



University of
Massachusetts
Amherst

Probes and their Horizons

Item Type	dissertation
Authors	Keine, Stefan
DOI	10.7275/8969235.0
Download date	2025-04-29 14:34:05
Link to Item	https://hdl.handle.net/20.500.14394/19968

PROBES AND THEIR HORIZONS

A Dissertation Presented

by

STEFAN KEINE

Submitted to the Graduate School of the
University of Massachusetts Amherst in partial fulfillment
of the requirements for the degree of

DOCTOR OF PHILOSOPHY

September 2016

Linguistics

© Copyright by Stefan Keine 2016
All Rights Reserved

PROBES AND THEIR HORIZONS

A Dissertation Presented

by

Stefan Keine

Approved as to style and content by:

Rajesh Bhatt, Co-Chair

Kyle Johnson, Co-Chair

Lyn Frazier, Member

Brian Dillon, Member

Charles Clifton, Jr., Member

John Kingston, Department Head
Department of Linguistics

ACKNOWLEDGMENTS

This dissertation represents the end of five stimulating and happy years of graduate school. Like most dissertations, it would not exist without the people who I have been fortunate to share these years with. My foremost thanks goes to my committee: Rajesh Bhatt, Kyle Johnson, Lyn Frazier, Brian Dillon, and Chuck Clifton.

It is by no means an exaggeration that none of this thesis would exist without Rajesh Bhatt. After all, it was my interest in long-distance agreement in Hindi which, over time, after many dead ends, demolitions, and rebuilding has grown into this dissertation. Rajesh has been, and continues to be, an amazing advisor, collaborator, and friend in all things linguistic and non-linguistic. His enthusiasm for virtually everything is contagious. I thank him for his unwavering support and encouragement throughout these five years, for his eagerness to discuss every half-baked idea that I was throwing around in my head (and understanding it better than I did!), for his insistence on keeping me honest and not glossing over the details, and for his constant willingness to share not just his knowledge of Hindi, but also myriads of judgments. I have learned a lot from him, not just about linguistics but also about how to be a linguist, and I am extremely grateful for that.

Kyle Johnson has been a fantastic mentor, teacher, advisor, and friend. I am glad he agreed to be my co-chair. Kyle has a wonderful knack for immediately understanding everything that I am clumsily trying to say, detecting all the problems with it, and challenging me with them or offering solutions. I have discovered and learned a lot in the process. Many parts of this dissertation would simply not exist without Kyle. Kyle has also never failed to keep my eyes on the big questions and not to get lost in the details. I believe there has not been a single meeting with him that I left without the feeling of having made progress. Plus, he is incredibly charming and witty, so I got lots of laughs on top of that.

Working with Lyn Frazier has been a fantastic experience. Lyn and her work are one of the main reasons why I became interested in psycholinguistics in the first place. I thank her for taking me under her wing so gladly when I first approached her with plans to pursue experimental work. From there, her support and guidance has been unwavering. Her conviction that evidence from processing can and should inform theoretical syntax has had a profound impact on me and my work. I thank Lyn for making me a more well-rounded linguist as a result. Lyn has always been enthusiastic about my work as a whole, building bridges to processing and acquisition, and constantly suggested new ways of thinking about particular that problems I was grappling with.

Brian Dillon is the other main reason that I became interested in processing and experimental evidence. He once taught a class with Kyle that both masterfully weaved together theoretical and processing questions and sparked my endeavors into processing, some of which made it into chapter 7 of this dissertation. I have learned an immense amount from him, not only about processing in general, but also about the nitty-gritty of running an experiment and doing stats. His generous guidance and constant enthusiasm was utterly indispensable in allowing me to venture into experimental territory and I am very grateful for that.

I am very happy that Chuck Clifton agreed to be my outside member. My meetings with Chuck were always both insightful and fun. His comments on my processing work have been consistently helpful and uplifting. They also never failed to raise new questions for me to explore, and help me to better understand or solve problems I was wrestling with. I have profited a lot from his feedback and expertise.

A very special thanks is due to all the people who have shared their grammaticality judgments with me. I am indebted to Rajesh Bhatt, Ayesha Kidwai, Sakshi Bhatia, and Jyoti Iyer for Hindi judgments, to Ilaria Frana for Italian, to Ekaterina Vostrikova for Russian, to Jon Ander Mendia for Spanish, to Saurov Syed for Bangla, and to many people for the English judgments (especially Lyn Frazier and Ethan Poole). You all have made this work possible! Needless to say, none of them should be held responsible for what I have made of their judgments.

I am also indebted to the members of the UMass Cognitive Science of Language lab for essential help and guidance in coding, experimental logistics, and analysis. In addition to Brian Dillon, Ethan Poole and Shayne Sloggett deserve particular thanks.

This dissertation would not have been possible without the pioneering and groundbreaking work of Klaus Abels, Gereon Müller, Wolfgang Sternefeld, and Edwin Williams. I would like to expressly acknowledge the profound impact that this body of work has had on my thinking about selective opacity. I was very fortunate to be able to discuss some of the ideas presented here with Klaus Abels, Gereon Müller, and Edwin Williams. These discussions have helped me towards a better understanding of the problems that we are after and have greatly improved this thesis.

The intellectual debt I owe to Gereon Müller is very hard to put into words. He was my first syntax teacher, my advisor throughout my undergrad days in Leipzig, and my advisor on a DFG-funded project I was a part of before making the move over the Pond. Anyone reading this dissertation will quickly realize that his influence on me has by no means stopped there. He has paved the way for me in how I think about linguistics, how I approach problems, and the kinds of questions I feel drawn to. Gereon also has always been extremely supportive of me and appreciative of my work, both during my time in Leipzig and after I left. This support has always helped to drive me forward.

While I was developing the ideas presented here, I had various opportunities to present and discuss them, and I have profited tremendously from these opportunities. Various stages of this work have been presented at GLOW 36, FASAL 3, CUNY 2015, WCCFL 33, the 2016 LSA Annual Meeting, the University of Connecticut, the University of Leipzig, UMass Amherst, the University of Southern California, McGill University, and UC Santa Cruz. More people than I can hope to remember have provided me with valuable and insightful feedback. I would like to acknowledge (in addition to the above) Jonathan Bobaljik, Željko Bošković, Sandy Chung, Jessica Coon, Veneeta Dayal, Janet Dean Fodor, Jakub Dotlačil, Donka Farkas, Doreen Georgi, Grant Goodall, Claire Halpert, Fabian Heck, Anke Himmelreich, Norbert Hornstein, Khalil Iskarous, Ayesha Kidway, Anoop Mahajan, Emily Manetta, Jim McCloskey, Barbara Partee, David Pesetsky, Ethan Poole, Omer Preminger, Martin Salzmann, Junko Shimoyama, Andrew Simpson, Shayne Sloggett, Adrian Staub, Lisa Travis, Coppe van Urk, Matt Wagers, Michael Wagner, Susi Wurmbrand, and Maria Luisa Zubizarreta. This list is likely to be incomplete, and I apologize to everyone I may have inadvertently forgotten!

I also thank all the faculty of the UMass Linguistics department for making the department what it is and teaching me what I know. I have been fortunate to be able to take classes at UMass

with Michael Becker, Rajesh Bhatt, Seth Cable, Brian Dillon, Lyn Frazier, Lisa Green, Iaria Frana, Kyle Johnson, Angelika Kratzer, Joe Pater, Tom Roeper, and Ellen Woolford. A big thank you to all of you! Special thanks go to Kathy Adamczyk, Michelle McBride, and Tom Maxfield for always being there to help me out with administrative pitfalls, for critical deadline reminders, and for chit-chats and laughs.

Many people have helped make the past five years so much more enjoyable. I would like to thank the members of my cohort Mike Clauss, Hannah Greene (who craftily managed to find me a beer I liked), J r my Pasquereau, and Shayne Sloggett. I would also like to thank Athulya Aravind, Michael Becker, Sakshi Bhatia, Elizabeth Bogal-Allbritten, Sakshi Bhatia, Thomas Borer, Thuy Bui, Tracy Conner, Elmira Elvazova, Minta Elsmann, Iaria Frana, Alex Goebel, Jake Grush, Mikee Guzman, Chris Hammerly, Lisa Hofmann, Ailia Hopkins, Hsin-Lun Huang, Rodica Ivan, Jyoti Iyer, Laura Kalin, Allie Kleber, Nick LaCara, Jon Ander Mendia, Claire Moore-Cantwell, Andy Murphy, Mythili Menon, Tian Hui Ng, Jason Overfelt, Deniz  zyildız, Yangsook Park, Amanda Rysling, Dar o S nchez Gonz lez, Anisa Schardl, Georgia Simon, Sandip Southehal, Walter Sistrunk, Megan Somerday, Ulli Steindl, Saurov Syed, Katia Vostrikova, Martin Walkow, Andrew Weir, Becky Woods, and Joanna Zaleska.

I thank Caf  Evolution for their award-winning vegan cupcakes.

My family deserves a special thank you. I know it has not always been easy for them being an ocean apart from me. They have always believed in me and supported me in my choices, and I am extremely grateful for that.

For Ethan Poole, there are no words to express my thanks.

ABSTRACT

PROBES AND THEIR HORIZONS

SEPTEMBER 2016

STEFAN KEINE

Magister Artium (M.A.), UNIVERSITÄT LEIPZIG

Ph.D., UNIVERSITY OF MASSACHUSETTS AMHERST

Directed by: Professor Rajesh Bhatt and Professor Kyle Johnson

This dissertation develops a comprehensive theory of *selective opacity*, syntactic configurations in which one and the same syntactic domain (typically a clause) is transparent to some operations, but opaque to others. The prime example of selective opacity are finite clauses in English, which are transparent to \bar{A} -movement, but opaque to A-movement. Following and extending the previous literature, this thesis argues that selective opacity is much more widespread than it is usually taken to be in that it extends beyond the A/\bar{A} -distinction and even to syntactic dependencies that do not involve movement.

From an empirical point of view, I argue that selective opacity exhibits intriguing meta-generalizations, which become evident once selective opacity across constructions and languages is treated as a uniform phenomenon. These two meta-generalizations are what I call *Upward Entailment* and the *Height–Locality Connection*. Upward Entailment states that if a clause of a given structural size is opaque to some operation, then structurally larger clauses are likewise opaque to this operation. The Height–Locality Connection states that the locality of a movement type is related to the height of the landing site of that movement type within the clausal spine in that

movement types that land in a structurally high position are able to escape more domains than movement types that target a structurally lower position.

The core theoretical proposal of the dissertation is that selective opacity is the manifestation of a constraint on the locality of probes. I propose that probes have characteristic *Horizons*, which delimit their search space. The crucial aspect of horizons is that they can differ between probes. As a result, the opacity of a domain can be relative to the probe conducting the search. I argue that this is what underlies locality differences between movement types, between movement and agreement, and between different types of agreement dependencies that do not involve movement. I demonstrate how a wide array of selective opacity effects and complex interactions between them can be derived from this account. I also demonstrate how meta-generalizations of selective opacity are derived in this framework.

Finally, I explore the consequences of horizons for more familiar concepts of syntactic locality like phases. I show that horizons coexist with CP phases, but that they are incompatible with *v*P phases. Independent experimental evidence for this conclusion is provided and I reassess previous arguments in support of *v*P phases.

TABLE OF CONTENTS

	Page
ACKNOWLEDGMENTS	iv
ABSTRACT	viii
LIST OF TABLES	xv
LIST OF FIGURES	xvi
LIST OF ABBREVIATIONS USED IN GLOSSES	xvii
 CHAPTER	
1 INTRODUCTION	1
1.1 Selective opacity: The phenomenon	1
1.2 The analytical challenge of selective opacity	13
1.3 Meta-generalizations of selective opacity	16
1.3.1 The Height–Locality Connection	17
1.3.2 Upward Entailment	21
1.4 Selective opacity beyond movement	27
1.5 Horizons: An overview of the proposal	29
1.6 Deriving the meta-generalizations	38
1.7 Consequences for the distribution of selective opacity	40
1.8 Consequences for phase theory	44
1.9 Overview of this dissertation	45
2 SELECTIVE OPACITY BEYOND MOVEMENT: A HINDI CASE STUDY	49
2.1 Introduction	49
2.2 The locality profile of finite vs. nonfinite clauses	52
2.2.1 A- and \bar{A} -movement	54
2.2.1.1 Some background on Hindi movement	54
2.2.1.2 Diagnosing A-movement	57
2.2.1.3 Movement out of finite clauses	61
2.2.1.4 Movement out of nonfinite clauses	63

2.2.2	Local and long-distance ϕ -agreement	65
2.2.2.1	Some background on local agreement	66
2.2.2.2	Long-distance agreement (LDA)	68
2.2.2.3	Long-distance agreement is long-distance	72
2.2.2.4	ϕ -Agreement and finite clauses	76
2.2.3	Section summary	76
2.3	The structure of finite and nonfinite clauses	77
2.3.1	The size of finite vs. nonfinite clauses	78
2.3.2	Two types of nonfinite clauses	87
2.3.3	The A-movement–agreement connection	93
2.3.3.1	Weak crossover	93
2.3.3.2	Reciprocal binding	102
2.3.3.3	Quantifier scope	104
2.3.3.4	Section summary	106
2.3.4	Implications for selective opacity	107
2.4	<i>Wh</i> -licensing	111
2.4.1	The absence of <i>wh</i> -movement: Evidence from focus intervention	111
2.4.2	The locality of <i>wh</i> -licensing	118
2.4.3	Detour 1: More <i>wh</i> -constructions in Hindi	122
2.4.4	Detour 2: Structure vs. positions – The case of Bangla	130
2.4.5	Section summary	137
2.5	Regularities of selective opacity	138
2.5.1	Upward entailment	139
2.5.2	The Height–Locality Connection	140
2.5.2.1	The landing site of A-movement	141
2.5.2.2	The landing site of \bar{A} -movement	144
2.5.2.3	The location of the ϕ -probe	148
2.5.2.4	The location of the <i>wh</i> -licensor	154
2.5.2.5	Consequences for the Height–Locality Connection	154
2.6	Chapter summary	156
3	HORIZONS	159
3.1	Introduction	159
3.2	The puzzle for standard principles	161
3.2.1	Phases	161
3.2.2	A-over-A Principle	165
3.2.3	Case and the Activity Condition	169
3.2.4	Section summary	170
3.3	The proposal	171

3.3.1	Horizons	171
3.3.2	Deriving Upward Entailment	178
3.3.2.1	The puzzle	179
3.3.2.2	Proposal: Category inheritance	181
3.3.2.3	The Horizon Inheritance Theorem	188
3.4	Accounting for selective opacity	191
3.4.1	Selective opacity in Hindi	191
3.4.1.1	Finite (CP) clauses	193
3.4.1.2	Large nonfinite (TP) clauses	199
3.4.1.3	Small nonfinite (vP) clauses	199
3.4.1.4	Summary	206
3.4.2	The syntax of superraising and other movement asymmetries	207
3.4.3	Accounting for movement–agreement mismatches	217
3.5	Constraining selective opacity: The Height–Locality Connection	220
3.5.1	Deriving the Height–Locality Connection	220
3.5.2	Variation in horizons	232
3.5.3	Topic islands in English: A problem for the Height–Locality Theorem?	237
3.5.4	Default-horizons	241
3.6	Chapter summary	245
4	SELECTIVE OPACITY IN GERMAN	248
4.1	Introduction	248
4.2	Clause types and their locality profiles	251
4.2.1	Coherent and non-coherent infinitives	252
4.2.1.1	Scrambling	253
4.2.1.2	Relativization	257
4.2.1.3	<i>Wh</i> -movement into V-final clause	259
4.2.1.4	Movement into V2 clause	259
4.2.1.5	The structure of infinitival clauses	261
4.2.2	V-final finite clauses	265
4.2.2.1	Scrambling	265
4.2.2.2	Relativization	265
4.2.2.3	<i>Wh</i> -movement into V-final clause	266
4.2.2.4	Movement into V2 clause	266
4.2.2.5	The structure of V-final clauses	267
4.2.3	V2 clauses	267
4.2.3.1	Scrambling	268
4.2.3.2	Relativization	268
4.2.3.3	<i>Wh</i> -movement into V-final clause	269

4.2.3.4	Movement into V2 clause	270
4.2.3.5	The structure of V2 clauses	271
4.2.4	Interim summary	275
4.3	Regularities of selective opacity	276
4.3.1	Opacity layering	276
4.3.2	Upward Entailment	277
4.3.3	The Height–Locality Connection	278
4.3.3.1	The landing site of scrambling	279
4.3.3.2	The landing site of relativization	279
4.3.3.3	The landing site of <i>wh</i> -movement in V-final clauses	281
4.3.3.4	The landing site of topicalization and <i>wh</i> -movement in V2 clauses	282
4.3.3.5	Section summary	282
4.4	A horizon-based account	284
4.4.1	The probes	284
4.4.2	Probe locality	286
4.4.3	Item-based vs. domain-based approaches	293
4.5	The role of phases	305
4.6	Constraining selective opacity: The Height–Locality Connection	309
4.7	Detour: Topic islands and <i>wh</i> -islands	313
4.8	Chapter summary	316
5	A COMPARISON OF HORIZONS TO PREVIOUS ACCOUNTS	318
5.1	Introduction	318
5.2	Case-based approaches	319
5.3	Ban on improper movement	330
5.4	Unambiguous binding	333
5.5	Operational ordering	339
5.6	Buffers	347
5.7	Derivational clock	354
5.8	Previous approaches to the Height–Locality Connection	356
5.8.1	Reducing locality to height (Williams 2003, 2011, 2013)	358
5.8.2	Reducing height to locality (Abels 2012a)	360
5.9	Chapter summary	365
6	THE ROLE AND DISTRIBUTION OF PHASES	367
6.1	Introduction	367
6.2	The necessity of CP phases	369
6.2.1	The successive cyclicity of movement	369
6.2.2	Evidence from selective opacity for CP phases: Revisiting German V2 clauses	379
6.2.3	Evidence from ϕ -agreement for CP phases	383

6.2.4	Section summary	386
6.3	Against <i>vP</i> phases	387
6.3.1	Selective opacity as an argument against <i>vP</i> phases	387
6.3.2	Nonlocal agreement as an argument against <i>vP</i> phases	398
6.4	A reassessment of previous arguments for <i>vP</i> phases	404
6.4.1	Reconstruction	406
6.4.2	Successive cyclicity in Dinka	408
6.4.3	<i>meN</i> -Deletion in Indonesian	412
6.4.4	Section summary	414
6.5	Chapter summary	415
7	PROCESSING EVIDENCE FOR THE DISTRIBUTION OF PHASES	417
7.1	Introduction	417
7.2	Successive cyclicity and a bias problem	418
7.3	Successive cyclicity in online sentence parsing	421
7.3.1	Previous results	421
7.3.2	Limitations	424
7.4	Experiment 1	427
7.4.1	Method	428
7.4.2	Results	432
7.4.3	Discussion	437
7.5	Experiment 2	441
7.5.1	Method	446
7.5.2	Results	450
7.5.3	Discussion	454
7.6	General discussion	459
7.6.1	The distribution of phases	459
7.6.2	Implications for theories of parsing	461
7.6.3	Limitations of the experiments	464
7.7	Chapter summary	467
APPENDICES		
A.	Plausibility norming	468
B.	Statistical analysis of non-crucial regions	470
C.	Materials used	475
BIBLIOGRAPHY		484

LIST OF TABLES

Table	Page
1.1 Examples of selective opacity	12
7.1 Regioning of stimuli in Experiment 1 ([±mv] = [±move])	431
7.2 Mean answer accuracy in Experiment 1	433
7.3 Mean raw reading times by condition and region in ms computed over participants for Experiment 1	434
7.4 Coefficient estimates and corresponding <i>t</i> -value for linear mixed effects model analyses of log reading times in critical regions of Experiment 1	436
7.5 Regioning of stimuli in Experiment 2 ([±mv] = [±move])	449
7.6 Mean answer accuracy by condition in Experiment 2	451
7.7 Mean raw reading times by condition and region in ms computed over participants for Experiment 2	452
7.8 Coefficient estimates and corresponding <i>t</i> -values for the linear mixed effects model analyses of log reading times in critical regions of Experiment 2	453
7.9 Mean plausibility rating (and by-participant standard errors) by condition for stimuli used in Experiment 1	469
7.10 Mean plausibility rating (and by-participant standard errors) by condition for stimuli used in Experiment 2	470
7.11 By-participant standard error of raw reading times in ms by condition and region for Experiment 1	471
7.12 Coefficient estimates and corresponding <i>t</i> -value for linear mixed effects model analyses of log reading times in Experiment 1 for all regions	472
7.13 By-participant standard error of raw reading times in ms by condition and region for Experiment 2	473
7.14 Coefficient estimates and corresponding <i>t</i> -values for the linear mixed effects model analyses of log reading times in Experiment 2 for all regions	474

LIST OF FIGURES

Figure	Page
7.1 Residual reading times by condition and region in Experiment 1 in ms with by-participant standard errors	435
7.2 Mean increase in residual reading times between [–move] and [+move] conditions and by-participant standard errors in region 6 of Experiment 1 by intervener and verb type	437
7.3 Residual reading times by condition and region in ms with by-participant standard errors	452
7.4 Mean increase in residual reading times between [–move] and [+move] conditions in gap and spillover regions and by-participant standard errors in Experiment 2 by intervener	454

ABBREVIATIONS USED IN GLOSSES

ABS	absolute
ACC	accusative
ANO	animate object
ASP	aspect
ASS	associative
AUG	augment
AUX	auxiliary
CF	classifier
CL	clitic
CL	clitic
COMP	complementizer
CONJ	conjunct inflection
DAT	dative
DIR	direct voice
DUB	dubitative
ERG	ergative
EVID	evidential
FOC	focus
FUT	future
FV	final vowel
F	feminine
GEN	genitive
IND	indicative
INF	infinitive
INSTR	instrumental
IPFV	imperfective
LOC	locative
M	masculine
NMLZ	nominalizer
NOM	nominative
NPI	negative polarity item

NS	nonsubject voice
N	neuter
OBJ	object
OBL	oblique
OBV	obviative third person
PASS	passive
PASTPART	past participle
PERS	persistive
PFV	perfective
PL	plural
PRES	present
PROG	progressive
PRTAGR	participle agreement
PRTCL	particle
PRT	participle
PST	past
REFL	reflexive
SA	subject agreement
SG	singular
SM	subject marker
SUBJ	subjunctive
TOP	topic
TRANS	transitive

CHAPTER 1

INTRODUCTION

1.1 Selective opacity: The phenomenon

The focus of this dissertation is a type of syntactic locality that has traditionally received little systematic attention in the literature. This type of locality involves syntactic domains that are transparent to some syntactic operations, but at the same time opaque to others. To have a handy terminology, I will refer to such configurations as **SELECTIVE OPACITY**. Selective opacity is defined in (1). The underlying intuition is that in these cases, a given syntactic domain (typically a clause) is not completely transparent to all operations or completely opaque to all operations. Rather, its opacity is selective and crucially depends on the type of operation involved. In other words, in selective opacity configurations, the domain's opacity is relative to specific operations.

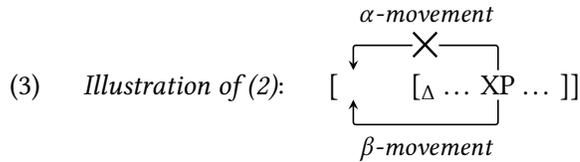
(1) **Selective Opacity**

A syntactic domain Δ is selectively opaque for an operation α if Δ is opaque to α , but transparent to some other operation β .

Historically, selective opacity effects have been investigated primarily in the domain of movement, i.e., in configurations where a given domain is transparent to some movement types, but not others.

(2) *Selective opacity in movement dependencies*

A syntactic domain Δ is selectively opaque for α -movement if Δ prohibits α -movement but allows β -movement out of it, where α and β are different types of movement.



The domain Δ in (3) is selectively opaque because it blocks α -movement out of it, but it allows β -movement out of it. The hallmark property of selective opacity configurations is thus that there is a **LOCALITY MISMATCH/DIFFERENCE** between different operations. Different operations have different locality profiles.

The notion of selective opacity is related to weak islands (see, e.g., Szabolcsi 2006 for an overview of the latter). But the term ‘weak island’ is usually applied to environments that allow argument extraction but not adjunct extraction (e.g. Lasnik & Saito 1992) or DP gaps but not PP gaps (e.g. Cinque 1990). In standard cases of weak islands, the opacity or transparency of a domain is modulated by either internal properties of the moving element (such as categorial or semantic status) or alternatively by the syntactic context of the launching site of movement (argument vs. adjunct). By contrast, I reserve the term ‘selective opacity’ to refer to instances where it is the type of the movement itself or, roughly equivalently, the landing site of the movement that affects barrierhood.

