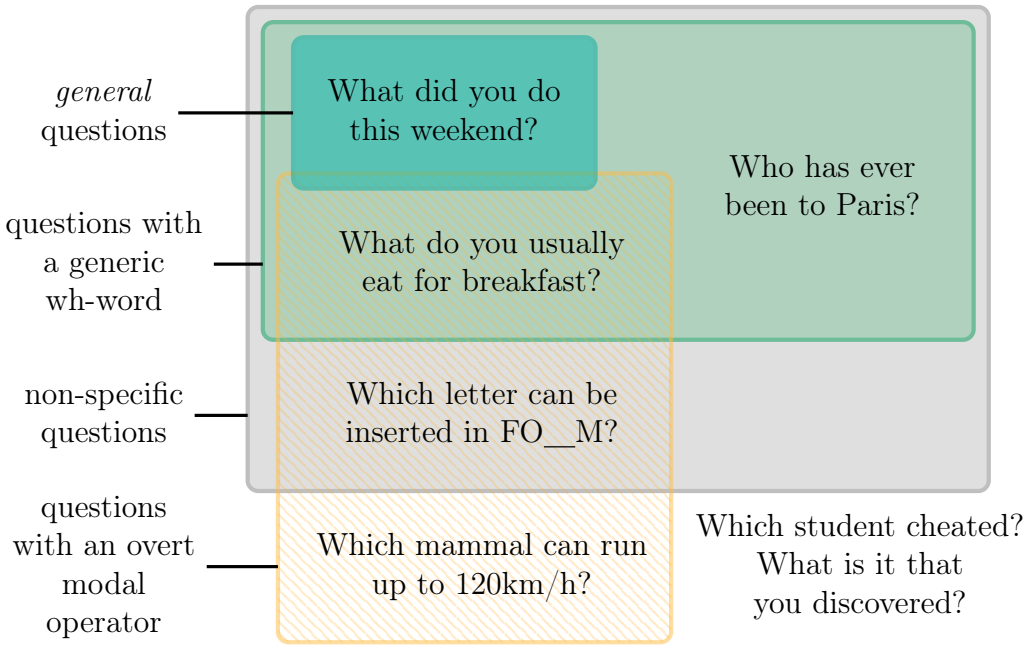


Figure 7.1 – Categories of questions with examples

require further investigation. As final remarks, I would like to discuss two of these points, as well as research prospects for improving their treatment.

7.2.1 More finely-grained properties for wh-words

Refining generic questions

The specific vs. generic duality is a significant tool I bring to the table. The notions of specificity and genericity are well-studied in the literature. However, they are not necessarily the exact opposite of each other. For simplicity, I mixed the concepts of *generic* questions (in the sense of *general* questions), non-specificity, generic terms, and generic mentions (in anaphora theory). I provided evidence that all of these concepts are relevant for what I call *generic question* / *generic who*$_G$. However, there is probably room for refinements.

Figure 7.1 shows a glimpse of some question categories discussed in this dissertation. I assume that wh-words are either specific or generic. Nevertheless, there is a gray space of questions with intermediary existential presuppositions. As seen in Chapter 4, this happens when specific wh-words scope below a possibility modal, thus weakening the existential presupposition. Those questions are not specific because they cannot be rephrased as "*There exists f, and I want to know the identity of f*". Nor are they generic because they are presuppositional. Understanding this class of questions may better define the boundary between generic and specific questions.

A trail towards more finely-grained properties for generic wh-word requires a deeper look at indefinites. Non-specific indefinites exhibit many meanings, e.g. generic, predicative, attributive, or NPI (von Heusinger 2002). Precise tests are required to tease apart these readings, especially in anticipation of cross-linguistic comparisons. Then, these tests can be broadened to wh-words. The questionnaire set up by Dayal (to appear) is a promising starting point for such an investigation.

Reference to wh-words under attitudes

Similarly, von Heusinger 2002 draws attention to the existence of several kinds of specificities. A subsequent analysis could help better pinpoint which one is at stake in *specific questions*.

A trail can be drawn by studying the interpretation of discourse referents introduced by wh-words under attitudes. There exist interesting discrepancies that, to my knowledge, have not been mentioned before. Let me expand on that briefly.

Factive attitude verbs project the wh-referent to the matrix level. Moreover, this referent is construed as the true witness evaluated in the enunciation world. For example, in discourse (1-a), *this*$_u$ *student* is the real cheater.

(1) a. Mary knows who^u cheated. This$_u$ student will get 0 points.
$\rightsquigarrow$ *u is the real cheater*
b. Mary is certain who^u cheated. This$_u$ student will get 0 points.
$\rightsquigarrow$ *u is the person Mary thinks cheated*

However, P-to-Q-distributive non-veridical responsive verbs [1], like *agree*, behave differently. Although the discourse referent introduced by the wh-word can be accessible at the matrix level in subsequent assertions, it refers to the intensional witness relative to the attitude holder's mental state. For example, in (1-b), *this*$_u$ *student* refers to the student that Mary thinks cheated, even if this might not be the real cheater.

The accessibility of wh-referents under non-veridical attitudes and their intensional interpretation is a puzzle for a compositional dynamic approach. Although this discourse referent is introduced and evaluated in a local context, it must be accessible in the global context. Understanding how this phenomenon is licensed can shed light on the nature of specific *who*$_S$. Extending Modal Dynamic Inquisitive Semantics to model (several kinds of) attitudes is also an ambitious task that this thesis opens.

1. I refer to Özyıldız et al. 2023 for this terminology.

7.2.2 Scope and instantiation of wh-words

This dissertation only scratched the surface of the interactions between modals and questions. Plenty of scope-related issues remain unanswered.

Scope of null answers

Section 2.4.2 describes a method to compute null answers. In questions containing an operator able to rise above the question operator, this method correctly predicts the two possible readings: wide scope and narrow scope. They are calculated by deriving an exhaustification operator $\text{EXH}_\star$ above or below the other operator.

However, there are sentences where this method overgenerates. For example, with a factive predicate or a sentential negation, the narrow scope seems unavailable. For example, short answer (2-b) to question (2-a) can mean (2-c) but not (2-d), and similar for (3).

(2) a. A: Who did Mary discover that John met?
b. B: Nobody.
c. There is nobody that Mary discovered that John met. (= Among the persons that John met, Mary discovered none of them.) ($\text{EXH}_\star$ > discover)
d. #Mary discovered that John met 0 persons. (discover > $\text{EXH}_\star$)

(3) a. A: What didn't Jack read?
b. B: Nothing.
c. There is nothing that Jack didn't read. (= Jack read everything.) ($\text{EXH}_\star$ > $\neg$)
d. #It is false that Jack read no books (= Jack read some books.) ($\neg$ > $\text{EXH}_\star$)

Bylinina and Nouwen 2018 also observed that the word *zero* can sometimes only have a surface scope. The wh-word instantiating the null individual has a wide scope, but in their case, *zero* has a surface narrow scope.

Factive and negation arguments are considered weak islands. Therefore, it is not surprising that generic questions are affected by these constructions. However, contrary to the cases studied in sections 2.5.3, questions (2-a) and (3-a) are felicitous. Only some of their answers are unlicensed. Consequently, additional investigation is required to predict precisely when null answers are derived and how.

Higher-order quantification and the generic reading

Spector 2007 observes that the same weak islands block the narrow scope of the higher-order reading, e.g. when the question is answered by a disjunction.

For example, short answer (4-b) to (4-a) can be interpreted as (4-c), but not as the stronger (4-d).

(4) a. A: What didn't Jack read?
b. B: Moby Dick or Madame Bovary.
c. Jack didn't read Moby Dick, or he didn't read Madame Bovary. $(\vee > \neg)$
d. #Jack read neither Moby Dick nor Madame Bovary. $(\neg > \vee)$

The parallel between the higher-order and the generic readings is probably a sign that they are related, as suggested in section 2.7.5. Moreover, the existence of two dualities (type-flexibility and specific/generic) is a high cost for any theory of question semantics. For these reasons, inspecting the relationships between higher-order reading and generic wh-words appears to me as one of the most important next steps after this thesis.

Appendix A

Extracting French Interrogatives from Treebanks

In Chapter 5, I use an annotated corpus to help solve a semantic question. Extensive data is helpful to provide quantitative evidence in favor or disfavor of a hypothesis. However, such large corpora are not always available, especially if we are interested in uncommon phenomena, like certain embedded interrogatives. This chapter presents a solution to this issue by showing how we can extract a high-quality corpus of sentences containing an interrogative clause.

Corpora containing clauses of a certain type must either be curated by expert annotators or automatically extracted from larger corpora. Gold corpora are of high quality but may be limited in size. In contrast, extraction from raw text can gather large amounts of examples but at the risk of limited accuracy. Automatic extraction is challenging in languages like French, where interrogatives exhibit a wide diversity in their morphosyntax, and many other clauses share common structures with interrogatives.

I advocate for an intermediate approach, taking advantage of large corpora annotated in syntax and automatically extracting interrogative clauses from them. Unlike most other retrieval methods used in the literature, I suggest using internal distribution properties to identify interrogatives. In particular, I do not rely on a predefined list of interrogative-embedding predicates, which allows me to discover interrogative-embedding predicates absent from expert lists.

This chapter presents an extended version of a poster presented at GdR LIFT 2023 (Richard 2023).

A.1 Introduction

French is famous for exhibiting a large set of possible ways to build interrogatives (Larrivée and Guryev 2021), e.g. for embedded constituent interrogatives: fronting with or without subject-verb inversion (1-a) (1-b), insertion of an inter-

rogative marker (1-c), cleft (1-d) or *in situ* interrogative word (1-e). Sentences (1-b) (1-c) and (1-d) mean "*He knows where she is*".[1]

(1) a. Il sait où est Marie.
He knows where is Marie.
'*He knows where Marie is.*'

b. Il sait où elle est.
He knows where she is.

c. ! Il sait où est -ce qu' elle est.
He knows where is -it that she is.

d. % Il sait c' est où qu' elle est.
He knows it is where that she is.

e. ! Il sait c' est où.
He know it is where.
'*He knows where it is.*'

As a consequence, identifying embedded interrogatives is a complex task, especially if we aim to capture all French variants, including spoken French. Corpus studies on French interrogatives often extract and annotate relevant clauses by hand or using simple heuristics on raw text. More generally, interrogative detection is usually based on plain text and primarily focuses on questions only.

This chapter takes advantage of structured datasets and designs a program that automatically identifies interrogative clauses. We present a rule-based graph rewriting tool named FUDIA (French UD Interrogative Annotator)[2], relying on Grew[3] (Guillaume 2021) and Universal Dependencies. FUDIA is able to detect root interrogatives as well as embedded (aka. indirect) interrogatives, yes-no, as well as constituent interrogatives, and different variants (especially all those in (1)). Run on a UD corpus, FUDIA adds supplementary annotations: `ClauseType=Int` on the head of interrogatives, `PhraseType=Int` on the head of interrogative phrases and some missing pronoun types. It also reannotates some French grammaticalized expressions. The sentences bearing `ClauseType=Int` are extracted from French UD corpora to provide a separate corpus named French Interrogative Bank (FIB).

The feature proposition `ClauseType=Int` is a step towards providing UD corpora with a way to express clause types (e.g. declarative, interrogative, exclamative, etc.). This kind of information may be helpful in some studies, for instance, studies on universals (Choi, Guillaume, and Fort 2022) or grammar extraction (Herrera, Kahane, and Guillaume 2022).

1. Some sentences are regional (%) and some are non-standard (!). See also appendix B.2 for an overview of the diversity of French matrix questions.

2. The source code and the complete documentation is available online: `https://github.com/Valentin-D-Richard/FUDIA`

3. https://grew.fr/

The key contributions of this chapter are:

— the FUDIA program, a new tool to detect interrogatives from any French UD corpus

— the gathering of French UD-annotated sentences containing interrogatives extracted from existing UD treebanks and statistics about the distribution of their structures

In this chapter, when using the word *interrogative*, we refer to syntactic objects: phrases with a particular morphosyntactic marking and a coherent distribution. We do not define this construction based on semantic or pragmatic features. For example, interrogative root clauses can be information-seeking as well as rhetorical questions.

In section A.2, we mention other approaches used by corpus studies on French interrogatives, question detection in general, and competing proposals for a UD clause-type feature. Section A.3 presents our proposal and the evaluation method. We give quantitative results in section A.4 and discuss them. Finally, section A.6 concludes.

A.2 Related works

A.2.1 Corpus studies on French interrogatives

Up to now, linguistic corpus studies on French interrogatives have retrieved them from a corpus using a concordancer (Rossi-Gensane et al. 2021; Benzitoun 2022; Gillet 2022), simple heuristics on raw text (Reinhardt 2019b; Eshkol-Taravella et al. 2022), regular expressions (Lefeuvre and Rossi-Gensane 2017) or by hand (Reinhardt 2019a; Bally 2022). Note that for most of them, it is not well-stated how interrogatives were extracted (Maury 1990; Quillard 2001; Ledegen 2007; Defrancq 2005; Elsig and Poplack 2006; Queffélec 2006; Myers 2007; Farmer 2013; Dekhissi 2016). We assume that it was done by hand (especially for spoken data) or with simple heuristics as well. Morphosyntactic annotations, typically concerning the interrogative structure, were mainly added by hand (Guryev and Delafontaine 2015; Reinhardt 2016). This method may be necessary with noisy data like text messages or children's productions. However, each newly created corpus consumes much expert time.

To our knowledge, only Lefeuvre 2021 benefited from syntactically annotated data. The French Interrogative Bank provides a qualitative and diverse corpus for such statistical studies.

A.2.2 Question detection

In general, automated interrogative recognition is usually either rule-based, learning-based, or a combination of both (B. Li et al. 2011; Tian, Qiu, and Huang

2013; Özger, Dìrì, and Gìrgìn 2014).

Rule-based approaches consider heuristics to spot as many interrogatives as possible, e.g. a list of interrogative words (Bélanger and Lapalme 2004; B. Wang et al. 2009), the presence of a question mark (Efron and Winget 2010), a list of interrogative-embedding verbs (Bélanger and Lapalme 2004; Efron and Winget 2010), word order or interrogative markers (Bélanger and Lapalme 2004; B. Wang et al. 2009).

These techniques have several drawbacks. Firstly, interrogative words may be lexically ambiguous with other functional words, e.g. relative pronouns. As emphasized by Shrestha and McKeown 2004 and K. Wang and Chua 2010, relying on question marks has many flaws. They may be absent from interrogatives due to segmentation mistakes, non-questioning intention, or transcription choices for spoken corpora. All of this can lead to false negatives. Using a list of interrogative-embedding predicates takes the risk of being overly restrictive, e.g. due to a lack of attestations or over-normative judgments from the authors. Finally, some languages, like French or Chinese, may exhibit a wide range of interrogative structures (e.g. because of *in situ* interrogative words), so it is pretty challenging to list them all.

Learning-based approaches may circumvent these flaws thanks to classifiers (Hong and Davison 2009; Sun et al. 2010; Margolis and Ostendorf 2011), pattern recognition techniques using lexical and/or syntactic features (Cong et al. 2008; K. Wang and Chua 2010) or other methods (Shrestha and McKeown 2004; Ding et al. 2008).

Most of these studies work on raw written texts or use parsing preprocessing (K. Wang and Chua 2010). Contrary to them, we already have structured data at our disposal. Therefore, we do not need statistical learning techniques or naive heuristics. Moreover, the program we present also detects embedded interrogatives. Most of the other approaches cannot do that.

A.2.3 Clause type features in UD

Universal Dependencies version 2.12 lacks a universal feature for clause types.

A sentence-level feature `s_type` was introduced in UD_English-GUM (Zeldes 2017). It can take the values `decl` for declaratives, `wh` for constituent question, `q` for other questions, and many others [4]. Sentences combining a declarative and an interrogative are annotated with `s_type = multiple`. The first drawback of this approach is the ambiguity of the `multiple` value. It does not specify what the different sub-values are. Moreover, such a sentence-level feature prevents us from specifying the type of dependent clauses, e.g. interrogatives embedded in declarative or vice versa.

4. The other values that `s_type` can take are: `frag`, `intj`, `sub`, `imp`, `ger`, `inf`, `multiple` and `other`.

An issue on the documentation GitHub project of Universal Dependencies [5] discusses the possibility of including information about clause types. A feature `Mood=Int` on the clause head was suggested. While this better captures the fact that being declarative or interrogative is a clause-level feature, calling it "mood" may lead to confusion with (grammatical or semantic) verbal mood. As such, interrogatives may have various verbal moods on their main verb.

Our proposal is a `ClauseType` token feature that can take the values `Decl` (declarative), `Int` (interrogative), `Exc` (exclamative), and `Des` (desiderative). See for example Fig. A.1. Every (potentially embedded) clause's head would be labeled with this feature. FUDIA provides a first step toward this annotation scheme by adding `Int`-valued `ClauseType` features to French UD corpora.

The closest competing proposal is the token feature `Stype` in UD_Hindi-HDTB and UD_Urdu-UDTB (Bhat et al. 2017; Palmer et al. 2009; Tandon et al. 2016). [6] It can take the values `declarative`, `imperative`, `interrogative` or `interjective`. Contrary to our proposition, this feature is borne by the main verb of the clause, even if it is not the head but a copula. If a sentence contains multiple clauses, each bears a (potentially distinct) `Stype` feature. However, this annotation was most likely performed by hand.

Contrary to the `Stype` word feature, we choose to put a `ClauseType` feature on QU phrases which are not finite clauses, e.g. on *pourquoi* '*why*' in (3-b). We follow the idea that these phrases could be interpreted as elliptical clauses, given that the matrix verb does not allow any regular proform in object position, but only interrogative ones.

A.3 Method

We first present how FUDIA was developed in section A.3.1. In section A.3.2, we list its modules and the added and changed annotations. The evaluation phase is described in section A.3.3. Finally, section A.3.4 tells how the FIB was created using FUDIA.

A.3.1 Development phase

Our development set Dev is the set of the current open source French UD corpora (29,762 sentences, 635,145 tokens):

— FQB (written) (Seddah and Candito 2016; Bonfante, Guillaume, and Perrier 2018; Judge, Cahill, and van Genabith 2006)

— GSD (written) (Guillaume, Marneffe, and Perrier 2019)

— ParisStories (spoken) (Kahane et al. 2021)

5. `https://github.com/UniversalDependencies/docs/issues/877`
6. `https://universaldependencies.org/misc#stype`

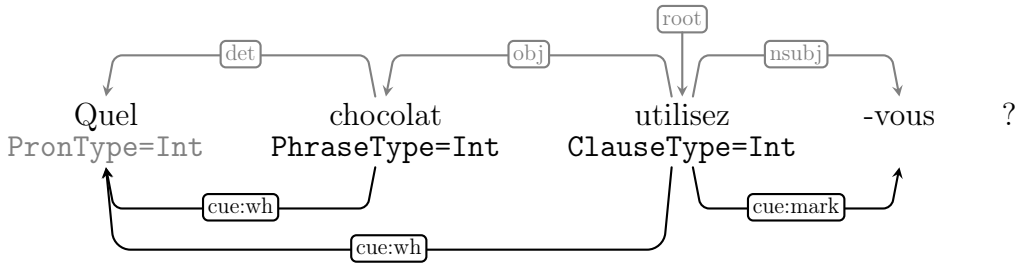


Figure A.1 – Annotation of sentence (3-a) with FUDIA. Original features and edges are in gray. The `cue` edges are only present in the enriched version of the FIB (see section A.3.2).

— ParTUT (written) (Sanguinetti and Bosco 2015, 2014; Bosco and Sanguinetti 2014)
— PUD (written) (McDonald et al. 2013)
— Rhapsodie (spoken) (Bawden et al. 2014; Lacheret et al. 2014; Lacheret-Dujour, Kahane, and Pietrandrea 2019)
— Sequoia (written) (Candito and Seddah 2012; Candito, Perrier, et al. 2014; Perrier et al. 2014; Bonfante, Guillaume, and Perrier 2018)

As Universal Dependencies (UD) is a rich framework, we can elaborate more complex rules than simple heuristics. We do not implement interrogative mark detection. We also avoid basing FUDIA on a list of open class terms. For example, unlike Lefeuvre and Rossi-Gensane 2017, we detect embedded interrogatives via their internal structure rather than thanks to a list of interrogative-embedding predicates. The goal behind this decision is to allow the program to "discover" any interrogative-embedding predicate that could have been missed otherwise.

There is already a rich literature on the syntax of French interrogatives: Delaveau, Cappeau, and Dagnac 2021; Larrivée and Guryev 2021; Coveney 2011 among many others. We used the online Grew-match interface [7] to search for the patterns described in that literature and see how they were annotated in Dev. From these observations, we wrote rules that were as general as possible. We regularly tested the rules on Dev during this development phase.

As Grew is not suited for rule disjunction, we used Python as a meta-program to automatically generate a rule file from case descriptions.

A.3.2 Description of FUDIA modules

FUDIA contains eight modules (i.e. rule packages) executed one after the other. Each module is run until it obtains a unique normal form. Thus, FUDIA is deterministic. In execution order, the modules are: `constr`, `prontype`, `ecq`, `quoted`, `wh`, `cl_head_pull`, `conj`, `mark`. We describe them all briefly.

7. `https://universal.grew.fr/`

prontype This module runs a heuristics to add missing pronoun type features `PronType`. Some relative pronouns can be identified as markers of a clause governed by `acl:relcl` (relative clause) or `advcl:cleft` (cleft). The other QU-words [8] are mostly interrogative words.

constr and ecq These modules reannotate three grammaticalized expressions as fixed: *n'importe + QU* in `constr`, and *est-ce que* and *qu'est-ce que* in `ecq`. See appendix B.1 for linguistic motivations to consider them as lexicalized.

The UD guidelines (de Marneffe et al. 2021) stipulate that whenever an expression is proven to be grammaticalized, to behave as a function word, and to have no internal syntactic structure (except from a historical perspective), its elements should be linked with the `fixed` relation, its head being the first token. It corresponds to the "fixed" multiword-expression category of Sag et al. 2002.

The expression *est-ce que* lit. '*is it that*' is a lexicalized conjunction (Delaveau, Cappeau, and Dagnac 2021, §XII-3.1.2 p.1403). In the French UD corpora, we counted 6 different annotations of *est-ce que* over 36 occurrences. None of them used `fixed`. As a consequence, changing for a fixed *est-ce que* would follow three converging imperatives: (i.) bringing annotations as close as possible to current linguistic knowledge; (ii.) simplifying and unifying annotation schemes; (iii.) sticking to standardizing UD guidelines. According to the Grande Grammaire du Français (GGF) (Delaveau, Cappeau, and Dagnac 2021, §XII-3 p.1402) and Orféo guidelines [9] , we analyze *est-ce que* as a subordination conjunction (`SCONJ`). As advocated by Gerdes et al. 2019 for SUD, we add `ExtPos=SCONJ` (external POS) to its head.

The expression *qu'est-ce que* lit. '*what is it that*' is a lexicalized interrogative pronoun (Delaveau, Cappeau, and Dagnac 2021, Tab. XII-5, p.1405). In the French UD corpora, we counted 10 different annotations of this expression over 133 occurrences. As stated above, reannotating it to `fixed` everywhere would be beneficial in several aspects. Encoding such a change requires a shift in the sentence head. In Fig. A.2, we display the output of FUDIA on sentences (2). The head of the embedded clause is shifted from the first *qu'* (original analysis as a cleft) to the main verb *passait*.

(2) [...] je savais pas [qu' est -ce qu' il se passait] .
[...] I knew not [what is -it that EXPL REFL happened] .
I didn't know [what was happening]. (ParisStories)

The locution *n'importe + QU* '*no matter + wh*' is considered lexicalized by

8. The most common French QU-words with are *quoi/que* '*what*', *qui* '*who*', *quel* '*which*', *lequel* '*which one*', *quand* '*when*', *comment* '*how*', *combien* '*how many/much*', *où* '*where*', and *pourquoi* '*why*'

9. `https://repository.ortolang.fr/api/content/cefc-orfeo/11/documentation/site-orfeo/guide-dannotation-syntaxique-du-corpus-orfeo/index.html`

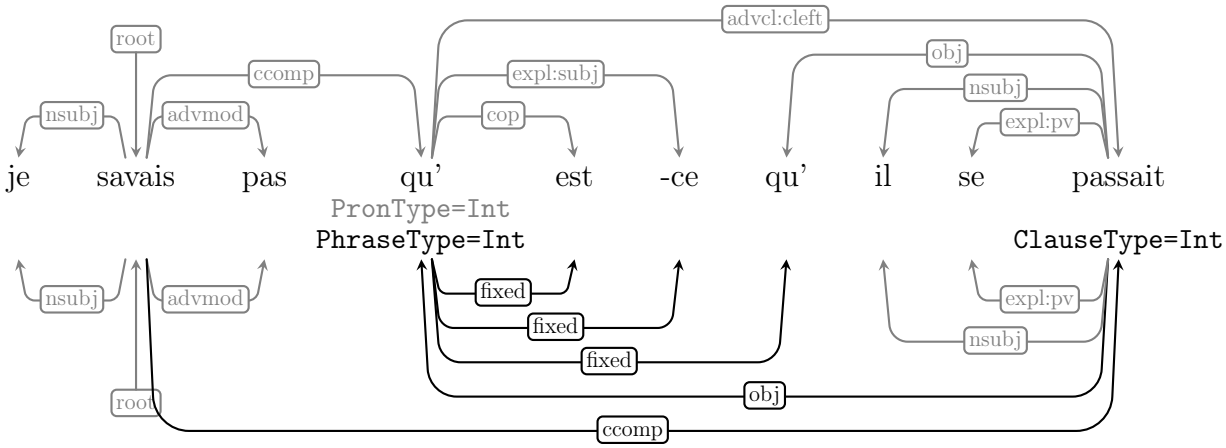


Figure A.2 – Annotation of sentence (2). The original edges are depicted on top of the sentence and in gray. FUDIA's output is depicted below it. The `cue` edge between *passait* and the first *qu'* is not shown for legibility.

the GGF (Jayez 2021, §V-6.3 p.633) and is analyzed as a single token in the Orféo guidelines. Only GSD currently uses the `fixed` relation for it. We generalize this and add an `ExtPos` feature depending on the POS of QU (determiner, pronoun, or adverb).

quoted This module identifies quoted segments: titles, reported speech and parenthesized segments. They help us spot some phrase boundaries and filter out subject-verb inversion specific to speech report constructions.

wh and cl_head_pull These are the main modules for QU-interrogative identification. Given a sentence like (3-a), we add two features. The clause type feature `ClauseType=Int` is put on the head of the interrogative clause, here *utilisez* (see Fig. A.1). A feature `PhraseType=Int` is put on the head of the interrogative phrase.[10] Here, the interrogative phrase *Quel chocolat* is fronted, and its head is *chocolat*. These features may be put on the same word. A typical case is short questions (aka. sluicing) or embedded short interrogatives, e.g. on *pourquoi* in (3-b).