By far the most well-known instance of selective opacity is the ban on **SUPERRAISING** in languages like English, also known as *hyperraising* or *improper movement*, terms that I will use interchangeably here.¹

Finite clauses in English are transparent for \bar{A} -movement (4a), but at the same time they are opaque to A-movement (4b), a fact originally observed by Chomsky (1973, 1977).

¹ The two terms ‘superraising’ and ‘hyperraising’ are sometimes reserved for distinct constructions. Ura (1994) uses the terms ‘superraising’ to refer to movement of a subject past the subject of another clause into an A-position in a higher clause, as in (i.a). The term ‘hyperraising’ he reserves for movement of a subject of a finite clause to the subject position of a higher finite clause, as in (i.b).

(i) a. *John_i seems [that it was told t_i [that Mary is a genius]]

b. *They_i seem [that t_i like Mary]

(Ura 1994: 5,64)

On the account developed in this thesis, both structures in (i) are ruled out by the same principle and so I will not terminologically distinguish between them.

(4) *Movement out of finite clauses in English*

a. \bar{A} -movement:

Who_i do you think [t_i likes oatmeal]?

b. *A*-movement

*Sue_i seems [t_i likes oatmeal].

There is thus a locality mismatch between *A*- and \bar{A} -movement. Using the terminology just introduced, I will say that finite clauses are selectively opaque to *A*-movement, but not \bar{A} -movement.

The ungrammaticality of (4b) does not simply reduce to the fact that *Sue* receives case in both the embedded and the matrix clause. The ungrammaticality persists even if the source of case in either the launching site (5a,b) or the landing site (5c,d) is eliminated, as is well-known:

- (5) a. *John_i [seems that it is certain [t_i to like ice cream]] (Chomsky 1981: 58)
b. *John_i seems [that it was told t_i [that Mary is a genius]] (Lasnik & Saito 1992: 192)
c. *Mary's belief [John_i to be likely [t_i will win]] (Lasnik & Boeckx 2006: 118)
d. *It is certain [Rhoda_i to be likely [t_i is intelligent]]. (Nevins 2005: 292)

The locality contrast between *A*- and \bar{A} -movement in (4) has been investigated in great detail in the literature, and a variety of accounts have been explored. The standard account, due to Chomsky (1973, 1977, 1981) and May (1979), involves a conspiracy of two constraints: (i) a constraint like subjacency or phases that requires movement to be successive-cyclic and proceed through the edge of a finite clause, an \bar{A} -position, and (ii) a prohibition against moving from an \bar{A} - to an *A*-position (often referred to as the *Ban on Improper Movement*). A more recent approach that has been suggested by Chomsky (2001) claims that DPs whose case feature is valued become invisible to all subsequent *A*-processes. I will defer a discussion of these accounts until chapter 5.

A striking discovery of the previous literature on selective opacity effects is that locality mismatches are not limited to difference between *A*- and \bar{A} -movement, but that they may be considerably more fine-grained than that (see Williams 1974, 2003, 2011, 2013, Sternefeld 1992, Müller & Sternefeld 1993, Müller 1995, 2014a,b, Abels 2007, 2009, 2012a,b, Neeleman & van de Koot 2010). As an example, consider nonfinite clauses in English. As (6a) shows, they are transparent

to A-movement. Yet at the same time, they are opaque to extraposition in that they do not allow extraposition out of them (6b).²

(6) *Nonfinite clause in English*

a. *A-movement:*

John_i is believed [t_i to be certain [that Fred is crazy]] (by everyone).

b. *Extraposition:*

*[John_i is believed [t_i to be certain t_j] by everybody [that Fred is crazy]_j].

(Baltin 1978: 144)

There are thus at least two layers of selective opacity in English as \bar{A} -movement, A-movement, and extraposition all differ with respect to their locality properties. Standard accounts of improper movement like the ban on A-movement from an \bar{A} -position are narrowly confined to the A/ \bar{A} -distinction, and they hence do not extend to the locality of extraposition, which is therefore usually attributed to a separate constraint, Ross' (1967) *Right Roof Constraint*.

Because selective opacity has received comparatively little systematic attention outside of the ban on superraising, it is instructive to consider a number of illustrative examples from a variety of constructions and languages to showcase the diversity as well as the pervasiveness of selective opacity effects. Some of the examples to be given will be taken up again in greater detail in the course of this dissertation. For the time being, the following list primarily serves the purpose of illustrating the phenomenon and highlighting its scope and diversity.

Interestingly, selective opacity effects are not limited to interactions between A- and \bar{A} -dependencies, but also extend to locality mismatches between different types of \bar{A} -extractions. A second example of selective opacity are finite clauses in German, which illustrate this point. As shown in (7), finite V-final clauses allow *wh*-movement out of them, but curiously they are opaque to relativization (Bayer & Salzmann 2013, Müller 2014b) and scrambling (Bierwisch 1963, Ross 1967).

² Van Riemsdijk & Williams (1986: 30) explicitly relate the problem of the upward boundedness of extraposition to the problem of superraising as in both cases the challenge is to block a successive-cyclic derivation.

(7) *Finite V-final clauses in German*

a. *Wh-movement:*

Wen_i glaubst du [dass die Maria t_i liebt]
who believe you that the Maria loves
'Who do you believe Maria loves?'

b. *Relativization:*

*der Mann, den_i Fritz glaubt [dass Maria t_i mag]
the man who Fritz believes that Maria likes
Intended: 'the man who Fritz believe that Maria likes'

c. *Scrambling:*

*Ich glaube den Fritz_i [dass die Maria t_i liebt]
I believe the Fritz that the Maria loves
Intended: 'Fritz, I believe that Maria loves.'

A third example of selective opacity comes from V2 clauses in German. As Haider (1984) has first observed, embedded V2 clauses in German are transparent for *wh*-movement that lands in a higher V2 clause, but they disallow *wh*-movement out of them that lands inside a higher V-final clause.

(8) *V2 clauses in German*

a. *Wh-movement into V2 clause:*

[_{CP_{V2}} Wen_i meinst du [_{CP_{V2}} hat sie t_i getroffen]]?
who think you has she met
'Who do you think that she met?'

b. *Wh-movement into V-final clause:*

*(Ich weiß nicht) [_{CP_{V-final}} wen_i du meinst [_{CP_{V2}} hat sie t_i getroffen]]?
I know not who you think has she met
Intended: '(I don't know) who you think that she met.'

Both German examples are discussed in much greater detail in chapter 4.

A fourth example is Russian, where nonfinite clauses disallow A-movement out of them (see, e.g., Stepanov 2007), but are transparent to topicalization/scrambling. As (9b) and (10) illustrate, subject-to-subject raising out of a nonfinite clause is impossible.

(9) *Russian*

- a. Kažetsja [čto èti studenty znajut tri jazyka].
seem.3SG that these students know.3PL three languages
'It seems that these students know three languages.'
- b. *Èti studenty_i kažutsja [_{t_i} učit' tri jazyka].
these students seem.3PL learn.INF three languages
Intended: 'These students seem to be learning three languages.'

(Ekaterina Vostrikova, p.c.)

(10) *Russian*

- *Ivan_i sčitaetsja [_{t_i} byt' bol'nym]
Ivan is-considered to-be sick.INSTR
Intended: 'Ivan is considered to be sick.'

(Stepanov 2007: 84)

Moreover, the example in (11b) is derived from (11a) by scrambling/topicalization of *kitajskij jazyk* 'Chinese' out of the embedded clause and the result is grammatical. (11c) involves A-movement (*tough*-movement), diagnosed by agreement with the adjective, and is ungrammatical, in line with (9) and (10).

(11) *Russian*

- a. Trudno [učit' kitajskij jazyk].
difficult learn.INF Chinese language.M
'It is hard to learn Chinese.'
- b. Kitajskij jazyk_i trudno [učit' _{t_i}].
Chinese language.M difficult learn.INF
- c. *Kitajskij jazyk_i trudnyj [učit' _{t_i}].
Chinese language.M difficult.M.SG learn.INF
Intended: 'Chinese is hard to learn.'

(Ekaterina Vostrikova, p.c.)

Finite clauses in Russian exhibit the same asymmetry. See (33b) for an example of impossible A-movement out of a finite clause and (344) for grammatical scrambling out of a finite clause.

A fifth example of selective opacity is the well-known locality difference between clitic climbing and phrasal movement. In Spanish, for instance, it is possible to *wh*-move or topicalize out of a finite clause:

(12) *Spanish*

a. A quién_i piensa Juan [que María ha visto *t_i*]?
A who thinks Juan that María AUX.3SG see.PRT
'Who does Juan think that María saw?'

b. A Pedro piensa Juan [que María ha visto *t_i*].
A Pedro thinks Juan that María AUX.3SG see.PRT
'Pedro, Juan thinks that María saw.'

(Jon Ander Mendia, p.c.)

Moreover, as is well-known, it is possible for a clitic cross-referencing an object to appear on a higher verb if the embedded clause is nonfinite and the higher verb is a restructuring verb ('clitic climbing', see Aissen & Perlmutter 1976 for discussion).

(13) *Spanish*

a. Juan quiere ver le a Pedro.
Juan want.3SG see CL.DAT.3SG A Pedro

b. Juan le quiere ver a Pedro.
Juan CL.DAT.3SG want.3.SG see.INF A Pedro
'Juan wants to see Pedro.'

(Jon Ander Mendia, p.c.)

However, clitic climbing is not possible out of finite clauses, despite the fact in (12) that these clauses allow *wh*-extraction and topicalization out of them:

(14) *Spanish*

a. Juan piensa que María le ha visto a Pedro.
Juan thinks that María CL.DAT.3SG AUX.3SG see.PRT A Pedro

b. *Juan le piensa [que María ha visto a Pedro].
Juan CL.DAT.3SG thinks that María AUX.3SG see.PRT A Pedro
Intended: 'Juan thinks that María saw Pedro.'

(Jon Ander Mendia, p.c.)

A sixth illustrative instance of selective opacity comes from widely observed differences in the locality of scrambling and movement that target a left-peripheral position. As Wurmbrand (2015) discusses, Polish has three types of infinitival complementation: (i) infinitives in which the complementizer *żeby* is obligatory, (ii) infinitives in which *żeby* is impossible, and (iii) infinitives in which the complementizer is optional. Different verbs select for different types of infinitives. Crucially,

the presence or absence of the complementizer correlates with the possibility of scrambling out of the infinitive. If the complementizer is obligatory, crossclausal scrambling is blocked (15a). If the complementizer is impossible, such scrambling is licit (15b). Finally, if the complementizer is optional, crossclausal scrambling is possible only if the complementizer is absent (15c).

(15) *Polish scrambling*

- a. Jan { *pieniądze/*je } nalegał *(żeby) { pieniądze/je } zostawić.
 Jan { *money / *them } insisted *(so.that) { money / them } leave.INF
 ‘Jan insisted on leaving the money/them.’
- b. Jan { książkę } / { ja } zdołał (*żeby) { ja } przeczytać { książkę }.
 Jan { book } { it } managed (*so.that) { it } read.INF { book }
 ‘Jan managed to read a/the book.’
- c. i. Jan postanowił (żeby) { ja } przeczytać { książkę }.
 Jan decided (so.that) { it } read.INF { book }
 Jan decided to read a/the book/it.’
- ii. Jan książkę / ja postanowił (*żeby) przeczytać.
 Jan book / it decided (*so.that) read.INF
 ‘Jan decided to read a/the book/it.’ (Wurmbrand 2015: 229–230)

Crucially, the presence of *żeby* does not restrict the possibility of topicalization out of an embedded clause, as (16) demonstrates, where topicalization of *żabę* ‘frog’ out of the lower clause and over the complementizer is grammatical.

(16) *Polish topicalization*

- Żabę, to Jan chciałby żeby tylko Maria pocałowała.
 frog.ACC TOP Jan want.SUBJ that.SUBJ only Maria.NOM kissed
 ‘The frog, John would like only Mary to kiss.’ (Wurmbrand 2015: 230)

This pattern constitutes another instance of selective opacity because clauses containing *żeby* are opaque to scrambling, but transparent to topicalization. Analogous restrictions are observed in various other languages, including Slovenian, also reported in Wurmbrand (2015).

Kikuyu provides a seventh example of selective opacity in that topicalization and *wh*-movement differ in their locality properties. As discussed by Schwarz (2007), topicalization in the language is clause-bounded. From the baseline in (17a), clause-internal topicalization of the locative adjunct is

possible, as in (17b). Topicalization that leaves the embedded clause, by contrast, is impossible, as (17c) demonstrates.

(17) *Kikuyu topicalization*

- a. *abdul ne uy- ir- ε ate nyina ne- ɔɔn- ir- ε i-βuku mberε ya*
 Abdul FOC- say- ASP- FV that mother FOC- see- ASP- FV 5-book in-front 9.ASS
nyomba
 9.house
- b. *abdul ne uy- ir- ε ate mberε ya nyomba nyina ne- ɔɔn- ir-*
 Abdul FOC- say- ASP- FV that in-front 9.ASS 9.house mother FOC- see- ASP-
 ε i-βuku
 FV 5-book
- c. *#mberε ya nyomba abdul ne uy- ir- ε ate nyina ne- ɔɔn- ir-*
 in-front 9.ASS 9.house Abdul FOC- say- ASP- FV that mother FOC- see- ASP-
 ε i-βuku
 FV 5-book
 ‘Abdul said his mother saw the book in front of the house.’ (Schwarz 2007: 153)

In contrast to topicalization, *wh*-movement may cross clause boundaries:

(18) *Kikuyu wh-movement*

- ne- ko ngoye a- uy- ire ate kamau ne ɔ- ɔn- ire kanakε*
 FOC- where Ngoge SM- say- ASP that Kamau FOC SM see- ASP Kanake
 ‘Where did Ngoge say that Kamau saw Kanake?’ (Schwarz 2007: 154)

The eighth and final illustration of selective opacity is provided by Italian. As Abels (2012a) discusses, different types of \bar{A} -movement in Italian exhibit various locality mismatches. For example, fronting of an adverb (a process called ‘adverb preposing’ by Rizzi 2004 and ‘modifier fronting’ by Abels 2012a) may not cross a finite clause boundary, as (19a) shows (unless the adverb is a topic or focus, see fn. 3), while topicalization, focus fronting, and relativization are not similarly constrained, as illustrated in (19b) for topicalization.³

³ All else equal, the contrast between (19a) and (19b) could be taken not as a difference between movement types but as an argument–adjunct asymmetry. However, Rizzi (2004) shows that once a preposed adverb is mentioned in the preceding discourse and therefore a topic, it may leave a clause:

(19) *Italian*

a. *No crossclausal adverb fronting*

Rapidamente_i, (*Gianni dice che) hanno risolto il problema t_i
'Rapidly, (Gianni says that) they solved the problem.'

b. *Crossclausal topicalization*

Il problema, (Gianni dice che) lo hanno risolto rapidamente
'The problem, (Gianni says that) they solved it rapidly.' (Rizzi 2004: 249n10)

In addition, within the group of topicalization, focus movement, and relativization, additional locality mismatches can be observed. As Abels (2012a) discusses, *wh*-islands are transparent for relativization and topicalization out of them, but they are opaque to focus movement.⁴ The claim that *wh*-islands allow relativization out of them is illustrated in (20a). The transparency of *wh*-islands for topicalization is demonstrated in (20b).

(20) *Italian wh-islands*

a. *Relativization possible*

Tuo fratello, a cui mi domando che storie abbiano raccontato, era molto preoccupato.
'Your brother, to whom I wonder which stories they told, was very troubled.'
(Rizzi 1982: 50)

b. *Topicalization possible*

?A Gianni, non so come pensi che gli dovremmo parlare.
to Gianni I don't know how you think that to him we should talk
(Ilaria Frana, p.c.)

This contrasts with focus movement, which is impossible out of *wh*-island:

(21) *No focus movement out of wh-islands*

*QUESTO mi domando a chi hanno detto.
THIS I wonder to whom they have said (Abels 2012a: 241)

(i) a. C'è qualche problema che hanno risolto rapidamente?
'Is there a problem that they solved rapidly?'

b. Rapidamente, Gianni dice che hanno risolto il primo problema, ma non gli altri.
'Rapidly, Gianni says that they solved the first problem, but not the others.' (Rizzi 2004: 249n10)

⁴ Though see Sprouse, Caponigro, Greco & Cecchetto (2016) for recent discussion.

Interestingly, the opacity of clauses in which *wh*-movement has taken place for focus movement holds for all embedded elements, even the *wh*-element at the edge of the clause. This point is illustrated in (22), based on Rizzi (2006). In (22a), *quale ragazza* ‘which girl’ is *wh*-moved inside the lower clause and contrastively stressed, a well-formed structure. In (22b), this *wh*-element then undergoes further focus movement into the matrix clause and the result is ungrammatical. Rizzi (2006) concludes that a *wh*-moved element may not undergo further focus movement. He emphasizes that this is in all likelihood a syntactic constraint because there is evidently no general ban against focusing *wh*-elements (cf. (22)). Furthermore, because focus movement in Italian is not otherwise clausebounded, it is clear that the ungrammaticality of (22b) stems from the interaction with *wh*-movement.

- (22) a. Mi domandavo quale RAGAZZA avessero scelto, non quale ragazzo
 I wondered which GIRL they had chosen not which boy
- b. *Quale RAGAZZA mi domandavo ___ avessero scelto, non quale ragazzo
 which GIRL I wondered they had chosen not which boy
- (Rizzi 2006: 113)

The ungrammaticality of (22b) is arguably a manifestation of the same constraint that rules out (21) (this is, however, not Rizzi’s analysis). In both cases, a clause in which *wh*-movement has taken place is completely opaque to focus movement out of it, regardless of whether this focus movement applies to an element at the edge of this clause, as in (22b) or to an element more deeply embedded, as in (21).

This cursory inspection suggests that Italian exhibits several layers of selective opacity and specifically that locality mismatches are not limited to a binary distinction between A- and \bar{A} -movement.

The various examples just discussed are summarized in Table 1.1. What I hope this brief survey to have demonstrated is that locality mismatches between movement types and selective opacity is a widespread phenomenon that arises in a considerable range of constructions and languages.

Analytically, selective opacity effects usually receive treatments that are specifically tailored towards particular instances and are therefore construction-specific. For example, as noted above, the classical account of the prohibition against superraising in English involves (i) a requirement

Table 1.1. Examples of selective opacity

<i>Domain</i>	<i>transparent for ...</i>	<i>opaque for ...</i>
<i>English</i>		
nonfinite clauses	A-movement, \bar{A} -movement	extraposition
finite clauses	\bar{A} -movement	A-movement, extraposition
<i>German</i>		
V-final finite clauses	<i>wh</i> -movement	relativization, scrambling
V2 clauses	<i>wh</i> -movement into V2 clause	<i>wh</i> -movement into V-final clause
<i>Russian</i>		
nonfinite clauses	topicalization/scrambling	A-movement
<i>Spanish</i>		
finite clauses	topicalization	clitic climbing
<i>Polish</i>		
nonfinite <i>żeby</i> clauses	topicalization	scrambling
<i>Kikuyu</i>		
finite clauses	<i>wh</i> -movement	topicalization
<i>Italian</i>		
finite clauses	focus movement, topicalization, relativization	adverb preposing
<i>wh</i> -islands	relativization, topicalization	focus movement

for crossclausal movement to move through the edge of the clause, an \bar{A} -position and (ii) a ban on movement from an \bar{A} - to an A-position (May 1979, Chomsky 1981). It is clear from the various examples just discussed that such an account is narrowly restricted to superraising and therefore has nothing to say about the various other instances of selective opacity like the ones just reviewed. Similarly, the stricter locality of extraposition in English (recall (6b)) is standardly attributed to a designated constraint that restricts the locality of rightward movement but not leftward movement (Ross 1967). Accounts of this sort are essentially piecemeal as every instance of selective opacity is attributed to a designated and narrowly focused constraint whose domain of application does not extend beyond the construction it was devised for.

The vantage point I will adopt in this dissertation is that selective opacity is a *uniform* phenomenon in the sense that the various locality mismatches just mentioned are all manifestations of the same underlying phenomenon and hence of a central and fundamental aspect of syntactic locality. Specifically, I will argue for the view that syntactic domains may indeed be selectively

opaque in that they may be transparent to one operation, but opaque to others. In other words, I will argue that selective opacity effects are theoretically real and in fact central to the theory of locality. An immediate consequence of this view is that at least some locality restrictions are relativized to specific operations. That is, domains are not necessarily completely transparent to all operations or completely opaque to all operations. Rather, locality may be non-binary, with one and the same domain exhibiting transparent and opaque behavior. In the view that locality mismatches are real, I follow the lead of Williams (1974, 2003, 2011, 2013), Sternefeld (1992), Müller & Sternefeld (1993), Müller (1995, 2014a,b), Grewendorf (2003, 2015), Abels (2007, 2009, 2012a,b), and Neeleman & van de Koot (2010), among others, although the theoretical proposal I will explore differs significantly from all of these.

The recognition of selective opacity effects in their variety of guises has far-reaching theoretical consequences. The reason is that standard principles of syntactic locality have very little to say about selective opacity because they do not have the right structure, a point I will elaborate on in the next section. Selective opacity effects thus call for a novel view on locality. This thesis is an attempt to motivate and develop such a view in detail, to lay out its implications, and to explore its relationship to more standard locality principles.

1.2 The analytical challenge of selective opacity

Selective opacity is theoretically challenging for standard approaches to syntactic locality like subjacency (Chomsky 1973, 1977), barriers (Chomsky 1986) or phases (Chomsky 2000, 2001) for a very general reason. Despite considerable differences in the scope and theoretical underpinning behind these principles, they have in common that they are *strictly binary* in nature. They all specify, in one way or another, that certain syntactic domains are impenetrable from the outside and that, as a result, all syntactic operations across such domains are impossible. While there has been considerable debate on what the relevant domain boundaries are (e.g., the size and distribution of phase heads), the very nature of these constraints is fundamentally *unselective*: If a given domain is opaque to some operation due to subjacency or phasehood, it is also invariably opaque to all other operations, simply because the locality domains defined by these constraints do not discriminate between different types of movement.

To illustrate this point, consider as an example the contrast between scrambling and *wh*-movement out of a finite clause in German. We saw in (7) that finite clauses allow *wh*-movement out of them, but are opaque to scrambling.

(23) *Finite V-final clauses in German*

a. *Wh-movement:* [=(7a)]

Wen_i glaubst du [dass die Maria t_i liebt]
 who believe you that the Maria loves
 ‘Who do you believe Maria loves?’

b. *Scrambling:* [=(7c)]

*Ich glaube den Fritz_i [dass die Maria t_i liebt]
 I believe the Fritz that the Maria loves
Intended: ‘Fritz, I believe that Maria loves.’

On the standard assumption that CP is a phase, extraction out of the embedded clause has to proceed successive-cyclically through the specifier of that CP. Phases in and of themselves only require that extraction be successive-cyclic. They do not lend themselves to an account of why scrambling out of the lower clause is impossible. The derivation in (24) should be well-formed, contrary to fact.

(24) *Ich glaube den Fritz_i [t_i dass die Maria t_i liebt]

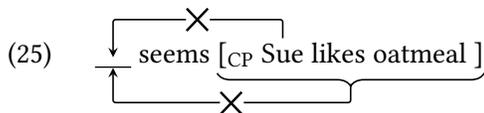
I believe the Fritz that the Maria loves

To express the locality difference between *wh*-movement and scrambling, we need to be able to state that an element in the embedded SpecCP can undergo *wh*-movement into the higher clause, but that it cannot undergo scrambling. Phases do not provide us with the technology to make such a statement, simply because phases only make the blanket statement that elements in SpecCP are visible and others are not, period. Selective opacity lies beyond the scope of phases precisely because selective opacity requires reference to the type of the operation, a distinction that phases are fundamentally blind to.

Similar remarks apply to subjacency or barriers. Like phases, these constraints do not have the right structure to account for selective opacity effects because they treat all operations alike. Locality mismatches are hence beyond the purview of these constraints as well. A consequence

of these considerations is that a successful account of selective opacity effects requires a locality principle that is quite distinct in nature from standard principles like phases in that it must be sensitive to the type of movement involved.

Selective opacity effects also pose interesting problems for an account in terms of Relativized Minimality (Rizzi 1990, *et seq.*). The reason is that minimality requires that an operation cannot proceed past an element that would itself be a suitable candidate for this operation. Locality arises from the requirement to pick the closest element that could undergo this operation. In most examples of selective opacity, this line of account has little to say because the embedded clause that defines the locality domain for a process is typically not itself a possible target for this process. This renders the domain irrelevant for the computation of minimality. To give just one illustrative example, we have seen above that finite clauses in English are opaque to A-movement out of them. Interestingly, this restriction holds even if that finite clause cannot itself undergo A-movement. In (25), A-movement of *Sue* over the CP is impossible. But A-movement of the CP is likewise ruled out. Moreover, Iatridou & Embick (1997) argue that CPs lack the ϕ -features necessary for agreement with T. As a consequence, the CP in (25) does not qualify as a closer target for either A-movement or ϕ -agreement. In other words, the CP blocks A-movement out it despite not being a viable goal for A-movement or ϕ -agreement itself. Because minimality principles only focus on the set of viable targets for an operation, the opacity of CPs cannot be attributed to regular minimality.



- a. *CP movement* → * $[_{CP}$ Sue likes oatmeal] seems.
- b. *DP movement* → *Sue seems $[_{CP}$ t likes oatmeal].

A more detailed discussion of the problems that selective opacity poses for traditional approaches to syntactic locality can be found in section 3.2 of chapter 3.

These considerations do not, of course, entail that phases or minimality are necessarily superfluous as constraints on syntactic representation. My point here is merely that what underlies selective opacity effects must be some locality constraint other than phases or their predecessors and that a novel line of approach is called for. As a matter of fact, I will argue in chapter 6 that phases do indeed play a crucial role for cases other than selective opacity. Interestingly, however,

selective opacity effects will be shown to have direct implications for the distribution of phases. Specifically, I will argue that CPs are phases, but that *v*Ps are not, a conclusion that will receive support from novel evidence from sentence processing in chapter 7.

1.3 Meta-generalizations of selective opacity

As mentioned above, the starting point of this thesis is the view that selective opacity effects are not due to totally unrelated construction-specific locality constraints, but rather are manifestations of a fundamental type of syntactic locality, which manifests itself in a wide variety of constructions and languages. The goal, hence, is to develop a unified theory of selective opacity effects, a goal that has, to varying degrees, also been pursued by works such as Williams (1974, 2003, 2011, 2013), Sternefeld (1992), Müller & Sternefeld (1993), Müller (1995, 2014a,b), Grewendorf (2003, 2015), Abels (2007, 2009, 2012a,b), and Neeleman & van de Koot (2010).

What is gained by treating selective opacity as a uniform phenomenon? The first reason to explore this line of inquiry is the standard methodological point that a unified account of a range of apparently unrelated phenomena is to be preferred over an account that essentially invokes a separate syntactic constraint for each instance of selective opacity. This move allows us to face heads-on the pervasive fact that syntactic domains can be selectively opaque. Second, treating selective opacity effects as a class of phenomena allows us to meaningfully ask the question of what this class has to tell us about the nature of syntactic locality, particularly in light of the fact that standard principles are ill-equipped to deal with selective opacity.

The third, and most surprising reason to take selective opacity effects to be a natural class is the fact that intriguing empirical generalizations emerge once selective opacity is treated as a phenomenon that manifests itself in a wide array of configurations. The recent literature on selective opacity effects has unearthed a number of meta-generalizations of selective opacity. Crucially, these meta-generalizations become apparent only once locality mismatches in a wide range of constructions and languages are considered. Viewing selective opacity as a natural class thus has an empirical payoff as well, in that it allows us to identify patterns that remain hidden on a piecemeal approach. I will briefly preview these meta-generalizations in this section. They will be corroborated by in-depth case studies of selective opacity of Hindi in chapter 2 and German

in chapter 4. The meta-generalizations are of crucial importance because they demonstrate that locality mismatches are not distributed randomly, but follow overarching patterns, which any account of selective opacity effects has to capture in one way or another.

1.3.1 The Height–Locality Connection

The first meta-generalization is what I will call the **HEIGHT–LOCALITY CONNECTION**. I will use the term *height* to refer to the structural height of the landing site of a movement type in the clausal spine. *Locality* will refer to the locality restrictions on that movement type. One of the key discoveries in the recent literature on selective opacity is that the two are related to each other: The higher the landing site of a movement type in the clausal spine, the more domains are transparent to this movement type. Conversely, movement types that target a structurally low position are typically subject to stricter locality constraints than movements targeting high positions. This connection has been argued for in slightly different forms by Williams (2003, 2011, 2013), Abels (2007, 2009, 2012a), and Müller (2014a,b), and the case studies of selective opacity in this dissertation will provide further support for it.

(26) **Height–Locality Connection**

Movement types differ in their landing sites. The higher the landing site of a movement type is in the clausal structure, the more kinds of structures are transparent to this movement type.

(26) may be illustrated with the familiar A/\bar{A} -movement distinction in English. It is uncontroversial that A - and \bar{A} -movement target different landing sites in English and specifically that A -movement targets a position that is structurally lower than \bar{A} -movement (SpecTP vs. SpecCP, respectively). Importantly, this difference in the height of the landing site correlates with their respective locality profiles. Finite clauses are opaque to A -movement, as we have seen on the basis of (4), but they are transparent to \bar{A} -movement out of them. Nonfinite clauses, on the other hand, are transparent to both. Thus, the movement type that targets a structurally high position (i.e., \bar{A} -movement) is able to leave more types of clauses (finite and nonfinite clauses) than movement that lands in a low position (i.e., A -movement). In other words, movement that lands low is subject to stricter locality

constraints than the movement type that lands high. According to (26), this correlation is not a coincidence, but a systematic and general property of selective opacity effects.⁵

More instances of (26) are easy to find. We saw on the basis of (7) that finite clauses in German allow *wh*-movement out of them, but that they disallow scrambling. Again, this locality difference correlates with the relative height of the respective landing sites of *wh*-movement and scrambling. *Wh*-movement transparently lands in a higher position than scrambling because *wh*-moved elements invariably appear to the left of scrambled material. The connection between the relative height of the landing site of a movement type and its locality profile again conforms to (26): The movement type that lands in a high position (i.e., *wh*-movement) is able to leave finite clauses, which are opaque to the movement type that lands in a low position (i.e., scrambling). The locality differences between clitic climbing and phrasal movement in Romance languages also corresponds to (26) as clitic climbing evidently targets a position lower than the landing site of movement to the left periphery.

Abels (2012a) argues in detail for a connection between height and locality in line with the Height–Locality Connection on the basis of the Italian left periphery. As we have seen in (20), (21) and (22) above, relativization and topicalization may leave a *wh*-island, whereas focus movement may not. This locality difference again correlates with a difference in the height of landing site of these movement types. As (27) demonstrates, a relative pronoun has to precede a focused element in the same clause, the inverse order is ungrammatical. This makes it clear that relativization targets a position higher than focus movement.

(27) *Italian: Relativization lands higher than focus movement*

- a. Ecco un uomo a cui IL PREMIO NOBEL dovrebbero dare (non
here is a man to whom THE NOBEL PRIZE they should give not
il premio X).
prize X

⁵ Williams (2013), extending ideas in Williams (2003), discusses a similar connection between the height of the landing site of a movement type and its locality for relativization, *wh*-movement and topicalization in English, but the data are somewhat controversial in this domain (see section 3.5.3 in chapter 3).

- b. *Ecco un uomo IL PREMIO NOBEL a cui dovrebbero dare (non
 here is a man THE NOBLE PRIZE to whom they should give not
 il premio X).
 prize X (Rizzi 1997: 298)

Like relativization, topicalization can also land in a position higher than that targeted by focus movement, though in this case the ordering is flexible.⁶

(28) *Italian: Topicalization can land higher than focus movement*

Credo che a Gianni QUESTO gli dovremmo dire.
 I believe that to Gianni THIS we should say to him (Abels 2012a: 237)

Just as in the other examples just seen, the movement type that may (in the case of topicalization) or must (in the case of relativization) land in a structurally high position is able to escape more kinds of structures/clauses than a movement type that lands in a lower position (focus fronting), even if all of these movements land in the left periphery. In the example at hand, relativization and topicalization are able to leave *wh*-islands, but focus movement is not ((20), (21) and (22)). It is this connection between a movement type's landing site and its locality properties that the Height–Locality Connection (26) expresses.

A similar illustration of the Height–Locality Connection is provided by a comparison of focus movement and modifier fronting in Italian. As Abels (2012a: 237–238) shows, focused elements have to appear to the left of a fronted modifier. Focus fronting thus targets a position higher than modifier fronting. This height difference again correlates with a locality difference. Modifier fronting is clausebounded (Rizzi 2004: 249n10, Abels 2012a: 236–238), as we saw in (19), but focus fronting is not. That is, finite clauses are opaque to modifier fronting but not to focus movement. Once again, then, the movement type that targets a position relatively low compared to another movement type is also subject to stricter locality boundaries.

Additional illustrations of the Height–Locality Connection are quite easy to find. Müller (2014a,b) points to German pronoun fronting, Icelandic object shift, and English extraposition: Each of these movement types targets a low position and is not able to cross a CP, unlike movement

⁶ Rizzi (1997) proposes that there are multiple landing sites for topicalization in Italian. (28) demonstrates that topicalization can land higher than focus movement. By (26), this difference in the height of the two movement types correlates with topicalization being potentially less local than focus movement.

types that target a structurally high position. Abels (2007, 2009) notes a connection between height and locality for various movement types in German.

Chapters 2 and 4 of this dissertation explore in detail selective opacity effects in Hindi and German, respectively, and I will argue that these case studies provide further support for the Height–Locality Connection (26). The discovery that the locality profiles of various movement types may not only differ, but that they are related to the height of the landing site of that movement type provides compelling evidence that locality differences between operations are not arbitrary or random. Rather, once selective opacity is treated as a systematic phenomenon that arises across a variety of constructions and languages, meta-generalizations like the Height–Locality Connection become evident. The systematicity that such generalizations reveal calls for a uniform approach to selective opacity effects, simply because only a unified account that handles selective opacity in a variety of constructions will be able to capture overarching generalizations that arise across constructions and languages. The analytical challenge that we are faced with, then, is not only to devise a theory that allows for locality differences between operations. At the same time, such a theory must also impose limits on possible locality differences. In particular, it has to relate differences in locality to differences in the height of the landing site of the operation.

It is worth noting at the outset that the recent literature on selective opacity effects has engaged in a debate about the precise link between height and locality. Williams (2003, 2011, 2013), for example, presents an account on which the two are correlated and locality is a direct function of a movement type’s landing site. Abels (2012a) presents a theory, which, at least in its strongest form, would reduce height to locality, hence also result in a correlation between the two. Müller (2014a,b), on the other hand, argues that height and locality are connected, but not in the form of a strict correlation. The Hindi and German case studies in this thesis will support the latter view. In particular, I will show that operations that target the same position may nonetheless differ in their locality. Moreover, while there is an empirical connection between height and locality, it is nonetheless impossible to predict precisely what domains are transparent and opaque to an operation based on that operation’s height. For these reasons I will refer to the empirical generalization in (26) as the Height–Locality *Connection*, rather than the Height–Locality *Correlation*.

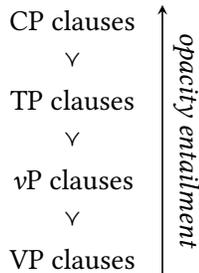
1.3.2 Upward Entailment

A second important meta-generalization across selective opacity effects that I will motivate is **UPWARD ENTAILMENT** in (29). Upward entailment states a generalization across embedded clauses of various sizes. As a background, I will adopt here the mainstream view that embedded clauses can come in different structural sizes (e.g., nonfinite TP clauses vs. finite CP clauses in English). Clauses of different sizes may differ in their locality properties. For example, nonfinite TP clauses in English are transparent for A- and \bar{A} -movement, but opaque to extraposition, as shown above, whereas finite CP clauses allow only \bar{A} -movement out of them, not A-movement or extraposition. Against this general background, Upward Entailment states that there is an entailment relation between structurally smaller and structurally larger clauses: If a clause of a given size is opaque to some movement type, than structurally larger clauses are also opaque to this movement type.

(29) **Upward Entailment**

If a clause of a certain structural size is opaque to an operation, then clauses that are structurally larger are also opaque to this operation.

(30) *Illustration of Upward Entailment:*



In English, TP clauses are opaque to extraposition, as shown in (31), repeated from (6b). Based on this fact, Upward Entailment states that CP clauses, in virtue of being CPs and therefore structurally larger than TP clauses, are also opaque to extraposition, as in (32).

(31) *TP clauses are opaque to extraposition in English*

* [John_i is believed [TP t_i to be certain t_j] by everybody [that Fred is crazy]_j].

(32) *CP clauses are opaque to extraposition as well*

*[It is believed [_{CP} that John is certain t_j] by everybody [that Fred is crazy]_j].

One might wonder whether the ungrammaticality of (32) is not simply due to the fact that extraposition must cross a TP inside the CP, hence the same reason that (31) is out. I will show shortly that it is technically possible using fairly standard machinery to devise systems in which TP clauses are opaque, but CP clauses are not. Upward Entailment is therefore not a theoretical necessity.

Another illustrative example of Upward Entailment comes from Russian. As we saw in (9)–(11) above, nonfinite clauses are opaque to A-movement out of them, again illustrated by (33a). This opacity to A-movement is shared by finite clauses, as shown in (33b). The structure is grammatical only in the absence of A-movement, as in (33c). See Stepanov (2007) for analogous examples using *sčitat’cja* ‘to be considered’. On the view that finite clauses are structurally larger than nonfinite clauses, this entailment is non-accidental, but rather instantiates Upward Entailment.⁷

(33) *Russian*

- a. *Èti studenty_i kažutsja [_{nonfinite} t_i znat’ tri jazyka].
 these students seem.3PL know.INF three languages
- b. *Èti studenty_i kažutsja [_{finite} (čto) t_i znajut tri jazyka].
 these students seem.3PL (that) know.3PL three languages
- c. Kažetsja [_{finite} čto èti studenty znajut tri jazyka].
 seem.3SG that these students know.3PL three languages
 ‘It seems that these students know three languages.’ (Ekaterina Vostrikova, p.c.)

The detailed case studies of Hindi and German in chapters 2 and 4 will not only provide converging evidence for the validity of Upward Entailment as an empirical generalization, I will also show that Upward Entailment extends beyond the contrast between finite and nonfinite clauses. Specifically, I will present evidence that nonfinite clause can differ in their clause sizes (e.g., Chomsky 1981,

⁷ Halpert (2015) argues that in Zulu finite clauses, but not nonfinite ones, allow A-movement out of them, in apparent violation of Upward Entailment. However, she also provides evidence that nonfinite clauses are nominalized, whereas finite clauses are not. If this is correct, Zulu does not, after all, constitute evidence against (29), because finite clauses in Zulu are not then structurally larger than nonfinite ones. Put differently, what appears to block A-movement in Zulu is the presence of nominal structure. Because only nonfinite clauses contain such nominal structure, the Zulu pattern is fully consistent with Upward Entailment.

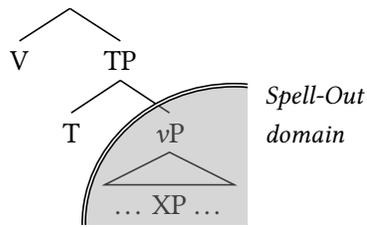
Rochette 1988, Wurmbrand 2001) and that Upward Entailment holds for nonfinite clauses of different sizes as well.

The inverse of Upward Entailment is, to the best of my knowledge, unattested. There is for instance, no movement type for which finite (CP) clauses are transparent, but nonfinite (TP) clauses are opaque. Such a locality type would of course be perfectly conceivable *a priori*. A hypothetical example for English would be the inverse of extraposition, with movement impossible out of nonfinite clauses, but possible out of finite clauses. All else equal, such a locality profile could be derived rather straightforwardly using standard technology like phases and the A-over-A Principle, as I will show momentarily. Upward Entailment encodes the empirical generalization that such a locality profile nonetheless does not exist.

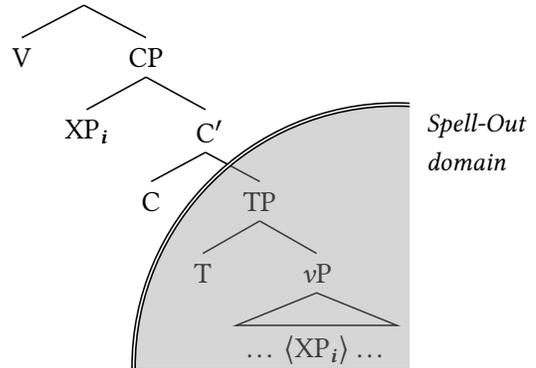
To further illustrate the challenge posed by Upward Entailment, it is instructive to consider how configurations that violate it could arise from standard principles of syntactic locality. For the sake of concreteness, I will present two illustrative examples of how such configurations could arise. The main point here is a proof of concept. The focus is not on these specific analyses, but on the general question of what kind of situations need to be blocked to derive Upward Entailment. The first example is based on Bošković's (2014) proposal that the highest projection in an extended projection defines a phase. Assume furthermore that (complete) opacity is implemented as the impossibility of a phase head to take a specifier (hence, absence of an edge feature). To give a concrete example for the sake of the argument, suppose that T cannot take a specifier, but C can. This state of affairs produces a violation of Upward Entailment. If the embedded clause is a TP, the TP will constitute a phase. Because it cannot take a specifier, no extraction out of a TP clause will be possible, as shown in (34a). By contrast, if the embedded clause is a CP, then the CP will be the phase instead of the TP. Because CP by assumption allows specifier, elements may move to its edge and hence be accessible from the outside. As a result, CP clauses are transparent to extraction, as (34b) illustrates. The result is again one where a structurally smaller TP clause is opaque, but a structurally larger CP clause is transparent. The addition of a projection thus makes available an escape hatch and hence allows extraction that a TP clause blocks. The result would be an unattested violation of Upward Entailment.

(34) *Hypothetical violation of Upward Entailment: Example 2*

a. *TP clauses → opaque*



b. *CP clauses → transparent*



The second illustrative example is based on a recent proposal by Wurmbrand (2015) that addresses a specific instance of selective opacity, namely the locality difference between topicalization and scrambling in cases such as (15)–(16). As shown there, nonfinite clauses containing the complementizer *żeby* are opaque to scrambling in Polish, but transparent to topicalization.

(35) *Polish*

a. *No scrambling over ‘żeby’*

[≅(15a)]

Jan { *pieniądze/*je } nalegał żeby { pieniądze/je } zostawić.
 Jan { *money /*them } insisted so.that { money /them } leave.INF
 ‘Jan insisted on leaving the money/them.’

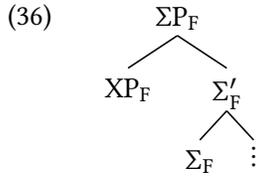
b. *Topicalization over ‘żeby’*

[=(16)]

Żabę, to Jan chciałby żeby tylko Maria pocałowała.
 frog.ACC TOP Jan want.SUBJ that.SUBJ only Maria.NOM kissed

‘The frog, John would like only Mary to kiss.’ (Wurmbrand 2015: 230)

Wurmbrand (2015) proposes an account of this asymmetry based on the A-over-A Principle. She suggests that scrambling is featurally licensed in a projection ΣP , where scrambling targets the specifier position of ΣP . Because a scrambled constituent XP in $\text{Spec}\Sigma P$ shares with ΣP the feature F that gave rise to scrambling, the A-over-A Principle prevents further scrambling of XP out of ΣP because ΣP constitutes a closer goal to any higher probe. This is schematized in (36).



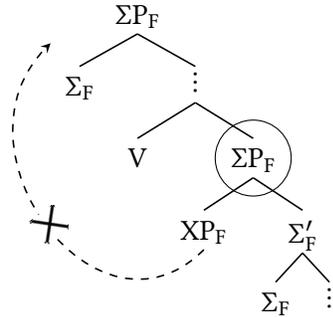
If an embedded clause contains a ΣP , it is therefore opaque to scrambling. Wurmbrand (2015) furthermore assumes that topicalization is triggered by a generic EPP/edge feature, which is invisible for the further computation. Topicalization can hence access XP in (36) and it does not itself induce opacity.

There are general problems that an account of selective opacity in terms of the A-over-A Principle faces, which I discuss in section 3.2.2 of chapter 3, but these do not need to concern us here. What is crucial for our present concerns is that Wurmbrand's (2015) account would allow for systems that violate Upward Entailment. To see this, suppose that topicalization targets SpecCP and that ΣP is located lower than that, a possibility that Wurmbrand (2015) acknowledges. Suppose also that a clause can be either a full-fledged CP or a pruned ΣP . If the clause is a ΣP , as in (37a), it will be opaque to scrambling, due to the A-over-A Principle. By contrast, if it is a CP, as in (37b), the following derivation becomes available. XP scrambles to Spec ΣP of the embedded clause and then topicalizes to SpecCP of that clause. Topicalization is possible in this case precisely because it does not involve the feature F . In SpecCP, XP is no longer encapsulated in a projection bearing F and hence accessible to ΣP in the higher clause, allowing scrambling. The net result is that CP clauses are transparent to scrambling, while structurally smaller ΣP clauses are not.

(37) *Hypothetical violation of Upward Entailment: Example 1*

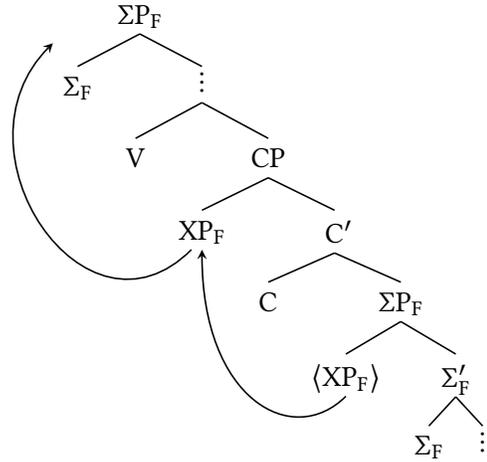
a. ΣP clause

→ scrambling impossible



b. CP clause

→ scrambling possible



The result would be a violation of Upward Entailment as a structurally small clause is opaque to a movement type that a structurally larger clause is transparent to.

Just as in the discussion of Bošković (2014), my point here is not to evaluate Wurmbrand's (2015) proposal, but to give a proof of concept. The two examples have made it clear that Upward Entailment is by no means conceptually necessary or trivial. Upward Entailment can be violated if the addition of clause structure makes available an escape hatch that allows an item to escape a configuration that is otherwise a barrier to it. To the extent that Upward Entailment is indeed a valid empirical generalization, as I argue, situations in which additional clause structures provides additional escape hatches have to be ruled out. See section 1.6 below for an overview of my proposal and section 3.3.2 in chapter 3 for an in-depth discussion.

A key consequence of Upward Entailment is that it provides additional justification for a unified approach to selective opacity effects because Upward Entailment arises as a generalization only once selective opacity is viewed as a uniform phenomenon. Moreover, like the Height–Locality Connection, Upward Entailment attests to the fact that locality mismatches are not distributed randomly, but follow specific patterns. Capturing these patterns is the key challenge and benchmark for any account of selective opacity.

1.4 Selective opacity beyond movement

Selective opacity effects like the ones presented in the previous section have been primarily analyzed in the domain of *movement dependencies*, i.e., as locality discrepancies between types of movement. Consequently, the vast majority of accounts of selective opacity are couched in terms of movement (Müller & Sternefeld 1993, Müller 1995, 2014a,b, Abels 2007, 2009, 2012a,b, Neeleman & van de Koot 2010). In chapter 2, I will argue, based on evidence from Hindi-Urdu (henceforth Hindi) that selective opacity is in fact considerably more general as a phenomenon and that it is not restricted to movement dependencies. In particular, I will argue that selective opacity also arises in the domain of *in-situ dependencies*, specifically ϕ -agreement and *wh*-licensing. In this chapter, I will investigate in detail the locality properties of A-movement, \bar{A} -movement, ϕ -agreement, and *wh*-licensing in Hindi and argue that all four operations partake in selective opacity. There is, in other words, every reason to believe that selective opacity applies to *movement and non-movement operations alike*. I will conclude from these considerations that selective opacity in the domain of movement is merely a special case of a much more general and abstract constraint. In other words, I will argue that the characterization in (1) is empirically more adequate than the purely movement-based characterization in (2). This discovery will then set the stage for the account of selective opacity to be developed in chapter 3.