(3) a. Quel chocolat utilisez -vous ?
Which chocolate use -you ?
'What kind of chocolate do you use?' (GSD)

b. [...] euh je sais plus pourquoi .
[...] er I know no_more why .
'er, I don't know why any more.' (ParisStories)

10. By interrogative phrase, we mean the nominal or prepositional phrase containing the QU-word. See appendix B.3 for a discussion on this definition.

In the enriched version of the FIB, we add three specific edges called cue edges. Statistical studies on French interrogatives classify interrogatives by their internal structure (e.g. absence/presence of *est-ce que*, etc.) and the QU-word used. The cue edges provide a quick access to these features.

A `cue:wh` edge is added between the clause head and the interrogative word, and one more between the interrogative phrase head and the interrogative word. A `cue:mark` edge is added between the clause head and any morphosyntactic structure that marks the clause as interrogative. In (3-a), we would have an edge `cue:mark` from *utilisez* to *-vous* to account for subject-verb inversion. The enriched annotation of (3-a) with FUDIA is displayed in Fig. A.1. Due to graphical interface constraints, such edges are not added when the source node is the target node. Note that these supplementary edges break the tree structure of sentences.

conj This module handles conjuncted interrogative phrases.

marker This final module adds cue edges to morphosyntactic markers: subject-verb inversion, *est-ce que* '*is it that*', *que* '*that*', *si* '*whether*' and *-tu* / *-ti* (regional interrogative particles).

A.3.3 Evaluation phase

As a binary classification task on sentences, we measure the accuracy of FUDIA on an evaluation set Eval.

Evaluation set The evaluation set contains 200 sentences aiming to represent the kind of data present in UD. The sentences are randomly picked from Annodis (Péry-Woodley et al. 2011; Afantenos et al. 2012), TenNovels (Reinhardt 2019b), Defrancq_written and Defrancq_spoken (extracted from examples in the book by Defrancq 2005), OFROM (Avanzi et al. 2012) and Maya (Reinhardt 2016). Half of the sentences are from spoken corpora and half from written corpora, with similar genres. Orthogonally, half of the sentences are from non-specific corpora, and the other half from question or interrogative corpora. The number of sentences from each corpus used to build Eval is detailed in Tab. A.1.

The two classes to predict are has_interrogative (the sentence contains at least one interrogative) and no_interrogative. We asked 12 French native speakers to annotate Eval according to this schema. The majority judgment is taken as the gold standard on Eval. The evaluation set, its gold labels, and all annotator's labels are freely accessible on GitHub. [11]

To compute FUDIA's accuracy on Eval, we first parsed it with ArboratorGrew (Guibon et al. 2020). The written part is parsed with fine-tuning on French GSD `dev` (1476 sentences, obtaining LAS = 0.922), and the spoken part with

11. `https://github.com/Valentin-D-Richard/FUDIA/tree/main/Eval`

Corpus	Medium	Type	#sentences
Annodis	written	non-specific	50
TenNovels	written (novels)	questions	25
Defrancq_written	written (newspapers)	embedded interrogatives	25
OFROM	spoken	non-specific	50
Maya	spoken (TV cartoon)	questions	25
Defrancq_spoken	spoken (radio)	embedded interrogatives	25

Table A.1 – Number of sentences per origin in the evaluation corpus Eval

fine-tuning on Rhapsodie `train` + ParisStories `train` (total: 2675 sentences, LAS = 0.818). We also report precision (ratio of true positives over positives) and recall (ratio of true positives over actual interrogatives) to better understand FUDIA's characteristics.

Baselines We compare these results with two rule-based baselines, whose codes are given in appendix B.4.

The first baseline, called QUECQ?, simply assigns the class has_interrogative to a sentence if and only if it contains a question mark, the *est-ce que* sequence, or a French QU-word. We expect this last heuristics to perform badly, given that many French QU-words can have other parts-of-speech (e.g. *que* '*what/that*' and *quand* '*when*' can be complementizers) or can be used in other constructions (e.g. *où* '*where*' and *lequel* '*which one*' can be relative proforms).

Contrary to the first baseline, working on raw text, the second baseline, called SimpleFUDIC (Simple French UD Interrogative Classifier), takes advantage of basic UD features and dependencies. It assigns the class has_interrogative to a sentence if and only if it contains a question mark, the *est-ce que* sequence, subject-verb inversion with a personal pronoun (except in imperative mood), or a word with feature `PronType=Int`. Comparing FUDIA to SimpleFUDIC may help us shed light on the utility of the different modules of our program.

A.3.4 Extraction

FUDIA was run on the French UD corpora. The sentences containing at least one node with `ClauseType=Int` were extracted, forming the French Interrogative Bank (FIB). [12]

The FIB is split according to the original splits, i.e. FIB `dev` contains the sentences extracted from GSD `dev`, ParTUT `dev`, etc., and similar for FIB `test` and FIB `train`. In total, FIB contains 2,972 sentences (38,884 tokens). They contain a total of 3,013 nodes labeled as interrogative heads. Due to license incompatibility, different dataset parts fall under different licenses.

12. The FIB is available at `https://github.com/Valentin-D-Richard/UD_French-FIB`

Corpus	#sentences in FIB	% in FIB	% extracted
FQB	2230	75.0	97.4
GSD	322	10.8	2.0
ParisStories	131	4.4	4.7
Rhapsodie	130	4.4	4.1
Sequoia	96	3.2	3.1
ParTUT	39	1.3	3.8
PUD	24	0.8	2.4
All corpora	2972	100	10.0

Table A.2 – Proportion of sentences from origin corpus and proportion of sentences extracted to the FIB.

The most significant part of the FIB comes from the FQB. The French Question Bank (FQB) (Seddah and Candito 2016) is a UD corpus of French questions. Initially created to help train syntactic parsers, it provides an interesting variety of French questions. 75% of FIB sentences are from the FQB, and 97.4% of the FQB sentences were extracted to the FIB. [13] However, as we show in section A.4.4, adding interrogatives from other corpora improves the syntactic diversity of examples. Table A.2 summarizes the proportion of sentences from origin corpora and their proportion of extracted sentences.

As stated above, we provide an enriched version of the FIB with cue edges. These edges can be used to define requests targeting specific constructions. The FIB suite contains a Python script `stats.py`, which can count the number of occurrences according to a classification scheme. It also includes visualization options and access to raw data. To test it, we implemented the classification of root interrogatives proposed by Coveney 2011. We report the results in section A.4.4.

A.4 Results

A.4.1 Annotation task

The annotation task obtained a good inter-annotator agreement: Fleiss $\kappa = 0.780$, Cohen κ: $\mathsf{min} = 0.613$, $\mathsf{mean} = 0.781$, $\mathsf{max} = 0.924$. Nevertheless, some patterns triggered disagreement.

We spot 11 sentences of Eval with high standard deviation over annotators ($\mathsf{std} > 0.47$). The first difficulty concerns embedded interrogatives, which seem to be harder to identify. The other most debated sentences involve the distinc-

13. The FQB contains monoclausal declarative sentences which where not included in the FIB.

tion between root interrogatives and question-raising declaratives. In particular, annotators did not know how to handle "question tags", like *hein ?* 'right?' or *non ?* 'no?', e.g. (4-a).

(4) a. Tu me fais confiance, non ?
You to_me do trust, no ?
'*You trust me, don't you?*' (TenNovels)

b. Que ce n' est pas quelqu'un qui vous fait une mauvaise blague
that it NEG is not someone who to_you do a bad trick
?
?
'*That it's not someone playing a bad joke on you?*' (TenNovels)

The last source of difficulty concerns two sentences having the word *que* '*what/that*' fronted plus a question mark. The majority of participants annotated them as interrogatives. But we think that they are actually question-raising declarative complementizer clauses, as illustrated by the gloss on sentence (4-b). Accessing the discursive context of this sentence would have probably helped participants better recognize its syntactic structure.

A.4.2 Evaluation results

The accuracy, precision, and recall of FUDIA and our two baselines are displayed in Tab. A.3. As expected, the accuracy of FUDIA is better than the two baselines.

QUECQ? Baseline QUECQ? has a very low precision. This was expected, given that French QU words are very polysemous. Nevertheless, QUECQ? has a high recall. Only 8 sentences containing an interrogative were not detected by QUECQ?. All but one correspond to yes-no embedded interrogatives. This shows that yes-no embedded interrogatives are hard to detect.

In the literature (e.g. Rossi-Gensane et al. 2021), detecting embedded yes-no interrogatives is often performed by listing interrogative-embedding verbs and looking for the interrogative conjunction *si* '*whether / if*'. Here, this technique might fail on some examples for two reasons. First, *si* does not always directly follow the verb (e.g. in coordinated interrogative clauses). Second, some interrogative-embedding verbs might be missed.

In FUDIA, we address that problem by analyzing the internal structure of the interrogative. However, in UD, no feature distinguishes the subordination conjunction *si* introducing a interrogative clause from its homophone *si* introducing a conditional clause. We rely on heuristics based on several characteristics to tell them apart (e.g. governing relation or verb form).

Model	accuracy	precision	recall
QUECQ?	0.660	0.524	**0.892**
SimpleFUDIC	0.755	0.667	0.676
FUDIA, on Eval	**0.905**	**0.966**	0.770
on written	0.860	0.917	0.647
on spoken	0.950	1.00	0.875

Table A.3 – Scores of FUDIA and the baselines on French interrogative detection.

SimpleFUDIC The baseline SimpleFUDIC circumvents the synonymy of French QU-words by resorting to the feature `PronType=Int`. However, many interrogative pronouns do not bear `PronType=Int` in UD. This makes the recall of SimpleFUDIC drop. This fact comforts us with the necessity of the FUDIA module `prontype`.

The precision of SimpleFUDIC is relatively low. This is primarily due to the high number of subject-verb inversions wrongly classified as interrogative. Even outside of imperative mood and interrogatives, subject-verb inversion with a personal pronoun is not so rare in French, especially: (i.) after some adverbs (e.g. *ainsi* '*so*'), (ii.) as speech reporting phrase or (iii.) in other inserts.

In FUDIA, case (i.) is handled by ruling out first-place adverbs. In contrast, cases (ii.) and (iii.) are hard to tackle. No feature in UD@2.12 could help us differentiate subject-verb inversion in inserts from interrogative subject-verb inversion. We could use a list of speech-reporting predicates. However, this would go against our rule design policy. Moreover, interrogative inserts are attested, like "*faut-il le préciser*" in (5-a).

(5) a. [...] et cela, faut -il le préciser, sans donner un
[...] and that, must -EXPL it.ACC clarify, without give a
blanc-seing [...]
blank-check
'*and this, should it be clarified, without giving a blank check*' (Sequoia)

To solve this issue, we rely on a linguistic observation: speech-reporting predicates and other inserts usually do not have any NP object (e.g. unlike *le* in (5-a)), as their semantic argument is the speech itself (Kahane et al. 2021). We advocate for adding to UD the `parataxis:insert` relation suggested by Kahane et al. 2021 to identify inserts better and, among them, speech-reporting predicates.

FUDIA The precision of FUDIA is high: 96,6% of sentences detected are indeed interrogatives. This score presumably results from the many heuristics that efficiently filter out non-interrogatives. However, its recall is lower than that QUECQ?. To investigate why, we proceeded to a qualitative evaluation.

Out of the 17 false negatives, we estimate that 13 are due to the parser. The

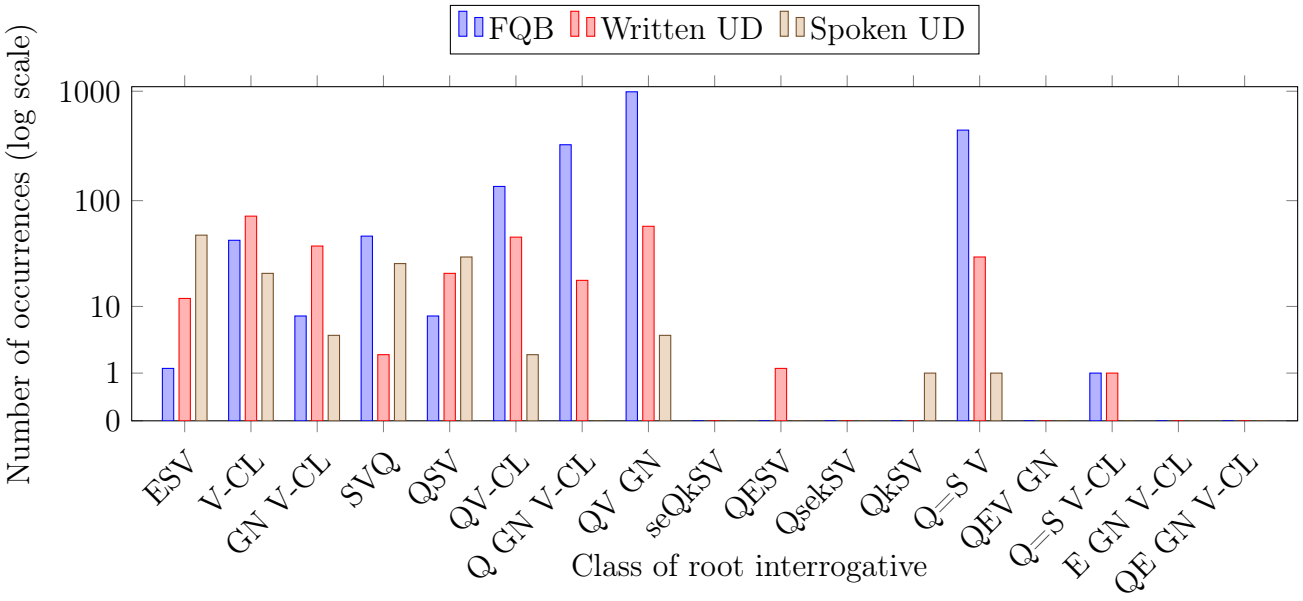


Figure A.3 – Bar plot of number of occurrences of root interrogatives according to Coveney 2011's classification. Corpora are grouped as such: FQB, Written UD (GSD, Sequoia, ParTUT, PUD), Spoken UD (Rhapsodie, ParisStories). The list of class acronyms is given in Appendix B.2.

most common parser mistake is confusing interrogative words with relative words or subordination conjunctions. Notably, the parser fine-tuned on GSD wrongly annotated 6 out of 8 occurrences of *comment* '*how*' as subordination conjunction. However, this word is always an adverb in French (and all occurrences in the training data are adverbs). The written part of Eval contains more parser errors. This explains why FUDIA performs worse on the written part of Eval than on its spoken part. However, we do not know the reason for this lower performance. We expected the parser fine-tuned on GSD to be better. In particular, the GSD trees contain, on average, more precise features than ParisStories, and the LAS score was higher than for the parser fine-tuned on Rhapsodie and ParisStories.

This evaluation shows that, assuming a good parser, even a rule-based program can perform well on raw text and automatically identify French interrogatives. The efficiency of this method mainly depends on the quality of the parser, especially for subtle distinctions, like pronoun types.

A.4.3 Other difficulties

FUDIA is not only a classifier. It identifies the interrogative head and links it to the morphosyntactic interrogative marking(s). In this section, we discuss issues related to this challenge.

Since colloquial French allows *in situ* QU-words, finding the phrase head and

the interrogative clause head is complex. To do so, FUDIA proceeds as follows. After detecting an interrogative word, it must first decide whether its governor belongs to the same phrase. For example, the object relation `obj` can link a verb with both an *in situ* QU-word (6) or an embedded sluicing (3-b). When the QU-word is a pronoun (e.g. *quoi* '*what*', *qui* '*who*'), there is no easy way to determine in which case we are. Given that almost all sluicing with an interrogative pronoun observed in our development corpus are governed by *savoir* '*know*', we decided to use the lemma of the matrix verb as a categorizing feature. This is the only case where we had to violate our rule design policy.

(6) [...] vous en pensez quoi ici ?
[...] you of-it think what here ?
'What do you think of this, here?' (Rhapsodie)

A.4.4 Statistics on the French Interrogative Bank

In Fig. A.3 we plot the statistics of root interrogatives in the FIB according to Coveney 2011's classification[14] (given in appendix B.2). The distribution of interrogative types in the FQB has many similarities with the other French UD corpora. The graphic yet shows that some classes with few representatives in FQB gain from including the other UD corpora, e.g `ESV` (*est-ce que + subject + verb*). Some examples of rare types are also found, e.g. two `QESV` (*QU-word + est-ce que + subject + verb*) from ParTUT.

The other UD corpora also bring many subordinated interrogatives: 117 were found in the other written corpora, plus 77 in spoken corpora, compared to only 5 embedded in FQB. Retrieving them would not have been possible if we had only relied on a list of French interrogative embedding predicates. As an illustration, if we used the lists given by the GGF (Delaveau, Cappeau, and Dagnac 2021, Tab. XXII-7 p.1414) and Defrancq 2005, chap. 1 fn. 11, some sentences would have been missed out. Out of the 28 predicates embedding an interrogative found in the FIB, 6 are listed neither by the GGF nor by Defrancq 2005. These are: *connaitre* '*know*', *enseigner* '*teach*', *interroger* '*question*', *mesurer* '*measure*', *souligner* '*emphasize*' and *tester* '*test*'.

A.5 Limits

Here, we present the limits of our work.

14. Here is a brief explanation of the class letters: E = *est-ce que*, S = subject, V = finite verb, CL = subject clitic, GN = noun phrase, Q = QU-word, se...k = *c'est ... que* '*it's ... that*' (cleft), and Q=S means subject QU-word.

A.5.1 Limits of the rule design policy

FUDIA only works on a specific annotation schema: Universal Dependencies. Furthermore, although some pattern descriptions occur in other similar languages, most of the structures listed in the rules are specific to French. Extending it to another language is not trivial. However, the additional annotation proposal (`ClauseType=Int` feature, `cue` edges) aims to be compatible with any language.

We chose not to rely on question marks and a list of open class words (e.g. interrogative-embedding predicates). However, due to some syntactical characteristics that UD could not provide, we decided to violate this rule design policy at some rare specific points, as mentioned in section A.4.3. It is unclear how changing this policy would benefit FUDIA. It might have a limited impact, as most false negatives are due to the parser.

FUDIA is highly dependent on the quality of input annotations. The confusion in pronoun types shows it (section A.4.2). Moreover, non-standard or unexpected typographic situations prevent FUDIA from predicting the expected result, e.g. unexpected word segmentation with dashes and euphonic *t* in subject clitic inversion. We sometimes used the `form` feature (in the original text) instead of `lemma` to try to solve some of these issues. Even if we ensured it correctly detects all-caps words (as found in the development corpus), it might fail on typos and, more generally, on noisy texts.

A.5.2 Evaluation limits

FUDIA includes rules to account for regional variants, e.g. *-tu* (Quebec) or *-ti* interrogative particles (7). However, none of the French UD corpora or evaluation subcorpora includes extracts of spoken Quebec or African varieties of French. Therefore, even if these cases are theoretically implemented in the program, it is unclear how people would annotate these variants in practice and, thus, whether FUDIA would correctly detect them.

(7) On peut -tu être constructifs ?
One can -PART be constructive ?
'*Can we be constructive?*' (*La Presse* 2014, from Bally 2022)

This issue also concerns colloquial variants. For example, *qu'est-ce que* /kɛsk(ə)/ exists under a non-standard apocoped variant *qu'est-ce* pronounced /kɛs/ like in (8). As it is present in Dev, we took example (8) into account. However, other kinds of non-documented non-standard phenomena might fail to be detected correctly.

(8) [...] qu' est -ce vous comptez faire dans un avenir proche ?
[...] what is -it you intend do in a future near ?
'*what do you intend to do in a near future?*' (GSD)

A.6 Conclusion

We presented FUDIA (French UD Interrogative Annotator), a rule-based program that retrieves and adds annotations to interrogatives from French sentences annotated in Universal Dependencies (UD). It adds the feature `ClauseType=Int` to any interrogative clause head. Additional modules add some missing `PronType=Int` and reannotate some fixed expressions related to interrogatives.

FUDIA obtains good results when evaluated on unseen parsed raw text. Its efficiency is highly dependent on the quality of the input syntactic annotations. This comforts us with the idea that rule-based programs are still relevant for some syntax NLP tasks, provided these programs can access structured enough qualitative data.

The French Interrogative Bank (FIB), a treebank of UD-annotated French interrogatives, was extracted from existing French UD corpora (FQB, GSD, ParisStories, ParTUT, PUD, Rhapsodie and Sequoia) and enriched with interrogative-specific dependencies using FUDIA. We illustrated how FIB enriches the FQB by bringing some little-represented interrogative structures and embedded interrogatives.

Appendix B

Appendix of Chapter A

B.1 Grammaticalized QU-expressions: some linguistic background

Here, we expose linguistic literature that motivates our classification of the QU-expressions mentioned in section A.3.2 as fixed.

The expression *est-ce que* lit. '*is it that*' was shown to be grammaticalized and even lexicalized by Druetta 2003, 2004, chap. 5 and other linguists before (Foulet 1921; Kayne 1972; Rooryck 1994). It is also reported as fixed in the Grande Grammaire du Français (GGF) (Delaveau, Cappeau, and Dagnac 2021, §XII-3.1.2 p.1403) and the Encyclopédie Grammaticale du Français (EGF) (Berrendonner 2023). Contrary to common belief, it is not a clefting device (anymore), but just marks a clause as interrogative (Obenauer 1976).

Similarly, *qu'est-ce que*[1] lit. '*what is it that*' is claimed to be a complex lexical unit by Obenauer 1976, the GGF (Delaveau, Cappeau, and Dagnac 2021, Tab. XII-5 p.1405) and Orféo. Its grammaticalization may have been more recent than *est-ce que*. Arguments to consider *qu'est-ce que* as different from other *QU + est-ce que* constructions are the following. Hulk 1996's study on interrogative acquisition (reported by Zuckerman 2001, chap. 5) shows that *qu'est-ce que* is acquired earlier to other *QU + est-ce que* constructions. Farmer 2013's sociopragmatic corpus study on French movies emphasizes the gap between their respective uses. Farmer claims that "the interrogative word *que* so often occurs with *est-ce que* across speaker class and sex—in every style and in every decade—that it appears to be lexicalized". Finally, *qu'est-ce que* can be used as an exclamative adverb (Dekhissi 2016), (GGF, §IX-10.4.3 p.1119).

We believe that the reasons these expressions are not currently annotated as fixed probably include:

1. The unit *qu'est-ce que* has a variant *qu'est-ce qui* when it introduces a sentence lacking its subject.

1. French spelling being very conservative, thus giving wrong intuitions about the morphology of these expressions
2. A potential lack of specific linguistic knowledge from annotators
3. A possible tendency of parsers to decompose strings with determined relations (and thus, to avoid `fixed` relations)
4. A desire always to syntactically decompose historically decomposable phrases (especially for corpora automatically converted from SUD (Gerdes et al. 2019))

B.2 Coveney's classification of French interrogative forms

Coveney 2011 proposes a syntactic classification of French root interrogatives into classes, 3 primary classes for yes-no interrogatives, 10 primary classes for constituent interrogatives, and 4 hybrids (rare and non-standard). Only finite verbal clauses are considered.

Yes-No interrogatives

— `ESV`: *est-ce que* marker

(1) Est -ce que les autres sont partis ?
Is -it that the others are gone ?
'*Are the others gone?*'

— `V-CL`: clitic inversion

(2) Sont -ils partis ?
Are -they gone ?
'*Are they gone?*'

— `GN V-CL`: complex inversion

(3) Les autres sont -ils partis ?
The others are -they gone ?
'*Are the others gone?*'

Constituent interrogatives

— `SVQ`: in-situ QU-word

(4) Ils sont partis où ?
They are gone where ?

'*Where are they gone?*'

— `QSV`: fronted QU-word

(5) Où ils sont partis ?
Where they are gone ?
'*Where are they gone?*'

— `QV-CL`: QU-word + clitic inversion

(6) Où sont -ils partis ?
Where are -they gone ?
'*Where are they gone?*'

— `Q GN V-CL`: QU-word + complex inversion

(7) Où les autres sont -ils partis ?
Where the others are -they gone ?
'*Where are the others gone?*'

— `QV GN`: QU-word + stylistic inversion

(8) Où sont partis les autres ?
Where are gone the others ?
'*Where are the others gone?*'

— `seQkSV`: clefted QU-word

(9) C' est où qu' ils sont partis ?
It is where that they are gone ?
'*Where are they gone?*'

— `QESV`: QU-word + *est-ce que*

(10) Où est -ce qu' ils sont partis ?
Where is it that they are gone ?
'*Where are they gone?*'

— `QsekSV`: QU-word + *est-ce que* variant

(11) Où c' est qu' ils sont partis ?
Where it is that they are gone ?
'*Where are they gone?*'

— `QkSV`: QU-word + *que*

(12) Où qu'ils sont partis ?
Where that they are gone
'*Where are they gone?*'

— `Q=S V`: subject QU-word

(13) Lesquels sont partis ?
Which_ones are gone ?
'*Which ones are gone?*'

Hybrids

— `QEV GN`: QU-word + *est-ce que* + stylistic inversion

(14) Avec qui est -ce que travaille Nicole Dupont ?
With who is -it that work Nicole Dupont ?
'*Who is Nicole Dupont working with?*'

— `Q=S V-CL`: subject QU-word + clitic inversion

(15) De ces fillettes, lesquelles sont -elles les tiennes ?
Of these little_girls, which_ones are -they the yours ?
'*Which of these little girls are yours?*'

— `E GN V-CL`: *est-ce que* + complex inversion

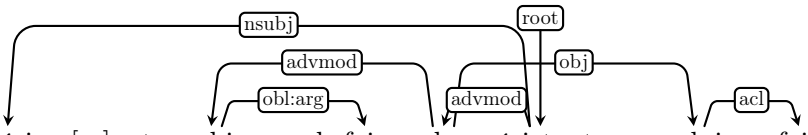
(16) Est -ce que demain les sauveteurs pourront -ils s'
Is -it that tomorrow the rescuers can -they REFL
approcher des alpinistes en détresse ?
approach to mountaineers in distress ?
'*Will rescuers be able to approach climbers in distress tomorrow?*'

— `QE GN V-CL`: QU-word + *est-ce que* + complex inversion

(17) Pourquoi est -ce que le rédacteur de la rubrique des chats
Why is -it that the author of the section of cats
écrasés veut -il créer une pétition ?
crushed wants -he create a petition ?
'*Why does the editor of the miscellaneous news section want to set up a petition?*'

B.3 Linguistic conceptual limits

Contrary to interrogative clauses, we found no clear definition of an *interrogative phrase* in the literature. In interrogatives containing a fronted phrase, the



(18) Le nouveau matériau [...] est combien de fois plus résistant que celui fait de cellulose ?
The new material [...] is how_many of times more resistant than the_one made of cellulose ?
'*How many times stronger than the one made of cellulose is the new material?*' (FQB)

whole fronted phrase is considered to be the interrogative phrase. However, when the QU-word is *in situ*, linguistic tests must be used to determine the interrogative phrase. Definition (19) attempts to capture this intuitive notion.