Locality mismatches between movement and ϕ -agreement have occasionally been attested in the literature, but have not, to my knowledge, been linked to locality mismatches within the class of movement dependencies. One illustrative example comes from the Chukotko-Kamchatkan language Itelmen. Bobaljik & Wurmbrand (2005) show that in Itelmen a matrix verb can agree with the object of an embedded nonfinite clause (long-distance agreement). Crucially, if such agreement takes place, the object has to take scope over the matrix predicate. An illustrative example is (38), where the matrix verb *t'-antxa-če?n* 'forget' agrees with the embedded object *miš okno-?n* 'all windows'. As a result of this agreement, *miš okno-?n* 'all windows' has to take scope over the matrix verb.

(38) *Itelmen*

t'-əntxa-čeʔn [miʔ okno-ʔn sop-es]
1SG-forget-3PL.OBJ all window-PL close-INF

'I forgot to close all the windows.'

($\forall > \textit{forget}$; * $\textit{forget} > \forall$)

(Bobaljik & Wurmbrand 2005: 849)

Bobaljik & Wurmbrand's (2005) account crucially involves a locality mismatch between movement and ϕ -agreement. They propose that the embedded clause is transparent to movement, but opaque to ϕ -agreement. Agreement with *miʔ okno-ʔn* 'all windows' in (38) thus requires that it move into the matrix clause, resulting in it taking scope over the predicate *forget*. Central to this account is that the embedded clause allows (A-)movement out of it, but blocks ϕ -agreement into it. This is indeed the conclusion that Bobaljik & Wurmbrand (2005) draw, albeit as a general locality discrepancy between A-movement and Agree. I suggest, by contrast, that the pattern is nothing other than selective opacity, but arising not as a difference between two movement types, but between movement and agreement.

The inverse locality mismatch between movement and ϕ -agreement is attested in Tsez, as discussed and analyzed by Polinsky & Potsdam (2001). As they show, Tsez prohibits all crossclausal movement:

(39) *Tsez*

a. kid-bā [už-ā hibore-d bikori žāk'-ru-li] esis
girl-ERG boy-ERG stick-INSTR snake hit-PASTPART-NMLZ said
'The girl said that the boy had hit the snake with a stick.'

b. *bikori kidbā [užā hibored žāk'ruʔi] esis
snake girl boy stick hit said (Polinsky & Potsdam 2001: 590)

ϕ -Agreement, on the other hand, may access at least the edge of the lower clause, as in (40), where the embedded object *magalu* 'bread' triggers gender (III) agreement on the matrix verb.

(40) *Tsez*

enir [užā magalu bāc'ruʔi] b-iyxo
mother boy bread.III.ABS ate III-know

'The mother knows the boy ate the bread.'

(Polinsky & Potsdam 2001: 584)

Embedded clauses (or their edges) in Tsez are thus transparent to ϕ -agreement, but they are completely opaque to movement dependencies.

One of the core proposals to be made in this dissertation is that movement–agreement asymmetries like those in Itelmen and Tsez are instances of selective opacity just like locality asymmetries between movement types. That means that selective opacity is not confined to movement despite the fact that it has traditionally been investigated in this domain.

Based on an in-depth case study of Hindi, chapter 2 will argue at length for this conclusion by demonstrating that locality mismatches hold between different movement types, between movement and in-situ relations, and between different in-situ relations. I will argue based on this discovery that selective opacity is a phenomenon that holds for syntactic *operations*, irrespective of whether they involve movement or are genuinely long-distance. On an empirical level, then, this dissertation advances the thesis that selective opacity effects are even more widespread and fundamental than they are usually taken to be.

This extension of the scope of selective opacity effects has direct analytical repercussions. It shows that selective opacity cannot be the result of a constraint on movement per se, for otherwise the account would have nothing to say about in-situ relations. Rather, the constraint that gives rise to selective opacity must be more abstract and apply to syntactic operations more generally, regardless of whether they involve movement or not. The next section briefly lays out my proposal.

1.5 Horizons: An overview of the proposal

Chapter 3 will develop the core proposal of this thesis, based on the investigation of selective opacity in Hindi in chapter 2 and the key conclusion that selective opacity is not limited to movement types, but is an abstract property of syntactic operations more generally. First of all, in order to capture the fact that selective opacity effects are observable in the domains of movement and non-movement operations alike, I propose that selective opacity is the result of a constraint on the operation Agree (Chomsky 2000, 2001). Adopting the standard view that Agree is a precondition for movement (Chomsky 2000, 2001, 2004), a constraint on Agree is abstract enough to restrict movement as well as non-movement operations. This move towards a more abstract characterization of selective opacity effects thereby yields the desired unification of movement and in-situ relations.

To develop this general analytical direction, I will introduce and motivate a novel concept of locality: **HORIZONS**. The core intuition is that probes have characteristic horizons that delimit how far they can search. Everything beyond a probe's horizon is out of sight for this probe and therefore outside of its search space. Just like real-world horizons, a probe's horizons designate the outer limit of what is visible to a probe. The crucial aspect of horizons is that horizons may differ between probes, just like real-world horizons are not absolute but relative to the observer. What constitutes the horizon for one probe may not be the horizon for another probe in another location.

Let me give just one example. Suppose that a probe π has CP as its horizon. In this case, π can see and search syntactic structure that lies between π and a CP node it c-commands. But nothing that lies beyond the CP will be visible to π and π is thus unable to search into any structure that is embedded inside a CP. A CP node effectively delimits π 's search space. Agree relations between π and material inside a CP are therefore ruled out for systematic reasons. π 's inability to establish an Agree relation into CP manifests itself in CP opacity for the process triggered by π . If π gives rise to movement, CPs will be opaque to such movement. If π is an agreement probe, CPs will be opaque for agreement, and so on.

Recall that horizons are not absolute but relative to the probe. CPs might not be a horizon for some other probe π_2 , in which case π_2 's search space is not delimited by a CP node. Syntactic processes triggered by π_2 will hence be able to cross a CP. In a nutshell, then, selective opacity follows from the fundamental relativity of horizons.

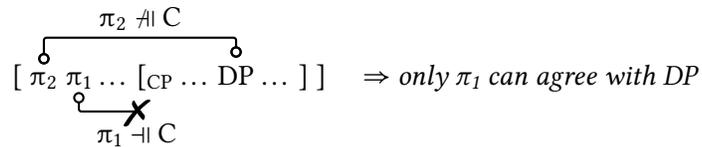
To develop this intuition, I propose that category features (like D, T, or C) are what potentially define a horizon for a probe. In the example at hand, the probe π would have the category C as its horizon. Any node that bears this category feature constitutes a horizon for π , blocking search past it. Because the CP node bears the C feature as part of its label, it will delimit π 's search space. I will use the symbol '–|' to designate horizons. For instance, the claim that C constitutes π 's horizons is expressed as ' π –| C'.

(41) **HORIZONS**

If a category label X is a horizon for probe π (notated as ' π –| X'), then a π -initiated search terminates at a node bearing X.

The relativity of horizons is key here. If a probe π_1 has C as its horizon ($\pi_1 \dashv\!\! \dashv C$), but another probe π_2 does not ($\pi_2 \dashv\!\! \dashv C$), then π_1 is unable to search into an embedded CP clause, whereas π_2 may. Consequently, π_2 is able to enter into Agree relationships with elements inside that CP clause, but π_1 is not, cf. (42). The result is a locality mismatch between π_1 and π_2 , and hence selective opacity.

(42) *Schematic illustration of horizons:*



The defining property of selective opacity is that the opacity or transparency of a given domain is relative to the *type* of the movement or the operation involved. Horizons make sense of this property because the transparency of a domain is determined relative to a probe. On the standard assumption that different movement types are triggered by specific probes, which are located on specific heads (e.g., *wh*-movement is triggered by a probe [*uwh*] on C, A-movement in English is triggered by an EPP-probe on T, etc.), locality differences between movement types is one manifestation of horizons. For example, we saw in (7) above that finite clauses in German are opaque to scrambling and relativization, but not to *wh*-movement. On the view that different probe features underlie these movement types, horizons provide a rationale for why the opacity of a domain should depend on the type of the dependency involved. The case of German is discussed in great detail in chapter 4.

The concept of horizons differs substantially from standard principles of locality in a number of ways. A first key difference to subadjacency, barriers or phases is that the locality domains defined by horizons are relative to the probe and that one and the same domain may at the same time be transparent and opaque. I have argued in section 1.2 above that this feature is a crucial component in accounting for selective opacity effects. As a result, horizons constitute a major departure from standard ways of conceiving of syntactic locality, in that they are crucially non-binary.

A second key difference between horizons and phases is that locality domains that arise from horizons are complete. In standard phase theory (Chomsky 2000, 2001), phases do *not* define opaque domains, because their edge remains accessible to the outside. It is only the complement of the phase head that is rendered inaccessible. As a result, phases enforce successive-cyclic movement, but unless they are supplemented with an independent constraint, they do not, in fact, render a

domain opaque for extraction (see, e.g., Boeckx & Grohmann 2007 and Abels 2012b for discussion). Horizons, on the other hand, do not attribute any special role to the edge of a domain. If, e.g., CP is a horizon for a probe, then all material inside CP, including material at its edge, are inaccessible to the probe:

$$(43) \quad \left[\begin{array}{c} \pi \dots \\ \circ \quad \swarrow \quad \searrow \\ \pi_1 \text{---} \text{H} \text{---} \text{C} \end{array} \left(\left[\text{CP} \text{ XP} \left[\text{C}' \text{ C}^0 \left[\dots \text{YP} \dots \right] \right] \right] \right) \right]$$

The relative nature of horizons is reminiscent of Relativized Minimality (Rizzi 1990). The crucial difference between the two is that nodes that constitute horizons for a probe are not necessarily themselves licit goals for that probe. In other words, the effects of horizons cannot be reduced to a general requirement for syntactic dependencies to be as short as possible. This characteristic of horizons avoids the pitfall for minimality-based accounts discussed on the basis of (25) above: A domain can be opaque to a process even if that domain cannot itself undergo that process.

Horizons constitute a significant departure from the previous literature on selective opacity effects in a number of ways. First, selective opacity now follows from a constraint on Agree and syntactic probes, rather than movement dependencies themselves. On the horizons view, improper movement and other instances of selective opacity are really instances of ‘improper agreement’. The advantage of this perspective is that locality mismatches between movement and agreement like the Itelmen and Tsez data discussed in (38)–(40) can be integrated into a general theory of locality mismatches. The conclusion reached in chapter 2 that there are furthermore locality mismatches between different in-situ operations in Hindi will likewise fall within the purview of the horizons account. This dissertation thus adds to a general trend in the recent literature that has argued that locality mismatches are more pervasive than they have been traditionally taken to be. Recall from section 1.1 that differences between the locality of A- and \bar{A} -movement have emerged as specific instances of locality mismatches between movement types more generally. The core claim of this thesis is that even general locality mismatches between movement types are specific instances of a more general pattern that comprises both movement and non-movement operations. As a result, the most familiar example of selective opacity, superraising in English, emerges as merely the tip of the iceberg.

A second fundamental difference between the horizons accounts and other approaches to selective opacity is that horizons do not involve any direct interaction between movement types and in this regard horizons constitute a major departure from the existing literature. Virtually all accounts of selective opacity are derivational in the sense that they involve an interaction between various movement steps of an element. This core approach goes back to Chomsky's (1973, 1977, 1981) account of superraising in English. Recall that this account invokes two constraints: First, subjacency forces an element that is extraction out of a finite clause to move the edge of that clause, an \bar{A} -position. Second, the ban on improper movement crucially prohibits A-movement from an \bar{A} -position.

(44) *Traditional account of superraising*

*Sue seems likes oatmeal.

a. *One-fell-swoop movement:*

*Sue_i seems [_{CP} C⁰ t_i likes oatmeal]

→ ruled out by subjacency/phases

b. *Successive-cyclic movement:*

*Sue_i seems [_{CP} t_i C⁰ t_i likes oatmeal]

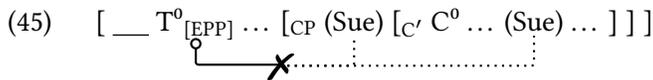
→ ruled out by ban on \bar{A} -movement followed by A-movement

The standard account of the selective opacity of finite clauses for A-movement is thus *item-based* and *operationally mediated*. It is item-based because the ban on improper movement makes reference only to the movement derivation of a single item (i.e., *Sue* in (44)). The account is operationally mediated in that the opacity of finite clauses for A-movement is expressed as a constraint on operations: No constituent may undergo \bar{A} -movement and then A-movement. On this standard account, selective opacity of a domain is the result of a conspiracy between (i) an unselective constraint that regulates that only the edge of a domain is visible, and (ii) an interaction between different movement types in that one movement type (\bar{A} -movement) directly bleeds the application of another movement type (A-movement).

Despite substantial innovations in the subsequent literature on selective opacity, the core approach has remained largely unchanged. Accounts in terms of *unambiguous binding* (Müller & Sternefeld 1993, Müller 1995), *operational ordering* (Abels 2007, 2009, Neeleman & van de Koot 2010), and *buffers* (Müller 2014a,b) all adhere to traditional analyses in that they rule out, in one way

or another, particular sequences of operations (see chapter 5 for a thorough comparison between horizons and previous approaches).⁸

The horizons approach constitutes a major departure from this line of approach. Horizons are thoroughly *domain-based* rather than item-based and they do not make reference to possible and impossible sequences of operations. What horizons state is that a given syntactic domain (usually a clause) is transparent for some probes but not for others. Crucially, if, e.g., a CP node is a horizon for a probe, any element dominated by that CP will be out of reach for that probe, irrespective of the location of that element inside the CP. The two derivations in (44) are *both* uniformly ruled out by horizons. Assume that the EPP-probe on T has CP as its horizon in English. As (45) shows, the one-fell-swoop derivation in (44a) is ruled out because it would require this probe to search past its CP horizon. The successive-cyclic derivation in (44b) is ruled out for exactly the same reason, as it too would require the probe to search past its horizon. Horizons provide us with the machinery to state that a *domain*, not just an item, is opaque to a probe. As I will show in detail in this dissertation, horizons thus derive selective opacity without the need for a conspiracy of constraints.



Unlike virtually all previous accounts of selective opacity, a horizon-based account is entirely unrelated to the internal structure of an opaque domain clause and specifically which movements have taken place within it. In fact, on the horizons account, the moving element in superraising configurations does not play any role whatsoever. The probe in (45) is unable to contact the DP due to an intervening horizon, a fact that is irrespective of the internal features and properties of that DP. All that matters are probes and their horizons, hence the title of this dissertation.

A direct consequence of this aspect of the horizons account is that no specific reference to \bar{A} -positions or their interactions with A-movement is necessary. In fact, finite clauses are opaque for extraction both from A- and \bar{A} -positions (see (45)) and horizons allow us to directly express this

⁸ Williams' (2003, 2011, 2013) account takes a special place in this dichotomy. On the one hand, it is strictly derivational in nature in that it heavily relies on possible and impossible sequences of operations. On the other hand, it is not item-based, like horizons.

fact, as elements inside a CP are invisible to the EPP on T probe regardless of whether they are located in an A- or an \bar{A} -position.

Moreover, movement type interactions such as between \bar{A} - and A-movement emerge as entirely *epiphenomenal* on the horizons account. The fact that \bar{A} -movement may not feed A-movement is not due to a designated constraint prohibiting such a sequence. Rather, \bar{A} -movement entails the presence of a CP in the embedded clause. This CP constitutes a horizon for an EPP-probe in the higher clause and thus renders everything inside it invisible to this probe, including the \bar{A} -moved element. \bar{A} -positions themselves do not factor into the explanation of selective opacity. Horizons thus constitute a radical departure from the conceived wisdom on selective opacity: While the previous literature has overwhelmingly treated domain opacity as a result of constraints on operations and derivations, horizons treat constraints on operations and derivations as the result of representational constraints on domain opacity.

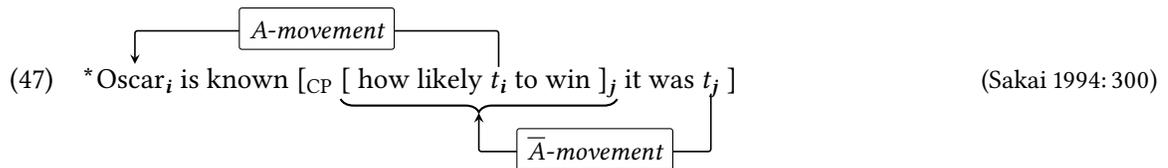
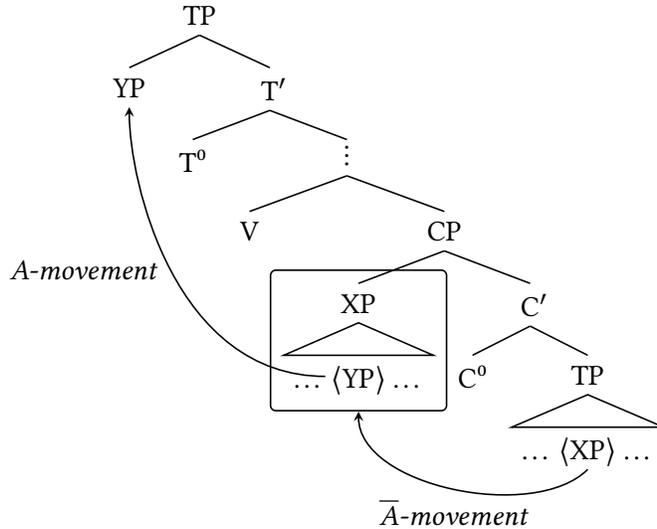
I will justify this shift in perspective away from an operationally-mediated account of selective opacity to a domain-based one in chapter 2. There I show based on Hindi that interactions between operations arise even if the elements that undergo these operations are *not the same* (section 2.3.3). Further evidence for this shift towards a domain-based view comes from a range of cases, discussed by Sakai (1994), Grewendorf (2003, 2015), Williams (2003), Abels (2007, 2009), and Neeleman & van de Koot (2010), in which interactions between movement types are observable *even if the elements undergoing the two movement steps are distinct*. Such non-identity cases are discussed in greater detail in section 3.4 of chapter 3 and section 4.4 of chapter 4. I will show that they are puzzling for a standard item-based view, but straightforwardly accounted for on the domain-based view of horizons.

The virtues of a domain-based account can be briefly illustrated with a related observation made by Sakai (1994), Grewendorf (2003, 2015), Williams (2003), Abels (2007, 2009), and Neeleman & van de Koot (2010). These authors have argued that interactions between movement types arise even if the elements undergoing the two movement steps are distinct. Particularly instructive are ‘smuggling’ configurations⁹ (a term due to Collins 2005a,b), in which an element XP moves the edge of a domain, followed by subextraction of YP out of XP. In the schematic structure in (46),

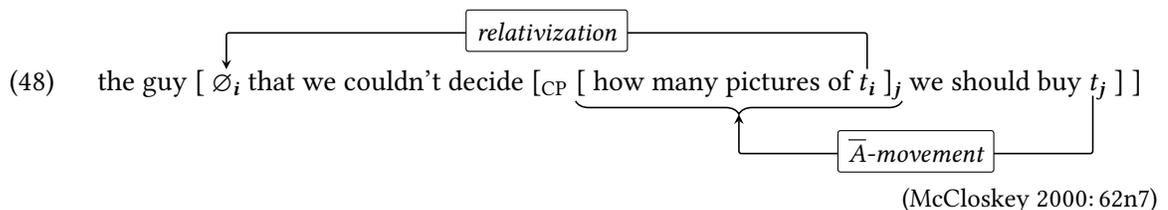
⁹ Sauerland (1999) uses the term ‘surfing’ for derivations of the type in (46).

XP is \bar{A} -moved to the edge of a CP clause, followed by A-subextraction of XP out of it. Such a derivation is impossible, as Sakai (1994) has emphasized on the basis of (47).

(46) *Illicit smuggling derivation*



Sakai (1994), Abels (2007, 2009), and Neeleman & van de Koot (2010) point out that this problem cannot be solved by simply ruling out smuggling derivations altogether because smuggling derivations themselves display selective opacity effects. To give just one example, in contrast to A-movement, relativization out of an \bar{A} -moved element is possible, as (48) illustrates:



Moreover, Abels (2007, 2009) and Neeleman & van de Koot (2010) have shown that inverting A- and \bar{A} -movement in (46) results in a grammatical structure. The key conclusion is that smuggling derivations exhibit movement type interactions similar to improper movement. Just like \bar{A} -movement of XP cannot feed subsequent A-movement of that XP (improper movement), so \bar{A} -movement of

an XP blocks A-subextraction out of that XP. An item-based account that merely restricts that an \bar{A} -moved item cannot A-move does not extend to smuggling configurations because the elements that undergo the two movement steps are distinct. An item-based account of selective opacity therefore has no means of handling indirect movement type interactions as they arise in smuggling derivations.

On the horizons account, on the other hand, constraints on smuggling derivations fall out without further ado. Recall that the EPP-probe in English has C as its horizon and is therefore unable to search into a CP. Nothing else needs to be said about smuggling as in (47) because this structure would require the EPP-probe to search into a CP clause, hence violating its horizon. The possibility of smuggling in (48) is likewise accounted for because CPs are never horizons to relativization, either in the identity or the non-identity (i.e., smuggling) case. As a result, due to their domain-based nature, horizons are much more general in their empirical scope and in the correlations they predict across different constructions than operationally mediated accounts. This results in an account that is conceptually as well as empirically superior to standard item-based accounts of selective opacity. See section 3.4.2 in chapter 3 and section 4.4 in chapter 4 for further discussion of such configurations. Chapter 4 also shows that selective opacity in smuggling configurations is not limited to the A/ \bar{A} -distinction but much more pervasive.

Finally, it is worth noting that the horizon account converges in an interesting way with the core insight of van Urk (2015). In a nutshell, van Urk (2015) argues that the traditional distinction between A- and \bar{A} -positions should be recast as a distinction between A- and \bar{A} -features. Movement type properties like crossover obviation, Principle C amnesty, anaphor binding, etc., are then reassessed as properties of the features underlying these movements, not features of the positions targeted by these movements. The horizon account makes essentially the same claim for locality, but with the caveat that the position of a probe imposes restrictions for its possible locality profiles, as I will show in the next section. On the horizons account, rather than attributing the locality of a movement type directly to the position that this movement type targets, it is the probe that triggers this movement that determines locality. There is thus a remarkable convergence between the general conclusions reached here and by van Urk (2015).

1.6 Deriving the meta-generalizations

As mentioned in section 1.3 and discussed at length in chapters 2 and 4, selective opacity effects are not distributed randomly, but adhere to meta-generalizations, which an adequate account of selective opacity has to capture. I show in chapter 3 how horizons shed light on these generalizations by imposing a systematic limit on possible selective opacity patterns. The two main generalizations motivated in section 1.3 above and further corroborated in section 2.5 of chapter 2 and section 4.3 of chapter 4 are the Height–Locality Connection and Upward Entailment, repeated in (49) and (50), respectively.

- (49) **Height–Locality Connection** [= (26)]
Movement types differ in their landing sites. The higher the landing site of a movement type is in the clausal structure, the more kinds of structures are transparent to this movement type.
- (50) **Upward Entailment** [= (29)]
If a clause of a certain structural size is opaque to an operation, then clauses that are structurally larger are also opaque to this operation.

I will argue in section 3.3.2 of chapter 3 that Upward Entailment is the result of an interaction between horizons and general properties of extended projections, building on earlier work by van Riemsdijk (1988, 1998) and Grimshaw (1991, 2000), among others. The basic intuition is that extended projections are *endocentric*, i.e., TPs and CP are verbal, DPs are nominal. Based on independent evidence from Grimshaw (1991, 2000), I propose that category features are inherited up through an extended projection even beyond the immediate projection line of a head. That is, CPs are quite literally verbal because they carry a V feature, inherited up from an embedded VP, and so on. This mechanism, **CATEGORY INHERITANCE**, is previewed in (51). I will notate extended projections as, e.g., $\langle \text{CP} > \text{TP} > \text{vP} > \text{VP} \rangle$, but the principle is independent of the precise number of projections that make up the clausal spine.

- (51) **CATEGORY INHERITANCE**
Given an extended projection $\Phi = \langle \Pi_n > \Pi_{n-1} > \dots > \Pi_1 \rangle$, where Π_x 's are all phrases, the categorial features of Π_m are inherited up to Π_{m+1} .

According to category inheritance, the category labels are cumulative within an extended projection: *v*P_s contain a *v* and a *V* category feature, TP_s contain a *T*, *v*, and a *V* feature and so on. Crucially, because category features also define horizons, category inheritance has the consequence that horizons also percolate up within an extended projection:

(52) **HORIZON INHERITANCE THEOREM**

Given a probe [*uF*] and an extended projection $\Phi = \langle \Pi_n > \Pi_{n-1} > \dots > \Pi_1 \rangle$, if $\Pi_m \in \Phi$ is a horizon for [*uF*], then all projection Π_{m+1}, \dots, Π_n are likewise horizons for [*uF*] (due to category inheritance (51)).

(52) then derives Upward Entailment as a necessary property of selective opacity effects. Importantly, this result is achieved without imposing any designated constraints on the distribution of horizons. Their distribution across probes is free, but independent constraints on the distribution of category features within extended projections will have the effect of Upward Entailment. The account thus offers a rationale for why Upward Entailment should exist in the first place by treating it as a consequence of independently motivated properties of extended projections.

The second crucial meta-generalization that figures prominently in this thesis is the Height–Locality Connection (49), the claim that a movement type’s locality profile is related to the structural height of its landing site (see section 1.3.1 above and sections 2.5.2 and 4.3.3 in chapters 2 and 4, respectively). This connection is particularly striking because standard locality principles do not give rise to a connection between the two. The explanation for the Height–Locality Connection should thus come from an account of selective opacity effects. The treatment of the Height–Locality Connection within the horizons framework is addressed in detail in section 3.5 of chapter 3. I will show there that the horizons account in fact derives a version of the connection without the need for a designated stipulation to its effect. In a nutshell, I will show that certain pairings of a probe’s location and its horizons give rise to probes which are inherently unable to trigger long-distance operations like movement or agreement. I will call these probes **VACUOUS**. Importantly, there is no designated stipulation that rules out vacuous probes. But because they cannot give rise to long-distance operations, all such dependencies must be triggered by non-vacuous probes. I will show that the horizons account has as a consequence the Height–Locality Theorem in (53). What (53) states is that only probes for which location and horizon stand in a specific relationship are non-

vacuous. It follows, then, that for all movement, agreement, and other long-distance dependencies, height and locality are necessarily connected. This will then immediately generalize to locality with respect to different types of embedded clauses. This derives the Height–Locality Theorem in (53) and hence a version of the Height–Locality Connection in (49).

(53) **HEIGHT–LOCALITY THEOREM**

Given an extended projection $\Phi = \langle \Pi_n > \Pi_{n-1} > \dots > \Pi_1 \rangle$, for any non-vacuous probe $[uF]$,

- a. If $[uF]$ is located on Π_m , then a projection $\in \{\Pi_{m-1}, \dots, \Pi_1\}$ cannot be a horizon for $[uF]$.
- b. If $[uF]$ has Π_m as a horizon, $[uF]$ cannot be located on a projection $\in \{\Pi_n, \dots, \Pi_{m+1}\}$.

The central aspect of this account is that it does not involve a stipulated constraint on possible pairings of location and horizons on a probe. Constraints on such pairings *emerge* from the interplay of various aspects of the horizons system. In this sense, the observation that there is an empirical connection between height and locality is *derived* and horizons provide an explanation for why such a connection should hold in the first place.

Horizons hence not only offer an account of locality mismatches, but they also provide a principled upper bound on locality mismatches. This upper bound explains the meta-generalizations that selective opacity is subject to. Finally, section 5.8 in chapter 5 compares the horizon account of the Height–Locality Connection to approaches in the previous literature, in particular Williams (2003, 2011, 2013) and Abels (2012a). I argue there that the horizons account is preferable on empirical ground in that it more accurately delimits the extent to which height and locality are correlated.

1.7 Consequences for the distribution of selective opacity

The horizons account developed in detail in this thesis has an analytical property that is worth highlighting. Virtually all accounts of (instances of) selective opacity aim at deriving that a given opacity pattern is necessary given some other aspect of the system. For example, the standard objective of theories of superraising in English is to derive that A-movement out of finite clause is necessarily impossible. A notorious problem for such an approach is the fact that there are a number of languages in which A-movement out of a finite clause is attested, e.g., Bantu languages (Carstens 2010, 2011, Diercks 2012, Halpert 2012), Greek (Alexiadou & Anagnostopoulou 2002),

Brazilian Portuguese (Nunes 2008), Rumanian (Grosu & Horvath 1984), and several other languages documented by Ura (1994). Various attempts have been undertaken in the literature to correlate the (im)possibility of superraising in a language with some independent property of this language. Despite significant insights, it seems fair to me to say that no such independent property has been identified. For example, Ura (1994) claims that, typologically, if a language allows the multiple subject construction, then that language will also allow superraising to take place. As Ura (1994: 106–107) himself notes, Hindi constitutes a counterexample to this entailment because Hindi does not allow superraising, but it allows A-scrambling of the object over the subject. Ura (1994) argues that such movement in Hindi is akin to passivization and claims that in these constructions the subject loses its subjecthood because it surfaces in an inherent case and cannot control verb agreement. That is, he claims that there is only a single subject position. However, these claims are empirically incorrect. Object A-scrambling does not in any way affect the case or agreement of the subject, as detailed in chapter 2 and the literature cited there. I therefore conclude that Hindi falsifies the entailment proposed by Ura (1994).

A second entailment argued for by Ura (1994) is that if a language allows hyperraising, it also allows null subjects. Interestingly, Nevins (2005: 298) points out that Brazilian Portuguese has hyperraising without having null subjects. Again, the purported entailment does not withstand closer scrutiny.

Chomsky (2000, 2001) proposes that the distribution of superraising follows from case properties. Hindi falsifies this connection, as discussed in detail in section 5.2 of chapter 5.

Carstens (2010, 2011) explores the view that the existence of uninterpretable gender in a language allows for superraising, the connection between the two being that such gender keeps a DP active for A-processes. Hindi again constitutes a counterexample, as it clearly has uninterpretable gender but no superraising. While one might save this connection by distinguishing between different types of uninterpretable gender, in the absence of independent evidence for such a difference, the entailment is reduced to empirical vacuity.

My intention here is not to belittle these efforts. What this discussion is intended to emphasize is that to date the literature has not been able to identify an analytical correlate of superraising that would allow us to deduce whether a language allows superraising or not. Matters get only worse once we move beyond the A/\bar{A} -distinction. As observed in (7), finite clauses in German

allow *wh*-movement out of them, but block relativization. Moreover, V2 clauses are transparent to *wh*-movement that lands inside a higher V2 clause, but opaque to *wh*-movement that targets a higher V-final clause. In these cases, it is even less clear what an independent correlate of the locality asymmetries between these movement types could be.

Rather than attempting to identify a novel predictor of a movement type's locality profile, this thesis will explore a very different line of approach. I will adopt the perspective that the locality properties of a probe are arbitrary as far as the theoretical axioms are concerned. That is, the choice of horizon for a probe is not deterministic, but free, with no designated stipulation on possible horizon settings. On this view, locality differences of movement types across languages, such as the possibility of superraising, cannot be predicted from some other aspect of these languages, because it is a parametric choice. One choice of horizons yields a language that allows superraising, another yields a system that does not. This arbitrariness extends to movement type asymmetries within a language. Thus, in German the locality contrast between *wh*-movement and relativization reduces to a variation in the choice of horizon. This view affords a straightforward account of crosslinguistic variation in horizons, which I discuss in section 3.5.2 of chapter 3.

We have seen above that selective opacity is not distributed randomly, but instead subject to overarching generalizations, like Upward Entailment and the Height–Locality Connection. A key question that the horizon account tries to answer is how a system that has arbitrary horizons can give rise to overarching empirical patterns of this type. As I have discussed in the preceding sections, I will argue that these patterns are the result of interactions between arbitrary horizons and independent properties of syntactic structures, in particular extended projections. As a result, constraints on possible selective opacity patterns emerge from the horizons account. As far as the analytical stipulations are concerned, horizon settings are arbitrary, but because certain settings will result in vacuous probes, meta-generalizations of selective opacity will arise indirectly. As I will show, despite the fact that horizons do not impose any designated connection between height and locality, such a connection nevertheless emerges in the output of the system.

The result of this approach is that a probe's location restricts the space of nonvacuous horizon settings for this probe, but it nonetheless does not uniquely determine a horizon settings. Within the set of the remaining options, the choice is free. The resulting theoretical picture is an unorthodox one: Syntactic domains exhibit selective opacity *precisely because* horizon settings are not subject

to designated stipulations, but rather arbitrary and hence free to vary across probes. Overarching patterns in the extent of the attested variation emerge from the interplay of free horizons and extended projections.

This perspective amounts to the claim that there is no property of a language that uniquely determines a movement type's locality profile in a given language. The horizon account thus expresses the view that it is simply not possible to uniquely determine a movement type's exact locality profile in a language on the basis of other properties of this movement type. All there is are free choices and emergent constraints on the range of these choices.

Another consequence of the claim that the locality of a movement type cannot be uniquely determined from its other properties is that children have to acquire locality properties directly from their input. On the face of it, this looks like a substantial challenge given that acquisition of locality would appear to require negative evidence, i.e., information about impossible syntactic dependencies. I take on this task in section 3.5.4 of chapter 3. There I propose that horizons naturally lead to *default horizons*, default horizon settings for a probe given the location of that probe. Once the notion of default horizons is granted, locality differences between operations can be acquired on the basis of purely positive evidence. This resolves the apparent acquisition paradox.

This general approach to selective opacity also implies that I will treat interpretive properties of movement types as analytically unrelated to that movement type's locality profile. For example, it is well-known that A- and \bar{A} -movement in English do not only differ in their locality, but also in their ability to obviate weak crossover, to feed anaphor binding, and to amnesty Principle C effects. I will treat these properties as analytically divorced from the locality properties of these movement types. This is a standard approach, in line with all previous accounts of selective opacity. Thus, the traditional account of the impossibility of superraising in English in terms of the ban on improper movement does not analytically tie this constraint to interpretive properties of A-movement. Conversely, accounts of the differential interpretive features of A- and \bar{A} -movement (see, e.g., Sauerland 1998 for binding and Takahashi & Hulsey 2009 for Principle C amnesty) do not entail any locality differences between A- and \bar{A} -movement. The analytical dissociation between locality and interpretive properties of movement types is supported by a clear lack of evidence that the two cluster together crosslinguistically. To give just two examples, there is to my knowledge no evidence that in German *wh*-movement into a V2 clause differs from *wh*-movement into a

V-final clause in their interpretative properties, but they evidently contrast in their locality (see (7)). Furthermore, A-movement differs across languages in whether or not it is able to leave a finite clause. As a result, I conclude from this absence of any implicational relationship between locality and interpretive properties that the two should be analytically separated.

1.8 Consequences for phase theory

In light of the generality of horizons as a constraint on syntactic locality, questions arise with respect to its relation to standard principles of locality, in particular phases. The distribution of phases is the subject of chapters 6 and 7. In chapter 6, I will show that horizons do not replace the need for phases. Specifically, I will show that horizons do not give rise to successive-cyclic movement. They simply determine whether a given movement dependency out of a lower clause is possible or not. They do not, however, condition the successive-cyclic path of licit movement dependencies. As mentioned in section 1.2 and discussed in greater detail in section 3.2.1 of chapter 3, phases on the other hand give rise to successive cyclicity, but not complete domain opacity. Based on these observations, I propose in chapter 6 that phases coexist with horizons. Both are analytically and empirically distinct enough that a system that incorporates both does not give rise to a redundancy. The empirical scope of both is thus clearly delimited and complex movement patterns emerge from the interplay of the two.

That said, I will also show in chapter 6 that not all conceptions of phases are compatible with this picture. In particular, I will show that there is ample motivation for the standard view that CPs are phases. On the other hand, *v*P phases are categorically incompatible with the horizons system. I will furthermore show in this chapter that this incompatibility is not limited to horizons, but in fact holds for virtually all accounts of selective opacity. The problem is thus quite general. In a nutshell, it arises because if *v*P is a phase, accounts of selective opacity would have to have access to information properly contained inside an already spelled-out phase, in direct violation of phase locality. To my knowledge, this problem has gone largely unnoticed in the literature. Selective opacity as an empirical phenomenon thus provides evidence against *v*P phases. I will furthermore argue that there is evidence independent of selective opacity that corroborates the conclusion that *v*Ps are not phasal. In particular, I will show that non-movement operations like ϕ -agreement in

Hindi can proceed past an unbounded number of ν Ps, directly disconfirming the phasal status of these ν Ps. This conclusion prompts a major reassessment of the distribution of phases: Only CPs are phasal, ν Ps are not.

Converging evidence for this distribution of phases is presented in chapter 7. In this chapter, I start out by observing that traditional diagnostics for successive-cyclic movement, like morpho-syntactic reflexes of intermediate landing sites, are *biased* in that they potentially allow us to diagnose the *presence* of an intermediate gap, but are inherently unable to diagnose the *absence* of an intermediate landing site. This bias may well have had the effect of exaggerating the apparent empirical support for ν P phases. To address this bias problem, I explore online sentence processing as a novel testing ground for theories of successive cyclicity. Building on pioneering work by Gibson & Warren (2004), I report the results of two self-paced reading experiments that aim to find processing evidence for the presence and distribution of intermediate landing sites. The results of these experiments support the conclusions reached on entirely independent grounds in chapter 6: CPs host an intermediate landing sites, but ν Ps crucially do not. On the standard view that intermediate gaps diagnose the presence of a phase head, these results provide support for the view that CPs are phasal, but ν Ps are not.

1.9 Overview of this dissertation

The empirical groundwork that will lead us to the concept of horizons is laid in **Chapter 2**, where I investigate in detail selective opacity effects in Hindi. I focus on four operations (i) A-movement, (ii) \bar{A} -movement, (iii) ϕ -agreement, and (iv) *wh*-licensing. I will investigate in particular the locality mismatches that exist between them, arguing that all four of them partake in selective opacity. Crucially, I will argue that ϕ -agreement and *wh*-licensing in Hindi do *not* involve movement, but are established long-distance, via pure Agree. The discovery that syntactic operations can partake in selective opacity irrespective of whether they involve movement or not will then lead us to the conclusion that a unified account of selective opacity in Hindi cannot be based on movement, but must be more abstract. Another key conclusion of chapter 2 is that selective opacity is a non-binary phenomenon. I have argued in section 1.1 above that selective opacity is not confined to a binary distinction between A- and \bar{A} -locality. The Hindi evidence will provide direct evidence for this

claim because it contains three locality types. Any account that is limited to just the A/\bar{A} -distinction will hence be insufficient. Furthermore, I will show in detail that the two meta-generalizations of selective opacity briefly discussed in section 1.3 above – Upward Entailment and the Height–Locality Connection – are fully borne out in Hindi, thus corroborating their validity. Finally, I will argue that the transparency or opacity of a clause is a function of its structural size, a conclusion that will guide the technical implementation of the horizon-based account developed here.

With these empirical conclusions of chapter 2 in place, **Chapter 3** will develop the account of selective opacity that lies at the heart of this dissertation. An important shift in perspective advocated here is that the relevant constraint is not about movement per se, but rather targets the operation Agree. Improper movement, on this view, is really improper agreement. I will introduce and develop the notion of a syntactic horizon. As outlined in section 1.5, the core intuition behind horizons is that syntactic nodes can delimit a probe’s search space. Crucially, probes can differ in which nodes confine their search, leading to mismatches between probes in which structures are searchable and which ones are not. I will show that horizons offer a comprehensive account of selective opacity effects, applying it to Hindi and a variety of other systems, some of which were discussed above. An important part of chapter 3 will be to demonstrate how an account in terms of horizons allows us to derive and explain the meta-generalizations that hold of selective opacity effects. I will propose that Upward Entailment follows from general properties of extended projections. We will also see that the account manages to impose restrictions on possible pairings of locations and horizons on probes, thus deriving a version of the Height–Locality Connection. The result is appealing because the horizons account not only offers a theory of locality mismatches, but also a theory of their limits.

In **Chapter 4**, I will investigate selective opacity in German. German is instructive because it has played a prominent role in the existing literature on selective opacity. German exhibits a rather striking number of locality mismatches, with striking differences between scrambling, relativization, *wh*-movement into a V-final clause, and *wh*-movement into a V2 clause. Chapter 4 lays out the relevant generalizations and argues that they offer additional support for the two meta-generalizations identified above: Upward Entailment and the Height–Locality Connection. I will then develop a comprehensive account of the German facts in the framework of horizons.

Having thus developed the account and applied it to a variety of instances of selective opacity, **Chapter 5** takes a step back and compares horizons to previous treatments of selective opacity. As alluded to above, a central point of divergence is that horizons eschew a designated ban on particular interactions between operations (e.g., that \bar{A} -movement may not feed A-movement). Selective opacity is derived from probes and syntactic domains, without any appeal to movement types, types of positions, etc. Various other points of comparison are discussed, in particular regarding the empirical scope and generality of the accounts. In light of the wealth of accounts of selective opacity, a central goal of chapter 5 is to highlight the empirical and theoretical consequences of specific analytical choices and explore how the choices made by various accounts place them with respect to the various empirical issues of selective opacity.

Chapter 6 is dedicated to exploring the relationship between horizons and phases. I will argue that horizons do not replace phases and that both are necessary as independent locality constraints on syntactic representations. Interestingly, the locality effects of horizons and of phases are clearly distinguishable. While horizons determine whether a given domain is opaque or transparent to a syntactic process, phases regulate the successive-cyclic dependency path of operations allowed by horizons. As a result, an account that incorporates both horizons and phases does not give rise to a redundancy. Moreover, for any empirical locality effect, it will be possible to clearly distinguish whether this effect is the result of horizons or of phases. However, while the general notion of a phase is compatible with horizons and indeed required, not all proposals about the specific distribution of phases are. I will argue that CP phases are necessary, but that selective opacity presents evidence *against* v P phases. This conclusion will be supported by independent evidence from the locality of ϕ -agreement and *wh*-licensing. A core conclusion of chapter 6 is thus that the standard view of phases is in need of revision: CPs are phases, but v Ps are not. I will then show how several purported arguments in favor of v P phases are too weak to establish the presence of v P phases. While they are compatible with v P phases, they do not require them.

Chapter 7 provides independent supporting evidence for this distribution of phase heads. I will start out by noting a bias problem inherent in standard diagnostics for intermediate landing sites. Standard sources of evidence like morpho-syntactic reflexes are able to diagnose the *presence* of an intermediate trace, but they lack the ability to detect the *absence* of such a trace. Positive evidence for the claim that v P is not a phase hence cannot come from such traditional diagnostics.

Based on pioneering work by Gibson & Warren (2004), chapter 7 motivates and explores the use of a novel type of evidence, namely online sentence processing. Two self-paced reading experiments are reported, whose results (i) provide evidence for an intermediate landing site in SpecCP, but (ii) do not show comparable evidence for an intermediate landing site in SpecvP. This asymmetry follows straightforwardly if CPs are phases and vPs are not, a remarkable convergence with the results from unrelated domains reached in chapter 6.

CHAPTER 2

SELECTIVE OPACITY BEYOND MOVEMENT: A HINDI CASE STUDY

2.1 Introduction

As I have discussed in chapter 1, previous investigations into selective opacity effects have almost exclusively focused on movement dependencies. This holds of approaches to the classical ban on superraising (e.g., Chomsky 1973, 1977, 1981, May 1979, Obata & Epstein 2011) and approaches that go beyond the A/\bar{A} -distinction (Sternefeld 1992, Müller & Sternefeld 1993, 1996, Müller 1995, 2014a,b, Grewendorf 2003, 2015, Abels 2007, 2009, 2012a, Neeleman & van de Koot 2010). As a consequence, analyses of selective opacity standardly pursue a movement-based approach to selective opacity. In this section, I will argue that selective opacity effects are not, in fact, restricted to movement. Instead, completely analogous effects can also be observed for *in-situ* relations that do not involve movement. To establish this claim, this chapter will investigate in detail four operations in the syntax of Hindi, focusing in particular on their locality profile with respect to embedded clauses of varying sizes. These four operations are given in (54).

- (54) *Syntactic operations in Hindi*
- a. A-movement
 - b. \bar{A} -movement
 - c. ϕ -agreement
 - d. *wh*-licensing

These four operations are all independent of each other in Hindi. In contrast to English A-movement of an element does not entail ϕ -agreement of this element, nor is it entailed by it. Similarly, *wh*-licensing does not require \bar{A} -movement and \bar{A} -movement can take place in the absence of *wh*-licensing. All four processes are therefore distinct syntactic operations.

Of these four operations, only the first two (A- and \bar{A} -movement) involve movement. I will argue that ϕ -agreement and *wh*-licensing, on the other hand, are established via long-distance Agree, without the necessary mediation of movement.

I will furthermore argue that embedded clauses in Hindi are variable in size and that there is evidence for distinguishing at least three types of clauses: First, finite embedded clauses are full-fledged CPs. Second, nonfinite clauses can be, and in Hindi are, functionally deficient and lack higher clausal projections (see Chomsky 1981, Tappe 1984, Rochette 1988, Fanselow 1989, Moore 1989, 1996, Li 1990, Rosen 1990, Rooryck 1994, Wurmbrand 2001, 2007, Williams 2003, Chung 2004, among others, for analyses along these lines). I will present evidence that nonfinite clauses are structurally ambiguous between a bare TP and a bare ν P structure in Hindi and that the locality of the four operations in (54) is crucially sensitive to this distinction.

The locality pattern that emerges when the four operations in (54) are crossed with the three clause sizes is previewed in (55), where the checkmarks and crosses indicate that a given clause is transparent or opaque to an operation, respectively. \bar{A} -movement has the widest domain of application in that \bar{A} -movement is possible out of clauses of all three structural sizes. *Wh*-licensing, by contrast, is unable to proceed across a finite CP clause boundary, but may cross a TP or ν P clause boundary. Finally, A-movement and ϕ -agreement have the most confined domain of application: Both are possible only out of ν P clauses.

(55) *Preview: Selective opacity across Hindi clauses*

<i>Operation</i>	<i>Landing site/ probe location</i>	<i>Size of embedded clause</i>		
		CP	TP	ν P
\bar{A} -movement	C	✓	✓	✓
<i>wh</i> -licensing	C	✗	✓	✓
A-movement	T	✗	✗	✓
ϕ -agreement	T	✗	✗	✓

(✓ : transparent; ✗ : opaque)

The distribution of opacity in (55) will then lead us to a number of important conclusions regarding the nature of selective opacity effects. First, the Hindi evidence corroborates the conclusion in chapter 1 that selective opacity is *not a binary distinction* between A- and \bar{A} -processes. Rather, three or even more types of locality may coexist in a system. The processes in (55) instantiate three distinct locality profiles and any binary approach to selective opacity will fail to capture this pattern.

Second, both movement and non-movement dependencies partake in selective opacity. The two movement operations in (55) (A- and \bar{A} -movement) and the two non-movement operations (ϕ -licensing and *wh*-licensing) do not form a natural class with respect to their locality. That is, operations that do not involve movement exhibit locality mismatches analogous to those between different types of movement. Moreover, A-movement and ϕ -agreement in fact exhibit the same locality profile, despite the fact that they are separate syntactic operations and one involves movement while the other does not. Overall, these properties of selective opacity in Hindi demonstrate that selective opacity is not confined to movement, but a considerably more general feature of syntactic *operations*. Because previous approaches to selective opacity are couched in terms of movement, these approaches do not extend to selective opacity beyond movement, and hence miss a generalization.

Third, the two crucial generalizations over selective opacity effects discussed in chapter 1 – Upward Entailment and the Height–Locality Connection – also hold for Hindi, thus corroborating their validity. Furthermore, they hold for movement and non-movement operations alike. This result provides further support for the claim that selective opacity is more widespread than it is standardly taken to be in that it is not limited to movement.

In gradually building up towards the summary (55), I will proceed as follows: Section 2.2 introduces a fundamental distinction between embedded finite and nonfinite clauses in Hindi, and illustrates how both interact with A-movement, \bar{A} -movement and ϕ -agreement. Section 2.3 then considers the clause structure of finite and nonfinite clauses and presents novel evidence that nonfinite clauses in Hindi are ambiguous between a TP and a vP structure. Crucially, the locality profiles of the two types of nonfinite clauses differ. Section 2.4 will then bring *wh*-licensing into the picture and show how its domain of application interacts with the three clause sizes previously identified. Finally, section 2.5 will turn to the meta-generalizations identified in chapter 1, namely

Upward Entailment and the Height–Locality Connection. I will show that both generalizations also hold for the four operations under consideration in Hindi.