(19) In a constituent interrogative clause C, the interrogative phrase of the interrogative word W is the largest phrase containing W and which can be fronted in C.

Let us illustrate this definition on examples. Consider sentence (18), containing the QU word *comment*, abstracting over the multiplicity factor of a comparative. We can perform fronting tests, as in (20), to identify the extractable phrase. Even if the judgment differences between the items of (20) are not very sharp, it seems like the interrogative phrase is *combien de fois*.

(20) a. [Combien de fois]$_i$ est-il [t_i plus résistant [que celui-ci]] ?
b. ?[[Combien de fois] plus [que celui-ci]]$_i$ est-il [t_i résistant] ?
c. ?[[Combien de fois] plus résistant [que celui-ci]]$_i$ est-il t_i ?

In contrast, more than one phrases can be fronted in some configurations exhibiting recursively embedded prepositional phrases (PPs). This holds in (21): both *de quel dragon* and *en quête de quel dragon* are candidate for the interrogative phrase. In this case, the maximality requirement of (19) selects *en quête de quel dragon*.

(21) a. Elle est [en quête [de quel dragon]] ?
She is [in quest [of which dragon]] ?
'*Which dragon is she after?*'
b. [De quel dragon]$_i$ est-elle [en quête t_i] ?
c. [En quête [de quel dragon]]$_i$ est-elle t_i ?

In question with recursively embedded PPs, we observe variation regarding which sub-PP can be fronted (Le Goffic 2024, § 3.2). For instance, question (22) seems to have the same syntactic structure as question (21), but behaves differently regarding fronting.

(22) a. Il est [en retard [de combien de jours]] ?
He is [in lateness [of how-many of days]] ?
'*How many days late is he?*'

b. [De combien de jours]$_i$ est-il [en retard t_i] ?
c. ?[En retard [de combien de jours]]$_i$ est-il t_i ?

As fronting cannot be tested in FUDIA, detecting interrogative phrases is a theoretical limit. Furthermore, definition (19) fails to capture the intuition behind interrogative phrases when no phrase containing the QU-word can be fronted, e.g. with extraction islands. Therefore, the `PhraseType=Int` annotations added by FUDIA are approximations of a pre-theoretical concept.

B.4 Baseline codes

```
rule wh { %
  pattern {
    WH[lemma="que"|"quoi"|"qui"|"lequel"|
    "quand"|"où"|"comment"|"combien"|
    "pourquoi"|"quel", !ClauseType]
  }
  commands {
    WH.ClauseType = "Int"
  }
}
rule int_mark { %
  pattern { P[lemma="?", !ClauseType] }
  commands { P.ClauseType = "Int" }
}
rule ecq { %
  pattern {
    N1[lemma="être", !ClauseType] ;
    N2[lemma="ce"] ; N3[lemma="que"] ;
    N1 < N2 ; N2 < N3
  }
  commands { N1.ClauseType = "Int" }
}
strat main { Seq(Onf(wh),Onf(int_mark),
  Onf(ecq)) }
```

Listing B.1 – Code of QUECQ?

```
rule wh { %
  pattern {
    WH[PronType="Int", !ClauseType]
  }
  commands { WH.ClauseType = "Int" }
}
rule int_mark { %
  pattern { P[lemma="?", !ClauseType] }
  commands { P.ClauseType = "Int" }
}
rule ecq { %
```

```
    pattern {
      N1[lemma="être", !ClauseType] ;
      N2[lemma="ce"] ; N3[lemma="que"] ;
      N1 < N2 ; N2 < N3
    }
    commands { N1.ClauseType = "Int" }
  }

  package pkg_inv {
    rule inv_1 { %
      pattern {
        S[upos="PRON"] ;
        S[form="ce"|"je"|"tu"|"il"|"elle"|
        "on"|"nous"|"vous"|"ils"|"elles"|
        "-ce"|"-je"|"-tu"|"-il"|"-elle"|
        "-on"|"-nous"|"-vous"|"-ils"|
        "-elles"|"-t-il"|"-t-elle"|
        "-t-on"|"-t-ils"|"-t-elles"] ;
        s: CL_HEAD -[1=expl|nsubj]-> S
      }

      pattern {
        CL_HEAD[upos="VERB"|"AUX",
        VerbForm="Fin", !ClauseType] ;
        CL_HEAD < S
      }
      without { V[Mood="Imp"] }
      commands {
        CL_HEAD.ClauseType = "Int"
      }
    }
    rule inv_2 { %
      pattern {
        S[upos="PRON"] ;
        S[form="ce"|"je"|"tu"|"il"|"elle"|
        "on"|"nous"|"vous"|"ils"|"elles"|
        "-ce"|"-je"|"-tu"|"-il"|"-elle"|
        "-on"|"-nous"|"-vous"|"-ils"|
        "-elles"|"-t-il"|"-t-elle"|
        "-t-on"|"-t-ils"|"-t-elles"] ;
        CL_HEAD[!ClauseType] ;
        s: CL_HEAD -[1=expl|nsubj]-> S
      }
      pattern {
        CL_HEAD -[1=cop|aux]-> V ;
        V[upos="VERB"|"AUX",
        VerbForm="Fin"] ; V < S
      }
      without { V[Mood="Imp"] }
      commands {
        CL_HEAD.ClauseType = "Int"
```

```
63        }
64      }
65    }
66    strat inv { Seq(pkg_inv) }
67
68    strat main { Seq(Onf(wh),Onf(int_mark),
69      Onf(ecq),Onf(inv)) }
```

Listing B.2 – Code of SimpleFUDIC

Appendix C

Corpus extracts

In section 5.2.2, I present a corpus study to provide real-life examples of wh-anaphora and argue for a different annotation scheme. This appendix presents the 23 dialogues from the corpus of study where the first co-referring mentions after the antecedent is a pronoun. The name of the speakers (SPK1,...) are copied from the corpus. Indications about potential annotations mistakes are added. For a quantitative review about ANCOR's annotation of chains containing an interrogative word, see Richard 2025a.

I use the same notational conventions as in the rest of the manuscript. To illustrate how I transcribe ANCOR's annotations into indices, consider an extract of the form (1).

(1) A^{j} ... B_{j} ... $C_{j'}$ ... $D_{j''}$ ... E_{j} ... $F_{j'}$

This index arrangement means the following:

— Expressions A, B, C, D, E and F are annotated as mentions.
— Mention A has the feature NEW=YES, i.e. it is the antecedent.
— Mentions A, B and E are in a co-reference chain.
— Mentions A and C are in a bridging chain.
— Mentions A and D are in a different bridging chain.
— Mentions C and F are in a co-reference chain.

Extracts

(2) a. SPK1: et vous aviez [quel âge]^{j} à cette époque -là à peu près
SPK1: and you had [what age]^{j} at this time there to few close
?
?
'*And how old were you at that time, approximately?*' (026_C-4)

b. SPK2: ça$_j$ devait être oh là oh là là je vais retrouver ça$_j$ devait être [treize ans]$_{j'}$
'*It$_j$ must have been... oh dear oh dear. I'm going to remember. It$_j$ must have been [thirteen years]$_{j'}$*'

In the two following extracts, the NPs *l'orthographe* ('*(the) spelling*') are not annotated in a chain in ANCOR. This is probably an annotation mistake.

(3) a. SPK2: euh dans [quelle matière]^j étiez-vous le plus fort à l'école ? (CO2_ESLO_001_C)
'*[Which subjec]^j t were you best at in school?*'
b. SPK1: bah c$_j$'est peut-être ça quand même l'orthographe
'*Well, maybe that$_j$'s the spelling after all.*'

(4) a. SPK1: et [quel genre de choses]^j y cherchez-vous le plus souvent ? l'orthographe et [le sens]$_{j'}$ ou [des renseignements]$_{j''}$ [l'histoire]$_{j'''}$ euh ? (026_C-5)
'*And [what kind of things]^j do you look up most often? Spelling and [meaning]$_{j'}$, or [information]$_{j''}$, [history]$_{j'''}$, uh?*'
b. SPK2: [n' importe quoi]$_j$ au fond
'*anything$_j$ really*'

In the following extract, the occurrences of *lesquels* ('*which ones*') are not part of a chain in ANCOR.

(5) a. SPK2: est-ce que vous faites des mots-croisés ? (012_C-3)
'*Do you do crosswords?*'
b. SPK1: oui bien sûr
'*Yes of course.*'
c. SPK2: oui lesquels ?
'*Which ones?*'
d. SPK1: lesquels dans [quel journal]^j dans [quel journal]^j ?
'*Do you mean "In which^j newspaper"?*'
e. SPK2: oui
'*Yes.*'
f. SPK1: oh [n'importe lequel]$_j$ et de préférence dans [le Canard Enchaîné]$_{j'}$
'*Any$_j$ (of them) and preferably in [the Canard Enchaîné]$_{j'}$*'

(6) a. SPK1: madame lorsque vous étiez encore à l'école dans [quelle matière]^j étiez vous le plus fort ? (021_C-6)
'*Madam, when you were still at school, [which subject]^j did you do best in?*'
b. SPK2: ah c$_j$'était littérature$_{j'}$.
'*It$_j$ was literature$_{j'}$*'

(7) a. SPK1: dans [quelle matière]^j étiez-vous le plus fort à l'école ? (025_C-2)

'*[Which subject]^{j} were you best at in school?*'

b. SPK2: oh là c'est tellement vieux ah c$_{j}$'était [la géographie]$_{j'}$ puis [la science]$_{j''}$
'*Oh, that's so old. Ah, it$_{j}$ was geography$_{j'}$ and then science$_{j''}$.*'

(8) a. SPK1: oui enfin dans [quelle matière français‡]^{j} étiez-vous le plus fort vous même à l'école monsieur ? (006_C-3)
'*Yes, well, [which subject]^{j} were you best at in school, sir?*'

b. SPK2: géographie$_{j'}$ histoire$_{j''}$ géographie$_{j'}$
'*Geography$_{j'}$, history$_{j''}$ and geography$_{j'}$.*'

c. SPK1: a oui histoire$_{j''}$ histoire$_{j''}$ géographie$_{j'}$
'*Oh yes, history$_{j''}$, history$_{j''}$ and geography$_{j'}$.*'

d. SPK2: oui
'*Yes.*'

e. SPK1: oui
'*Yes.*'

f. SPK2: mais à l'époque c$_{j}$'était pas poussé comme maintenant
'*But back then it$_{j}$ wasn't as advanced as it is now*'

(9) a. SPK1: vous commencez vers [quelle heure]^{j} par exemple le matin ? (013_C-1)
'*[What time]^{j} do you start in the morning, for example?*'

b. SPK2: ça dépend des jours euh trois jours par semaine à cette époque-ci c$_{j}$'est [trois heures moins le quart]$_{j'}$ [deux heures et demie]$_{j''}$ [trois heures moins lequart‡]$_{j'}$ de nuit
'*It depends on the day, um, three days a week at this time of year, it$_{j}$'s [quarter to three]$_{j'}$, [two and a half hours]$_{j''}$, [quarter to three]$_{j'}$ at night.*'

c. SPK1: hm hm
'*Hm hm.*'

d. SPK2: ma femme c$_{j}$'est plus tard elle se lève à [six heures]$_{j'''}$ [six heures et quart]$_{j''''}$
'*For my wife, that$_{j}$'s later, she wakes up at [six o'clock]$_{j'''}$, [quarter past six]$_{j''''}$.*'

(10) a. SPK1: alors dans [quelle matière]^{j} aimeriez-vous que les enfants soient forts ? est-ce que vous avez des des z-‡ des avis là-dessus ? (019_C-1)
'*So, in [what subject]^{j} would you like the children to excel? Do you have any thoughts on this?*'

b. SPK2: non
'*No.*'

c. SPK1: non ?
'*No?*'

d. SPK2: non

'*No.*'

e. SPK1: pas spéciale‡
'*Not particularly.*'

f. SPK2: pas spécialement
'*Not particularly.*'

g. SPK1: non
'*No.*'

h. SPK2: non hm hm
'*No hm hm.*'

i. SPK1: enfin
'*Anyway.*'

j. SPK2: aujourd'hui c_j'est peut-être [la chimie et la physique]$_{j'}$ peut-être
'*Today, it$_j$ might be [chemistry and physics]$_{j'}$, perhaps.*'

(11) a. SPK2: et vous personnellement dans [quelle matière]^j vous aimeriez que vos enfants soient forts à l'école ? (096_C-1)
'*And personally, in [which subject]^j would you like your children to excel at school?*'

b. SPK1: c_j'est [les mathématiques]$_{j'}$ euh je crois n-‡ oui
'*It$_j$'s mathematics$_{j'}$, I think, yes.*'

(12) a. SPK1: et vous à l'école vous étiez le plus fort à [quelle matière]^j ? (024_C-4)
'*And at school, [what subject]^j were you best at?*'

b. SPK2: ah c_j'était l'ort-‡ euh [l'orthographe]$_{j'}$
'*Ah, it$_j$ was [the spelling]$_{j'}$.*'

(13) a. SPK1: et dans [quelle matière]^j étiez-vous le plus fort à l'école ? (019_C-3)
'*And [which subject]^j were you best at in school?*'

b. SPK2: euh moi c_j'était dans le euh [l'orthographe]$_{j'}$
'*Well, for me it$_j$ was in, um, spelling$_{j'}$.*'

In the following dialogue, *qui* ('*who*') is annotated as a mention but belongs to no chain in ANCOR.

(14) a. SPK1: pour bien apprendre [quels gens]^j est-ce qu'il faudrait fréquenter ? qui est-ce qu'il faudrait voir ? (015_C-2)
'*To learn properly, [which people]^j should you get to know? Who should you see?*'

b. SPK2: tout$_j$‡
'*Everybody$_j$.*'

(15) a. SPK2: on vous l'avait offert ? ou on‡ à [quelle occasion]^j ? (096_C-2)
'*Was it offered to you? On [what occasion]^j?*'

b. SK1: oh c$_j$'était pour [un Noël]$_{j'}$ ou [le Jour de l'an]$_{j''}$ comme ça
'*Oh, it$_j$ was for Christmas$_{j'}$ or [New Year's Day]$_{j''}$, something like that.*'

In (16), the mention indexed by j'' is discontinuous: *une école ... privée* ('*a private school*'). Moreover, the NP *une petite école de campagne* ('*a small country school*') is annotated as a mention but doesn't belong to any chain in ANCOR. This is probably an annotation mistake.

(16) a. SPK2: et ils avaient été à l'école dans une des dans quel euh dans [quel type d'école]^j c'était [[une école publique]$_{j'}$ privée]$_{j''}$? (096_C-2)
'*And they had been to school in one of the... [What kind of school]^j was it? Was it [a public school]$_{j'}$ or [a private school]$_{j''}$?*'
b. SPK1: c$_j$'était une petite école de campagne sûrement
'*It$_j$ was certainly a small country school.*'

(17) a. SPK1: et alors [quel genre de choses]^j y cherchez-vous le plus souvent ? c$_j$'est [l'orthographe]$_{j'}$ [le sens du mot]$_{j''}$ s-‡ ou quoi ? (019_C-2)
'*So [what kind of things]^j do you look up most often? Is it spelling$_{j'}$, [the meaning of a word]$_{j''}$, or what?*'
b. SPK2: oui [le sens du mot]$_{j''}$ ou bien des [des indications]$_{j'''}$ sur les les pays les des choses comme ça
'*Yes, [the meaning of the word]$_{j''}$ or information$_{j'''}$ about countries, things like that.* '

(18) a. SPK4: et dans [quelle matière]^j est-il bon qu'un enfant soit fort ? (008_C-1)
'*And in [what subject]^j is it good for a child to be strong?*'
b. SPK1: ben maintenant moi à mon avis je pense que euh ce$_j$ serait pour l'avenir je pense pour les les enfants de maintenant ce$_j$ serait plutôt [les mathématiques]$_{j'}$ plus [1] qu'ils se dirigent vers la science
'*Well, nowadays, in my opinion, I think that, um, it$_j$ would, be for the future... I think that for today's children, it$_j$ would rather be mathematics$_{j'}$, so that they can move towards science.*'

(19) a. SPK3: ah oui ce serait [quel nom]^j euh ? (023_C)
'*What^j name would it be?*'
b. SPK2: c$_j$'était Ligneaux$_j$
'*It$_j$ was Ligneaux$_j$.*'
c. et puis du reste y a la rue [aux Ligneaux]$_j$ au bout de la un peu plus bas
'*And then there's Rue [aux Ligneaux]$_j$ at the end of the street, a little further down.*'

1. Listening to the audio (9:11) prompts me to consider that this token should actually be transcribed as *pour* instead.

d. SPK3: ah bon
'*Oh, really?*'

e. SPK2: il y avait beaucoup de Ligneaux$_j$ par ici
'*There were many Ligneaux$_j$ around here.*'

f. SPK3: oui
'*Yes.*'

g. SPK2: et ce$_j$ nom-là euh figurent‡ dans les registres de des paroisses
'*This$_j$ name appears in parish registers.*'

(20) a. SPK2: dans [quelle matière]^j étiez-vous le plus fort à l'école ? (078_C-4)
'*[Which subject]^j were you best at in school?*'

b. SPK1: oh j'étais fort en maths$_{j'}$ c'est surtout en maths$_{j'}$ que j'étais ‡
'*Oh, I was good at maths$_{j'}$. It$_j$ was mainly maths$_{j'}$ that I was good at.*'

(21) a. SPK1: et pour [quel genre de choses]^j surtout ? c$_j$'était des enfin [des fautes de français]$_{j'}$ ou (029_C-5)
'*And [what kind of things]^j in particular? Were they$_j$ [grammatical errors]$_{j'}$ or...?*'

b. SPK2: [des fautes de français]$_{j'}$
'*[Grammatical error]$_{j'}$.*'

In (22), the i index represents a co-reference chain that does not include *quelles matières* ('*which subjects*'). The antecedent *français* ('*French*') was introduced before in the dialogue.

(22) a. SPK1: euh enfin euh mettons que si vous aviez des enfants dans [quelles matières]^j est-ce que vous aimeriez qu'ils soient forts ? (015_C-1)
'*Well, let's say you had children. In [which subjects]^j would you like them to excel?*'

b. SPK2: en français$_i$
'*In French$_i$.*'

c. SPK1: en français$_i$?
'*In French$_i$?*'

d. SPK2: oui parce que je l'ai toujours été enfin plus douée en français$_i$ qu'en mathématiques et j'ai toujours eu horreur des mathématiques et des chiffres
'*Yes, because I've always been better at French$_i$ than maths, and I've always hated maths and numbers.*'

e. SPK1: ah bon ?
'*Oh, really?*'

f. SPK2: sous quelque forme que ce soit les chiffres euh je me trompe avec les zéros n'importe quoi euh j'en suis encore restée au stade le

plus bas
'*In any form whatsoever,.. the figures, um... I get confused with the zeros, whatever, um... I'm still at the lowest level.*'

g. SPK1: ah bon
'*I see.*'

h. SPK2: alors je crois que c$_j$'est le [le français]$_i$
'*So I think it$_j$'s French$_i$.*'

In (23), the audio (59:58) was used to reconstruct the segment order. The mentions *quelles circonstances* ('*what circumstances*') and *quel âge* ('*what age*') are both annotated as antecedents and give rise to respective bridging chains with *un anniversaire* ('*a birthday*') and *première communion* ('*a first communion*'). However, these two referential expressions do not belong to a common co-reference chain and only *quel âge* is in relation with *mes douze ans* ('*my twelve years*'), *onze ans* ('*eleven years*') and *quelque chose* ('*something*').

(23) a. SPK1: à la suite de [quelles circonstances]^j avez-vous possédé pour [la première fois]^i un stylo à plume ? c'est-à-dire à [quel âge]^m ? si vous pouvez (133_C-4)
'*Under [what circumstances]^j did you own a fountain pen for [the first time]^i? That is, at [what age]^j? If you can...*'

b. SPK1: c$_i$'est c$_i$'est [un anniversaire]$_{j',m'}$ ou [première communion]$_{j'',m''}$
'*It's [a birthday]$_{j',m'}$ or [a first communion]$_{j'',m''}$*'

c. SPK2: un stylo encre [la première]$_{j''}$ ou c'était `(inaudible)`
'*A pen ink, [the first]$_{j''}$... It was* `(inaudible)`'

d. SPK2: oui [mes douze ans]$_{m'''}$ quoi [onze ans]$_{m''''}$ quoi [quelque chose]$_m$ comme ça
'*Yes, [my twelve years]$_{m'''}$, well, [eleven years]$_{m''''}$, something$_m$ like that.*'

(24) a. SPK1: toujours la même suite dans les questions [quel genre de pièces]^j ou quels auteurs ? (CO2_ESLO_002_C)
'*Always the same sequence of questions: [what kind of plays]^j or which authors?*'

b. SPK2: hum je sais pas bon ben ça je euh en général je je regarde un peu sur [les journaux]^i les critiques oui oui [les pièces horribles]$_{j'}$ [les pièces noires]$_{j'}$ euh truc euh
'*Um, I don't know. Well, I usually look at [the reviews]^i in the newspapers, [the horrible plays]$_{j'}$, [the dark plays]$_{j'}$, um, stuff...*'

c. SPK1: où$_i$ est ‡ qu'il y a des hm qu'est-ce que c'est pour vous [une pièce horrible]$_{j'}$?
'*Where$_i$ are there some hm...? What is [a horrible play]$_{j'}$ for you?*'

d. SPK2: oh où$_{j'}$ il y a énormément de morts de des situations dramatiques euh des choses comme ça euh moi je vais au théâtre c'est pour

passer une soirée agréable euh
'*Oh, where there are lots of deaths and dramatic situations, things like that. I go to the theater to have an enjoyable evening.*'

e. SPK1: oui
'*Yes.*'

f. SPK2: essayer de voir [quelque chose de valable]$_j$
'*To try to see [something worthwhile]$_j$.*'

Bibliography

Abusch, Dorit (Feb. 1, 2010). "Presupposition Triggering from Alternatives". In: *Journal of Semantics* 27.1, pp. 37–80. ISSN: 0167-5133. DOI: 10.1093/jos/ffp009. URL: https://doi.org/10.1093/jos/ffp009 (visited on 10/10/2024) (cit. on pp. 11, 12, 51, 53).

Afantenos, Stergos, Nicholas Asher, Farah Benamara, Myriam Bras, Cécile Fabre, Mai Ho-dac, Anne Le Draoulec, Philippe Muller, Marie-Paule Péry-Woodley, Laurent Prévot, Josette Rebeyrolles, Ludovic Tanguy, Marianne Vergez-Couret, and Laure Vieu (May 2012). "An Empirical Resource for Discovering Cognitive Principles of Discourse Organisation: The ANNODIS Corpus". In: *Proceedings of the Eighth International Conference on Language Resources and Evaluation (LREC'12)*. LREC 2012. Istanbul, Turkey: European Language Resources Association (ELRA), pp. 2727–2734. URL: http://www.lrec-conf.org/proceedings/lrec2012/pdf/836_Paper.pdf (visited on 04/20/2023) (cit. on p. 151).

Agüero Bautista, Calixto (2001). "Cyclicity and the Scope of Wh-Phrases". PhD thesis. Massachusetts Institute of Technology. URL: https://dspace.mit.edu/handle/1721.1/8200 (visited on 02/11/2025) (cit. on p. 37).

Aloni, Maria (2001). "Quantification under Conceptual Covers". PhD thesis. Amsterdam: ILLC, University of Amsterdam. 219 pp. (cit. on p. 32).

— (2003). "Free Choice in Modal Contexts". In: *Proceedings of Sinn und Bedeutung* 7, pp. 25–37. ISSN: 2629-6055. DOI: 10.18148/sub/2003.v7i0.789. URL: https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/789 (visited on 02/13/2025) (cit. on p. 55).

— (June 2, 2022). "Logic and Conversation: The Case of Free Choice". In: *Semantics and Pragmatics* 15, 5:1–60. ISSN: 1937-8912. DOI: 10.3765/sp.15.5. URL: https://semprag.org/index.php/sp/article/view/sp.15.5 (visited on 01/17/2025) (cit. on p. 49).

Aloni, Maria and Floris Roelofsen (Oct. 1, 2011). "Interpreting Concealed Questions". In: *Linguistics and Philosophy* 34.5, pp. 443–478. ISSN: 1573-0549. DOI:

10.1007/s10988-011-9102-9. URL: https://doi.org/10.1007/s10988-011-9102-9 (visited on 03/05/2025) (cit. on p. 125).

Alonso-Ovalle, Luis and Vincent Rouillard (Aug. 30, 2023). "Spanish Bare Interrogatives and Number". In: *Journal of Semantics* 40.2–3, pp. 289–309. ISSN: 0167-5133. DOI: 10.1093/jos/ffad008. URL: https://doi.org/10.1093/jos/ffad008 (visited on 08/15/2025) (cit. on pp. 59, 78).

AnderBois, Scott (2011). "Issues and Alternatives". PhD thesis. University of California at Santa Cruz (cit. on p. 55).

Asher, Nicholas and Alex Lascarides (1998). "Questions in Dialogue". In: *Linguistics and Philosophy* 21.3, pp. 237–309. ISSN: 0165-0157. JSTOR: 25001706. URL: https://www.jstor.org/stable/25001706 (visited on 05/07/2024) (cit. on p. 19).

— (2003). *Logics of Conversation*. Cambridge University Press (cit. on pp. 80, 104).

Asher, Nicholas and Elin McCready (May 1, 2007). "Were, Would, Might and a Compositional Account of Counterfactuals". In: *Journal of Semantics* 24.2, pp. 93–129. ISSN: 0167-5133. DOI: 10.1093/jos/ffl013. URL: https://doi.org/10.1093/jos/ffl013 (visited on 04/23/2024) (cit. on pp. 79, 80).

Avanzi, Mathieu, Marie-José Béguelin, Gilles Corminboeuf, Frederica Diémoz, and Laure Anne Johnsen (2012). *OFROM – Corpus Oral de Français de Suisse Romande*. Université de Neuchâtel. URL: http://ofrom.unine.ch/uploads/Documents/AM-MJB-FD_GC_LAJ_OFROM_23.pdf (cit. on p. 151).

Bally, Anne-Sophie (2022). "Les interrogatives totales en français québécois dans l'écrit SMS : à la croisée de l'oral et de l'écrit". In: 8e Congrès Mondial de Linguistique Française. Ed. by Frank Neveu, Sophie Prévost, Agnès Steuckardt, Gabriel Bergounioux, and Badreddine Hamma. Vol. 138. Orléans: SHS Web of Conferences, p. 12006. DOI: 10.1051/shsconf/202213812006. URL: https://www.shs-conferences.org/articles/shsconf/abs/2022/08/shsconf_cmlf2022_12006/shsconf_cmlf2022_12006.html (visited on 09/22/2022) (cit. on pp. 145, 158).

Barwise, Jon and Robin Cooper (1981). "Generalized Quantifiers and Natural Language". In: *Linguistics and Philosophy* 4.2, pp. 159–219. ISSN: 0165-0157. JSTOR: 25001052. URL: https://www.jstor.org/stable/25001052 (visited on 05/19/2021) (cit. on p. 98).

Baunaz, Lena (June 1, 2016). "French "Quantifiers" in Questions: Interface Strategies". In: *Acta Linguistica Hungarica* 63.2, pp. 125–168. ISSN: 1216-8076, 1588-2624. DOI: 10.1556/064.2016.63.2.1. URL: https://akjournals.com/view/journals/064/63/2/article-p125.xml (visited on 08/14/2025) (cit. on pp. 24, 26, 39, 47).

Bawden, Rachel, Marie-Amélie Botalla, Kim Gerdes, and Sylvain Kahane (May 2014). "Correcting and Validating Syntactic Dependency in the Spoken French Treebank Rhapsodie". In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. LREC 2014. Reykjavik,

Iceland: European Language Resources Association (ELRA), pp. 2320–2325. URL: http://www.lrec-conf.org/proceedings/lrec2014/pdf/766_Paper.pdf (visited on 03/10/2023) (cit. on p. 148).

Beaver, David I. (June 2001). *Presupposition and Assertion in Dynamic Semantics: A Critical Review of Linguistic Theories of Presupposition and a Dynamic Alternative*. Studies in Logic, Language, and Information. Center for the Study of Language and Information. 250 pp. ISBN: 978-1-57586-120-3. URL: https://press.uchicago.edu/ucp/books/book/distributed/P/bo3639443.html (visited on 11/04/2024) (cit. on pp. 79, 82).

Beaver, David I. and Brady Z. Clark (2008). *Sense and Sensitivity: How Focus Determines Meaning*. Malden, MA: Blackwell (cit. on p. 126).

Bélanger, Luc and Guy Lapalme (2004). "Identification de questions pour traiter les courriels par une méthode question-réponse". In: *Le Poids des mots, Actes des 7es Journées internationales d'analyse statistique des données textuelles*. Analyse statistique des données textuelles (JADT'2004). Ed. by Cédric Fairon, Anne Dister, and Gérald Purnelle. Vol. 1. Louvain: Presses universitaires de Louvain, pp. 128–135. ISBN: 978-2-930344-49-2. URL: http://lexicometrica.univ-paris3.fr/jadt/jadt2004/pdf/JADT_011.pdf (cit. on p. 146).

Belnap, Nuel D. (1963). *An Analysis of Questions: Preliminary Report*. Santa Monica, CA, USA: System Development (cit. on p. 11).

Benzitoun, Christophe (2022). "Évolution des interrogatives partielles directes en français : le cas de combien". In: 8e Congrès Mondial de Linguistique Française. Ed. by Frank Neveu, Sophie Prévost, Agnès Steuckardt, Gabriel Bergounioux, and Badreddine Hamma. Vol. 138. Syntaxe. Orléans: SHS Web of Conferences, p. 13003. DOI: 10.1051/shsconf/202213813003. URL: https://www.shs-conferences.org/articles/shsconf/abs/2022/08/shsconf_cmlf2022_13003/shsconf_cmlf2022_13003.html (visited on 09/22/2022) (cit. on p. 145).

Berrendonner, Alain (Feb. 2023). *L'opposition grammatical/lexical*. In: *Encyclopédie Grammaticale du Français*. DOI: https://nakala.fr/10.34847/nkl.466233b1. URL: http://www.encyclogram.fr/notx/061/061_Notice.php (visited on 06/28/2023) (cit. on p. 161).

Bhat, Riyaz Ahmad, Rajesh Bhatt, Annahita Farudi, Prescott Klassen, Bhuvana Narasimhan, Martha Palmer, Owen Rambow, Dipti Misra Sharma, Ashwini Vaidya, Sri Ramagurumurthy Vishnu, and Fei Xia (2017). "The Hindi/Urdu Treebank Project". In: *Handbook of Linguistic Annotation*. Ed. by Nancy Ide and James Pustejovsky. Dordrecht: Springer Netherlands, pp. 659–697. ISBN: 978-94-024-0881-2. DOI: 10.1007/978-94-024-0881-2_24. URL: https://doi.org/10.1007/978-94-024-0881-2_24 (visited on 03/09/2023) (cit. on p. 147).

Bonfante, Guillaume, Bruno Guillaume, and Guy Perrier (Apr. 2018). *Application of Graph Rewriting to Natural Language Processing*. Vol. 1. ISTE Wi-

ley. 272 pp. ISBN: 978-1-78630-096-6. URL: https://hal.inria.fr/hal-01814386 (visited on 12/06/2022) (cit. on pp. 147, 148).

Bosco, Cristina and Manuela Sanguinetti (2014). "Towards a Universal Stanford Dependencies Parallel Treebank". In: *Proceedings of the 13th Workshop on Treebanks and Linguistic Theories (TLT-13)*. Workshop on Treebanks and Linguistic Theories. Ed. by Verena Henrich, Erhard Hinrichs, Daniël de Kok, Petya Osenova, and Adam Przepiórkowski. Tübingen (Germany), pp. 14–25. ISBN: 978-3-9809183-9-8. URL: https://clic2014.fileli.unipi.it/proceedings/vol1/CLICIT2014161.pdf (cit. on p. 148).

Brasoveanu, Adrian (2007a). "Donkey Pluralities: Plural Information States vs. Non-atomic Individuals". In: *Proceedings of Sinn und Bedeutung* 11, pp. 106–120. ISSN: 2629-6055. DOI: 10.18148/sub/2007.v11i0.634. URL: https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/634 (visited on 07/11/2023) (cit. on p. 67).

— (2007b). "Structured Anaphora to Quantifier Domains: A Unified Account of Quantificational and Modal Subordination". In: *Logic, Language, Information and Computation*. Ed. by Daniel Leivant and Ruy de Queiroz. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 101–119. ISBN: 978-3-540-73445-1. DOI: 10.1007/978-3-540-73445-1_8 (cit. on pp. 91, 93, 99, 100).

— (Nov. 1, 2010). "Decomposing Modal Quantification". In: *Journal of Semantics* 27.4, pp. 437–527. ISSN: 0167-5133. DOI: 10.1093/jos/ffq008. URL: https://doi.org/10.1093/jos/ffq008 (visited on 07/11/2023) (cit. on pp. 79, 80).

Buccola, Brian and Benjamin Spector (June 1, 2016). "Modified Numerals and Maximality". In: *Linguistics and Philosophy* 39.3, pp. 151–199. ISSN: 1573-0549. DOI: 10.1007/s10988-016-9187-2. URL: https://doi.org/10.1007/s10988-016-9187-2 (visited on 03/19/2025) (cit. on p. 36).

Büring, Daniel (Sept. 2, 2004). "Focus Suppositions". In: *Theoretical Linguistics* 30.1, pp. 65–76. ISSN: 1613-4060. DOI: 10.1515/thli.2004.007. URL: https://www.degruyterbrill.com/document/doi/10.1515/thli.2004.007/html (visited on 05/15/2025) (cit. on p. 54).

Bylinina, Lisa and Rick Nouwen (Sept. 7, 2018). "On "Zero" and Semantic Plurality". In: *Glossa: a journal of general linguistics* 3.1 (1). ISSN: 2397-1835. DOI: 10.5334/gjgl.441. URL: https://www.glossa-journal.org/article/id/5066/ (visited on 01/18/2025) (cit. on pp. 36, 141).

Candito, Marie, Guy Perrier, Bruno Guillaume, Corentin Ribeyre, Karën Fort, Djamé Seddah, and Eric De La Clergerie (Apr. 2014). "Deep Syntax Annotation of the Sequoia French Treebank". In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. Ed. by Nicoletta Calzolari (Conference Chair), Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno,

Jan Odijk, and Stelios Piperidis. Reykjavik, Iceland: European Language Resources Association (ELRA). ISBN: 978-2-9517408-8-4 (cit. on p. 148).

Candito, Marie and Djamé Seddah (June 2012). "Le corpus Sequoia : annotation syntaxique et exploitation pour l'adaptation d'analyseur par pont lexical". In: *Actes de la 19e conférence sur le Traitement Automatique des Langues Naturelles*. Grenoble, France: Association pour le Traitement Automatique des Langues, pp. 321–334. URL: http://www.atala.org/taln_archives/TALN/TALN-2012/taln-2012-long-024 (cit. on p. 148).

Caponigro, Ivano and Kathryn Davidson (Dec. 1, 2011). "Ask, and Tell as Well: Question–Answer Clauses in American Sign Language". In: *Natural Language Semantics* 19.4, pp. 323–371. ISSN: 1572-865X. DOI: 10.1007/s11050-011-9071-0. URL: https://doi.org/10.1007/s11050-011-9071-0 (visited on 01/21/2025) (cit. on p. 56).

Chierchia, Gennaro (Jan. 1, 1993). "Questions with Quantifiers". In: *Natural Language Semantics* 1.2, pp. 181–234. ISSN: 1572-865X. DOI: 10.1007/BF00372562. URL: https://doi.org/10.1007/BF00372562 (visited on 07/07/2023) (cit. on p. 37).

— (1995). "Individual Level Predicates as Inherent Generics". In: *The Generic Book*. Ed. by Gregory N. Carlson and Francis J. Pelletier. University of Chicago Press, pp. 176–223 (cit. on pp. 4, 24, 34).

— (July 2013). *Logic in Grammar: Polarity, Free Choice, and Intervention*. Oxford University Press. ISBN: 978-0-19-969797-7. DOI: 10.1093/acprof:oso/9780199697977.001.0001. URL: https://doi.org/10.1093/acprof:oso/9780199697977.001.0001 (cit. on pp. 34, 43).

Chierchia, Gennaro and Ivano Caponigro (2013). "Questions on Questions and Free Relatives". Talk. Sinn Und Bedeutung 18 (Vitoria-Gasteiz). URL: https://scholar.harvard.edu/files/chierchia/files/frs_and_qs_iii_2013-9-11.pdf (cit. on p. 53).

Chierchia, Gennaro, Danny Fox, and Benjamin Spector (Dec. 12, 2012). "Scalar Implicature as a Grammatical Phenomenon". In: *Handbücher Zur Sprach- Und Kommunikationswissenschaft / Handbooks of Linguistics and Communication Science*. Ed. by Claudia Maienborn, Klaus Von Heusinger, and Paul Portner. Berlin, Boston: De Gruyter, pp. 2297–2331. ISBN: 978-3-11-025337-5. DOI: 10.1515/9783110253382.2297. URL: https://library.oapen.org/handle/20.500.12657/23783 (visited on 01/08/2025) (cit. on p. 52).

Choi, Hee-Soo, Bruno Guillaume, and Karën Fort (Mar. 21, 2022). "Corpus-Based Language Universals Analysis Using Universal Dependencies". In: *Proceedings of the Second Workshop on Quantitative Syntax (Quasy, SyntaxFest 2021)*. SyntaxFest Quasy 2021 - Quantitative Syntax. Sofia, Bulgaria: Association for Computational Linguistics. URL: https://hal.inria.fr/hal-03501774 (visited on 03/31/2023) (cit. on p. 144).

Chomsky, Noam (1993). *Lectures on Government and Binding: The Pisa Lectures*. Walter de Gruyter. 824 pp. ISBN: 978-3-11-014131-3. Google Books: l08tpkOOdNQC (cit. on p. 73).

Ciardelli, Ivano (2016). "Questions in Logic". PhD thesis. Amsterdam: University of Amsterdam. 398 pp. (cit. on pp. 81, 82).

— (2018). "Dependence Statements Are Strict Conditionals". In: *Advances in Modal Logic*, pp. 123–142 (cit. on pp. 115, 117).

Ciardelli, Ivano, Jeroen Groenendijk, and Floris Roelofsen (Nov. 27, 2018). *Inquisitive Semantics*. Oxford Surveys in Semantics and Pragmatics. Oxford, New York: Oxford University Press. 240 pp. ISBN: 978-0-19-881478-8 (cit. on pp. 16, 63, 80, 213).

Ciardelli, Ivano, Floris Roelofsen, and Nadine Theiler (Feb. 1, 2017). "Composing Alternatives". In: *Linguistics and Philosophy* 40.1, pp. 1–36. ISSN: 1573-0549. DOI: 10.1007/s10988-016-9195-2. URL: https://doi.org/10.1007/s10988-016-9195-2 (visited on 03/15/2021) (cit. on pp. 37, 80).

Cinque, Guglielmo (Dec. 27, 1990). *Types of Ā-Dependencies*. Linguistic Inquiry Monographs. The MIT Press. 242 pp. ISBN: 978-0-262-53089-7 (cit. on p. 47).

— (1996). "'Long' WH-Movements and Referentiality". In: *Current Issues in Comparative Grammar*. Ed. by Robert Freidin. Dordrecht: Springer Netherlands, pp. 226–248. ISBN: 978-94-009-0135-3. DOI: 10.1007/978-94-009-0135-3_11. URL: https://doi.org/10.1007/978-94-009-0135-3_11 (visited on 02/11/2025) (cit. on p. 18).

Cole, Peter (May 1, 1987). "The Structure of Internally Headed Relative Clauses". In: *Natural Language & Linguistic Theory* 5.2, pp. 277–302. ISSN: 1573-0859. DOI: 10.1007/BF00166587. URL: https://doi.org/10.1007/BF00166587 (visited on 02/12/2025) (cit. on p. 48).

Comorovski, Ileana (1996). *Interrogative Phrases and the Syntax-Semantics Interface*. Red. by Gennaro Chierchia, Pauline Jacobson, and Francis J. Pelletier. Vol. 59. Studies in Linguistics and Philosophy. Dordrecht: Springer Netherlands. ISBN: 978-90-481-4641-3. DOI: 10.1007/978-94-015-8688-7. URL: http://link.springer.com/10.1007/978-94-015-8688-7 (visited on 07/24/2023) (cit. on pp. 11, 15, 18).

Cong, Gao, Long Wang, Chin-Yew Lin, Young-In Song, and Yueheng Sun (July 20, 2008). "Finding Question-Answer Pairs from Online Forums". In: *Proceedings of the 31st Annual International ACM SIGIR Conference on Research and Development in Information Retrieval*. SIGIR '08. New York, NY, USA: Association for Computing Machinery, pp. 467–474. ISBN: 978-1-60558-164-4. DOI: 10.1145/1390334.1390415. URL: https://doi.org/10.1145/1390334.1390415 (visited on 03/09/2023) (cit. on p. 146).

Coveney, Aidan (2011). "L'interrogation directe". In: *Travaux de linguistique* 63.2, pp. 112–145. ISSN: 0082-6049. URL: https://www.cairn.info/revue-travaux-de-linguistique-2011-2-page-112.htm (visited on 02/18/2022) (cit. on pp. 148, 153, 156, 157, 162).

Cremers, Alexandre and Emmanuel Chemla (Sept. 1, 2017). "Experiments on the Acceptability and Possible Readings of Questions Embedded under Emotive-Factives". In: *Natural Language Semantics* 25.3, pp. 223–261. ISSN: 1572-865X. DOI: 10.1007/s11050-017-9135-x. URL: https://doi.org/10.1007/s11050-017-9135-x (visited on 03/13/2025) (cit. on p. 40).

Cresti, Diana (1998). "Some Considerations on Wh-decomposition and Unselective Binding". In: *Sprachtheoretische Grundlagen für Computerlinguistik*. Arbeitspapiere Des Sonderforschungsbereichs 340, pp. 155–183 (cit. on p. 34).

Dayal, Veneeta (1995). "Quantification in Correlatives". In: *Quantification in Natural Languages*. Ed. by Emmon W. Bach, Eloise Jelinek, Angelika Kratzer, and Barbara H. Partee. Kluwer Academic Publishers, pp. 2–179 (cit. on p. 27).

— (1996). *Locality in WH Quantification*. Red. by Gennaro Chierchia, Pauline Jacobson, and Francis J. Pelletier. Vol. 62. Studies in Linguistics and Philosophy. Dordrecht: Springer Netherlands. ISBN: 978-0-7923-5478-9. DOI: 10.1007/978-94-011-4808-5. URL: http://link.springer.com/10.1007/978-94-011-4808-5 (visited on 07/11/2023) (cit. on pp. 3, 11, 53, 56, 77).

— (Nov. 3, 1997). "Free Relatives and "Ever": Identity and Free Choice Readings". In: *Semantics and Linguistic Theory* 7, pp. 99–116. ISSN: 2163-5951. DOI: 10.3765/salt.v7i0.2787. URL: https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/2787 (visited on 02/05/2025) (cit. on pp. 27, 57).

— (2016). *Questions*. Vol. 4. Oxford Surveys in Semantics and Pragmatics. Oxford University Press (cit. on pp. 16, 42, 52).

— ed. (to appear). *The Open Handbook of (In)Definiteness: A Hitchhiker's Guide to Interpreting Bare Arguments*. Red. by Heidi Harley. Open Handbooks in Linguistics. MIT Press (cit. on pp. 140, 215).

De Cornulier, Benoît (Apr. 2013). "Les subordonnées dites interrogatives comme compléments de dépendance". In: *Philologia*. Linguistique comparée des langues romanes (LiCoLaR) (La subordination à travers les langues romanes, Hommage au Prof. H.-J. Deulofeu). Ed. by Sophie Saffi, Sandrine Caddéo, Stefan Gencarau, and Romana Timoc-Bardy, pp. 131–143. URL: https://hal.science/hal-01024873 (cit. on p. 115).

Defrancq, Bart (2005). *L'interrogative enchâssée. Structure et interprétation*. Champs linguistiques. Louvain-la-Neuve: De Boeck Supérieur. 256 pp. ISBN: 978-2-8011-1365-3. URL: https://www.cairn.info/l-interrogative-enchassee--9782801113653.htm (cit. on pp. 145, 151, 157).

De Groote, Philippe and Valentin D. Richard (Dec. 13, 2021). "Embedding Intensional Semantics into Inquisitive Semantics". In: *MOL 2021 - 17th Meeting on the Mathematics of Language*. Ed. by Henrik Björklund and Frank Dewes. Montpellier (online), France. URL: https://aclanthology.org/2021.mol-1.6/ (visited on 06/10/2022) (cit. on p. 65).

Dekhissi, Laurie (2016). "Qu'est-ce t'as été te mêler de ça ?! Une « nouvelle » structure pour les questions rhétoriques conflictuelles". In: *Journal of French*

Language Studies 26.3, pp. 279–298. DOI: 10.1017/S0959269515000253 (cit. on pp. 145, 161).

Dekker, Paul (Jan. 1, 1992). *An Update Semantics for Dynamic Predicate Logic*. Report. URL: https://eprints.illc.uva.nl/id/eprint/1177/ (visited on 11/07/2024) (cit. on pp. 70, 79, 99).

— (Nov. 1, 1998). "Speaker's Reference, Descriptions and Information Structure". In: *Journal of Semantics* 15.4, pp. 305–334. ISSN: 0167-5133. DOI: 10.1093/jos/15.4.305. URL: https://doi.org/10.1093/jos/15.4.305 (visited on 03/10/2025) (cit. on p. 50).

Delaveau, Annie, Paul Cappeau, and Anne Dagnac (2021). "Les phrases interrogatives". In: *La Grande Grammaire du Français*. Ed. by Anne Abeillé and Danièle Godard. 1st ed. Vol. 2. Arles: Actes Sud/Imprimeries nationales Éditions, pp. 1402–1437. ISBN: 978-2-330-14239-1 (cit. on pp. 148, 149, 157, 161).

De Marneffe, Marie-Catherine, Christopher D. Manning, Joakim Nivre, and Daniel Zeman (July 13, 2021). "Universal Dependencies". In: *Computational Linguistics* 47.2, pp. 255–308. ISSN: 0891-2017. DOI: 10.1162/coli_a_00402. URL: https://doi.org/10.1162/coli_a_00402 (visited on 03/14/2023) (cit. on p. 149).

Destruel, Emilie (Nov. 1, 2017). "The Pragmatics of French (Non-)Prototypical Clefts: Influence of the Type of Question on Naturalness and Interpretation". In: *Journal of Pragmatics* 121, pp. 58–75. ISSN: 0378-2166. DOI: 10.1016/j.pragma.2017.08.014. URL: https://www.sciencedirect.com/science/article/pii/S0378216617302801 (visited on 04/22/2025) (cit. on p. 24).

Ding, Shilin, Gao Cong, Chin-Yew Lin, and Xiaoyan Zhu (June 2008). "Using Conditional Random Fields to Extract Contexts and Answers of Questions from Online Forums". In: *Proceedings of ACL-08: HLT*. ACL-HLT 2008. Columbus, Ohio: Association for Computational Linguistics, pp. 710–718. URL: https://aclanthology.org/P08-1081 (visited on 03/09/2023) (cit. on p. 146).

Donnellan, Keith S. (1966). "Reference and Definite Descriptions". In: *The Philosophical Review* 75.3, pp. 281–304. ISSN: 0031-8108. DOI: 10.2307/2183143. JSTOR: 2183143. URL: https://www.jstor.org/stable/2183143 (visited on 01/20/2025) (cit. on p. 27).

Dotlačil, Jakub and Floris Roelofsen (2019). "Dynamic Inquisitive Semantics: Anaphora and Questions". In: *Proceedings of Sinn Und Bedeutung 23*. Universitat Autònoma de Barcelona. DOI: 10.18148/sub/2019.v23i1.538. URL: https://dare.uva.nl/search?identifier=e767de8b-ac55-4848-af4f-7bf63dfbfb95 (visited on 06/25/2021) (cit. on pp. 67, 80, 98).

— (Mar. 2, 2021). "A Dynamic Semantics of Single-Wh and Multiple-Wh Questions". In: *Semantics and Linguistic Theory* 30.0, pp. 376–395. ISSN: 2163-5951. DOI: 10.3765/salt.v30i0.4839. URL: https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/30.376 (visited on 08/09/2021) (cit. on pp. 39, 67, 80).

Druetta, Ruggero (Nov. 24, 2003). “"Qu'est-ce tu fais?" État d'avancement de la grammaticalisation de "est-ce que". Première partie”. In: *Linguæ & - Rivista di lingue e culture moderne* 1.2 (2), pp. 67–88. ISSN: 1724-8698. URL: https://www.ledonline.it/index.php/linguae/article/view/154 (visited on 07/07/2022) (cit. on p. 161).

— (Mar. 15, 2004). “"Qu'est-ce tu fais?" État d'avancement de la grammaticalisation de "est-ce que": Deuxième partie”. In: *Linguæ & - Rivista di lingue e culture moderne* 2.1 (1), pp. 21–35. ISSN: 1724-8698. URL: https://www.ledonline.it/index.php/linguae/article/view/157 (visited on 07/07/2022) (cit. on p. 161).

Duží, Marie and Martina Číhalová (Dec. 2015). “Questions, Answers, and Presuppositions”. In: *Computación y Sistemas* 19.4, pp. 647–659. ISSN: 1405-5546. DOI: 10.13053/CyS-19-4-2327. URL: http://www.scielo.org.mx/scielo.php?script=sci_abstract&pid=S1405-55462015000400647&lng=es&nrm=iso&tlng=en (visited on 07/28/2023) (cit. on p. 35).

Eckardt, Regine (2007). “Inherent Focus on Wh-Phrases”. In: *Proceedings of Sinn und Bedeutung* 11, pp. 209–228. ISSN: 2629-6055. DOI: 10.18148/sub/2007.v11i0.641. URL: https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/641 (visited on 01/20/2025) (cit. on p. 30).

Efron, Miles and Megan Winget (Oct. 22, 2010). “Questions Are Content: A Taxonomy of Questions in a Microblogging Environment”. In: *Proceedings of the 73rd ASIS&T Annual Meeting on Navigating Streams in an Information Ecosystem - Volume 47*. ASIS&T ’10. USA: American Society for Information Science, pp. 1–10 (cit. on p. 146).

Elliott, Patrick D, Andreea C. Nicolae, and Uli Sauerland (Aug. 16, 2022). “Who and What Do Who and What Range Over Cross-Linguistically?” In: *Journal of Semantics* 39.3, pp. 551–579. ISSN: 0167-5133. DOI: 10.1093/jos/ffac008. URL: https://doi.org/10.1093/jos/ffac008 (visited on 02/01/2025) (cit. on pp. 12, 51, 52, 56–58).

Elsig, Martin and Shana Poplack (2006). “Transplanted Dialects and Language Change: Question Formation in Québec”. In: *U. Penn Working Papers in Linguistics* 12.2. Ed. by Michael L. Friesner and Maya Ravindranath, pp. 77–90. URL: https://www.semanticscholar.org/paper/Transplanted-dialects-and-language-change%3A-question-Elsig-Poplack/b3a794d0e1c151c84d8db6fe9de6b5a3c3d2bb89 (visited on 02/18/2022) (cit. on p. 145).

Engdahl, Elisabet (Sept. 1, 1980). “The Syntax and Semantics of Questions in Swedish”. In: *Doctoral Dissertations 1896 - February 2014*. DOI: 10.7275/bk1a-ns45. URL: https://scholarworks.umass.edu/dissertations_1/936 (cit. on p. 37).