2.2 The locality profile of finite vs. nonfinite clauses

Hindi is an SOV language with very flexible word order and a rich case and agreement system. Hindi displays a split-ergativity system: External arguments of transitive and some unergative predicates are marked with ergative case in the perfective, but appear without overt case marking in the imperfective/habitual aspect (see, e.g., Pandharipande & Kachru 1977, Davison 1991, 1999, Mohanan 1994, Woolford 2001, Butt & King 2004, de Hoop & Narasimhan 2005, Anand & Nevins 2006, and Keine 2010 for discussion). In addition, Hindi employs differential object marking. Because the precise generalizations are irrelevant for what is to come, it suffices to note that pronouns, proper names and certain animate and/or specific DPs bear the marker *-ko* in object positions. I will gloss *-ko* as ‘accusative’ here, but hasten to add that this is merely a terminological choice and should not be taken as an analytical commitment. Other direct objects, such as weak indefinites, do not bear an overt marker in object position (see, e.g., Mohanan 1994, Bhatt & Anagnostopoulou 1996, Butt & King 2004, and de Hoop & Narasimhan 2005 for discussion of differential object marking in Hindi). Incidentally, the same marker *-ko* is used to mark indirect objects, in which case it is obligatory. The aspect split in the ergativity and the system of differential object marking are independent.

Like many other languages, Hindi invokes a general bipartition between finite and nonfinite embedded clauses. Finite embedded clauses obligatorily occur to the right of the matrix verb, and may be introduced by the complementizer *ki*, which is, however, not obligatory. An example of an embedded finite clause is provided in (56).^{1,2}

- (56) siitaa soc-tii hai [(ki) raam-ne prataap-ko dekh-aa]
Sita think-IPFV.F.SG be.PRES.3SG that Ram-ERG Pratap-ACC see-PFV.M.SG
‘Sita thinks that Ram saw Pratap.’

¹ Unless indicated otherwise, Hindi judgments are due to my informants.

² For the purposes of this section, verbal agreement is irrelevant. I will nonetheless gloss it in detail for the sake of consistency with subsequent sections, where agreement will become of prime importance.

As Hindi is generally a verb-final language, the position of embedded finite clauses to the right of the matrix verb has attracted considerable attention in the literature on Hindi word order. Mahajan (1990) has proposed that finite clauses are base-generated in a preverbal position and obligatorily extraposes to the right. A related analysis is pursued by Bhatt & Dayal (2007), who argue that elements that are ostensibly rightward moved are in fact embedded inside a (remnant) matrix VP, which undergoes extraposition.³ On the other hand, Mahajan (1997), following Kayne's (1994) antisymmetry program, proposes that finite clauses are base-generated in their post-verbal position (see also Simpson & Bhattacharya 2003 for a similar account of the related language Bangla). Finally, Manetta (2012) argues that finite clauses appear in their base position, which is linearized to the right of the verb. As such Manetta (2012) does not subscribe to the antisymmetric view that all internal arguments are base-generated to the right of a verb. For the sake of concreteness, I will follow here Manetta's (2012) analysis that finite clause are base-generated in the same position as nominal objects but linearized to the right of the verb at PF. As such, finite clauses do not move on this account, though little of what I will have to say hinges on this view.

Nonfinite embedded clauses in Hindi exhibit somewhat different properties than finite clauses. In contrast to finite verbs, the infinitival verb lacks tense and aspect marking. With respect to their external syntax, nonfinite clauses appear in the canonical preverbal object position in the unmarked case, though they may also be extraposed under certain information-structural conditions. An example is given in (57):

- (57) siitaa [prataap-ko dekh-naa] caah-tii thii
 Sita Pratap-ACC see-INF.M.SG want-IPFV.F.SG be.PST.F.SG
 'Sita wanted to see Pratap.'

Finite and nonfinite clauses differ in the operations that they are transparent to. The next section illustrates differences in the domain of movement. Following that, section 2.2.2 demonstrates that the two also differ with respect to ϕ -agreement.

³ As Manetta (2012) points out, the fact that finite clauses obligatorily appear to the right would require obligatory VP extraposition if this VP contains an embedded clause on Bhatt & Dayal's (2007) analysis.

2.2.1 A- and \bar{A} -movement

This section will provide some background on A- and \bar{A} -movement in Hindi and then turn to their respective locality profiles.

2.2.1.1 Some background on Hindi movement

As mentioned above, Hindi exhibits a great freedom of the ordering of clausal constituents, first studied in detail by Gambhir (1981). All the permutations in (58) are possible. For a discussion of some of the information-structural factors that influence the appropriateness of the various orders, see Gambhir (1981). Kidwai (2000) provides extensive recent discussion and analysis.

- | | | | |
|------|----|-----------------------|-----|
| (58) | a. | raam-ne kelaa khaayaa | SOV |
| | | Ram-ERG banana ate | |
| | | ‘Ram ate a banana.’ | |
| | b. | raam-ne khaayaa kelaa | SVO |
| | c. | kelaa raam-ne khaayaa | OSV |
| | d. | kelaa khaayaa raam-ne | OVS |
| | e. | khaayaa raam-ne kelaa | VSO |
| | f. | khaayaa kelaa raam-ne | VOS |

(Mahajan 1990: 19–20)

While Hindi thus displays an extremely flexible word order, the discussion in this chapter will focus on movement of an object over the subject and its locality and interpretive properties. Interestingly, inverting the order between direct and indirect object in a ditransitive structure exhibits somewhat different properties and might plausibly be derived by a different mechanism (see Bhatt & Anagnostopoulou 1996 and Bhatt 2016 for discussion).

I will follow here the standard view that divergences from the SOV base order are the result of movement, following much of the previous literature on the issue (Déprez 1989, Mahajan 1990, 1994, 1997, Gurtu 1992, Dayal 1994a, Kidwai 2000, Anand & Nevins 2006, Bhatt & Dayal 2007, Manetta

2011).⁴ Suggestive evidence for this position comes from scope. As (59) illustrates, surface scope is required in the SOV order. (60) shows that the OSV order exhibits scopal ambiguity.⁵

(59) *Scope rigidity without movement*

kisii laṛkii-ne har laṛke-ko dāāṭ-aa
 some girl-ERG every boy-ACC scold-PFV.M.SG
 ‘Some girl scolded every boy.’ (∃ > ∀; *∀ > ∃)

(60) *Movement and scope*

a. sab tiin ciizē khariid-ēge
 everyone three things buy-FUT.M.PL
 ‘Everyone will buy three things.’ (∀ > 3)

b. tiin ciizē_i sab t_i khariid-ēge
 three things everyone buy-FUT.M.PL
 ‘Everyone will buy three things.’ (3 > ∀; ∀ > 3)
(Mahajan 1997: 199)

If both the SOV order and the OSV order were simply base-generated, this asymmetry would be difficult to capture. On the other hand, if the OSV is derived from SOV order by means of moving the object over the subject, the different scope options in (60b) follow immediately from the possibility of reconstructing *har laṛke-ko* ‘every boy-ACC’ into its base position below the subject.⁶

A note on the terminology: Because movement of this type is generally optional in the sense that a version of the sentence without movement is also grammatical, this movement is sometimes

⁴ See, however, Jones (1993) for a differing view, according to which cross-clausal permutations, but not clause-internal ones, are the result of movement.

⁵ In this regard, Hindi is in line with other free word order languages. As discussed by Kiss (1987) and Szabolcsi (1997) for Hungarian, Bayer & Kornfilt (1994) for Turkish, and Frey (1993) and Krifka (1998) for German, base order displays scope rigidity in these languages. See also Hoji (1985) for Japanese.

Due to the entailment relations between the two readings in (60), (60a) does not by itself establish the scope rigidity of the base order.

⁶ Rajesh Bhatt has pointed out to me that scope reconstruction is somewhat more involved than depicted in the main text. In particular, if the object does not contain a numeral but a quantificational determiner like *koi* ‘some’, a reconstructed reading is much harder, if not impossible:

(i) koi kitaab har laṛkaa paṛh-egaa
 some book every boy read-FUT.M.SG
 ‘Some book, every boy will read.’ (∃ > ∀; *∀ > ∃)

Interestingly, a reconstructed reading becomes possible (and in fact forced) if reconstruction is required for some other purpose, like anaphor binding:

referred to as ‘scrambling.’ This term, however, carries a variety of connotations, not the least of which is that this movement lands in a structurally low position, akin to object shift. I will present evidence in section 2.5.2 that this movement lands in a structurally high position in Hindi. In hopes of avoiding inadequate connotations, I will refrain from using the label ‘scrambling’ here and simply refer to the permutations as ‘movement.’

Movement is possible not only within a minimal clause, but may also escape an embedded finite or nonfinite clause. This is illustrated in (61) and (62) for finite clauses.

- (61) dev- ne_i māĩ-ne sun-aa hai [ki t_i miiraa-ko fon nahĩ
 Dev-ERG I-ERG hear-PFV.M.SG be.PRES.3SG that Mira-ACC phone not
 ki-yaa]
 do-PFV.M.SG
 ‘I heard that Dev did not call Mira.’ (Gambhir 1981: 221)

- (62) a. mohan- ko_i raam-ne soc-aa [ki siitaa-ne t_i dekh-aa thaa]
 Mohan-ACC Ram-ERG think-PFV.M.SG that Sita-ERG see-PFV.M.SG be.PST.M.SG
 b. raam-ne mohan- ko_i soc-aa [ki siitaa-ne t_i dekh-aa thaa]
 Ram-ERG Mohan-ACC think-PFV.M.SG that Sita-ERG see-PFV.M.SG be.PST.M.SG
 ‘Ram thought that Sita had seen Mohan.’ (Mahajan 1990: 38)

(63) illustrates extraction out of nonfinite clauses. It is irrelevant whether the nonfinite clause appears in the pre-verbal or post-verbal position:

- (63) a. prataap- ko_i siitaa [t_i dekh-naa] caah-tii thii
 Pratap-ACC Sita see-INF.M.SG want-IPFV.M.SG be.PST.M.SG
 b. prataap- ko_i siitaa t_j caah-tii thii [t_i dekh-naa] _{j}
 Pratap-ACC Sita want-IPFV.M.SG be.PST.M.SG see-INF.M.SG
 ‘Sita wanted to see Pratap.’

-
- (ii) [apnii _{i} kooi kitaab] har laṛkaa _{i} paṛh-egaa
 self’s some book every boy read-FUT.M.SG
 ‘Every boy _{i} will read some of his _{i} books.’ (* $\exists > \forall$; $\forall > \exists$)

It is clear, then, that reconstruction is possible in principle even if the determiner is *kooi*, supporting the general claim that these structures too are formed via movement. Why reconstruction is impossible in (i) is an unresolved question.

Another complicating factor is explored by Anand & Nevins (2006), who argue that the scope rigidity in the base order is limited to cases in which the subject bears ergative case and that inverse scope is possible without overt movement if the subject bears a zero case. Their account is compatible with the movement-based view adopted here.

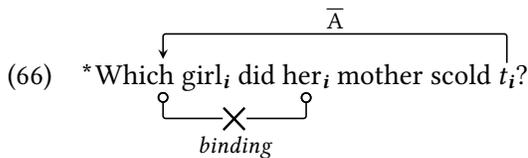
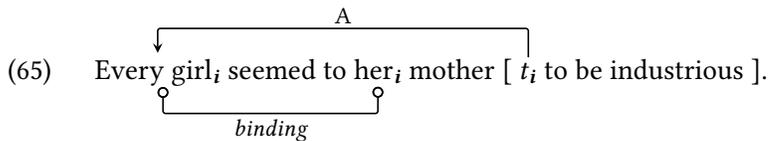
2.2.1.2 Diagnosing A-movement

It is widely known that movement displays somewhat different properties in these constructions (see Gurtu 1985, 1992, Déprez 1989, Mahajan 1990, 1994, Jones 1993, Dayal 1994a, Kidwai 2000, Bhatt 2016). Movement within a simple clause exhibits a variety of A-properties. One such property is **WEAK CROSSOVER**. A standard property of A-movement is that it can feed pronominal binding, whereas \bar{A} -movement cannot (Postal 1971, Wasow 1972):

(64) **Crossover**

An A-moved element can bind a pronoun from its landing site; an \bar{A} -movement cannot do so.

This distinction is exemplified for English in by the contrast between (65) and (66):



Returning to Hindi, the examples in (67) and (68) demonstrate that an object cannot bind a pronoun inside the local subject from its base position, due to lack of c-command. If the object is moved above the subject, however, such binding becomes possible (Gurtu 1985, 1992, Déprez 1989, Mahajan 1990, 1994, Jones 1993, Dayal 1994a, Kidwai 2000, Bhatt & Dayal 2007, Bhatt 2016). This generalization holds for a variety of non-referential expressions, including quantificational DPs such as *har laṛke-ko* ‘every boy’ in (67) and *kis-ko* ‘who’ in (68).⁷ The availability of pronominal binding from the landing site of clause-internal movement is thus on a par with A-movement in English (65).

(67) *Weak crossover obviation*

- a. [*us-kii*_{*i/j} *bahin-ne*] *har laṛke-ko*_i *dekh-aa*
 3SG-GEN sister-ERG every boy-ACC see-PFV.M.SG
 ‘His/her_{*i/j} sister saw every boy_i.’ (*bound reading impossible*)

⁷ Hindi is a *wh*-in-situ language (see section 2.4.1), which does, however, allow *wh*-elements and foci to scramble freely (Kidwai 2000: 36). The fronting *kis-ko* ‘who-ACC’ in (68b) is hence not an instance of *wh*-movement but the same movement that affects *har laṛke-ko* ‘every boy-ACC’ in (67b).

- b. har laṛke-ko_i [us-kii_i bahin-ne] t_i dekh-aa
 every boy-ACC 3SG-GEN sister-ERG see-PFV.M.SG
 ‘For every boy *x*, *x*’s sister saw *x*.’ (based on Kidwai 2000: 7)

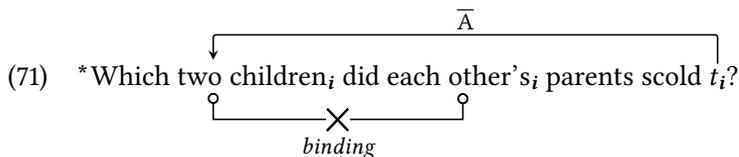
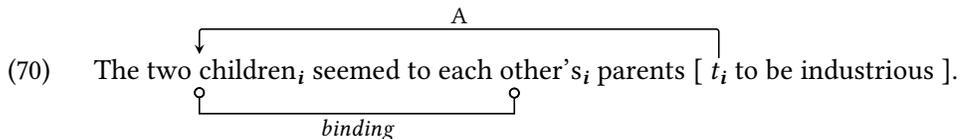
(68) *Weak crossover obviation*

- a. [us-kii_{i/j} bahin] kis-ko_i pyaar kar-tii thii?
 3SG-GEN sister who-ACC love do-IPFV.F.SG be.PST.F.SG
 ‘Who_i did his/her_{i/j} sister love?’ (bound reading impossible)
- b. kis-ko_i [us-kii_i bahin] t_i pyaar kar-tii thii?
 who-ACC 3SG-GEN sister love do-IPFV.F.SG be.PST.F.SG
 ‘For which *x*, *x*’s sister loved *x*?’ (based on Mahajan 1990: 25–26)

Pronominal binding is not the only diagnostic indicating that clause-internal A-movement is possible in Hindi. Converging evidence comes from **RECIPROCAL BINDING**. As the contrast between (70) and (71) demonstrates for English, A-movement can feed binding of a reciprocal, while \bar{A} -movement cannot.

(69) **Reciprocal binding**

An A-moved element may bind a reciprocal pronoun from its landing site; an \bar{A} -moved element may not.



Applying reciprocal binding as a test for A-movement to Hindi corroborates the conclusion drawn from pronominal binding above: Clause-internal movement can feed reciprocal binding and thus qualifies as A-movement (Jones 1993: 80, Bhatt & Dayal 2007: 289, Bhatt 2016: 515):⁸

⁸ Dayal (1994a) gives the example in (i), in which the reciprocal is itself the subject instead of embedded inside it. She judges binding marginal in this case:

(72) *Reciprocal binding*

- a. * [*ek-duusre-kii_i* bahinõ-ne] [raam aur prataap]-ko_i maar-aa
each other's sisters-ERG Ram and Pratap -ACC hit-PFV.M.SG
'*Each other's_i sisters hit [Ram and Pratap]_i.'
- b. [raam aur prataap]-ko_i [*ek-duusre-kii_i* bahinõ-ne] t_i maar-aa
Ram and Pratap -ACC each other's sisters-ERG hit-PFV.M.SG
'Ram and Pratap were hit by each other's sisters.'

Mahajan (1990, 1994) discusses a third A-diagnostic in Hindi, which is based on binding of the reflexive pronoun *apnaa*. He provides the example in (73):⁹

- (73) a. */??? [*apne_i* baccõ-ne] mohan-ko_i ghar se nikaal di-yaa
self's children-ERG Mohan-ACC house from throw give-PFV.M.SG
Intended: 'Her_i children threw Mohan_i out of the house.'
- b. ?mohan-ko_i [*apne_i* baccõ-ne] t_i ghar se nikaal di-yaa
Mohan-ACC self's children-ERG house from throw give-PFV.M.SG
'Mohan_i was thrown out of the house by her_i children.'

(examples and judgments from Mahajan 1990: 32–33)

The pattern of judgments in (73) is essentially equivalent to that in (72) for reciprocals and therefore would constitute further evidence for the A-character of clause-internal movement in Hindi. It

-
- (i) ? [jaun aur merii]-ko_i *ek-duusre-ne_i* t_i dekh-aa
John and Mary -ACC each other-ERG see-PFV.M.SG
'John and Mary saw each other.'

(Dayal 1994a: 252)

Kidwai (2000: 31–32) finds binding impossible in an example structurally analogous to (i). I will follow Bhatt (2016: 515n4) that the reciprocal is banned from occupying the subject position in Hindi. The example in (i) is then ruled out for reasons independent of binding. It is thus necessary to use slightly more complicated structures in which the reciprocal is embedded inside the subject to avoid this confound, as in (72). In such structures, binding is widely accepted if movement takes place. Bhatt (2016: 515) gives the example in (ii):

- (ii) a. * [*ek-duusre-ke_i* fans-ne] [Katrina aur Saif]-ko_i pehcaan li-yaa
each other-GEN fans-ERG Katrina and Saif -ACC recognize take-PFV.M.SG
'*Each other's fans recognized Katrina and Saif.'
- b. [Katrina aur Saif]-ko_i [*ek-duusre-ke_i* fans-ne] t_i pehcaan li-yaa
Katrina and Saif -ACC each other-GEN fans-ERG recognize take-PFV.M.SG
'Katrina and Saif were recognized by each other's fans.'

(ii) is analogous to the crucial data point in (72).

⁹ *Apnaa* surfaces as *apne* in (73) because it agrees in ϕ -features with the head noun *baccõ-ne* 'children-ERG'. I will refer to the reflexive in its citation form *apnaa* in the text.

should be noted, however, that the judgment in (73b) is contested in the literature. While a bound reading is marginally acceptable for Mahajan (1990, 1994), Dayal (1994a) gives the structurally analogous example in (74), which she judges as ungrammatical:

- (74) *mohan-ko_i [apne_i baccō-ne] t_i maar-aa
 Mohan-ACC self's children-ERG beat-PFV.M.SG
Intended: 'Her_i children beat Mohan_i.' (example and judgment from Dayal 1994a: 242)

The same judgment is reported by Jones (1993: 80). Dayal (1994a) takes the ungrammaticality of binding in (74) as evidence that the pre-subject position targeted by movement cannot be an A-position. I will not follow Dayal's (1994a) conclusion. First, the ability of such movement to obviate weak crossover and result in reciprocal binding are not contested in the literature, and both would be unaccounted for if this movement were \bar{A} -movement (as Dayal acknowledges). Second, the dialectal split in whether clause-internal movement can lead to anaphor binding interestingly appears to correlate with the coreference options of *apnaa* in the absence of movement. Consider the ditransitive structure in (75) (from Dayal 1994a: 250):

- (75) raam-ne_i mohan-ko_j [apnii kitaab] dii
 Ram-ERG Mohan-DAT self's book give.PFV.F.SG
 'Ram gave self's book to Mohan.' (Mahajan: *apnii_{i/j}*; Dayal: *apnii_{i/*j}*)

For Gurtu (1985, 1992) and Mahajan (1990, 1994) it is possible for the reflexive embedded within the direct object to be bound by the indirect object (*mohan-ko* in (75)). For other speakers, e.g., Dayal (1994a) and Kidwai (2000: 72), *apnaa* is subject-oriented and can only be bound by *raam-ne* 'Ram-ERG' in (75). As a consequence, in the latter variety, binding from an A-position is a *necessary* condition for binding of *apnaa*, but not a *sufficient* one. This fact provides us with an immediate explanation for why object movement cannot feed anaphor binding in this dialect, as exemplified in (74): *mohan-ko* is not a subject in (74) and is consequently unable to bind a subject-oriented anaphor, regardless of whether its landing site is an A- or an \bar{A} -position. Because (74) is therefore ruled out for reasons independent of the A/ \bar{A} -status of *mohan-ko*, reflexive binding is simply not an adequate testing ground for the nature of clause-internal movement, at least in Dayal's (1994a) variety, which appears to be the majority judgment. Due to this dialectal confound and because

binding of *apnaa* by non-subjects is degraded even for speakers who in principle allow it, I will put aside binding of *apnaa* in what is to follow.¹⁰

In sum, there is good evidence that clause-internal movement in Hindi has A-characteristics. The relevant evidence comes from weak crossover and reciprocal binding. A potential third diagnostic – reflexive binding – is confounded for many speakers and I will hence not make use of it.

Having thus established two general diagnostics for A-movement in Hindi, I will now turn to the properties of cross-clausal movement. We will first consider movement out of finite clauses and then move on to movement out of nonfinite clauses.

2.2.1.3 Movement out of finite clauses

Movement out of a finite clause contrasts with clause-internal movement in that it does not display A-behavior (Gurtu 1985, 1992, Déprez 1989, Mahajan 1990, 1994, Jones 1993, Bhatt 2016). I will illustrate this fact using weak crossover and reciprocal binding.

The example in (76) demonstrates extraction out of finite clause over a matrix subject that contains a pronoun. (76a) demonstrates that such movement is well-formed as long as the moved element *har laṛke-ko* ‘every boy-ACC’ does not bind the pronoun *uskii* ‘his/her’. In (76b), on the other hand, it is attempted to bind the pronoun from the landing site and ungrammaticality results. Extraction out of a finite clause is thus subject to weak crossover and hence \bar{A} -movement. This stands in stark contrast to clause-internal movement, which obviates weak crossover, as we saw in (67).

(76) *Movement out of finite clause: Weak crossover*

- a. har laṛke-ko_i [us-kii_j bahin] soc-tii hai [ki raam-ne t_i
 every boy-ACC 3SG-GEN sister think-IPFV.F.SG be.PRES.3SG that Ram-ERG
 dekh-aa]
 see-PFV.M.SG
 ‘His/her_j sister thinks that Ram saw every boy_i.’

¹⁰ See Bhatia & Poole (2015) for a recent treatment of the subject orientation of *apnaa* that is compatible with the assumptions made here.

- b. *har larke-ko_i [us-kii_i bahin] soc-tii hai [ki raam-ne t_i
 every boy-ACC 3SG-GEN sister think-IPFV.M.SG be.PRES.3SG that Ram-ERG
 dekh-aa]
 see-PFV.M.SG
Intended: ‘For every boy *x*, *x*’s sister thinks that Ram saw *x*.’

Another triplet contrasting the A-character of clause internal movement with the \bar{A} -nature of movement that leaves a finite clause is provided in (77):

- (77) a. *[us-kii_i bahin] harek-ko_i pasand kar-tii hai
 3SG-GEN sister everyone-ACC like do-IPFV.F.SG be.PRES.3SG
 ‘*His/her_i sister likes everyone_i.’
- b. harek-ko_i [us-kii_i bahin] t_i pasand kar-tii hai
 everyone-ACC 3SG-GEN sister like do-IPFV.F.SG be.PRES.3SG
 ‘For every *x*, *x*’s sister likes *x*.’
- c. *harek-ko_i [us-kii_i bahin-ne] kah-a [ki merii t_i pasand
 everyone-ACC 3SG-GEN sister-ERG say-IPFV.M.SG that Mary t_i like
 kar-tii hai]
 do-IPFV.F.SG be.PRES.3SG
Intended: ‘For every *x*, *x*’s sister said that Mary loves *x*.’

(based on Gurtu 1992: 99,100,103)

Converging evidence for the same conclusion comes from reciprocal binding. As (78) shows, movement out of *raam aur prataap-ko* ‘Ram and Pratap-ACC’ cannot result in binding of the reciprocal from the landing site in the higher clause:

- (78) *Movement out of finite clause: No reciprocal binding*
- *[raam aur prataap]-ko_i [ek-duusre-kii_i bahinõ-ne] soc-aa [ki
 Ram and Pratap -ACC each other’s sisters-ERG think-PFV.M.SG that
 sangiitaa-ne t_i maar-aa]
 Sangita-ERG hit-PFV.M.SG
 ‘*Each other’s_i sisters thought that Sangita had hit [Ram and Pratap]_i.’

The unavailability of binding in (78) again stands in direct contrast to the fact illustrated in (72) above that such binding is possible if movement is clause-internal.

2.2.1.4 Movement out of nonfinite clauses

While the different properties of movement within a clause and movement that leave a finite clause has received a considerable amount of attention in the syntactic literature on Hindi, movement out of nonfinite clauses has received very little systematic attention. To the best of my knowledge, the only attempt to apply A-diagnostics to such movement is Keine (2013), which I will extend here. As (79b) shows, movement of *har laṛke-ko* ‘every boy-ACC’ out of a nonfinite clause may feed pronominal binding from its landing site in the matrix clause. Hence, such movement can obviate weak crossover, hence diagnosing it as A-movement. The bound interpretation in (79b) is crucially a result of movement, as shown by the non-movement baseline in (79a), where the pronoun may only have a referential interpretation.

(79) *Movement out of nonfinite clause: No weak crossover*

- a. [*us-kii*_{*i/j*} bahin] [har laṛke-ko_{*i*} dekh-naa] caah-tii thii
 3SG-GEN sister every boy-ACC see-INF.M.SG want-IPFV.F.SG be.PST.F.SG
 ‘His/her _{*i/j*} sister wanted to see every boy_{*i*}.’
- b. har laṛke-ko_{*i*} [*us-kii*_{*i*} bahin] [_{*t*}_{*i*} dekh-naa] caah-tii thii
 every boy-ACC 3SG-GEN sister see-INF.M.SG want-IPFV.F.SG be.PST.F.SG
 ‘For every boy *x*, *x*’s sister wants to see *x*.’

A-extraction is also possible out of nonfinite clauses in a post-verbal, extraposed position:

- (80) a. har kuttaa_{*i*} [*us-ke*_{*i*} malik-ne] [_{*t*}_{*i*} ghumaa-naa] caah-aa
 every dog 3SG-SG owner-ERG walk-INF.M.SG want-PFV.M.SG
- b. har kuttaa_{*i*} [*us-ke*_{*i*} malik-ne] _{*t*}_{*j*} caah-aa [_{*t*}_{*i*} ghumaa-naa]_{*j*}
 every dog 3SG-GEN owner-ERG want-PFV.M.SG walk-INF.M.SG
 ‘For every dog *x*, *x*’s owner wanted to walk *x*.’

Reciprocal binding points towards the same conclusion, as illustrated in (81). Here movement of the embedded object *raam aur prataap-ko* ‘Ram and Pratap-ACC’ into the matrix clause results in a bound interpretation of the reciprocal inside the matrix subject (81b), an interpretation that is impossible in the absence of movement (81a).

(81) *Movement out of nonfinite clause: Reciprocal binding*

- a. * [*ek-duusre-kii_i* bahinõ-ne] [[raam aur prataap]-ko_i maar-naa]
each other's sisters-ERG Ram and Pratap -ACC hit-INF.M.SG
caah-aa
want-PFV.M.SG
'*Each other's_i sisters wanted to hit [Ram and Pratap]_i.'
- b. [raam aur prataap]-ko_i [*ek-duusre-kii_i* bahinõ-ne] [*t_i* maar-naa]
Ram and Pratap -ACC each other's sisters-ERG hit-INF.M.SG
caah-aa
want-PFV.M.SG
'[Ram and Pratap]_i each other's_i sisters wanted to hit (them)_i.'

The evidence from weak crossover in (79) converges with the evidence from reciprocal binding in (81) in showing that A-movement in Hindi is not simply clause-bounded, but can proceed out of nonfinite clauses.

I conclude from these considerations, essentially following Mahajan (1990, 1994), that movement in Hindi is not a uniform phenomenon. Rather, movement in Hindi can be either A- or \bar{A} -movement, which display distinct properties:

(82) *Movement properties*

- a. *A-movement*:
- (i) is not subject to weak crossover,
 - (ii) can feed reciprocal binding,
 - (iii) may not leave a finite clause
- b. \bar{A} -movement:
- (i) is subject to weak crossover,
 - (ii) may not feed reciprocal binding,
 - (iii) may leave a finite clause

While there are reliable diagnostics for A-movement in Hindi, \bar{A} -movement is much harder to diagnose in domains in which A-movement is also possible. We have seen that movement within a finite clause and out of a nonfinite clause can reconstruct, but it does not need to. This fact does not necessarily constitute a diagnostic for \bar{A} -movement, as there is ample evidence that reconstruction is a property of both A- and \bar{A} -movement (Barss 1986, Romero 1997, Fox 2000, Sportiche 2006,

Lebeaux 2009). Moreover, because Hindi allows free argument drop, it is difficult to diagnose parasitic gaps. Interestingly, Manetta (2016) argues for the existence of parasitic gaps in Hindi and the possibility of \bar{A} -movement within simple clauses. Thus, while only \bar{A} -movement may leave a finite clause in Hindi, there is no inverse requirement that \bar{A} -movement *must* leave a finite clause. It may be local too. Consequently, clause-internal movement and movement out of a nonfinite clause can be A- or \bar{A} -movement, whereas movement out of a finite clause is invariably \bar{A} -movement. The facts observed in this section then follow.

(83) *Locality of A- and \bar{A} -movement* (to be refined)

Nonfinite clauses allow A- as well as \bar{A} -movement out of them. Finite clauses only allow \bar{A} -extraction out of them.

In section 2.5.2 below, I will show that A- and \bar{A} -movement also differ in the position they target. Specifically, I will provide evidence that A-movement lands in SpecTP, while \bar{A} -movement targets SpecCP, much like their English counterparts. In the remainder of this thesis, I will focus on their landing sites and the locality properties of A- and \bar{A} -movement. I will use their behavior with respect to crossover and reciprocal binding as a diagnostic that allows us to distinguish between A- and \bar{A} -movement, but I will have nothing to say here about *why* A-movement is not subject to weak crossover and can bind reciprocal pronouns, but \bar{A} -movement cannot. See section 1.7 in chapter 1 for additional discussion.

We have thus arrived at the first instance of selective opacity in Hindi: Finite clauses are transparent to \bar{A} -movement, but opaque to A-movement, in a way that mirrors the English facts. The next section will bring ϕ -agreement into the picture.

2.2.2 Local and long-distance ϕ -agreement

We now turn to the verbal ϕ -agreement system in Hindi. I will start out by giving some relevant background on agreement in simple clauses and will then move on to the locality of cross-clausal agreement.

2.2.2.1 Some background on local agreement

Verbs in Hindi agree with the structurally highest unmarked argument (Pandharipande & Kachru 1977, Mahajan 1989, Butt 1993, Mohanan 1994). As an absolute rule, only nominals that do not bear an overt case marker can control verb agreement.¹¹ If the subject of a clause is not overtly case-marked, it obligatorily controls verb agreement. If the subject carries a case marker and the object does not, then the verb agrees with the object. Finally, if both the subject and the object are overtly case-marked, the verb shows masculine singular default agreement. This algorithm is stated in (84):

(84) **Hindi ϕ -agreement algorithm**

If the subject does not bear a case marker → agree with the subject

Otherwise: If object does not bear a case marker → agree with the object

Otherwise: Use masculine singular default agreement.

The main verb and all auxiliaries and light verbs all agree with the nominals determined by (84). Verbal agreement is principally for person, number, and gender, though verbal paradigms typically lack exponence for one or more of these features. Illustrative examples are provided in (85). In (85a), the subject is not overtly case-marked and therefore controls agreement on the matrix verb and the auxiliary. In (85b), the subject bears ergative case and is hence ineligible for agreement. Because the object is null marked, the verb agrees with it instead. Finally, in (85c) both the subject and the object are overtly case-marked. As a result, the verb agrees with neither, exhibiting default agreement instead.

(85) a. *Subject zero-marked* → *subject agreement*

laṛke is kitaab-ko paṛh-te hāĩ
boys.M.PL this book.F.SG-ACC read-IPFV.M.PL be.PRES.3PL
'The boys are reading this book.'

b. *Subject overtly marked, object zero-marked* → *object agreement*

laṛkō-ne yah kitaab paṛh-ii
boys.M.PL-ERG this book.F.SG read-PFV.F.SG
'The boys read this book.'

¹¹ For our present concerns, it is inconsequential whether this restriction is attributed to a case hierarchy (Bobaljik 2008, Preminger 2011, 2014), to the postpositional status of the case markers (Butt & King 2004, Spencer 2005) or to yet some other factor.

- c. *Both subject and object overtly marked* → *default agreement*

laṛkō-ne is kitaab-ko paṛh-aa
 boys.M.PL-ERG this book.F.SG-ACC read-PFV.M.SG
 ‘The boys read this book.’

There is no evidence that elements that trigger verbal agreement sit in a dedicated structural position. The familiar kind of word order evidence with respect to adverbs etc. in English or Romance does not indicate any difference in the position of agreeing and non-agreeing subjects and objects, and neither do familiar tests like reflexive binding, control, etc. I will hence take it that no movement for agreement takes place.

The agreement algorithm in (84) is deterministic. That is, there is never any optionality in clause-internal agreement. No agreement patterns for the sentences in (85) other than the ones indicated are grammatical.

Movement and ϕ -agreement are independent of each other. Consider the configurations in (86). Here, both the subject and the object are not overtly case marked and hence in principle eligible for verb agreement. By (84), the verb has to agree with the subject. This also holds if the object is moved over the subject, as in (86), and even if this movement is clearly A-movement (86b). Subject agreement is the only agreement option in (86).

- (86) *(A-)Movement does not affect local ϕ -agreement*

- a. kitaab_i laṛke t_i paṛh-te hãĩ
 book.F.SG boys.M.PL read-IPFV.M.PL be.PRES.3PL
 ‘The boys are reading a book.’
- b. har gaar_{ii} [us-kaa_i maalik (=hii)] t_i saaf kar-egaa
 every car.F 3SG-GEN owner.M(=only) clean do-FUT.M.SG
 ‘Every car *x* will be cleaned by *x*’s owner (not by anybody else).’

This point will be taken up in greater detail in section 2.5.2. For now, the independence of (A-)movement and ϕ -agreement shows that, unlike in English, A-movement and ϕ -agreement are *separate syntactic operations* in Hindi, which usually target distinct elements.

2.2.2.2 Long-distance agreement (LDA)

While the verb in Hindi standardly agrees with one of its own arguments, it is also possible in certain configurations for a verb to agree with the object of an embedded nonfinite clause (Mahajan 1989, Davison 1991, Butt 1993, 1995, Boeckx 2004, Frank 2004, Bhatt 2005, Chandra 2007, Keine 2013). Following standard terminology, I will refer to such cross-clausal agreement as **LONG-DISTANCE AGREEMENT (LDA)**. A representative example is provided in (87):

- (87) *Long-distance agreement (LDA)*
laṛkō-ne [roṭii khaa-nii] caah-ii
boys-ERG bread.F eat-INF.F.SG want-PFV.F.SG
'The boys wanted to eat bread.'

In (87), the object *roṭii* 'bread' of the embedded clause controls agreement on the matrix verb *caah* 'want'. Concomitant is agreement with the same DP on the infinitival verb *khaa* 'eat'. In striking contrast to clause-internal agreement, LDA is generally optional. Thus, the LDA version in (87) coexists with (88), in which both the matrix and the embedded verb show masculine singular default agreement.

- (88) *Default agreement*
laṛkō-ne [roṭii khaa-aa] caah-aa
boys-ERG bread.F eat-INF.M.SG want-PFV.M.SG
'The boys wanted to eat bread.'

As a general fact, there is a bidirectional implication between LDA on the matrix verb and agreement on the infinitival verb (Mahajan 1989: 234, Bhatt 2005: 761, Chandra 2007: 46). Thus, neither is it possible for the matrix verb to show LDA without being accompanied by infinitival agreement (89a), nor is it possible to have infinitival agreement without LDA (89b).

- (89) *LDA and infinitival agreement entail each other*¹²
a. *laṛkō-ne [roṭii khaa-naa] caah-ii
boys-ERG bread.F eat-INF.M.SG want-PFV.F.SG

¹² There is some variation with respect to the status of (89b). While Mahajan (1989: 234–235) and Bhatt (2005: 785) judge it as ungrammatical, Mahajan (1989: 235) notes that some speakers accept it. Bickel & Yadava (2000: 356) and Butt (1993: 77) report them as grammatical. This divergence is presumably an instance of dialectal variation, which I will put aside here.

- b. *laṛkō-ne [roṭii khaa-nii] caah-aa
 boys-ERG bread.F eat-INF.F.SG want-PFV.M.SG
Intended: ‘The boys wanted to eat bread.’

Despite the bidirectional entailment between LDA and infinitival agreement, there is evidence that agreement on the infinitive is not established independently, but rather as a side product of agreement with the matrix verb, a conclusion also reached by (Bhatt 2005). The evidence comes from configurations in which the matrix subject does not bear ergative case. In (90), the aspect of the matrix clause is switched from perfective, as in (87–89), to the imperfective. Due to Hindi’s split-ergativity system, the ergative marking of the matrix subject disappears. The absence of overt case marking then renders this subject visible for verbal agreement. In other words, both the local subject *laṛke* ‘boys’ and the embedded object *roṭii* ‘bread’ are unmarked and hence in principle able to trigger agreement. As (90) shows, the matrix verb has to locally agree with its subject in this case, a point already observed by Mahajan (1989), as illustrated in (90a) (we will turn to the agreement of the infinitival verb shortly). LDA with the embedded object *roṭii* ‘bread’ is thus impossible (90b), and so is default agreement (90c).

(90) *Local agreement preempts LDA*

- a. *Local subject agreement*
 laṛke [roṭii khaa-naa] caah-te hāī
 boys bread.F eat-INF.M.SG want-IPFV.M.PL be.PRES.3PL
 ‘The boys want to eat bread.’
- b. *LDA*
 *laṛke [roṭii khaa-naa] caah-tii hai
 boys bread.F eat-INF.M.SG want-IPFV.F.SG be.PRES.3SG
- c. *Default agreement*
 *laṛke [roṭii khaa-naa] caah-ta hai
 boys bread.F eat-INF.M.SG want-IPFV.M.SG be.PRES.3SG

Thus, LDA is possible only if there is no local DP that the matrix verb can agree with. This preference is of course not surprising in light of the general preference for subject over object agreement.

Having thus established that local agreement preempts LDA, let us now consider infinitival agreement in these configurations. Consider the paradigm in (91). Just as in (90), the matrix subject

is not overtly case marked. As a result, the matrix verb has to agree with it. In this case, the infinitival verb cannot agree at all in this configuration: it can agree with neither the embedded object *tehnii* ‘branch’, as in (91a), nor with the matrix subject, as in (91b), but instead has to appear in the masculine singular default agreement form as in (91c).

(91) *Infinitival agreement options in the absence of LDA*

a. *Object agreement on infinitive is impossible*

*larke [tehnii kaaṭ-nii] caah-te the
 boys branch.F cut-INF.F.SG want-IPFV.M.PL be.PST.M.PL

b. *Subject agreement on infinitive is impossible*

*larke [tehnii kaaṭ-ne] caah-te the
 boys branch.F cut-INF.M.PL want-IPFV.M.PL be.PST.M.PL

c. *Default agreement on infinitive is possible*

larke [tehnii kaaṭ-naa] caah-te the
 boys branch.F cut-INF.M.SG want-IPFV.M.PL be.PST.M.PL

These facts reveal that infinitival agreement is parasitic on LDA: the infinitival verb can agree with its object *only if* the matrix verb also agrees with it, a conclusion also reached by Boeckx (2004) and Bhatt (2005). Whenever the matrix verb does not agree with the embedded object – either because it agrees locally (91a) or because it does agree at all (89b) –, the embedded verb is incapable of agreeing with any element.

Additional support for the parasitic nature of the infinitival agreement comes from clauses in subject position. Unlike object clauses, subject clauses do not allow LDA into them in Hindi. Thus, in (92), *mehnat* ‘hard work’ cannot control agreement on the matrix predicate. What is critical for our present concerns is that agreement between *mehnat* and the infinitival verb *kar* ‘do’ is likewise ruled out in this case. All combinations of agreement other than masculine singular agreement on both verbs are grammatical.

(92) [mehnat kar-naa/*-nii] acch-aa/*-ii ho-taa/*-tii
 hardwork.F do-INF.M.SG/*-INF.F.SG good-M.SG/*-F.SG be-IPFV.M.SG/*-IPFV.F.SG
 hai
 be.PRES.3SG

‘It is good to work hard.’

(Bhatt 2005: 777)

Both pieces of evidence demonstrate that the decision of whether or not the infinitival verb agrees cannot be made locally, because the information that conditions the choice – i.e., whether or not the matrix verb agrees – is unavailable until LDA is established. Considerations of cyclicity thus lead to the conclusion that infinitival agreement is a side effect of agreement between the matrix verb and the embedded object and not a precondition for it.

This conclusion is empirically supported by properties of the ‘permissive’ construction (see Butt 1993, 1995, 2014, Davison 2014 for much discussion of this construction). In this construction, the matrix verb *de* ‘allow’ (lit. ‘let’) embeds a nonfinite clause and, like the examples just discussed, allows the matrix verb to agree with the embedded object.¹³ (93) provides an example. A distinctive feature of the permissive construction is that the infinitival verb does not share the agreement of the matrix predicate. Thus, in (93), the embedded verb cannot bear agreement morphology (i.e., *caaa-nii* ‘drive-INF.F.SG’). Instead, it appears in the invariant form *-ne*, regardless of whether the matrix verb shows LDA or default agreement.

- (93) a. anjum-ne saddaf-ko [gaar̥ii calaa-ne] d-ii
 Anjum-ERG Saddaf-dat car.F drive-INF give-PFV.F.SG
 ‘Anjum let Saddaf drive the car.’
- b. *anjum-ne saddaf-ko [gaar̥ii calaa-nii] d-ii
 Anjum-ERG Saddaf-dat car.F drive-INF.F.SG give-PFV.F.SG (Butt 1993: 60)

The fact that LDA is possible in the absence of infinitival agreement provides an additional argument that LDA is not derivationally dependent on infinitival agreement.

In sum, we have seen converging evidence that infinitival agreement is parasitic on LDA rather than the other way around. I will thus ignore for now the proper treatment of the infinitival agreement and focus on the agreement value of the matrix verb instead. Infinitival agreement can then be treated as a side effect of this agreement value for those verbs that do allow it. See Boeckx (2004) and Bhatt (2005) for proposals along these lines, using Multiple Agree and feature sharing, respectively.

¹³ Bhatt (2005: 795n11) reports some speaker variation with respect to whether LDA in the permissive is obligatory or optional.

2.2.2.3 Long-distance agreement is long-distance

I will now turn to the question of whether LDA in Hindi involves genuine agreement past a nonfinite clause boundary or whether, despite appearance, this agreement is established in a more local configuration. I will argue that LDA is genuinely non-local in that it can be established across a clause boundary. In order to establish this claim, I will consider two alternative treatments that have been proposed in the literature.

One line of analysis that treats LDA as local agreement has been proposed for Hindi by Mahajan (1989) and Chandra (2007). Both suggest that in LDA configurations the embedded object string-vacuously moves into the matrix and subsequently controls local agreement from its landing site. Agreement that genuinely crosses a clause boundary is ruled out on these accounts. I will refer to this line of analysis as the **MOVEMENT ACCOUNT**.

One prediction of the movement account is that overt movement of the embedded object into the matrix clause should render LDA obligatory. This is not the case. In (94), the embedded object *roṭii* ‘bread’ has been unambiguously moved into the matrix clause. This movement does not, however, affect the agreement options. Both long-distance agreement as well as default agreement are possible, the same options as in the absence of movement (compare (87–88)).

- (94) *roṭii_i* *laṛkō-ne* [*t_i* *khaa-naa/-nii*] *caah-aa/-ii*
bread.F boys-ERG eat-INF.M.SG/-INF.F.SG want-PFV.M.SG/PFV.F.SG
‘Bread, the boys wanted to eat.’

While the lack of movement and LDA does not conclusively rule out a movement account,¹⁴ (94) does show that there is no independent empirical motivation for linking LDA to cross-clausal movement.

The previous literature on Hindi LDA has accumulated a number of strong arguments in favor of the view that LDA does not depend on movement (at least in Hindi). Davison (1991) provides the example in (95), in which the embedded clause containing the embedded object *saaikal* ‘bicycle’ is extraposed to the right of the matrix verb. Despite the fact that *saaikal* is unambiguously inside the

¹⁴ A proponent of a movement analysis might, of course, invoke differences between movement types and claim that only some movement types can feed agreement, while others cannot. The lack of interaction between movement and agreement in (94) could then be attributed to ambiguity with respect to the movement involved. However, such an account would then have to demonstrate that movement dependencies that are unambiguous do in fact interact with LDA in the predicted manner. I am not aware of any such evidence.

embedded clause, it can control LDA. This indicates that LDA is not dependent on movement of the object into the matrix clause.

- (95) mujhe zaruur t_i aa-tii hai [saaikal calaa-nii]_i
 I.DAT surely come-IPFV.F.SG be.PRES.3SG bicycle.F ride-INF.F.SG
 ‘I certainly know how to ride a bicycle.’ (Davison 1991: (18))

Another argument against the movement account due to Bhatt (2005) comes from adverbs. In the example in (96), the embedded clause contains the adverb *phir-se* ‘again’, preceding the embedded object *mehnat* ‘hard work’. If *phir-se* is part of the embedded clause, then *mehnat* must likewise be. Nonetheless, LDA with *mehnat* is possible. As before, no object is implicated in this LDA configuration.

- (96) rohan-ne aaj [phir-se mehnat kar-nii] caah-ii
 Rohan-ERG today again hardwork.F do-INF.F.SG want-PFV.F.SG
 ‘Today Rohan wanted to work hard again.’ (Bhatt 2005: 766)

Another argument, due to Bhatt (2005), comes from weak indefinites and light verb constructions (see also Davison 1991). In (97a), LDA is controlled by *daal* ‘lentils’. Importantly, *daal* can be interpreted as a weak indefinite. That is, (97a) does not commit one to the claim that there are particular lentils that Usha had wanted to make. Rather, (97a) can be used to describe Usha’s intention of ‘daal-making’. If *daal* had to obligatorily raise into the matrix clause to control LDA, this weak interpretation would be surprising. The same reasoning holds for light verb constructions such as *madad kar* ‘help’ (lit. ‘help do’) in (97b), which are perfectly acceptable in LDA configurations.

- (97) a. usha-ne [potluck ke-liye daal banaa-nii] caah-ii thii
 Usha-ERG potluck for lentil.F make-INF.F.SG want-PFV.F.SG be.PST.F.SG
 ‘Usha had wanted to prepare lentils for the potluck.’
 b. akbar-ne [merii madad kar-nii] caah-ii thii
 Akbar-ERG my.F help.F do-INF.F.SG want-PFV.F.SG be.PST.F.SG
 ‘Akbar had wanted to help me.’ (Bhatt 2005: 798)

The final, and perhaps most compelling argument against a movement account comes from elements that resist movement. Bhatt & Keine (to appear) note that in the idiom *X-kii khuub marammat kar* ‘give X a good beating’ (lit. ‘do X’s many repairs’) the object *marammat* cannot move. (98a)

gives an example of this idiom. In (98b,c), the object part of this idiom have been moved and the idiomatic reading becomes deviant as a result.

(98) *No movement of ‘marammat’ when idiomatic*

- a. raam-ne prataap-kii khuub **marammat** kii
 Ram-ERG Pratap-GEN lot repair.F.SG do.PFV.F.SG
 ‘Ram gave Pratap a good beating.’ (lit. ‘Ram did Pratap’s many repairs.’)
- b. # [prataap-kii khuub **marammat**]_i raam-ne t_i kii
 Pratap-GEN lot repair.F.SG Ram-ERG do.PFV.F.SG
 (idiomatic reading deviant)
- c. # [khuub **marammat**]_i raam-ne prataap-kii t_i kii
 lot repair.F.SG Ram-ERG Pratap-GEN do.PFV.F.SG
 (idiomatic reading deviant)

On a movement account, LDA requires movement of the agreement controller. Because the object *marammat* is evidently unable to move if it is part of the idiom, this account predicts that the idiom should not be able to trigger LDA. As (99) shows, this prediction is not borne out. Instead, it is possible for *marammat* to control LDA even on the idiomatic reading.