Erteschik-Shir, Nomi (1986). “Wh-Questions and Focus”. In: *Linguistics and Philosophy* 9.2, pp. 117–149. ISSN: 0165-0157. JSTOR: 25001238. URL: https://www.jstor.org/stable/25001238 (visited on 03/20/2025) (cit. on p. 45).

Eshkol-Taravella, Iris, Angèle Barbedette, Xingyu Liu, and Valentin-Gabriel Soumah (2022). "Classification automatique de questions spontanées vs. préparées dans des transcriptions de l'oral". In: *Traitement Automatique des Langues Naturelles*. Ed. by Yannick Estève, Tania Jiménez, Titouan Parcollet, and Marcely Zanon Boito. Avignon, France: ATALA, pp. 305–314. URL: https://hal.archives-ouvertes.fr/hal-03701483 (visited on 07/04/2022) (cit. on p. 145).

Farmer, Kelly Lynne (2013). ""De Quoi Tu Parles?": A Diachronic Study of Sociopragmatic Interrogative Variation in French Films". In: *University of Pennsylvania Working Papers in Linguistics* 19.2, p. 8. DOI: https://repository.upenn.edu/pwpl/vol19/iss2/8 (cit. on pp. 145, 161).

Fitzpatrick, Justin (2005). "The Whys and How Comes of Presupposition and NPI Licensing in Questions". In: *Proceedings of the 24th West Coast Conference on Formal Linguistics*. 24th West Coast Conference on Formal Linguistics. Ed. by John Alderete, Chung-hye Han, and Alexei Kochetov. Somerville, MA: Cascadilla Proceedings Project, pp. 138–145. ISBN: 978-1-57473-407-2. URL: http://www.lingref.com/cpp/wccfl/24/abstract1216.html (visited on 09/12/2023) (cit. on pp. 12, 14–16, 35, 36, 39, 42, 52, 58).

Foulet, Lucien (1921). "Comment ont évolué les formes de l'interrogation". In: *Romania* 47.186, pp. 243–348. DOI: 10.3406/roma.1921.4437. URL: https://www.persee.fr/doc/roma_0035-8029_1921_num_47_186_4437 (visited on 05/12/2021) (cit. on p. 161).

Fox, Danny (Jan. 1, 2018). "Partition by Exhaustification: Comments on Dayal 1996". In: *ZAS Papers in Linguistics* 60, pp. 403–434. ISSN: 1435-9588. DOI: 10.21248/zaspil.60.2018.474. URL: https://zaspil.leibniz-zas.de/article/view/474 (visited on 02/20/2025) (cit. on pp. 30, 35, 57).

Frank, Anette (1997). "Context Dependence in Modal Constructions". PhD thesis. Stuttgart: Universität Stuttgart. 411 pp. (cit. on p. 80).

Gawron, Jean Mark (2001). "Universal Concessive Conditionals and Alternative NPs in English". In: *Logical Perspectives on Language and Information*. Ed. by Cleo A Condoravdi and Gerard Renardel De Lavalette. Stanford, CA: CSLI publications, pp. 73–106 (cit. on p. 11).

Gerdes, Kim, Bruno Guillaume, Sylvain Kahane, and Guy Perrier (Aug. 26, 2019). "Improving Surface-syntactic Universal Dependencies (SUD): Surface-Syntactic Relations and Deep Syntactic Features". In: *Proceedings of the 18th International Workshop on Treebanks and Linguistic Theories (TLT, SyntaxFest 2019)*. TLT 2019 - 18th International Workshop on Treebanks and Linguistic Theories. Ed. by Marie Candito, Kilian Evang, Stephan Oepen, and Djamé Seddah. Paris, France: Association for Computational Linguistics, pp. 126–132. DOI: https://aclanthology.org/W19-7814/ (cit. on pp. 149, 162).

Geurts, Bart (1999). *Presuppositions and Pronouns*. Current Research in the Semantics / Pragmatics Interface 3. Leiden, The Netherlands: Brill. ISBN: 978-0-08-043592-3 (cit. on p. 80).

Ghaddar, Abbas and Phillippe Langlais (May 2016). "WikiCoref: An English Coreference-annotated Corpus of Wikipedia Articles". In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*. LREC 2016. Ed. by Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Sara Goggi, Marko Grobelnik, Bente Maegaard, Joseph Mariani, Helene Mazo, Asuncion Moreno, Jan Odijk, and Stelios Piperidis. Portorož, Slovenia: European Language Resources Association (ELRA), pp. 136–142. URL: https://aclanthology.org/L16-1021 (visited on 11/28/2024) (cit. on p. 93).

Giannakidou, Anastasia and Lisa Lai-Shen Cheng (May 1, 2006). "(In)Definiteness, Polarity, and the Role of Wh-Morphology in Free Choice". In: *Journal of Semantics* 23.2, pp. 135–183. ISSN: 0167-5133. DOI: 10.1093/jos/ffl001. URL: https://doi.org/10.1093/jos/ffl001 (visited on 01/20/2025) (cit. on pp. 27–29).

Giannakidou, Anastasia and Alda Mari (2019). *Modalization and Bias in Questions*. Unpublished manuscript. University of Chicago and Institut Jean Nicod Paris (cit. on pp. 47, 82, 85).

— (Aug. 2021). *Truth and Veridicality in Grammar and Thought: Mood, Modality, and Propositional Attitudes*. Chicago, IL: University of Chicago Press. 312 pp. ISBN: 978-0-226-76334-7. URL: https://press.uchicago.edu/ucp/books/book/chicago/T/bo78676587.html (visited on 11/14/2024) (cit. on p. 106).

Gillet, Pauline (2022). "Développement du langage de l'enfant : l'exemple des interrogatives partielles". In: 8e Congrès Mondial de Linguistique Française. Ed. by Frank Neveu, Sophie Prévost, Agnès Steuckardt, Gabriel Bergounioux, and Badreddine Hamma. Vol. 138. Orléans: SHS Web of Conferences, p. 13002. DOI: 10.1051/shsconf/202213813002. URL: https://www.shs-conferences.org/articles/shsconf/abs/2022/08/shsconf_cmlf2022_13002/shsconf_cmlf2022_13002.html (visited on 09/22/2022) (cit. on p. 145).

Gillies, Anthony S. (2004). "Epistemic Conditionals and Conditional Epistemics". In: *Noûs* 38.4, pp. 585–616. DOI: 10.1111/j.0029-4624.2004.00485.x (cit. on pp. 84, 117).

Ginzburg, Jonathan (1992). "Questions, Queries, and Facts : A Semantics and Pragmatics for Interrogatives". PhD thesis. Stanford University. 710 pp. URL: https://philpapers.org/rec/GINQQA (visited on 12/05/2024) (cit. on pp. 3, 11, 15).

— (Oct. 1, 1995a). "Resolving Questions, I". In: *Linguistics and Philosophy* 18.5, pp. 459–527. ISSN: 1573-0549. DOI: 10.1007/BF00985365. URL: https://doi.org/10.1007/BF00985365 (visited on 11/26/2024) (cit. on pp. 19, 35).

Ginzburg, Jonathan (Dec. 1, 1995b). "Resolving Questions, II". In: *Linguistics and Philosophy* 18.6, pp. 567–609. ISSN: 1573-0549. DOI: 10.1007/BF00983299. URL: https://doi.org/10.1007/BF00983299 (visited on 11/26/2024) (cit. on p. 11).

Ginzburg, Jonathan and Ivan A. Sag (Apr. 2001). *Interrogative Investigations: The Form, Meaning, and Use of English Interrogatives*. 1st ed. Lecture Notes. Center for the Study of Language and Information: University of Chicago Press. 250 pp. ISBN: 978-1-57586-278-1. URL: https://press.uchicago.edu/ucp/books/book/distributed/I/bo3625102.html (visited on 12/04/2024) (cit. on p. 35).

Glikman, Julie, Frédéric Landragin, Amalia Todirascu, and Catherine Schnedecker (Oct. 13, 2022). "Un exemple de corpus annoté en diachronie longue : le corpus Democrat, enjeux et exploitations". In: Colloque ConCorDiaL, la constitution de corpus en diachronie longue : méthodologies, objectifs et exploitations linguistiques et stylistiques. URL: https://shs.hal.science/halshs-03767394 (visited on 08/20/2025) (cit. on p. 94).

Godar, Danièle (2021). "Les subordonnées relatives". In: *La Grande Grammaire du Français*. Ed. by Anne Abeillé and Danièle Godard. 1st ed. Vol. 2. Arles: Actes Sud/Imprimeries nationales Éditions, pp. 1402–1437. ISBN: 978-2-330-14239-1 (cit. on p. 27).

Groenendijk, Jeroen (1998). "Questions in Update Semantics". In: *Formal Semantics and Pragmatics of Dialogue: Proceedings of the Thirteenth Twente Workshop on Language Technology (Twendial '98)*. Twente: Universiteit Twente, Faculteit Informatica, pp. 125–137. URL: https://dare.uva.nl/search?identifier=e7f44061-4fc2-4924-8811-848a0c9630ed (visited on 07/24/2023) (cit. on pp. 67, 79, 89).

Groenendijk, Jeroen and Martin Stokhof (Jan. 1, 1984). "Studies on the Semantics of Questions and the Pragmatics of Answers". PhD thesis. Amsterdam, The Netherlands: Institute for Logic, Language and Computation (ILLC), University of Amsterdam (cit. on pp. 3, 11, 15, 16, 78, 80).

Groenendijk, Jeroen, Martin Stokhof, and Frank Veltman (1996). "Coreference and Modality". In: *The Handbook of Contemporary Semantic Theory*. Ed. by Shalom Lappin. Cambridge, Mass., USA: Blackwell Refernce, pp. 179–216. ISBN: 978-0-631-18752-3. URL: https://dare.uva.nl/search?identifier=c655089e-9fae-4842-90da-30e46f37825b (visited on 01/31/2024) (cit. on pp. 67, 70, 73, 79, 82, 210).

Grubic, Mira (2015). "Focus and Alternative Sensitivity in Ngamo (West-Chadic)". PhD thesis. University of Potsdam (cit. on p. 23).

Grubic, Mira and Malte Zimmermann (2011). "Conventional and Free Association with Focus in Ngamo (West Chadic)". In: *Proceedings of Sinn und Bedeutung* 15, pp. 291–306. ISSN: 2629-6055. URL: https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/382 (visited on 03/24/2025) (cit. on pp. 23, 25, 26).

Guerzoni, Elena and Yael Sharvit (June 1, 2007). "A Question of Strength: On NPIs in Interrogative Clauses". In: *Linguistics and Philosophy* 30.3, pp. 361–391. ISSN: 1573-0549. DOI: 10.1007/s10988-007-9014-x. URL: https://doi.org/10.1007/s10988-007-9014-x (visited on 01/30/2025) (cit. on pp. 40, 43).

Guibon, Gaël, Marine Courtin, Kim Gerdes, and Bruno Guillaume (May 2020). "When Collaborative Treebank Curation Meets Graph Grammars". In: *Proceedings of the Twelfth Language Resources and Evaluation Conference*. LREC 2020. Marseille, France: European Language Resources Association, pp. 5291–5300. ISBN: 979-10-95546-34-4. URL: https://aclanthology.org/2020.lrec-1.651 (visited on 09/28/2023) (cit. on p. 151).

Guillaume, Bruno (Apr. 19, 2021). "Graph Matching and Graph Rewriting: GREW Tools for Corpus Exploration, Maintenance and Conversion". In: EACL 2021 - 16th Conference of the European Chapter of the Association for Computational Linguistics: System Demonstrations. URL: https://inria.hal.science/hal-03177701 (visited on 01/11/2025) (cit. on p. 144).

Guillaume, Bruno, Marie-Catherine de Marneffe, and Guy Perrier (2019). "Conversion et améliorations de corpus du français annotés en Universal Dependencies". In: *Revue TAL* 60.2, p. 71. URL: https://hal.inria.fr/hal-02267418 (visited on 03/10/2023) (cit. on p. 147).

Guryev, Alexander and François Delafontaine (Jan. 1, 2015). "La variabilité formelle des questions dans les écrits SMS". In: *Travaux Neuchâtelois de Linguistique* 63, pp. 129–152. ISSN: 2504-205X. DOI: 10.26034/tranel.2015.2973. URL: https://www.revue-tranel.ch/article/view/2973 (visited on 03/20/2023) (cit. on p. 145).

Hagstrom, Paul Alan (1998). "Decomposing Questions". PhD thesis. Massachusetts Institute of Technology. URL: https://dspace.mit.edu/handle/1721.1/9649 (visited on 02/11/2025) (cit. on p. 15).

Haida, Andreas (2007). "The Indefiniteness and Focusing of Question Words". PhD thesis. Berlin: Humboldt University. URL: https://hu-berlin.hosted.exlibrisgroup.com/primo-explore/fulldisplay?docid=HUB_UB_ALMA_DS21618277570002882&context=L&vid=hub_ub&lang=en_US&search_scope=default_scope&adaptor=Local%20Search%20Engine&isFrbr=true&tab=default_tab&query=any,contains,haida&sortby=date&facet=frbrgroupid,include,1868078988&mode=Basic&offset=0 (cit. on pp. 15, 103).

— (Oct. 3, 2008). "The Indefiniteness and Focusing of Question Words". In: *Semantics and Linguistic Theory* 18.0, pp. 376–393. ISSN: 2163-5951. DOI: 10.3765/salt.v18i0.2510. URL: http://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/2510 (visited on 07/24/2023) (cit. on p. 15).

Hamblin, Charles Leonard (1973). "Questions in Montague English". In: *Foundations of Language* 10.1, pp. 41–53. ISSN: 0015-900X. JSTOR: 25000703. URL:

https://www.jstor.org/stable/25000703 (visited on 05/24/2021) (cit. on p. 16).

Han, Chung-hye and Laura Siegel (1996). "Syntactic and Semantic Conditions on NPI Licensing in Questions". In: *Proceedings of the Fifteenth West Coast Conference on Formal Linguistics*. Irvine, California, United States, pp. 177–192. URL: https://www.semanticscholar.org/paper/Syntactic-and-Semantic-Conditions-on-NPI-Licensing-Conferenc-Siegel/cd741f8ae75dfe8caf3bbbd4bd14e52234f787ad (visited on 01/18/2025) (cit. on pp. 15, 41–43, 48).

Hara, Yurie and Katsuhiko Sano (2017). "Conditional Questions Revisited". In: InqBnB2: Inquisitiveness Below and Beyond the Sentence Boundary. Amsterdam. URL: https://yuriehara.github.io/website/download/CQR_NINJALpdf.pdf (cit. on p. 80).

Haspelmath, Martin (1997). *Indefinite Pronouns*. Oxford University Press. ISBN: 978-0-19-829963-9. DOI: 10.1093/oso/9780198235606.001.0001. URL: https://library.oapen.org/handle/20.500.12657/30819 (visited on 03/12/2025) (cit. on pp. 31, 137).

Heim, Irene (1991). "Artikel Und Definitheit". In: *Semantik / Semantics: Ein Internationales Handbuch Zeitgenössischer Forschung*. Ed. by Arnim von Stechow and Dieter Wunderlich. De Gruyter Mouton, pp. 487–535. ISBN: 978-3-11-020336-3. DOI: 10.1515/9783110126969.7.487. URL: https://www.degruyter.com/document/doi/10.1515/9783110126969.7.487/html?lang=de&srsltid=AfmBOopIn9YmrrAEgiswvAKzCn__YSwW4iAlClb6ZZ_1Vt7l-PoOUld3 (visited on 02/01/2025) (cit. on p. 39).

— (1994). "Interrogative Semantics and Karttunen's Semantics for Know". In: *Proceedings of the Ninth Annual Conference and the Workshop on Discourse of the Israel Association for Theoretical Linguistics (IATL)*. Ed. by Rhonna Buchalla and Anita Mitwoch. Vol. 1. Jerusalem: Academon, pp. 128–144 (cit. on p. 54).

Hengeveld, Kees, Sabine Iatridou, and Floris Roelofsen (June 23, 2023). "Quexistentials and Focus". In: *Linguistic Inquiry* 54.3, pp. 571–624. ISSN: 0024-3892. DOI: 10.1162/ling_a_00441. URL: https://doi.org/10.1162/ling_a_00441 (visited on 01/21/2025) (cit. on p. 28).

Hénot-Mortier, Adèle (2024a). ""One Tool to Rule Them All"? An Integrated Model of the QuD for Hurford Sentences". In: *Proceedings of the 29th Sinn Und Bedeutung*. Noto, to appear. DOI: 10.18148/sub/2024.v29.1237 (cit. on p. 126).

— (2024b). "Scalarity, Information Structure and Relevance in Varieties of Hurford Conditionals". In: *Proceedings of the 24th Amsterdam Colloquium*. Ed. by Fausto Carcassi, Tamar Johnson, Søren Brinck Knudstorp, Sabina Domínguez Parrado, Pablo Rivas-Robledo, and Giorgio Sbardolini. Amsterdam, The Netherlands, pp. 145–151. URL: https://drive.google.com/file/d/1NDuEXqO4DylI3Ha5l7HFBM4Onz-YAoQN/view (cit. on p. 126).

Herbstritt, Michele (Oct. 23, 2014). *Why Can't We Be Surprised Whether It Rains in Amsterdam? A Semantics for Factive Verbs and Embedded Questions*. Report. URL: https://eprints.illc.uva.nl/id/eprint/931/ (visited on 02/14/2025) (cit. on p. 56).

Herrera, Santiago, Sylvain Kahane, and Bruno Guillaume (Nov. 14, 2022). "Extraction de règles de grammaire à partir de treebanks : développement d'un outil et premiers résultats". In: Journées Jointes des Groupements de Recherche Linguistique Informatique, Formelle et de Terrain (LIFT) et Traitement Automatique des Langues (TAL). Marseille, France: CNRS, pp. 93–98. URL: https://hal.science/hal-03846825 (visited on 06/13/2023) (cit. on p. 144).

Higginbotham, James (1996). "The Semantics of Questions". In: *The Handbook of Contemporary Semantic Theory*. Ed. by Shalom Lappin. Oxford: Oxford University Press (cit. on p. 45).

Higginbotham, James and Robert May (Jan. 1, 1981). "Questions, Quantifiers and Crossing". In: *The Linguistic Review* 1.1, pp. 41–80. ISSN: 1613-3676. DOI: 10.1515/tlir.1981.1.1.41. URL: https://www.degruyter.com/document/doi/10.1515/tlir.1981.1.1.41/html (visited on 07/27/2023) (cit. on p. 77).

Hirsch, Aron and Bernhard Schwarz (2019). "Singular Which, Mention-Some, and Variable Scope Uniqueness". In: *Semantics and Linguistic Theory* 29.0, pp. 748–767. ISSN: 2163-5951. DOI: 10.3765/salt.v29i0.4637. URL: https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/29.748 (visited on 09/05/2023) (cit. on pp. 38, 58, 77, 78, 138, 213).

Hofmann, Lisa (Dec. 20, 2019). "The Anaphoric Potential of Indefinites under Negation and Disjunction". In: *Proceedings of the 22nd Amsterdam Colloquium*. Amsterdam, pp. 181–190 (cit. on p. 80).

Hong, Liangjie and Brian D. Davison (July 19, 2009). "A Classification-Based Approach to Question Answering in Discussion Boards". In: *Proceedings of the 32nd International ACM SIGIR Conference on Research and Development in Information Retrieval*. SIGIR '09. New York, NY, USA: Association for Computing Machinery, pp. 171–178. ISBN: 978-1-60558-483-6. DOI: 10.1145/1571941.1571973. URL: https://doi.org/10.1145/1571941.1571973 (visited on 03/09/2023) (cit. on p. 146).

Horn, Laurence Robert (1972). "On the Semantic Properties of Logical Operators in English". In: University of California, Los Angeles. URL: https://www.semanticscholar.org/paper/On-the-semantic-properties-of-logical-operators-in-Horn/ebed1d0e52d5367f43e8fd8ca6b9b21e62286ada (visited on 01/06/2025) (cit. on pp. 11, 14).

Hovy, Eduard, Mitchell Marcus, Martha Palmer, Lance Ramshaw, and Ralph Weischedel (June 2006). "OntoNotes: The 90% Solution". In: *Proceedings of*

the Human Language Technology Conference of the NAACL, Companion Volume: Short Papers. NAACL-HLT 2006. Ed. by Robert C. Moore, Jeff Bilmes, Jennifer Chu-Carroll, and Mark Sanderson. New York City, USA: Association for Computational Linguistics, pp. 57–60. URL: https://aclanthology.org/N06-2015 (visited on 11/28/2024) (cit. on p. 93).

Hui, Terence (2021). "A Compositional Analysis of Dependence Statements". MSc. Thesis. Amsterdam: Universiteit van Amsterdam. 59 pp. (cit. on pp. 93, 115, 116, 125).

Hulk, Aafke (1996). "The Syntax of WH-questions in Child French". In: *Connecting Children's Language and Linguistic Theory*. Amsterdam Series in Child Language Development 5. Ed. by Frank Wijnen and William Philip, pp. 129–172 (cit. on p. 161).

Iatridou, Sabine and Spyridoula Varlokosta (1996). "A Crosslinguistic Perspective on Pseudoclefts". In: *North East Linguistics Society* 26, pp. 117–130. HDL: 20.500.14394/36828. URL: https://hdl.handle.net/20.500.14394/36828 (cit. on p. 57).

Isaacs, James and Kyle Rawlins (Aug. 1, 2008). "Conditional Questions". In: *Journal of Semantics* 25.3, pp. 269–319. ISSN: 0167-5133. DOI: 10.1093/jos/ffn003. URL: https://doi.org/10.1093/jos/ffn003 (visited on 04/18/2024) (cit. on p. 80).

Jakubíček, Miloš, Adam Kilgarriff, Vojtěch Kovář, Pavel Rychlý, and Vít Suchomel (2013). "The TenTen Corpus Family". In: *7th International Corpus Linguistics Conference CL*. Lancaster, pp. 125–127. URL: https://www.muni.cz/en/research/publications/1120431 (visited on 03/04/2025) (cit. on p. 116).

Jayez, Jacques (2021). "N'importe quel, tout et les expressions de choix libre". In: *La Grande Grammaire du Français*. Ed. by Anne Abeillé and Danièle Godard. 1st ed. Vol. 1. Arles: Actes Sud/Imprimeries nationales Éditions, pp. 632–645. ISBN: 978-2-330-14239-1 (cit. on p. 150).

Jayez, Jacques and Lucia M. Tovena (Aug. 1, 2006). "Epistemic Determiners". In: *Journal of Semantics* 23.3, pp. 217–250. ISSN: 0167-5133. DOI: 10.1093/jos/ffl002. URL: https://doi.org/10.1093/jos/ffl002 (visited on 03/06/2025) (cit. on p. 31).

Jeong, Sunwoo and Floris Roelofsen (May 11, 2023). "Focused NPIs in Statements and Questions". In: *Journal of Semantics* 40.1, pp. 1–68. ISSN: 0167-5133. DOI: 10.1093/jos/ffac014. URL: https://doi.org/10.1093/jos/ffac014 (visited on 10/11/2024) (cit. on pp. 18, 42, 44).

Judge, John, Aoife Cahill, and Josef van Genabith (July 2006). "QuestionBank: Creating a Corpus of Parse-Annotated Questions". In: *Proceedings of the 21st International Conference on Computational Linguistics and 44th Annual Meeting of the Association for Computational Linguistics*. COLING-ACL 2006. Sydney, Australia: Association for Computational Linguistics, pp. 497–

504. DOI: 10.3115/1220175.1220238. URL: https://aclanthology.org/P06-1063 (visited on 03/10/2023) (cit. on p. 147).

Kahane, Sylvain, Bernard Caron, Emmett Strickland, and Kim Gerdes (Dec. 2021). "Annotation Guidelines of UD and SUD Treebanks for Spoken Corpora: A Proposal". In: *Proceedings of the 20th International Workshop on Treebanks and Linguistic Theories (TLT, SyntaxFest 2021)*. TLT 2021. Sofia, Bulgaria: Association for Computational Linguistics, pp. 35–47. URL: https://aclanthology.org/2021.tlt-1.4 (visited on 02/03/2023) (cit. on pp. 147, 155).

Kamp, Hans (1981). "A Theory of Truth and Semantic Representation". In: *Formal Methods in the Study of Language*. Ed. by Jeroen Groenendijk, Theo M. V. Janssen, and Martin Stokhof. Amsterdam: Mathematisch Centrum, pp. 277–322. DOI: 10.1515/9783110867602.1. URL: https://www.degruyter.com/document/doi/10.1515/9783110867602.1/html (visited on 10/23/2024) (cit. on p. 118).

Karttunen, Lauri (1969). "Discourse Referents". In: *Proceedings of the 1969 Conference on Computational Linguistics (COLING '69)*. Sång-Säby, Sweden: Association for Computational Linguistics, pp. 1–38. DOI: 10.3115/990403.990490. URL: http://portal.acm.org/citation.cfm?doid=990403.990490 (visited on 07/07/2023) (cit. on p. 78).

— (1976). "Discourse Referents". In: *Syntax and Semantics Vol.* Ed. by J. D. McCawley. Academic Press, pp. 363–386 (cit. on pp. 90, 93).

— (Jan. 1, 1977). "Syntax and Semantics of Questions". In: *Linguistics and Philosophy* 1.1, pp. 3–44. ISSN: 1573-0549. DOI: 10.1007/BF00351935. URL: https://doi.org/10.1007/BF00351935 (visited on 01/05/2023) (cit. on pp. 16, 117).

Karttunen, Lauri and Stanley Peters (Jan. 1, 1976). "What Indirect Questions Conventionally Implicate". In: pp. 351–368 (cit. on pp. 11, 14).

Katz, Jerrold J. and Paul M. Postal (1964). *An Integrated Theory of Linguistic Descriptions*. Special Technical Report (Massachusetts Institute of Technology. Research Laboratory of Electronics) 9. Cambridge, England: The MIT Press. 178 pp. ISBN: 978-0-262-61021-6. URL: https://philpapers.org/rec/KATAIT-5 (visited on 09/01/2023) (cit. on pp. 3, 11).

Kaufmann, Magdalena (Sept. 1, 2016). "Free Choice Is a Form of Dependence". In: *Natural Language Semantics* 24.3 (3), pp. 247–290. ISSN: 1572-865X. DOI: 10.1007/s11050-016-9125-4. URL: https://link.springer.com/article/10.1007/s11050-016-9125-4 (visited on 06/10/2025) (cit. on pp. 116, 126).

Kaufmann, Stefan (1997). "What's a Stack in Discourse". In: *Proceedings of the 11th Amsterdam Colloquium*. Ed. by Paul Dekker, Martin Stokhof, and Yde Venema. Amsterdam, The Netherlands: ILLC, University of Amsterdam, pp. 187–192. ISBN: 90-74795-90-0 (cit. on p. 80).

— (2000). "Dynamic Context Management". In: Faller, Martina, Stefan Kaufmann, and Marc Pauly. *Formalizing the Dynamics of Information*. CSLI Lec-

ture Notes 92. Stanford, CA: CSLI Publications, pp. 171–181. ISBN: 1-68400-007-6. URL: https://scholar.google.com/citations?view_op=view_citation&hl=fr&user=gffJf60AAAAJ&citation_for_view=gffJf60AAAAJ:aqlVkmm33-oC (visited on 04/18/2024) (cit. on pp. 6, 79, 80, 83, 214).

Kayne, Richard (1972). "Subject Inversion in French Interrogatives". In: *Generative Studies in Romance Languages*. Ed. by J. Casagrande and B. Saciuk. Rowley, Mass.: Newbury House, pp. 70–126 (cit. on p. 161).

Keenan, Edward L. and Robert D. Hull (1973). "The Logical Presuppositions of Questions and Answers". In: *Präsuppositionen in Philosophie Und Linguistik - Presuppositions in Philosophy and Linguistics*. Ed. by J'Anos S. Petöfi and Dorothea Franck. Ahtenäum, pp. 441–466 (cit. on pp. 11, 12).

Kibble, Rodger (1998). "Modal Subordination, Focus, and Complement Anaphora". In: *The Tblisi Symposium on Logic, Language and Computation: Selected Papers*. Ed. by Jonathan Ginzburg, Zurab Khasidashvili, Carl Vogel, Jean-Jacques Levy, and Enric Vallduví. Stanford, USA: CSLI Publications, pp. 73–86. ISBN: 978-1-57586-098-5 (cit. on p. 80).

Klochowicz, Tomasz (Feb. 22, 2025). "Is There (Some) Exhaustification Present in Questions?" In: *25th Workshop on the Roots of Pragmasemantics*. Szklarska Poręba. URL: https://www.tklochowicz.com/assets/slides/szklarska.pdf (cit. on p. 107).

Kobayashi, Filipe Hisao and Vincent Rouillard (Mar. 2, 2021). "High and Low Uniqueness in Singular Wh-Interrogatives". In: *Semantics and Linguistic Theory* 30, p. 714. ISSN: 2163-5951. DOI: 10.3765/salt.v30i0.4844. URL: https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/30.714 (visited on 07/10/2023) (cit. on p. 77).

Kontinen, Juha, Julian-Steffen Müller, Henning Schnoor, and Heribert Vollmer (June 28, 2016). "Modal Independence Logic:" in: *Journal of Logic and Computation*, pp. 1333–1352. ISSN: 0955-792X, 1465-363X. DOI: 10.1093/logcom/exw019. URL: https://academic.oup.com/logcom/article-lookup/doi/10.1093/logcom/exw019 (visited on 03/14/2025) (cit. on p. 81).

Kratzer, Angelika (Jan. 1, 1977). "What 'Must' and 'Can' Must and Can Mean". In: *Linguistics and Philosophy* 1.3, pp. 337–355. ISSN: 1573-0549. DOI: 10.1007/BF00353453. URL: https://doi.org/10.1007/BF00353453 (visited on 01/26/2024) (cit. on p. 80).

— (1991). "Modality". In: *Semantik / Semantics: Ein internationales Handbuch zeitgenössischer Forschung*. Ed. by Arnim van Stechow and Dieter Wunderlich, pp. 639–650. DOI: 10.1515/9783110126969.7.639 (cit. on p. 80).

— (1998). "Scope or Pseudoscope? Are There Wide-Scope Indefinites?" In: *Events and Grammar*. Ed. by Susan Rothstein. Dordrecht: Springer Netherlands, pp. 163–196. ISBN: 978-94-011-3969-4. DOI: 10.1007/978-94-011-3969-4_8. URL: https://doi.org/10.1007/978-94-011-3969-4_8 (visited on 03/20/2025) (cit. on p. 47).

Kratzer, Angelika and Junko Shimoyama (2002). "Indeterminate Pronouns: The View from Japanese". In: *Contrastiveness in Information Structure, Alternatives and Scalar Implicatures*. Ed. by Chungmin Lee, Ferenc Kiefer, and Manfred Krifka. Cham: Springer International Publishing, pp. 123–143. ISBN: 978-3-319-10106-4. DOI: 10.1007/978-3-319-10106-4_7. URL: https://doi.org/10.1007/978-3-319-10106-4_7 (visited on 02/10/2025) (cit. on p. 35).

Krifka, Manfred (1989). "Nominal Reference, Temporal Constitution and Quantification in Event Semantics". In: *Semantics and Contextual Expression*. Ed. by Renate Bartsch, Johan van Benthem, and Peter van Emde Boas. De Gruyter Mouton, pp. 75–116. ISBN: 978-3-11-087733-5. DOI: 10.1515/9783110877335-005. URL: https://www.degruyter.com/document/doi/10.1515/9783110877335-005/html (visited on 10/29/2024) (cit. on p. 71).

— (1995). "The Semantics and Pragmatics of Polarity Items". In: *Linguistic Analysis* 25, pp. 209–257 (cit. on p. 44).

— (2011). "Questions". In: *Semantics: An International Handbook of Natural Language Meaning*. Vol. 2, pp. 1742–1785 (cit. on p. 17).

Kripke, Saul A. (1959). "A Completeness Theorem in Modal Logic". In: *The Journal of Symbolic Logic* 24.1, pp. 1–14. ISSN: 0022-4812. DOI: 10.2307/2964568. JSTOR: 2964568. URL: https://www.jstor.org/stable/2964568 (visited on 05/21/2021) (cit. on p. 80).

Kroch, Anthony (1998). "Amount Quantification, Referentiality, and Long Wh-Movement". In: *Current Work in Linguistics*. University of Pennsylvania Working Papers in Linguistics 5 (cit. on p. 18).

Lacheret, Anne, Sylvain Kahane, Julie Beliao, Anne Dister, Kim Gerdes, Jean-Philippe Goldman, Nicolas Obin, Paola Pietrandrea, and Atanas Tchobanov (2014). "Rhapsodie : un Treebank annoté pour l'étude de l'interface syntaxe-prosodie en français parlé". In: 4e Congrès Mondial de Linguistique Française. Ed. by Frank Neveu, Peter Blumenthal, Linda Hriba, Annette Gerstenberg, Judith Meinschaefer, and Sophie Prévost. Vol. 8. Berlin, Germany: SHS Web of Conferences, pp. 2675–2689. DOI: 10.1051/shsconf/20140801305. URL: https://www.shs-conferences.org/articles/shsconf/abs/2014/05/shsconf_cmlf14_01305/shsconf_cmlf14_01305.html (visited on 03/10/2023) (cit. on p. 148).

Lacheret-Dujour, Anne, Sylvain Kahane, and Paola Pietrandrea, eds. (2019). *Rhapsodie, A Prosodic and Syntactic Treebank for Spoken French*. Studies in Corpus Linguistics 89. John Benjamins Publishing Company. 396 pp. ISBN: 978-90-272-6292-9. URL: https://benjamins.com/catalog/scl.89 (visited on 03/10/2023) (cit. on p. 148).

Lahiri, Utpal (1991). "Embedded Interrogatives and Predicates That Embed Them". PhD thesis. Massachusetts Institute of Technology. URL: https://dspace.mit.edu/handle/1721.1/13520 (visited on 08/31/2023) (cit. on p. 34).

Lambrecht, Knud (1988). "Presentational Cleft Constructions in Spoken French". In: *Clause Combining in Grammar and Discourse*. John Benjamins Publishing Company, pp. 135–180. ISBN: 978-90-272-7859-3. URL: https://www.degruyterbrill.com/document/doi/10.1075/tsl.18.08lam/html (visited on 04/22/2025) (cit. on p. 23).

— (June 18, 2001). "A Framework for the Analysis of Cleft Constructions". In: 39.3, pp. 463–516. ISSN: 1613-396X. DOI: 10.1515/ling.2001.021. URL: https://www.degruyterbrill.com/document/doi/10.1515/ling.2001.021/html (visited on 04/22/2025) (cit. on pp. 24, 35).

Landman, Fred (Jan. 26, 2004). *Indefinites and the Type of Sets*. 1st edition. Hoboken: Wiley-Blackwell. 304 pp. ISBN: 978-1-4051-1630-5 (cit. on p. 36).

— (2011). "Boolean Pragmatics". In: *This Is Not a Festschrift*. Ed. by Jaap van der Does and Catarina Dulith Novaes. Amsterdam: Universiteit van Amsterdam (cit. on pp. 49, 50, 106).

Larrivée, Pierre and Alexander Guryev (Dec. 24, 2021). "Variantes formelles de l'interrogation. Présentation". In: *Langue francaise* 212.4, pp. 9–24. ISSN: 0023-8368. URL: https://www.cairn.info/revue-langue-francaise-2021-4-page-9.htm?ref=doi (visited on 02/15/2022) (cit. on pp. 143, 148).

Larrivée, Pierre and Alda Mari (Nov. 21–22, 2019). "Negative Biased Questions in the Modal Skeleton". In: Non-Canonical Questions Workshop. Konstanz (cit. on p. 85).

Lawler, John M. (1971). "Any Questions?" In: *Papers from the 7th Regional Meeting*. Vol. 7. Chicago Linguistic Society (cit. on p. 11).

Le Goffic, Pierre (2024). *Les mots qu- du français*. In: *Encyclopédie grammaticale du français*. Vol. 38. NAKALA - https://nakala.fr (Huma-Num - CNRS). DOI: 10.34847/NKL.FC1CYN4X. URL: https://nakala.fr/10.34847/nkl.fc1cyn4x (visited on 01/16/2025) (cit. on p. 165).

Ledegen, Gudrun (2007). "L'interrogative indirecte in situ à la Réunion : elle connaît elle veut quoi". In: *Le français parlé au 21ème siècle: Normes et variations géographiques et sociales. Actes du Colloque à l'Université d'Oxford (23 et 24 juin 2005)*. Vol. 1. Oxford: L'Harmattan, pp. 177–200. ISBN: 978-2-296-04407-4 (cit. on p. 145).

Lefeuvre, Florence (Dec. 24, 2021). "Les interrogatives averbales dans la presse, stratégies discursives récurrentes ?" In: *Langue francaise* 212.4, pp. 107–122. ISSN: 0023-8368. URL: https://www.cairn.info/revue-langue-francaise-2021-4-page-107.htm (visited on 02/16/2022) (cit. on p. 145).

Lefeuvre, Florence and Nathalie Rossi-Gensane (2017). "Les interrogatives indirectes en discours informel oral". In: *Langue française* 196.4, pp. 51–74. ISSN: 0023-8368. DOI: 10.3917/lf.196.0051. URL: https://www.cairn.info/revue-langue-francaise-2017-4-page-51.htm (visited on 04/03/2023) (cit. on pp. 145, 148).

Lewis, David (1975). "Adverbs of Quantification". In: *Formal Semantics of Natural Language*. Ed. by Edward L. Keenan. Cambridge: Cambridge University

Press, pp. 3–15. ISBN: 978-0-521-11111-9. DOI: 10.1017/CBO9780511897696.003. URL: https://www.cambridge.org/core/books/formal-semantics-of-natural-language/adverbs-of-quantification/282CF42EE7AE9CE7E00F2D09C72E17F3 (visited on 06/09/2025) (cit. on p. 34).

Li, Baichuan, Xiance Si, Michael R. Lyu, Irwin King, and Edward Y. Chang (Oct. 24, 2011). "Question Identification on Twitter". In: *Proceedings of the 20th ACM International Conference on Information and Knowledge Management*. CIKM '11. New York, NY, USA: Association for Computing Machinery, pp. 2477–2480. ISBN: 978-1-4503-0717-8. DOI: 10.1145/2063576.2063996. URL: https://doi.org/10.1145/2063576.2063996 (visited on 03/09/2023) (cit. on p. 145).

Li, Haoze (2020). "A Dynamic Semantics for Wh-Questions". PhD thesis. New York University. URL: https://www.proquest.com/openview/48d408f252370231cec764215532e6a5/1?cbl=18750&diss=y&pq-origsite=gscholar&parentSessionId=nDR7mVyN0AoeAGwiM%2F61ffs%2FONtPsOEtJUET550uTrs%3D (visited on 06/26/2024) (cit. on pp. 29, 89, 103, 104).

Lin, Jo-wang (July 1, 1998). "On Existential Polarity-WH-Phrases in Chinese". In: *Journal of East Asian Linguistics* 7.3, pp. 219–255. ISSN: 1572-8560. DOI: 10.1023/A:1008284513325. URL: https://doi.org/10.1023/A:1008284513325 (visited on 02/01/2025) (cit. on p. 28).

Maldonado, Mora (2020). "Plural Marking and D-Linking in Spanish Interrogatives". In: *Journal of Semantics* 37.1, pp. 145–170. URL: Paper=https://semanticsarchive.net/Archive/TBiNGExN/Maldonado-PluralSpanish-paper.pdf (cit. on pp. 39, 56).

Margolis, Anna and Mari Ostendorf (June 2011). "Question Detection in Spoken Conversations Using Textual Conversations". In: *Proceedings of the 49th Annual Meeting of the Association for Computational Linguistics: Human Language Technologies*. ACL-HLT 2011. Portland, Oregon, USA: Association for Computational Linguistics, pp. 118–124. URL: https://aclanthology.org/P11-2021 (visited on 03/09/2023) (cit. on p. 146).

Mari, Alda and Anastasia Giannakidou (2021). "Questions, Possibility Modals, and Negative Bias". In: Biased Questions: Experimental Results & Theoretical Modelling. Ed. by Manfred Krifka, Kazuko Yatsushiro, and Tue Trihn. Leibniz-Zentrum Allgemeine Sprachwissenschaft (ZAS). URL: https://osf.io/kxuzy (cit. on pp. 47, 85).

Maury, Nicole (1990). "Questions totales en français du Québec : le statut acoustique des morphèmes -tu et -ti". In: *Revue québécoise de linguistique*. La phonétique 19.2, pp. 111–134. ISSN: 1705-4591. DOI: 10.7202/602679AR (cit. on p. 145).

Mayr, Clemens (2013). "Downward Monotonicity in Questions". In: *Proceedings of Sinn und Bedeutung* 17, pp. 345–362. ISSN: 2629-6055. URL: https://ojs.

ub.uni-konstanz.de/sub/index.php/sub/article/view/349 (visited on 02/11/2025) (cit. on p. 42).

McDonald, Ryan, Joakim Nivre, Yvonne Quirmbach-Brundage, Yoav Goldberg, Dipanjan Das, Kuzman Ganchev, Keith Hall, Slav Petrov, Hao Zhang, Oscar Täckström, Claudia Bedini, Núria Bertomeu Castelló, and Jungmee Lee (Aug. 2013). "Universal Dependency Annotation for Multilingual Parsing". In: *Proceedings of the 51st Annual Meeting of the Association for Computational Linguistics (Volume 2: Short Papers)*. ACL 2013. Sofia, Bulgaria: Association for Computational Linguistics, pp. 92–97. URL: https://aclanthology.org/P13-2017 (visited on 03/10/2023) (cit. on p. 148).

Mittwoch, Anita (Oct. 1, 1979). "Final Parentheticals with English Questions — Their Illocutionary Function and Grammar". In: *Journal of Pragmatics* 3.5, pp. 401–412. ISSN: 0378-2166. DOI: 10.1016/0378-2166(79)90016-X. URL: https://www.sciencedirect.com/science/article/pii/037821667990016X (visited on 03/20/2025) (cit. on p. 15).

Muzerelle, Judith, Anaïs Lefeuvre, Emmanuel Schang, Jean-Yves Antoine, Aurore Pelletier, Denis Maurel, Iris Eshkol, and Jeanne Villaneau (May 2014). "ANCOR_Centre, a Large Free Spoken French Coreference Corpus: Description of the Resource and Reliability Measures". In: *Proceedings of the Ninth International Conference on Language Resources and Evaluation (LREC'14)*. LREC 2014. Ed. by Nicoletta Calzolari, Khalid Choukri, Thierry Declerck, Hrafn Loftsson, Bente Maegaard, Joseph Mariani, Asuncion Moreno, Jan Odijk, and Stelios Piperidis. Reykjavik, Iceland: European Language Resources Association (ELRA), pp. 843–847. URL: http://www.lrec-conf.org/proceedings/lrec2014/pdf/150_Paper.pdf (visited on 11/28/2024) (cit. on pp. 94, 214).

Myers, Lindsy Lee (2007). "WH-interrogatives in Spoken French: A Corpus-Based Analysis of Their Form and Function". PhD thesis. Austin, Texas: The University of Texas at Austin. 268 pp. URL: https://repositories.lib.utexas.edu/handle/2152/3235 (visited on 09/01/2021) (cit. on p. 145).

Nand, Parma and Wai Yeap (2013). "A Framework for Interpreting Bridging Anaphora". In: *Agents and Artificial Intelligence*. Ed. by Joaquim Filipe and Ana Fred. Berlin, Heidelberg: Springer, pp. 131–144. ISBN: 978-3-642-36907-0. DOI: 10.1007/978-3-642-36907-0_9 (cit. on p. 92).

Nicolae, Andreea C. (Mar. 1, 2015). "Questions with NPIs". In: *Natural Language Semantics* 23.1, pp. 21–76. ISSN: 1572-865X. DOI: 10.1007/s11050-014-9110-8. URL: https://doi.org/10.1007/s11050-014-9110-8 (visited on 01/18/2025) (cit. on p. 44).

Nishigauchi, Taisuke (Jan. 1, 1986). "Quantification in Syntax (Logical Form, Japanese, WH-Constructions)". PhD thesis. UMass Amherst. 290 pp. URL: https://scholarworks.umass.edu/dissertations/AAI8701206 (cit. on pp. 89, 118).

— (1990). *Quantification in the Theory of Grammar*. Red. by Gennaro Chierchia, Pauline Jacobson, and Francis J. Pelletier. Vol. 37. Studies in Linguistics and

Philosophy. Dordrecht: Springer Netherlands. ISBN: 978-0-7923-0644-3. DOI: 10.1007/978-94-009-1972-3. URL: http://link.springer.com/10.1007/978-94-009-1972-3 (visited on 09/11/2023) (cit. on pp. 74, 89, 118, 119).

Nouwen, Rick (Feb. 1, 2003). "Complement Anaphora and Interpretation". In: *Journal of Semantics* 20.1, pp. 73–113. ISSN: 0167-5133. DOI: 10.1093/jos/20.1.73. URL: https://doi.org/10.1093/jos/20.1.73 (visited on 05/07/2024) (cit. on p. 90).

Obenauer, Hans-Georg (1976). *Etudes de syntaxe interrogative du français: Quoi, combien et le complémenteur*. Max Niemeyer Verlag. ISBN: 978-3-11-134036-4. DOI: 10.1515/9783111340364. URL: https://www.degruyter.com/document/doi/10.1515/9783111340364/html (visited on 05/25/2022) (cit. on p. 161).

Onea, Edgar (Feb. 2, 2016). *Potential Questions at the Semantics-Pragmatics Interface*. Brill. ISBN: 978-90-04-21793-5. URL: https://brill.com/display/title/32659 (visited on 02/08/2025) (cit. on pp. 15, 55, 56).

Onea, Edgar and Dolgor Guntsetseg (2011). "Focus and Wh-Questions in Mongolian". In: *Proceedings of Sinn und Bedeutung* 15, pp. 467–482. ISSN: 2629-6055. URL: https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/393 (visited on 03/20/2025) (cit. on pp. 22, 25, 26, 34).

Onea, Edgar and Malte Zimmermann (Mar. 26, 2019). "Questions in Discourse: An Overview". In: *Questions in Discourse*. Brill, pp. 5–117. ISBN: 978-90-04-37830-8. DOI: 10.1163/9789004378308_003. URL: https://brill.com/display/book/edcoll/9789004378308/BP000002.xml (visited on 02/08/2025) (cit. on pp. 14, 52, 54).

Özger, Zeynep Banu, Banu Dirì, and Canan Girgìn (June 2014). "Question Identification on Turkish Tweets". In: *2014 IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA) Proceedings*. 2014 IEEE International Symposium on Innovations in Intelligent Systems and Applications (INISTA) Proceedings, pp. 126–130. DOI: 10.1109/INISTA.2014.6873608 (cit. on p. 146).

Özyıldız, Deniz, Ciyang Qing, Floris Roelofsen, Maribel Romero, and Wataru Uegaki (May 2023). "A Crosslinguistic Database for Combinatorial and Semantic Properties of Attitude Predicates". In: *Proceedings of the 5th Workshop on Research in Computational Linguistic Typology and Multilingual NLP*. SIGTYP 2023. Dubrovnik, Croatia: Association for Computational Linguistics, pp. 65–75. URL: https://aclanthology.org/2023.sigtyp-1.7 (visited on 07/10/2023) (cit. on p. 140).

Palmer, Martha, Rajesh Bhatt, Bhuvana Narasimhan, Owen Rambow, Dipti Misra Sharma, and Fei Xia (2009). "Hindi Syntax: Annotating Dependency, Lexical Predicate-Argument Structure, and Phrase Structure". In: *The 7th International Conference on Natural Language Processing*, pp. 14–17 (cit. on p. 147).

Partee, Barbara (2014). "Formal Semantics and Typology of Anaphora". Lecture notes of HSE Spring 2014. UMass Amherst. URL: https://people.umass.edu/partee/HSE_Web_14/index.html (cit. on p. 90).

Perrier, Guy, Marie Candito, Bruno Guillaume, Corentin Ribeyre, Karën Fort, and Djamé Seddah (July 2014). "Un schéma d'annotation en dépendances syntaxiques profondes pour le français". In: *Proceedings of TALN 2014 (Volume 2: Short Papers)*. JEP/TALN/RECITAL 2014. Marseille, France: Association pour le Traitement Automatique des Langues, pp. 574–579. URL: https://aclanthology.org/F14-2031 (visited on 03/10/2023) (cit. on p. 148).

Péry-Woodley, Marie-Paule, Stergos Afantenos, Lydia-Mai Ho-Dac, and Nicholas Asher (2011). "La ressource ANNODIS, un corpus enrichi d'annotations discursives". In: *Revue TAL* 52.3, pp. 71–101. URL: https://shs.hal.science/halshs-00935201 (visited on 04/20/2023) (cit. on p. 151).

Pesetsky, David (1987). "Wh-In-Situ: Movement and Unselective Binding". In: *The Representation of (In)Definiteness*. Ed. by Eric Reuland and Alice ter Meulen. Cambridge, England: MIT Press (cit. on pp. 18, 47–49).

Poesio, Massimo, Maris Camilleri, Paloma Carretero Garcia, Juntao Yu, and Mark-Christoph Müller (Mar. 2024). "The ARRAU 3.0 Corpus". In: *Proceedings of the 5th Workshop on Computational Approaches to Discourse (CODI 2024)*. Ed. by Michael Strube, Chloe Braud, Christian Hardmeier, Junyi Jessy Li, Sharid Loaiciga, Amir Zeldes, and Chuyuan Li. St. Julians, Malta: Association for Computational Linguistics, pp. 127–138. URL: https://aclanthology.org/2024.codi-1.12 (visited on 11/28/2024) (cit. on pp. 34, 92, 93, 96).

Postal, Paul M. (1971). *Cross-over Phenomena*. The Transatlantic Series in Linguistics. New York: Holt, Rinehart and Winston (cit. on p. 11).

Queffélec, Ambroise (2006). "Restructurations morphosyntaxiques en français populaire camerounais : L'expression des modalités injonctives et interrogatives dans le discours rapporté". In: *Le Français en Afrique* 21. Ed. by Ambroise Queffélec and Suzanne Lafage, pp. 1–24. ISSN: 1157-1454. URL: http://www.unice.fr/ILF-CNRS/ofcaf/21/Table.html (cit. on p. 145).

Quillard, Virginie (2001). "La diversité des formes interrogatives : comment l'interpréter ?" In: *Langage et société* 95.1, pp. 57–72. ISSN: 0181-4095. DOI: 10.3917/ls.095.0057. URL: https://www.cairn.info/revue-langage-et-societe-2001-1-page-57.htm (visited on 03/20/2023) (cit. on p. 145).

Reinhardt, Janina (2016). "Établir un corpus oral de questions : L'analyse semi-automatisée avec Praat et Perl à l'exemple de cinq épisodes de Maya l'Abeille". In: 5e Congrès Mondial de Linguistique Française. Ed. by Frank Neveu, Gabriel Bergounioux, Marie-Hélène Côté, Jean-Marie Fournier, Linda Hriba, and Sophie Prévost. Vol. 27. Tours: SHS Web of Conferences, p. 11007. DOI: 10.1051/shsconf/20162711007. URL: https://www.shs-conferences.

org/articles/shsconf/abs/2016/05/shsconf_cmlf2016_11007/shsconf_cmlf2016_11007.html (visited on 03/20/2023) (cit. on pp. 145, 151).

— (Feb. 18, 2019a). "La transmission du manque d'information dans la téléréalité française". In: Transmission, oubli et mémoire dans les sciences du langage, JC2017 - 20èmes Rencontres des jeunes chercheurs en Sciences du Langage. Paris, France. URL: https://www.semanticscholar.org/paper/La-transmission-du-manque-d%E2%80%99information-dans-la-Reinhardt/6789d2cca8326b85d02a71b199b3b8aef332a4d3 (visited on 02/16/2022) (cit. on p. 145).

— (May 21, 2019b). *Les interrogatives directes tirées de dix romans policier*. Corpus. Version 1. https://www.ortolang.fr/: ORTOLANG (Open Resources and TOols for LANGuage). DOI: https://hdl.handle.net/11403/interrogatives-in-novels/v1 (cit. on pp. 145, 151).

Richard, Valentin D. (Nov. 2023). "Est-ce que l'extraction des interrogatives du français peut-elle être automatisée ?" In: *5èmes journées du Groupement de Recherche CNRS " Linguistique Informatique, Formelle et de Terrain " (LIFT 2023)*. Ed. by Karën Fort, Claire Gardent, and Yannick Parmentier. Nancy, France: CNRS, pp. 69–76. URL: https://hal.science/hal-04359947 (cit. on pp. 7, 143, 216).

— (2024a). "« selon comment vous vous positionnez » : Étude des circonstancielles à interrogative". In: *9e Congrès Mondial de Linguistique Française*. Ed. by Frank Neveu, Sophie Prévost, Aurore Montébran, Agnès Steuckardt, Gabriel Bergounioux, Gilles Merminod, and Gilles Philippe. Vol. 191. Syntaxe. Lausanne: SHS Web of Conferences, p. 14010. DOI: 10.1051/shsconf/202419114010 (cit. on pp. 115, 116, 120, 215).

— (Dec. 2024b). "Dynamic Effects of Modalized Questions". In: *Proceedings of the 24th Amsterdam Colloquium*. Ed. by Fausto Carcassi, Tamar Johnson, Søren Brinck Knudstorp, Sabina Domínguez Parrado, Pablo Rivas-Robledo, and Giorgio Sbardolini. Amsterdam, The Netherlands, pp. 289–307. URL: https://hal.science/hal-04866506 (cit. on pp. 7, 38).

— (Oct. 2025a). "Évaluation d'un corpus annoté en anaphores : le cas des chaines contenant un mot interrogatif". In: *Lift2-2025 : Journées scientifiques du réseau thématique LIFT2 - linguistique informatique, formelle et de terrain*. Ed. by Timothée Bernard, Berthold Crysmann, Ali Dhibi, and Guillaume Wisniewski. Paris, France, pp. 1–6. URL: https://hal.science/hal-05326668 (cit. on pp. 7, 94, 169).

— (Nov. 2025b). "Raising Alternatives to Express Dependence: A Compositional Issue". In: *16th Syntax and Semantics Conference in Paris (CSSP 2025)*. Ed. by Charlotte Hauser, Céline Pozniak, and Laurent Roussarie. Paris, France, pp. 1–7. URL: https://conf.llf-paris.fr/cssp2025/content/abstracts/CSSP_2025_paper_5.pdf (cit. on p. 7).

Rizzi, Luigi (June 13, 1990). *Relativized Minimality*. Linguistic Inquiry Monographs. The MIT Press. 160 pp. ISBN: 978-0-262-68061-5 (cit. on p. 47).

Rizzi, Luigi (1997). "The Fine Structure of the Left Periphery". In: *Elements of Grammar: Handbook in Generative Syntax*. Ed. by Liliane Haegeman. Dordrecht: Springer Netherlands, pp. 281–337. ISBN: 978-94-011-5420-8. DOI: 10.1007/978-94-011-5420-8_7. URL: https://doi.org/10.1007/978-94-011-5420-8_7 (visited on 11/20/2024) (cit. on p. 73).

Roberts, Craige (Dec. 1, 1989). "Modal Subordination and Pronominal Anaphora in Discourse". In: *Linguistics and Philosophy* 12.6, pp. 683–721. ISSN: 1573-0549. DOI: 10.1007/BF00632602. URL: https://doi.org/10.1007/BF00632602 (visited on 07/26/2023) (cit. on pp. 4, 78, 84, 92).

— (1996). "Information Structure: Towards an Integrated Formal Theory of Pragmatics". In: *Semantics and Pragmatics* 5, 6:1–69. ISSN: 1937-8912. DOI: 10.3765/sp.5.6. URL: https://semprag.org/index.php/sp/article/view/sp.5.6 (visited on 12/01/2023) (cit. on p. 55).

Roelofsen, Floris (2015). "The Semantics of Declarative and Interrogative Lists". Report. URL: http://semanticsarchive.net/Archive/jdiNjljO/paper.pdf (cit. on p. 40).

— (2018). "NPIs in Questions". Presentation (New York University, Department of Linguistics) (cit. on pp. 40–43).

Roelofsen, Floris and Jakub Dotlačil (Dec. 1, 2023a). "Dynamic Inquisitive Semantics—Looking Ahead and Looking Back". In: *Theoretical Linguistics* 49.3–4, pp. 319–337. ISSN: 1613-4060. DOI: 10.1515/tl-2023-2017. URL: https://www.degruyter.com/document/doi/10.1515/tl-2023-2017/html (visited on 12/08/2023) (cit. on pp. 35, 67).

— (June 1, 2023b). "Wh-Questions in Dynamic Inquisitive Semantics". In: *Theoretical Linguistics* 49.1–2, pp. 1–91. ISSN: 1613-4060. DOI: 10.1515/tl-2023-2001. URL: https://www.degruyter.com/document/doi/10.1515/tl-2023-2001/html (visited on 12/13/2023) (cit. on pp. 5, 37, 67, 68, 70, 73, 77, 80, 99, 109, 124, 137, 213, 217, 219).

Roelofsen, Floris and Donka F. Farkas (2015). "Polarity Particle Responses as a Window onto the Interpretation of Questions and Assertions". In: *Language* 91.2, pp. 359–414. ISSN: 0097-8507. DOI: 10.1353/lan.2015.0017. URL: https://escholarship.org/uc/item/99k892mp (visited on 01/10/2025) (cit. on p. 55).

Roelofsen, Floris, Michele Herbstritt, and Maria Aloni (Mar. 26, 2019). "The *whether Puzzle". In: *Questions in Discourse*. Vol. 35. Current Research in the Semantics / Pragmatics Interface. Brill, pp. 172–197. ISBN: 978-90-04-37830-8. DOI: 10.1163/9789004378308_005. URL: https://brill.com/display/book/edcoll/9789004378308/BP000004.xml (visited on 02/08/2025) (cit. on pp. 42, 56).

Roelofsen, Floris and Sam van Gool (2010). "Disjunctive Questions, Intonation, and Highlighting". In: *Logic, Language and Meaning*. Ed. by Maria Aloni, Harald Bastiaanse, Tikitu de Jager, and Katrin Schulz. Berlin, Heidelberg:

Springer, pp. 384–394. ISBN: 978-3-642-14287-1. DOI: 10.1007/978-3-642-14287-1_39 (cit. on p. 55).

Rooryck, Johan (Jan. 1, 1994). "On Two Types of Underspecification: Towards a Feature Theory Shared by Syntax and Phonology". In: 6.2–3, pp. 207–234. ISSN: 1613-4079. DOI: 10.1515/prbs.1994.6.2-3.207. URL: https://www.degruyter.com/document/doi/10.1515/prbs.1994.6.2-3.207/html (visited on 09/02/2022) (cit. on p. 161).

Rossi-Gensane, Nathalie, Luisa Fernanda Acosta Córdoba, Biagio Ursi, and Margot Lambert (2021). "Les structures interrogatives directes partielles fondées sur *où* dans les dialogues de romans français du XXe siècle". In: *Journal of French Language Studies*. L'oralité mise en scène : syntaxe et phonologie 31.2, pp. 169–191. DOI: 10.1017/S0959269520000253 (cit. on pp. 145, 154).

Rullmann, Hotze and Sigrid Beck (Oct. 6, 1998). "Presupposition Projection and the Interpretation of Which-Questions". In: *Semantics and Linguistic Theory* 8, pp. 215–232. ISSN: 2163-5951. DOI: 10.3765/salt.v8i0.2811. URL: http://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/2811 (visited on 09/04/2023) (cit. on pp. 11, 20, 33).

Ruys, Eddy G. (Nov. 30, 2023). "On the Semantics of 'Wh'-". In: *Proceedings of Sinn und Bedeutung* 27, pp. 554–570. ISSN: 2629-6055. DOI: 10.18148/sub/2023.v27.1086. URL: https://ojs.ub.uni-konstanz.de/sub/index.php/sub/article/view/1086 (visited on 06/09/2025) (cit. on p. 30).

Sag, Ivan A., Timothy Baldwin, Francis Bond, Ann Copestake, and Dan Flickinger (2002). "Multiword Expressions: A Pain in the Neck for NLP". In: *Computational Linguistics and Intelligent Text Processing*. Ed. by Alexander Gelbukh. Lecture Notes in Computer Science. Berlin, Heidelberg: Springer, pp. 1–15. ISBN: 978-3-540-45715-2. DOI: 10.1007/3-540-45715-1_1 (cit. on p. 149).

Sanguinetti, Manuela and Cristina Bosco (2014). "Converting the Parallel Treebank ParTUT in Universal Stanford Dependencies". In: *Proceedings of the First Italian Conference on Computational Linguistics CLiC-it 2014 and of the Fourth International Workshop EVALITA 2014 9-11 December 2014, Pisa*. Pisa University Press. DOI: 10.12871/clicit2014161. URL: http://clic2014.fileli.unipi.it/proceedings/Proceedings-CLICit-2014.pdf (visited on 03/10/2023) (cit. on p. 148).

— (2015). "PartTUT: The Turin University Parallel Treebank". In: *Harmonization and Development of Resources and Tools for Italian Natural Language Processing within the PARLI Project*. Ed. by Roberto Basili, Cristina Bosco, Rodolfo Delmonte, Alessandro Moschitti, and Maria Simi. Studies in Computational Intelligence. Cham: Springer International Publishing, pp. 51–69. ISBN: 978-3-319-14206-7. DOI: 10.1007/978-3-319-14206-7_3. URL: https://doi.org/10.1007/978-3-319-14206-7_3 (visited on 03/10/2023) (cit. on p. 148).

Sauerland, Uli, Jan Anderssen, and Kazuko Yatsushiro (Aug. 22, 2008). "The Plural Is Semantically Unmarked". In: *Linguistic Evidence: Empirical, Theoretical and Computational Perspectives*. Ed. by Stephan Kepser and Marga Reis. De Gruyter Mouton, pp. 413–434. ISBN: 978-3-11-019754-9. DOI: 10.1515/9783110197549.413. URL: https://www.degruyter.com/document/doi/10.1515/9783110197549.413/html?srsltid=AfmBOoqHBBdQSm0Cr5l6ORw5Jd-mnmsf5t0rlnzYbU5HH01ZXAzVDWpt (visited on 01/08/2025) (cit. on p. 31).

Schlenker, Philippe (Aug. 1, 2011). "Donkey Anaphora: The View from Sign Language (ASL and LSF)". In: *Linguistics and Philosophy* 34.4, pp. 341–395. ISSN: 1573-0549. DOI: 10.1007/s10988-011-9098-1. URL: https://doi.org/10.1007/s10988-011-9098-1 (visited on 07/11/2023) (cit. on p. 90).

— (Dec. 1, 2012). "Maximize Presupposition and Gricean Reasoning". In: *Natural Language Semantics* 20.4, pp. 391–429. ISSN: 1572-865X. DOI: 10.1007/s11050-012-9085-2. URL: https://doi.org/10.1007/s11050-012-9085-2 (visited on 08/13/2025) (cit. on p. 39).

Schwarz, Bernhard (Dec. 6, 2017a). "Negative Polarity Items: A Case for Questions as Licensers". In: *Semantics and Linguistic Theory*, pp. 230–247. ISSN: 2163-5951. DOI: 10.3765/salt.v27i0.4113. URL: https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/27.230 (visited on 01/18/2025) (cit. on p. 42).

— (2017b). "On Question Exhaustivity and NPI Licensing". In: *Proceedings of the 21st Amsterdam Colloquium*. Amsterdam, The Netherlands: ILLC, University of Amsterdam, pp. 405–415 (cit. on pp. 42, 44).

Schwarzschild, Roger (Aug. 1, 2002). "Singleton Indefinites". In: *Journal of Semantics* 19.3, pp. 289–314. ISSN: 0167-5133. DOI: 10.1093/jos/19.3.289. URL: https://doi.org/10.1093/jos/19.3.289 (visited on 03/25/2025) (cit. on pp. 35, 104).

Seddah, Djamé and Marie Candito (May 2016). "Hard Time Parsing Questions: Building a QuestionBank for French". In: *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*. LREC 2016. Portorož, Slovenia: European Language Resources Association (ELRA), pp. 2366–2370. URL: https://aclanthology.org/L16-1375 (visited on 01/03/2023) (cit. on pp. 147, 153, 216).

Shrestha, Lokesh and Kathleen McKeown (Aug. 23, 2004). "Detection of Question-Answer Pairs in Email Conversations". In: *COLING 2004: Proceedings of the 20th International Conference on Computational Linguistics*. COLING 2004. Geneva, Switzerland: Association for Computational Linguistics, pp. 889–895. URL: https://aclanthology.org/C04-1128 (visited on 03/09/2023) (cit. on p. 146).

Socolof, Michaela, Aron Hirsch, and Bernhard Schwarz (2020). "Which-Questions, Uniqueness, and Answerhood: Evidence from Disjunction". In: *Proceedings of NELS*. Vol. 50 (cit. on p. 80).

Spector, Benjamin (Oct. 3, 2007). "Modalized Questions and Exhaustivity". In: *Semantics and Linguistic Theory* 17, pp. 282–299. ISSN: 2163-5951. DOI: 10.3765/salt.v17i0.2962. URL: https://journals.linguisticsociety.org/proceedings/index.php/SALT/article/view/2962 (visited on 06/11/2024) (cit. on pp. 56–58, 141, 216).

Spector, Benjamin and Paul Egré (June 1, 2015). "A Uniform Semantics for Embedded Interrogatives: An Answer, Not Necessarily the Answer". In: *Synthese* 192.6, pp. 1729–1784. ISSN: 1573-0964. DOI: 10.1007/s11229-015-0722-4. URL: https://doi.org/10.1007/s11229-015-0722-4 (visited on 07/19/2021) (cit. on p. 16).

Stalnaker, Robert C. (1981). "A Theory of Conditionals". In: *IFS: Conditionals, Belief, Decision, Chance and Time*. Ed. by William L. Harper, Robert Stalnaker, and Glenn Pearce. Dordrecht: Springer Netherlands, pp. 41–55. ISBN: 978-94-009-9117-0. DOI: 10.1007/978-94-009-9117-0_2. URL: https://doi.org/10.1007/978-94-009-9117-0_2 (visited on 06/06/2024) (cit. on p. 80).

Starke, Michal (2001). "Move Dissolves into Merge: A Theory of Locality". PhD thesis. University of Geneva (cit. on p. 45).

Sun, Lin, Bingquan Liu, Baoxun Wang, Deyuan Zhang, and Xiaolong Wang (Aug. 2010). "A Study of Features on Primary Question Detection in Chinese Online Forums". In: *2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery*. 2010 Seventh International Conference on Fuzzy Systems and Knowledge Discovery. Vol. 5, pp. 2422–2427. DOI: 10.1109/FSKD.2010.5569298 (cit. on p. 146).

Szabolcsi, Anna and Frans Zwarts (Jan. 1, 1993). "Weak Islands and an Algebraic Semantics for Scope Taking". In: *Natural Language Semantics* 1.3, pp. 235–284. ISSN: 1572-865X. DOI: 10.1007/BF00263545. URL: https://doi.org/10.1007/BF00263545 (visited on 02/12/2025) (cit. on pp. 48, 49).

Tandon, Juhi, Himani Chaudhry, Riyaz Ahmad Bhat, and Dipti Sharma (Aug. 2016). "Conversion from Paninian Karakas to Universal Dependencies for Hindi Dependency Treebank". In: *Proceedings of the 10th Linguistic Annotation Workshop Held in Conjunction with ACL 2016 (LAW-X 2016)*. LAW 2016. Berlin, Germany: Association for Computational Linguistics, pp. 141–150. DOI: 10.18653/v1/W16-1716. URL: https://aclanthology.org/W16-1716 (visited on 03/09/2023) (cit. on p. 147).

Theiler, Nadine (Aug. 22, 2014). "A Multitude of Answers: Embedded Questions in Typed Inquisitive Semantics". MSc. Thesis. Amsterdam, The Netherlands: Institute for Logic, Language and Computation (ILLC), Universiteit van Amsterdam. URL: https://eprints.illc.uva.nl/925/ (visited on 06/10/2021) (cit. on p. 122).

Theiler, Nadine, Floris Roelofsen, and Maria Aloni (June 1, 2019). "Picky Predicates: Why Believe Doesn't like Interrogative Complements, and Other Puzzles". In: *Natural Language Semantics* 27.2, pp. 95–134. ISSN: 1572-865X. DOI:

10.1007/s11050-019-09152-9. URL: https://doi.org/10.1007/s11050-019-09152-9 (visited on 07/29/2021) (cit. on pp. 116–118, 123, 124, 129, 215).

Tian, Le, Xipeng Qiu, and Xuanjing Huang (July 2013). "Question Identification in Chinese Micro-Texts". In: *2013 International Conference on Machine Learning and Cybernetics*. 2013 International Conference on Machine Learning and Cybernetics. Vol. 4, pp. 1910–1915. DOI: 10.1109/ICMLC.2013.6890907 (cit. on p. 145).

Uegaki, Wataru (Aug. 1, 2021). "The Existential/Uniqueness Presupposition of Wh-Complements Projects from the Answers". In: *Linguistics and Philosophy* 44.4, pp. 911–951. ISSN: 1573-0549. DOI: 10.1007/s10988-020-09309-4. URL: https://doi.org/10.1007/s10988-020-09309-4 (visited on 06/29/2023) (cit. on pp. 11, 20, 33, 58).

Väänänen, Jouko (2008). *Modal Dependence Logic*. Amsterdam: Amsterdam University Press. ISBN: 978-90-8964-057-4. URL: https://dare.uva.nl/search?identifier=eab2a54c-277c-4416-984d-640d88e4e8c6 (visited on 03/14/2025) (cit. on p. 81).

Van den Berg, Martin H. (Jan. 1, 1996). "Some Aspects of the Internal Structure of Discourse: The Dynamics of Nominal Anaphora". PhD thesis. Amsterdam: University of Amsterdam. 202 pp. URL: https://eprints.illc.uva.nl/id/eprint/1996/ (visited on 07/24/2023) (cit. on pp. 79, 90).

Van Rooij, Robert (1997). "The Context-Dependence of Questions and Answers". In: *The Proceedings of the 11th Amsterdam Colloquium*. Ed. by Paul Dekker, Martin Stokhof, and Yde Venema. Amsterdam, The Netherlands, pp. 61–66 (cit. on pp. 67, 89, 101).

— (1998). "Modal Subordination in Questions". In: *Formal Semantics and Pragmatics of Dialogue. Proceedings of Twendial '98*. Ed. by Jan Hulstijn and Anton Nijholt. Enschede: University of Twente, pp. 237–247. URL: https://dare.uva.nl/search?identifier=31017f54-cf94-4cfd-b5c1-a4aa9f433b55 (visited on 07/07/2023) (cit. on pp. 30, 79, 89, 95, 98, 99, 101, 103).

— (Aug. 2003). "Negative Polarity Items in Questions: Strength as Relevance". In: *Journal of Semantics* 20.3, pp. 239–273. ISSN: 0167-5133. DOI: 10.1093/jos/20.3.239. eprint: https://academic.oup.com/jos/article-pdf/20/3/239/4668345/200239.pdf. URL: https://doi.org/10.1093/jos/20.3.239 (cit. on p. 42).

Veltman, Frank (1996). "Defaults in Update Semantics". In: *Journal of Philosophical Logic* 25.3, pp. 221–261. ISSN: 0022-3611. JSTOR: 30226572. URL: https://www.jstor.org/stable/30226572 (visited on 05/07/2024) (cit. on pp. 79, 80).

Von Heusinger, Klaus (Aug. 1, 2002). "Specificity and Definiteness in Sentence and Discourse Structure". In: *Journal of Semantics* 19.3, pp. 245–274. ISSN: 0167-5133. DOI: 10.1093/jos/19.3.245. URL: https://doi.org/10.1093/

jos/19.3.245 (visited on 03/20/2025) (cit. on pp. 4, 29, 31, 32, 48, 60, 109, 140).

Wang, Baoxun, Bingquan Liu, Chengjie Sun, Xiaolong Wang, and Lin Sun (Oct. 2009). "Extracting Chinese Question-Answer Pairs from Online Forums". In: *2009 IEEE International Conference on Systems, Man and Cybernetics*. 2009 IEEE International Conference on Systems, Man and Cybernetics, pp. 1159–1164. DOI: 10.1109/ICSMC.2009.5345956 (cit. on p. 146).

Wang, Kai and Tat-Seng Chua (Aug. 23, 2010). "Exploiting Salient Patterns for Question Detection and Question Retrieval in Community-Based Question Answering". In: *Proceedings of the 23rd International Conference on Computational Linguistics*. COLING '10. USA: Association for Computational Linguistics, pp. 1155–1163 (cit. on p. 146).

Webber, Bonnie, Matthew Stone, Aravind Joshi, and Alistair Knott (2003). "Anaphora and Discourse Structure". In: *Computational Linguistics* 29.4, pp. 545–587. DOI: 10.1162/089120103322753347. URL: https://aclanthology.org/J03-4002/ (visited on 05/05/2025) (cit. on p. 101).

Xiang, Yimei (May 17, 2016). "Interpreting Questions with Non-Exhaustive Answers". PhD thesis. Cambridge, Massachusetts, USA: Harvard University, Graduate School of Arts & Sciences. 300 pp. URL: https://www.semanticscholar.org/paper/Interpreting-Questions-with-Non-Exhaustive-Answers-Xiang/3f3183c8150447d1cfdb5ab0faf6bc14bb1bf316 (visited on 09/05/2023) (cit. on p. 78).

Xiang, Yimei and Alexandre Cremers (2017). "Mention-Some Readings of Plural-Marked Questions: Experimental Evidence". In: *Proceedings of the Forty-Seventh Annual Meeting of the North East Linguistic Society (NELS 47)*. Ed. by Andrew Lamont and Katerina Tetzloff. Vol. 3. University of Massachusetts, Amherst, USA: GLSA (Graduate Linguistics Student Association), pp. 261–274. ISBN: 978-1-9763-4460-2. URL: https://www.semanticscholar.org/paper/Mention-some-readings-of-plural-marked-questions%3A-Xiang-Cremers/f29b8de6ff99142533c255fc4c831ac1fc4237bc (visited on 07/10/2023) (cit. on p. 78).

Zeldes, Amir (Sept. 1, 2017). "The GUM Corpus: Creating Multilayer Resources in the Classroom". In: *Language Resources and Evaluation* 51.3, pp. 581–612. ISSN: 1574-0218. DOI: 10.1007/s10579-016-9343-x. URL: https://doi.org/10.1007/s10579-016-9343-x (visited on 11/28/2024) (cit. on pp. 93, 146).

Zuckerman, Shalom (2001). "The Acquisition of "Optional" Movement". Thesis fully internal (DIV). Groningen: University of Groningen. 175 pp. URL: https://research.rug.nl/nl/publications/the-acquisition-of-optional-movement (cit. on p. 161).

List of symbols and abbreviations

! Inquisitiveness-elimination operator. 66, 71
; Conjunction in DInq, of type $k \to k$ or $\ell \to \ell$. 71
?u Witness request operator. 72
$[u/a]$ Assignment function assigning individual a to discourse referent u. 74
$[u]$ Referent introduction operator. 72
† Presuppotional closure operator. 73
:= Definitional equation. 65
$\neg$ Negation in DInq. 71
$\Diamond$ Possibility modal operator. 77
$\langle ? \rangle$ Conditional inquisitive closure operator. 73
$\leqslant$ Extension. 69
$\Subset$ "Belong to" subsistence. 69
$\sqcup$ Disjunction in DInq. 71
$\Box$ Necessity modal operator. 84
$\star$ Null individual. 35
$\sqsubseteq$ "Is equivalent to" subsistence. 69
$c[c' \vdash \mathcal{U}]$ Local conditionalization: c after learning that *if* c' *then* $\mathcal{U}$. Used in percolation. 84
‡ Disfluency or annoation mistake. 95, 118
$\rightsquigarrow$ Inference. 11

ACC Accusative case. 22
ADU Adjunct Dependence Utterance. 114
ALT Alternative operator. 64
atom Atom operator. 72

BM Background marker. 23

$c, \ldots$ A (local) context. 68
CL Clitic. 28
COM Comitative case. 21
COND Conditional tense. 24
CVB Converb. 21

D The set of plural non-null individuals. 64
ddom Definition domain. 71
DEC Type head of declarative clauses. 73
DInq Dynamic Inquisitive Semantics. 66

e Type of individuals. 64
EI Existential inference of a simplex wh-question. 11
EXH Exhaustification operator. 35
EXPL Expletive pronoun. 23

FC(I) Free Choice (Item). 17
FinP Finite Phrase. 73
FOC$_u$ Focus head. 73

FOC Focus marking (gloss). 22
FocP Focus Phrase. 73
FUT Future tense. 28

$g, ...$ An assignment function. 67
GSV Groenendijk, Stokhof, and Veltman 1996's dynamic semantics. 67

$\mathcal{I}, \mathcal{J}, ...$ An issue. 64
IND Indicative mood. 130
Inq Inquisitive Semantics. 63
Int Intensional Semantics. 65
INT Type head of interrogative clauses. 73
INT Interrogative word (gloss). 23

k Type of (local) contexts. 68

ℓ Type of macro-contexts, i.e. of stacks of local contexts. 83

M A modal base. 115
MA Mention-all reading. 19
max Individual maximization operator. 72
MDInq Modal Dynamic Inquisitive Semantics. 83
MS Mention-some reading. 19
μ A modal base function. 81

NEG Negation. 22
NPI Negative Polarity Item. 15
NPST Non-past tense. 22

$\wp$ Powerset operator. 65
PERC Stack percolation operator. 84
PFV Perfective aspect. 23
PL Plural agreement. 96
POP Stack popping operator. 84
rdom Domain of defined discourse referents. 67
PPRT Past participle. 47
$\overset{\textbf{presup}}{\Longrightarrow}$ Presupposes. 53
PROG Progressive aspect. 120
PST Past tense. 21
PUSH Stack push operator. 84

Q Question particle (gloss). 21

r Type of discourse referents. 67
REFL Reflexive pronoun. 24
$R, S, ...$ A logical predicate, usually restrictor and nuclear scope resp.. 70

s Type of possible worlds. 64
SBJ Subject allomorph. 23
SELON Expression introducing an adjunct forming an ADU: French prepositions *selon*, *suivant* or *en fonction de*, and English verb *depending on*. 113
SG Singular agreement. 24
SJV Subjunctive mood. 24, 130
$s, t, ...$ An information state. 68

t Type of truth values. 64
τ A stack. 80
t_i Wh-trace, i.e. trace of a wh-word. 40
TypeP/CP Syntactic Type Phrase / Complementizer phrase (for embedded clauses). 73

$\bigcup$ Union operator. 64
UPDATE Stack operator updating the topmost element. 84
$\mathcal{U}, \mathcal{V}, \mathcal{S}, \mathcal{T}, ...$ A context update function. 71
$u, v, ...$ Discourse referents, and constants of type r in DInq. 67

W The set of possible worlds. 64
$w, ...$ A possible world. 65
which+sg/pl Interrogative word *which* followed by a singular / plural noun. 18
$\boldsymbol{who}_G$ Generic simplex wh-word. 25
$\boldsymbol{who}_S$ Specific simplex wh-word. 25

$x, y, ...$ An individual. 65

Résumé

Aspects dynamiques et présuppositionnels des questions

Les questions sont une composante essentielle d'un dialogue. Contrairement aux assertions, elles permettent d'exprimer une requête d'information. Cette thèse de doctorat s'intéresse aux effets des questions partielles sur le dialogue, et à la manière dont ils s'articulent avec l'information demandée dans la question.

Les deux propriétés dynamiques étudiées sont les présuppositions et les anaphores. Les présuppositions des questions sont difficiles à étudier car on ne peut pas recourir à des tests de projection habituels à cause de la distribution limitée des interrogatives. Les présuppositions d'existence et d'unicité sont bien comprises pour certains mots interrogatifs dans les cas simples, par exemple avec *quel* au singulier. Cependant, d'autres mots interrogatifs ou d'autres environnements sémantiques produisent des inférences plus complexes qui ont reçu peu d'attention.

Les mots interrogatifs sont référentiels. Très peu d'études se sont intéressées aux anaphores incluant un mot interrogatif. La plupart des corpus annotés en coréférence ne considèrent pas qu'un mot interrogatif peut être antécédent d'une anaphore. Pourtant, les propriétés référentielles des mots interrogatifs sont liées aux présuppositions qui en sont générées.

L'objectif de cette thèse est donc de mener des recherches approfondies sur ces deux propriétés dynamiques dans le but de comprendre leur lien et ainsi mieux modéliser les inférences complexes des questions.

Introduction (Chapitre 1) L'introduction présente les concepts clés qui structurent l'ensemble de ce manuscrit. Ce doctorat porte sur les effets des questions

partielles sur le discours. En plus de la requête d'information, un énoncé questionnant peut poser des contraintes sur le savoir partagé énonciatif ou fournir du contenu référentiel disponible dans la suite du discours. Ces deux aspects dynamiques sont respectivement appelés présuppositions et anaphores. La problématique de cette thèse est de comprendre l'articulation de ces deux phénomènes, entre eux et en interaction avec un environnement modal. Ce chapitre introductif donne un aperçu des nouvelles données empiriques rassemblées dans cette thèse, des analyses envisagées et des résultats.

Chapitre 2 Ce chapitre s'attaque à l'inférence existentielle des syntagmes interrogatifs monomorphémiques (notamment *qui* et *quoi*). Par exemple, la question "*Qui a triché ?*" engendre spontanément l'inférence "*Quelqu'un a triché*". Dans la littérature scientifique, deux principales positions s'affrontent. Certain·e·s défendent qu'il s'agit d'une présupposition alors que d'autres optent pour une implicature, insistant sur le fait que la réplique "*Personne n'a triché*" est une réponse à part entière. Pour éclairer ce débat, différents contextes sémantiques et discursifs sont étudiés pour déterminer leur influence sur la présence et la force de cette inférence. Les données suggèrent que cette inférence d'existentialité est variable et instable, indiquant qu'une analyse plus complexe est nécessaire.

La théorie défendue est la suivante : les mots interrogatifs monomorphémiques sont ambigus sémantiquement entre une lecture spécifique et une lecture générique. L'interprétation spécifique déclenche une présupposition d'existentialité alors que l'interprétation générique ne génère pas de telle inférence.

Une question comme (25-b) est ambigüe. Dans la SITUATION A, elle est spécifique, dans la SITUATION B, elle est générique. En français, il existe des constructions qui désambigüisent cette phrase. Le clivage en *c'est* (25-d) n'est possible qu'avec la SITUATION A alors que le clivage en *y'a* (25-f) n'est compatible qu'avec la SITUATION B.

(25) a. CONTEXTE : *Gaby est allé dans la cave il y a deux jours. Il n'y avait pas de chaussette par terre ni d'odeur bizarre. Ce matin, il est retourné dans la cave. Puis, il pose cette question au reste de la famille :*
b. GABY: Qui est allé dans la cave hier ?
c. SITUATION A : *Gaby a trouvé une chaussette sur le sol de la cave ce matin. Il aimerait savoir à qui elle appartient.*
d. GABY: C'est qui qu'est allé dans la cave hier ? (spécifique)
e. SITUATION B : *Ce matin, la cave sentait bizarre. Gaby suspecte qu'il s'agit d'une fuite de gaz. Il aimerait savoir quand est-ce qu'elle a débuté. Pour cela, il a besoin de demander aux personnes qui sont allées dans la cave hier, s'il y en a, si elles y avaient déjà senti une odeur bizarre.*
f. GABY: Y'a qui qu'est allé dans la cave hier ? (générique)

Ces deux lectures ont des propriétés différentes en terme de focus, coréférence et domaine de quantification. De plus, elles ont des équivalents parmi d'autres éléments non interrogatifs. Le *qui* spécifique se rapproche du déterminant épistémique d'identification *certain* et de la lecture définie de *qui* introducteur de relatives libres. À l'inverse, le *qui* générique se rapproche de l'adjectif épistémique de non identification *quelconque* et de la lecture de choix libre de *qui* introducteur de relatives libres. Enfin, cette théorie permet d'expliquer des données nouvelles concernant les ilots faibles, les différences entre types de questions et les items à polarité négative (NPI) dans les questions. En étendant la dualité spécifique / générique des syntagmes nominaux non interrogatifs aux syntagmes interrogatifs, elle offre un cadre explicatif plus adéquat que les analyses précédentes.

Chapitre 3 La deuxième partie de ce manuscrit s'ouvre avec un chapitre introduisant la Sémantique Inquisitrice Dynamique (DInq) (Roelofsen et Dotlačil 2023b). Ce modèle formel est la base de la formalisation développée dans les chapitres suivants.

La Sémantique Inquisitrice (Inq) (Ciardelli, Groenendijk et Roelofsen 2018) est une théorie qui représente de manière uniforme la dénotation des assertions et des questions : par un ensemble d'états informationnels qui est non vide et clos par le bas, appelé *problématique* (en anglais : *issue*). Un état informationnel s (type $s \to t$) résout une problématique $\mathcal{I}$ (type $s \to (s \to t)$) si s appartient à $\mathcal{I}$. Intuitivement, cela signifie que s est une réponse possible à $\mathcal{I}$.

La Sémantique Inquisitrice Dynamique étend la Sémantique Inquisitrice de plusieurs manières. Premièrement, les états informationnels comportent des fonctions d'assignation permettant de garder une trace des référents du discours. Deuxièmement, les énoncés sont interprétés comme des fonctions d'un contexte (ensemble d'états informationnels) vers un autre contexte. Ces fonctions peuvent être partielles, modélisant une présupposition le cas échéant. Enfin, la théorie compositionnelle proposée décrit plus finement la périphérie gauche, permettant de décomposer les effets complexes des questions en une interaction d'opérateurs modulaires.

Chapitre 4 Ce chapitre étudie les questions comportant une modalité épistémique à grande portée, ex. "*Quelle lettre peut être ajoutée à P_RT pour former un mot français ?*". Ces questions subissent un affaiblissement de leur présupposition d'unicité et de leur exhaustivité (Hirsch et Schwarz 2019). De plus, comme pour les déclaratives, la référence au mot interrogatif est bloquée, sauf via subordination modale. L'objectif est de modéliser ces phénomènes.

La Sémantique Inquisitrice Dynamique Modale (MDInq) développée ici étend la Sémantique Inquisitrice Dynamique. Elle définit un opérateur modal existentiel externellement statique basé sur la sémantique modale Kratzerienne. Cet opérateur projette les alternatives en les élargissant, transformant ainsi des ques-

tions exhaustives (en anglais : *mention-all*) en questions non exhaustives (en anglais : *mention-some*). Les énoncés sont interprétés par des fonctions d'une pile de contextes vers une autre pile (S. Kaufmann 2000). Le bas d'une pile représente le savoir partagé énonciatif et les éléments plus hauts des contextes locaux, seulement accessibles par subordination modale. Grâce à ce modèle, les questions modales et conditionnelles reçoivent une représentation uniforme.

Chapitre 5 Ce chapitre plonge plus en profondeur dans l'analyse des propriétés référentielles des mots interrogatifs. Il montre qu'il existe trois types d'anaphores à un mot interrogatif. Premièrement, les réponses à une question offrent un cadre privilégié pour la coréférence et l'anaphore associative. En s'appuyant sur une étude du corpus ANCOR (Muzerelle et al. 2014), il est montré que la structure de ces relations anaphoriques est similaire à celle d'autres quantificateurs généralisés. Par exemple, pour qu'un item de choix libre puisse co-indexer un syntagme interrogatif, comme en (26), il faut que le référent u dénote un ensemble d'individus (type $e \to t$), et pas juste un individu (même pluriel) (type e).

(26) a. A: Dans [quel journal]^u faites-vous des mots croisés ?
b. B: [N'importe lequel]$_u$

Les questions spécifiques autorisent la coréférence par un pronom dans une assertion suivante à l'indicatif, ex. (27). À l'inverse, les questions génériques requièrent la subordination modale, comme illustré en (28).

(27) A: Qui^u été assis là ? Elle$_u$ a oublié son sac. (spécifique)

(28) a. A: Qui^u sait comment modéliser le choix libre dans les questions ? (générique)
b. B: Je ne sais pas. Mais [une telle personne]$_u$ (si elle$_u$ existe) doit travailler à Amsterdam.
c. B: #Je ne sais pas. Mais [une telle personne]$_u$ (si elle$_u$ existe) travaille à Amsterdam.

Je modélise ces phénomènes en Sémantique Inquisitrice Dynamique Modale (MDInq). Je fais l'hypothèse que les questions génériques contiennent un opérateur modal silencieux G qui induit la nécessité de subordination modale. De plus, la lecture générique d'un mot interrogatif contient dans son domaine l'individu nul $\star$, correspondant à l'ensemble vide. Cet individu génère la réponse négative, ex. "*Personne n'a triché*".

La coréférence spécifique est modélisée par l'existence d'un concept individuel (type $s \to e$) déterminé par un autre référent du discours, typiquement l'énonciateur. Cette coréférence est soumise à une condition de cohérence du discours qui, du fait de l'acte illocutoire questionnant, est plus restrictive que pour une assertion.

Chapitre 6 Ce chapitre est une étude de cas qui applique le modèle MDInq à une construction du français appelée conditionnelle à interrogative, ex. (29-a) (Richard 2024a). Ce tour syntaxique exprime une conjonction de dépendances conditionnelles implicites, tout comme son équivalent anglais (29-b).

(29) a. Selon [$_{\text{protase}}$ comment vous vous positionnez], [$_{\text{apodose}}$ vous n'aurez pas tous la même perception].
b. '*Depending on how you position yourself, you won't all have the same perception.*'

Deux problèmes à propos de ces propositions de dépendance sont identifiés. Premièrement, il est possible d'introduire un référent du discours dans la protase et d'y référer dans l'apodose ou dans une conditionnelle exemplifiant la dépendance fonctionnelle. Ce type de subordination modale est pris en charge par MDInq.

Le second problème concerne le décalage entre les contraintes syntaxiques et sémantiques. L'entrée sémantique de *depend on* proposée par Theiler, Roelofsen et Aloni 2019 présume que son sujet (correspondant à l'apodose de la conditionnelle) est interprété par une proposition inquisitrice. Or, seule des déclaratives sont possibles à cet emplacement. La solution consiste à attacher la circonstancielle bas sur l'arbre de la forme logique pour capturer les alternatives levées par le focus. Un attachement plus en amont autorise, au contraire, les questions de dépendance conditionnelle, ex. (30).

(30) Selon les âges, combien d'œufs peut-on consommer ?

Conclusion (Chapitre 7) Ce chapitre résume les points essentiels abordés dans cette thèse. Il conclut en abordant quelques limites du travail présenté. La théorie de l'ambigüité spécifique / générique constitue un premier pas vers une description plus fine des mots interrogatifs. Mais plus de recherche est nécessaire pour en comprendre toutes les subtilités. Plus précisément, il existe plusieurs types de spécificité et de généricité. Identifier lesquelles sont présents et comment ils se structurent serait un approfondissement intéressant. Un piste prometteuse constituerait à étendre la batterie de tests linguistiques proposée par Dayal to appear (initialement conçue pour les déterminants définis et indéfinis) aux mots interrogatifs.

Enfin, des pistes de réflexions pour remédier à ces lacunes et porter l'étude de la sémantique des mots interrogatifs plus loin sont proposées. Une possible suite au Chapitre 5 serait l'étude des relations anaphoriques dont l'antécédent est un mot QU issu d'une interrogative enchâssée. Comprendre la projection des référents du discours introduits par les mots interrogatifs permettrait d'appréhender plus clairement le lien entre lecture spécifique et portée. Une extension de la Sémantique Inquisitrice Dynamique Modale aux verbes d'attitude pourrait alors

être envisagée. Le deuxième point qui mérite plus d'attention est la portée des réponses courtes aux questions génériques. Tout comme SPECTOR 2007 observe pour les réponses quantifiées, les réponses négatives ne peuvent pas avoir de portée basse après une question contenant un ilot faible. Modéliser cette restriction demanderait de modéliser précisément les contraintes de portée des mots interrogatifs génériques. Ceci requiert donc de plus amples investigations. De manière plus générale, une comparaison détaillée entre la quantification d'ordre supérieur et la lecture générique semble être le sujet le plus prometteur ouvert par ce doctorat.

Annexes A et B Ces annexes présentent une version étendue de l'article RICHARD 2023. L'étude des interrogatives demande parfois un aperçu de la distribution en usage ou une recherche d'attestations ayant une structure spécifique. Or, l'obtention de grands corpus annotés d'interrogatives du français est une tâche complexe. D'une part, il existe en français un grand nombre des variantes dans la formation des interrogatives. D'autre part, de nombreuses structures syntaxiques sont similaires en surface (comme les conditionnelles ou les exclamatives) et compliquent l'identification des interrogatives, surtout enchâssées. L'objectif de ce travail est de combler cette lacune en développant un programme, nommé FUDIA, qui extrait et annote automatiquement les propositions interrogatives à partir de phrases préalablement annotées en dépendances universelles (UD). FUDIA est constitué de règles de réécriture de graphe. Grâce aux informations syntaxiques déjà présentes, FUDIA peut recourir à des règles plus complexes que les heuristiques simples souvent utilisées dans beaucoup d'articles en linguistique (présence de point d'interrogation, de *est-ce que*, etc.). En appliquant le programme sur les corpus UD francophones, le French Interrogative Bank ainsi créé étend le French Question Bank (SEDDAH et CANDITO 2016). Il contient des interrogatives enchâssées mais aussi des constructions plus diverses et moins normatives.

Annexe C Cette annexe rassemble les dialogues du corpus d'étude présenté dans le Chapitre 5.

Samenvatting

Dit proefschrift bestudeert de presupposities en anaforische eigenschappen van vraagwoorden. Deze fenomenen dragen beide bij aan de semantische en pragmatische effecten van vragen in een tekst of dialoog. Presupposities kunnen beperkingen opleggen aan de common ground. Vraagwoorden kunnen zogenaamde discoursereferenten oproepen waar anaforische uitdrukkingen naar kunnen verwijzen in volgende uitingen. Bovendien zijn deze twee fenomenen aan elkaar gerelateerd. De introductie van een naar een individu verwijzende referent is gecorreleerd aan de presuppositie dat het individu bestaat.

De hier voorgestelde theorieën bieden een nieuw perspectief op vragen. Ze pleiten voor een meer verfijnde semantiek van vraagwoorden en benadrukken de rol van modaliteit bij het vormen van sommige lezingen. In tegenstelling tot eerdere benaderingen worden de bestudeerde verschijnselen verklaard door analyses uit te breiden die oorspronkelijk waren ontworpen om determinatoren en voornaamwoorden te behandelen. In het bijzonder gaan de hoofdstukken van dit proefschrift in op de volgende onderwerpen: de dualiteit van specifieke en generieke vragen, Kratzeriaanse semantiek voor modale uitdrukkingen, en modale ondergeschiktheid. De extra expressiviteit van het huidige voorstel wordt gebruikt om de interactie tussen vragen en verschillende semantische omgevingen—zoals omgevingen gevormd door modale operatoren, focus omgevingen, zwakke eilanden, en omgevingen met negatieve polariteit—vast te leggen.

De genoemde uitbreiding van dynamische theorieën van voornaamwoorden en determinatoren naar die van vraagwoorden is mogelijk dankzij een bestaande uniforme semantiek voor vragen en beweringen: Dynamic Inquisitive Semantics (Roelofsen en Dotlačil 2023b). Dit raamwerk interpreteert beide uitingstypen als functies van context naar context. Presupposities worden geïmplementeerd als voorwaarden voor gedefinieerdheid. De contexten bevatten toewijzingsfuncties die discoursereferenten bijhouden; proposities kunnen op die manier referenten introduceren of hun waarden bijwerken. Dit proefschrift breidt Dynamic Inquisitive Semantics uit naar zinnen met modale werkwoorden en naar conditionele

zinnen. Door de genoemde functies te laten werken op objecten die complexer zijn dan de contexten van Dynamic Inquisitive Semantics (stapels lokale contexten in plaats van lokale contexten) kunnen subtielere discourseafhankelijkheden worden gecodeerd.

Abstract

This dissertation studies the presuppositions and anaphoric properties of interrogative words. These phenomena both contribute to the semantic and pragmatic effects of questions in a discourse. Presuppositions impose restrictions on the common ground. Interrogative words can raise discourse referents that anaphoric expressions can refer to in subsequent utterances. Moreover, these two phenomena are related to each other. Introducing a discourse referent interpreting an individual is correlated with the presupposition of its existence.

The theories proposed here offer a novel perspective on questions. They advocate for a more fine-grained semantics of interrogative words, while emphasizing the role of modality in shaping certain readings. Contrary to previous approaches, the studied phenomena are explained by extending analyses originally designed for determiners and pronouns in assertions. First, the chapters develop a specific vs. generic duality for wh-words. Second, Kratzerian semantics for modals and modal subordination are extended to modalized questions. These two interpretative frameworks are used to capture the interactions between questions and various semantic environments, including modal operators, focus, weak islands, and negative polarity items (NPIs).

Formalizing these two ideas for question is possible thanks to a uniform semantics for declaratives and interrogatives: Dynamic Inquisitive Semantics (Roelofsen and Dotlačil 2023b). This semantic model interprets both clause types as functions from context to context. Presuppositions are implemented as definedness conditions. The contexts contain assignment functions that keep track of discourse referents. Propositions can thus introduce or update the values assigned to discourse referents. This dissertation extends Dynamic Inquisitive Semantics to modals and conditionals. By allowing functions to act on more complex objects (stacks of local contexts), more subtle discourse dependencies can be encoded.

Titles in the ILLC Dissertation Series:

ILLC DS-2021-01: **Yfke Dulek**
Delegated and Distributed Quantum Computation

ILLC DS-2021-02: **Elbert J. Booij**
The Things Before Us: On What it Is to Be an Object

ILLC DS-2021-03: **Seyyed Hadi Hashemi**
Modeling Users Interacting with Smart Devices

ILLC DS-2021-04: **Sophie Arnoult**
Adjunction in Hierarchical Phrase-Based Translation

ILLC DS-2021-05: **Cian Guilfoyle Chartier**
A Pragmatic Defense of Logical Pluralism

ILLC DS-2021-06: **Zoi Terzopoulou**
Collective Decisions with Incomplete Individual Opinions

ILLC DS-2021-07: **Anthia Solaki**
Logical Models for Bounded Reasoners

ILLC DS-2021-08: **Michael Sejr Schlichtkrull**
Incorporating Structure into Neural Models for Language Processing

ILLC DS-2021-09: **Taichi Uemura**
Abstract and Concrete Type Theories

ILLC DS-2021-10: **Levin Hornischer**
Dynamical Systems via Domains: Toward a Unified Foundation of Symbolic and Non-symbolic Computation

ILLC DS-2021-11: **Sirin Botan**
Strategyproof Social Choice for Restricted Domains

ILLC DS-2021-12: **Michael Cohen**
Dynamic Introspection

ILLC DS-2021-13: **Dazhu Li**
Formal Threads in the Social Fabric: Studies in the Logical Dynamics of Multi-Agent Interaction

ILLC DS-2021-14: **Álvaro Piedrafita**
On Span Programs and Quantum Algorithms

ILLC DS-2022-01: **Anna Bellomo**
Sums, Numbers and Infinity: Collections in Bolzano's Mathematics and Philosophy

ILLC DS-2022-02: **Jan Czajkowski**
Post-Quantum Security of Hash Functions

ILLC DS-2022-03: **Sonia Ramotowska**
Quantifying quantifier representations: Experimental studies, computational modeling, and individual differences

ILLC DS-2022-04: **Ruben Brokkelkamp**
How Close Does It Get?: From Near-Optimal Network Algorithms to Suboptimal Equilibrium Outcomes

ILLC DS-2022-05: **Lwenn Bussière-Carae**
No means No! Speech Acts in Conflict

ILLC DS-2022-06: **Emma Mojet**
Observing Disciplines: Data Practices In and Between Disciplines in the 19th and Early 20th Centuries

ILLC DS-2022-07: **Freek Gerrit Witteveen**
Quantum information theory and many-body physics

ILLC DS-2023-01: **Subhasree Patro**
Quantum Fine-Grained Complexity

ILLC DS-2023-02: **Arjan Cornelissen**
Quantum multivariate estimation and span program algorithms

ILLC DS-2023-03: **Robert Paßmann**
Logical Structure of Constructive Set Theories

ILLC DS-2023-04: **Samira Abnar**
Inductive Biases for Learning Natural Language

ILLC DS-2023-05: **Dean McHugh**
Causation and Modality: Models and Meanings

ILLC DS-2023-06: **Jialiang Yan**
Monotonicity in Intensional Contexts: Weakening and: Pragmatic Effects under Modals and Attitudes

ILLC DS-2023-07: **Yiyan Wang**
Collective Agency: From Philosophical and Logical Perspectives

ILLC DS-2023-08: **Lei Li**
Games, Boards and Play: A Logical Perspective

ILLC DS-2023-09: **Simon Rey**
Variations on Participatory Budgeting

ILLC DS-2023-10: **Mario Giulianelli**
Neural Models of Language Use: Studies of Language Comprehension and Production in Context

ILLC DS-2023-11: **Guillermo Menéndez Turata**
Cyclic Proof Systems for Modal Fixpoint Logics

ILLC DS-2023-12: **Ned J.H. Wontner**
Views From a Peak: Generalisations and Descriptive Set Theory

ILLC DS-2024-01: **Jan Rooduijn**
Fragments and Frame Classes: Towards a Uniform Proof Theory for Modal Fixed Point Logics

ILLC DS-2024-02: **Bas Cornelissen**
Measuring musics: Notes on modes, motifs, and melodies

ILLC DS-2024-03: **Nicola De Cao**
Entity Centric Neural Models for Natural Language Processing

ILLC DS-2024-04: **Ece Takmaz**
Visual and Linguistic Processes in Deep Neural Networks: A Cognitive Perspective

ILLC DS-2024-05: **Fatemeh Seifan**
Coalgebraic fixpoint logic Expressivity and completeness result

ILLC DS-2024-06: **Jana Sotáková**
Isogenies and Cryptography

ILLC DS-2024-07: **Marco Degano**
Indefinites and their values

ILLC DS-2024-08: **Philip Verduyn Lunel**
Quantum Position Verification: Loss-tolerant Protocols and Fundamental Limits

ILLC DS-2024-09: **Rene Allerstorfer**
Position-based Quantum Cryptography: From Theory towards Practice

ILLC DS-2024-10: **Willem Feijen**
Fast, Right, or Best? Algorithms for Practical Optimization Problems

ILLC DS-2024-11: **Daira Pinto Prieto**
Combining Uncertain Evidence: Logic and Complexity

ILLC DS-2024-12: **Yanlin Chen**
On Quantum Algorithms and Limitations for Convex Optimization and Lattice Problems

ILLC DS-2024-13: **Jaap Jumelet**
Finding Structure in Language Models

ILLC DS-2025-01: **Julian Chingoma**
On Proportionality in Complex Domains

ILLC DS-2025-02: **Dmitry Grinko**
Mixed Schur-Weyl duality in quantum information

ILLC DS-2025-03: **Rochelle Choenni**
Multilinguality and Multiculturalism: Towards more Effective and Inclusive Neural Language Models

ILLC DS-2025-04: **Aleksi Anttila**
Not Nothing: Nonemptiness in Team Semantics

ILLC DS-2025-05: **Niels M. P. Neumann**
Adaptive Quantum Computers: decoding and state preparation

ILLC DS-2025-06: **Alina Leidinger**
Towards Language Models that benefit us all: Studies on stereotypes, robustness, and values

ILLC DS-2025-07: **Zhi Zhang**
Advancing Vision and Language Models through Commonsense Knowledge, Efficient Adaptation and Transparency

ILLC DS-2025-08: **Sophie Klumper**
The Gap and the Gain: Improving the Approximate Mechanism Design Frontier in Constrained Environments

ILLC DS-2026-01: **Bryan Eikema**
A Sampling-Based Exploration of Neural Text Generation Models

ILLC DS-2026-02: **Marten Folkertsma**
Empowering Quantum Computation with: Measurements, Catalysts, and Guiding States

Questions are essential in social interactions. They allow us to ask for information. However, questions can also put pragmatic constraints on a dialogue.

This PhD dissertation studies the impact of questions on the discourse. In particular, it focuses on the presuppositions and anaphoric properties of interrogative words.

The theories proposed here offer a novel perspective on questions. They advocate for a more fine-grained semantics of interrogative words, while emphasizing the role of modality in shaping certain readings. Contrary to previous approaches, the studied phenomena are explained by extending analyses originally designed for determiners and pronouns in assertions.

First, the chapters develop a specific vs. generic duality for wh-words. Second, Kratzerian semantics for modals and modal subordination are extended to modalized questions. These two interpretative frameworks are used to capture the interactions between questions and various semantic environments, including modal operators, focus, weak islands, and negative polarity items (NPIs).

Universiteit van Amsterdam

Université de Lorraine