(99) *LDA with ‘marammat’ possible even when idiomatic*

- raam-ne [prataap-kii khuub **marammat** kar-nii] caah-ii
 Ram-ERG Pratap-GEN lot repair.F.SG do-INF.F.SG want-PFV.F.SG
 ‘Ram wanted to give Pratap a good beating.’ (Bhatt & Keine to appear: (18))

In sum, there exists a considerable body of evidence that LDA in Hindi does not require the agreement trigger to undergo movement into the matrix clause.

A second line of account that is worth considering here is a **VERB CLUSTER ACCOUNT**. On such an account, LDA configurations are not in fact biclausal but involve a single clause that is projected over a complex verbal base. On such an account, agreement would thus never cross a clause boundary.¹⁵

¹⁵ Butt (1993, 1995) proposes a version of this analysis (a ‘complex predicate’ analysis in her terms) for the permissive construction illustrated in (93) above. Working within LFG, she proposes that the two verbs form a cluster at the level of *f*-structure. At the level of *c*-structure, the two verbs may or may not form a constituent, but in either case they give rise to a monoclausal structure. See Davison (2014) for arguments in favor of a biclausal analysis of the permissive and Butt (2014) for a reply to these arguments.

I will present two arguments against a verb cluster analysis along these lines. The first argument is based on (100). In this example, the infinitival verb *de-nii* ‘give’ and the finite verb *caah-ii* ‘want’ do not appear next to each other and therefore cannot form a cluster. Nonetheless, LDA is perfectly possible in this case, demonstrating that LDA configurations cannot presuppose a verb cluster.

(100) *LDA does not require verb cluster*

raam-ne **kitaab de-nii** siitaa-ko **caah-ii** thii yaa miinaa-ko
 ram-ERG book.F give-INF.F.SG sita-DAT want-PFV.F.SG be.PST.F.SG or Mina-DAT
 ‘Did Ram want to give a book to Sita or to Mina?’

A second argument against a verb cluster account comes from scope and is based on the sentence in (101). In this sentence, the matrix verb *caah* ‘want’ agrees with the quantificational object *har kitaab* ‘every book’. Crucial are the possible scope relations between *har kitaab* and *caah*. As indicated, either can take scope over the other.¹⁶ The possibility the ‘*want* > \forall ’ reading is entirely unexpected on a verb cluster analysis. The reason is that on this account, the sentence is headed by the internally complex verb *paṛh-nii caah-ii* ‘read want’. Because *har kitaab* ‘every book’ is an argument of this complex verb, it would invariably take scope over it. Consequently, a verb cluster analysis predicts that only the ‘ \forall > *want*’ is available in (101), contrary to fact.

(101) *Scope ambiguity under LDA*

naim-ne har **kitaab paṛh-nii** caah-ii thii
 Naim-ERG every book.F read-INF.F.SG want-PFV.F.SG be.PST.F.SG
 ‘Naim wanted to read every book.’

\forall > *want*: ‘For every book, Naim wanted to read it.’

want > \forall : ‘Naim’s desire: to read every book’

(Bhatt 2005: 799)

We can conclude from these considerations that *har kitaab* ‘every book’ must originate in a position lower than *caah* ‘want’ and therefore that (101) must have a bi-clausal structure. Because (101) involves LDA, agreement into a nonfinite clause must be possible in Hindi.

To summarize this section, ϕ -agreement in Hindi can enter an embedded nonfinite clause. I have considered two alternative accounts of LDA that do not involve cross-clausal agreement and

¹⁶ The two readings differ as follows: According to the ‘*want* > \forall ’ reading, Naim’s goal is to read every book in the library, regardless of what these books are. On the ‘ \forall > *want*’ reading, Naim intends to read a particular set of books, which happens to be the set of all books in the library, a fact that he may be oblivious to.

provided a number of converging arguments against such accounts. Agreement in Hindi, then, can be nonlocal. Certain clauses are transparent for it. The next section demonstrates that not all are.

2.2.2.4 ϕ -Agreement and finite clauses

In striking contrast to nonfinite clauses, finite clauses do not allow ϕ -agreement into them, a fact widely noted in the previous literature (e.g., Bhatt 2005: 776, Chandra 2007: 45). Thus, the matrix verb *soc* ‘think’ in (102) cannot agree with *ghazal* inside the embedded clause but must instead display default agreement.¹⁷

(102) *No ϕ -agreement into finite clause*

lar̥kō-ne soc-aa/*-ii [ki monaa-ne ghazal gaa-yii
 boys-ERG think-PFV.M.SG/*-PFV.F.SG that Mona-ERG ghazal.F sing-PFV.F.SG
 thii]
 be.PST.F.SG

‘The boys thought that Mona had sung ghazal.’

(Bhatt 2005: 776)

Agreement is likewise impossible with elements at the edge of a finite clause. Thus, even if *ghazal* is moved to the left-peripheral position of the finite clause, the matrix verb is still unable to agree with it:

(103) *No ϕ -agreement into edge of finite clause*

lar̥kō-ne soc-aa/*-ii [ghazal_i monaa-ne t_i gaa-yii thii]
 boys-ERG think-PFV.M.SG/*-PFV.F.SG ghazal.F Mona-ERG sing-PFV.F.SG be.PST.F.SG

‘The boys thought that Mona had sung ghazal.’

It is clear, then, that ϕ -agreement is not simply unbounded in Hindi. While nonfinite clauses are transparent to it, finite ones are not.

2.2.3 Section summary

This section has considered the locality of A-movement, \bar{A} -movement, and ϕ -agreement with respect to the basic distinction between finite and nonfinite clauses in the language. We have seen

¹⁷ Recall from the agreement algorithm in (84) that the ergative subject *lar̥kō-ne* ‘boys-ERG’ in (102) and (103) is invisible to verbal agreement due to its case marking. Thus, the matrix verb in (102) and (103) cannot enter into an agreement relationship with any DP and masculine singular default agreement results.

that nonfinite clauses allow A- as well as \bar{A} -movement out of them, and ϕ -agreement into them. This contrasts with finite clauses, which allow \bar{A} -extraction out of them, but block A-extraction and ϕ -agreement into them. This distribution of properties is summarized in (104).

(104) *Interim summary (to be extended)*

<i>Operation</i>	Finite clauses	Nonfinite clauses
\bar{A} -movement	✓	✓
A-movement	✗	✓
ϕ -agreement	✗	✓

(✓ : transparent; ✗ : opaque)

(104) instantiates selective opacity because finite clauses constitute impenetrable domains for A-movement and agreement, but are at the same time accessible to \bar{A} -movement. Furthermore, A-movement and ϕ -agreement are able to access nonfinite embedded clauses and are hence not simply clause-bounded. The next section will motivate a third layer of clause size in Hindi by presenting evidence that nonfinite clauses in Hindi come in two sizes, which differ in their locality profile.

2.3 The structure of finite and nonfinite clauses

In this section, I will provide evidence that the distinction between finite and nonfinite clauses in Hindi is correlated with a difference in clause size. That is, I propose that nonfinite clauses are structurally smaller in Hindi than finite clauses. Specifically, I propose that finite clauses are CPs, whereas nonfinite clauses are at most TPs. I will then present evidence that nonfinite clauses come in at least two varieties, with locality being one of the properties that distinguish them. I will argue that this difference between the two types of nonfinite clauses is one of clause structure (Wurmbrand 2001). Specifically, I will propose that nonfinite clauses are ambiguous between a ν P and a TP structure, a distinction that is not overtly manifested and thus not directly observable. Crucially, I will show that A-movement and ϕ -agreement are both sensitive to the categorial distinction between ν P and TP nonfinite clauses. The evidence for this claim comes from a hitherto unnoticed interaction between A-movement out of the lower clause and ϕ -agreement into that clause. As I will show, A-movement of any element out of a nonfinite clause renders this clause

obligatorily transparent for ϕ -agreement, while \bar{A} -movement has no such effect. A crucial aspect of this interaction is that it does not hold at the level of individual syntactic items, because the element that undergoes the A-movement step does not need to be the one that controls ϕ -agreement as a result. Rather, the entailment holds of the nonfinite clause as a whole. This set of facts can be straightforwardly accounted for if (i) nonfinite clauses are structurally ambiguous, (ii) A-movement and ϕ -agreement are sensitive to this distinction, and (iii) \bar{A} -movement is insensitive to it. The conclusion of this section will thus be that A-movement and ϕ -agreement exhibit the same locality profile, which crucially has to be distinct from the locality of \bar{A} -movement. The broader conclusion to be drawn from this pattern is that ϕ -agreement partakes in selective opacity in the same way that A-movement does. Any account of selective opacity that is confined to movement dependencies is unable to express this convergence.

2.3.1 The size of finite vs. nonfinite clauses

There is good reason to believe that nonfinite clauses are structurally smaller than finite clauses in Hindi. As shown in (105), finite embedded clause may optionally contain the complementizer *ki*. This is not the case for nonfinite clauses, which do not allow *ki*, as (106) illustrates. It is irrelevant for this generalization whether the nonfinite clause precedes or follows the matrix predicate.

(105) *Finite clauses can contain a complementizer*

siitaa soc-tii hai [(ki) raam-ne prataap-ko dekh-aa] [= (56)]
 Sita think-IPFV.F.SG be.PRES.3SG that Ram-ERG Prataap-ACC see-PPFV.M.SG
 ‘Sita thinks that Ram saw Prataap.’

(106) *No complementizer in nonfinite clauses*

- a. siitaa [(**ki*) prataap-ko dekh-naa] caah-tii thii
 Sita that Prataap-ACC see-INF.M.SG want-IPFV.F.SG be.PST.F.SG
- b. siitaa caah-tii thii [(**ki*) prataap-ko dekh-naa]
 Sita want-IPFV.F.SG be.PST.F.SG that Prataap-ACC see-INF.M.SG
 ‘Sita wanted to see Prataap.’

On the assumption that the complementizer is a realization of the C^0 head, this contrast between finite and nonfinite clauses follows without further ado if finite clauses contain a CP layer, while nonfinite clauses obligatorily do not.

Furthermore, it is a well-known fact in the literature on Hindi that finite clauses provide an interrogative scope position, but nonfinite clauses do not (Mahajan 1990, Srivastav 1991b, Dayal 1996, Bhatt & Dayal 2007, Manetta 2012). In (107), it is possible – and as we will see in section 2.4 in fact obligatory – for the *wh*-element *kyaa* ‘what’ to take scope within the embedded finite clause. By contrast, a *wh*-element inside a nonfinite clause can only take matrix scope in (108a). If the embedding predicate requires a *wh*-complement, ungrammaticality results, as in (108b).¹⁸

(107) *Wh-scope in finite clauses*

tum jaan-te ho [(ki) us-ne kyaa ki-yaa]
 you know-IPFV.M.PL be.PRES.2PL that he-ERG what do-PFV.M.SG
 ‘You know what he did.’ (Dayal 1996: 31)

(108) *No wh-scope in nonfinite clauses*

a. tum [kyaa kar-naa] jaan-te ho
 you what do-INF.M.SG know-IPFV.M.PL be.PRES.3SG
 ‘What do you know to do?’
 Not: ‘You know what to do.’

b. *vo [kyaa kar-naa] puuch rahaa thaa
 he what do-INF.M.SG ask PROG.M.SG be.PST.M.SG
 Intended: ‘He was asking what to do.’ (Dayal 1996: 23,30)

On the standard assumptions that interrogative scope is a property of C^0 , its obligatory absence in nonfinite clauses follows straightforwardly if these clauses obligatorily lack a CP layer. Evidence from the distribution of complementizers and *wh*-scope thus converges on the conclusion that finite

¹⁸ The combination of the verb *jaan* ‘know’ in (108a) with a nonfinite complement clause is somewhat degraded irrespective of issues pertaining to *wh*-construal. Another, semantically similar verb that effortlessly combines with nonfinite clauses is *aa* ‘know’ (lit. ‘come’), an example of which is (95). Embedding a *wh*-element inside it produces only matrix scope, just as in (108a):

(i) tumhe [kyaa kar-naa] aa-taa hai
 you.DAT what do-INF.M.SG come-IPFV.M.SG be.PRES.3SG
 ‘What do you know how to do?’
 Not: ‘You know what to do (how).’

Dayal’s (1996) point thus stands regardless of the choice of the embedding verb.

clauses contain a CP layer, whereas nonfinite clauses invariably lack it. I will consequently adopt a clause pruning/restructuring account (Tappe 1984, Rochette 1988, Fanselow 1989, Moore 1989, 1996, Li 1990, Rosen 1990, Rooryck 1994, Wurmbrand 2001, 2007, Williams 2003, Chung 2004), according to which nonfinite clauses may contain less functional material than finite clauses. For the sake of concreteness, I will treat nonfinite clauses as TPs for now. Little hinges on this label I will in fact refine this view in section 2.3.2. The claim that nonfinite clauses in Hindi are structurally deficient is by no means new: Dayal (1996) likewise treats finite clauses as CPs and nonfinite clauses as InfPs, Bhatt (2005) and Chandra (2007) treat nonfinite clauses as TPs, and Boeckx (2004) analyzes them as *v*Ps or VPs.

At the same time, there is a lower bound on the possible size of nonfinite clauses. Specifically, there is evidence for the obligatory presence of a PRO subject and for an accusative case assigner in nonfinite clauses. Following the standard assumption that both are properties of *v*⁰, these facts point to the conclusion that nonfinite clauses must be at least *v*Ps in Hindi. I will illustrate both points in turn.

Davison (2010, 2014) presents a number of arguments in favor of the view that the nonfinite clauses under investigation here obligatorily contain a PRO subject. One of her arguments is based on an independently observable ban on PRO occurring in a position associated with dative case (Davison 2008). There are a number of verbs in Hindi whose subject appears marked with lexical dative case (see Davison 2004a,b and also Poole 2015 for discussion).¹⁹ An example is the verb *mil* ‘get’, illustrated in (109), where the subject *baccō-ko* ‘children-DAT’ is lexically case-marked.

- (109) *baccō-ko* *miiṭhaaiyaa* *mil-ī*
 children-DAT sweet.F.PL get-PFV.F.PL
 ‘The children got sweets.’ (Davison 2010: 1)

Crucially, it is not possible for these verbs to appear in obligatory control environments if the PRO would replace the dative element (Davison 2008). The verb *caah* ‘want’ provides an example. In (110), embedding the predicate *mil* ‘get’ under this control predicate leads to ungrammaticality because PRO ends up in a position that receives dative case.

¹⁹ The subject status of the dative DP is motivated, for instance, by its ability to bind subject-oriented anaphors.

- (110) *baccō-ne [PRO miṭhaaiyaa mil-naa] caah-aa
 children-DAT PRO.DAT sweet.F.PL get-INF.M.SG want-PFV.M.SG
Intended: ‘The children wanted to get sweets.’ (Davison 2010: 1)

The proper analysis of this curious restriction do not need to concern us here. It suffices to note (111) as an empirical generalization:

- (111) **Dative restriction** (Davison 2008: 34)
 In contexts of obligatory control, the embedded verb may not assign its (null) subject dative case.

The ban on verbs like *mil* from appearing in obligatory-control configurations is independent of whether LDA takes place or not. Thus, (112) is just as ungrammatical as its default agreement counterpart in (110).

- (112) *baccō-ne [PRO miṭhaaiyaa mil-nii] caah-ī
 children-DAT PRO.DAT sweet.F.PL get-INF.F.PL want-PFV.F.PL
Intended: ‘The children wanted to get sweets.’ (Davison 2010: 2)

Davison (2008) furthermore argues that the ungrammaticality of (110) and (112) is due to the presence of a PRO subject inside the embedded clause. Raising predicates, for instance, allow dative-subject predicates embedded under them (see also Davison 2014). This set of facts thus provides evidence that nonfinite clauses embedded under verbs like *caah* ‘want’ obligatorily project a PRO subject. Assuming, as is standard, that external arguments are introduced by the functional head v^0 (Marantz 1984, Kratzer 1996, among many), we can conclude that nonfinite clauses invariably contain a vP layer in Hindi. This structural requirement holds irrespective of whether LDA takes place or not.²⁰

An independent argument for the obligatory presence of a v^0 head inside the nonfinite clauses of interest to us is based on case assignment and advanced by Bhatt (2005). The argument is

²⁰ Converging evidence for the syntactic presence of an external argument in nonfinite clauses comes from anaphor binding. As noted in section 2.2 above, the reflexive *apnaa* is subject-oriented for many speakers:

- (i) raam-ne_i mohan-ko_j [apnii_i/?_j kitaab] dii [≅(75)]
 Ram-ERG Mohan-DAT self’s book give.PFV.F.SG
 ‘Ram gave self’s book to Mohan.’

In the permissive construction in (ii), it is possible for an embedded reflexive to be coreferent with a dative argument of the matrix predicate, even for speakers for who the reflexive is subject-oriented. Davison (2010, 2014) concludes from this binding option in (ii) that *apne-ko* cannot be directly bound by *raadhaa-ko* in (ii) and that the binding must be established indirectly via an embedded PRO subject, which fulfills the licensing requirements of the reflexive:

based on Wurmbrand's (2001) work on the so-called **LONG PASSIVE** in German.²¹ Just like in English, passivization of a transitive predicate in German overrides this predicate's ability to assign accusative case to its object, which receives nominative case instead. The long passive is a construction in which passivization of a matrix predicate leads to nominative assignment to the object of an embedded predicate. This is illustrated in (113).

(113) *German long passive*

- a. weil er den Traktor zu reparieren vergessen hat
 because he.NOM the tractor.ACC to repair forgotten has
 'because he forgot to repair the tractor'
- b. weil der Traktor zu reparieren vergessen wurde
 because the tractor.NOM to repair forgotten was
 'because it was forgotten to repair the tractor'

(113a) is an active clause in which the verb *vergessen* 'forget' takes as its object the infinitival clause *den Traktor zu reparieren* 'to repair the tractor'. As (113b) demonstrates, passivization of the matrix predicate *vergessen* leads to nominative case assignment to the embedded object *der Traktor* 'the tractor'. This interaction is remarkable because it is the matrix predicate that undergoes passivization but the embedded predicate whose argument undergoes the change in case.

To resolve this paradox, Wurmbrand (2001) and Bobaljik & Wurmbrand (2005) propose that the infinitive in (113) is a mere VP, lacking all functional projections, including *vP*. Assuming that *v⁰* is the source of accusative case, they reason that VP predicates lack a source of accusative case. Consequently, the accusative case in (113a) must stem from the matrix clause, as schematized in (114a). If this higher clause is passivized, it is no longer capable of assigning accusative case and

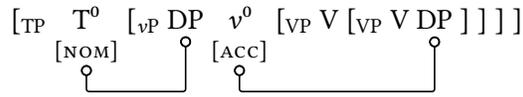
(ii) mǎā-ne_i raadhaa-ko_j [PRO_j apne-ko_{i/j} aaiine-mē dekh-ne] nahī di-yaa
 mother-ERG Radha-DAT self-ACC mirror-in see-INF not give-PFV.M.SG
 'Mother_i did not allow Radha_j to look at self_{i/j} in the mirror.' (Davison 2014: 146)

This argument is entirely consistent with the presence of a PRO subject in the embedded clause, but it does not establish whether this PRO is obligatorily or optionally present. The arguments in the main text are more conclusive in this regard.

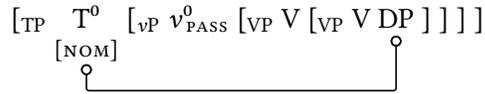
²¹ See von Stechow (1992), Haider (1993, 2010), Sabel (1996), Wöllstein-Leisten (2001), Bayer & Kornfilt (1990, 1994), Wurmbrand (2001), Bobaljik & Wurmbrand (2005), and Keine & Bhatt (to appear) for discussion and analyses of the long passive construction.

matrix T^0 has to assign nominative case to the embedded object instead (114b). This analysis thus captures why matrix passivization should affect case assignment to the embedded object.

(114) a. *Case assignment in (113a)*²² (Wurmbrand 2001, Bobaljik & Wurmbrand 2005)



b. *Case assignment in (113b)*



Crucial to this account, which is also adopted in Keine & Bhatt (to appear), is that the embedded clause is devoid of all functional structure. The long passive, then, acts as a diagnostic of bare VP clauses.

Against this background, Bhatt (2005) shows that Hindi lacks long passives. In Hindi, the case of an object argument is advanced under passivization, just like in English and German. As (115a) shows, the proper name *siitaa* has to be overtly case-marked if it is the object of a transitive predicate. In the corresponding passive structure (115b), *siitaa* can appear in the zero-marked nominative case form.²³ This follows if passivization discharges v^0 's ability to assign accusative case, just as in English.

(115) *Hindi passivization*

a. raam-ne mujhe/*māĩ ḍāāṭ-aa
 Ram-ERG me.ACC/I.NOM scold-PFV.M.SG
 'Ram scolded me.'

b. māĩ raam-dwaaraa ḍāāṭ-aa ga-yaa
 I.NOM Ram-by scold-PFV.M.SG PASS-PFV.M.SG
 'I was scolded by Ram.'

Importantly, passivization of a predicate that takes a nonfinite clause as its complement does not affect the case of the embedded object in Hindi. This is shown in (116) for the permissive

²² On Wurmbrand's (2001) and Bobaljik & Wurmbrand's (2005) analysis, (113a) is in fact structurally ambiguous between a VP structure of the infinitival clause and a vP structure. This complication is irrelevant for the point made here.

²³ Such advancement of the internal argument is only optional, for reasons that are not well understood. The underlying cause for this optionality is irrelevant for the argument and will be put aside here.

construction, which, as we have seen in (93), is an example of a predicate that allows LDA into its complement.

(116) *No long passive in Hindi*

[siitaa-ko mujhe/*māĩ piiṭ-ne] di-yaa ga-yaa
 Sita-DAT me.ACC/*I.NOM hit-INF give-PFV.M.SG PASS-PFV.M.SG

‘Sita was allowed to hit me.’ (= ‘Someone let Sita hit me.’) (Bhatt 2005: 782)

Despite the fact that the matrix predicate *di* ‘let’ (lit. ‘give’) is passivized in (116), the case of the embedded object cannot be advanced. In other words, Hindi does not allow for long passives analogous to (113b). From these considerations, Bhatt (2005) concludes that nonfinite clauses in Hindi *cannot* be bare VPs, but obligatorily contain at least a *vP* projection, which licenses accusative case. I will follow this conclusion.

To summarize this section, we have seen evidence that finite and nonfinite clauses differ in their structural size in Hindi. Finite clauses exhibit CP behavior in that they can contain a complementizer and may carry interrogative force. Nonfinite clauses, on the other hand, can do neither and thus plausibly lack a CP projection. I will therefore treat nonfinite clauses as maximally TPs. In addition to this upper bound on the structure of nonfinite clauses, we have also seen evidence for a lower size bound. Nonfinite clauses in Hindi (or at least those of interest to us here) are not bare VPs, but have to project a *vP*.

(117) *Clause size in Hindi*

- a. Finite clauses are CPs.
- b. Nonfinite clauses are structurally smaller than CPs and larger than VPs.

We have thus identified a lower and an upper bound for the size of nonfinite clauses in Hindi. Before moving on, it is instructive to consider a claim made by, e.g., Butt (1993, 1995) and Dayal (1996) that nonfinite clauses in Hindi are nominal in nature. While this view is compatible with the analysis developed in the next chapter, I will not adopt it here. The reason is that there are variants of nonfinite clauses that carry overt reflexes of nominal structure but whose locality properties differ from the nonfinite clauses considered here. As Bhatt (2005) shows, nonfinite clauses can (somewhat marginally) take a genitive subject, as shown in (118a), as long as the matrix verb shows

default agreement. LDA is impossible in this case (118b). If the genitive subject is removed, LDA is possible, as expected (118c). Bhatt (2005) emphasizes that genitive case is only ever assigned within the nominal domain in Hindi (i.e., there are no verbs that take genitive arguments, etc.). The presence of a genitive subject is hence indicative of the presence of nominal structure in the embedded clause, hence a gerund structure. The fact that such gerunds are completely opaque to ϕ -agreement suggests that nominal structure renders a clause opaque to ϕ -agreement. This in turn entails that regular nonfinite clauses lack such nominal structure.

(118) *Gerunds are opaque to LDA*

a. ?firoz-ne [shabnam-kaa rottii khaa-naa] caah-aa
 FIROZ-ERG Shabnam-GEN bread.F eat-INF.M.SG want-PFV.M.SG
 ‘Firoz wanted Shabnam’s eating bread.’

b. *firoz-ne [shabnam-kaa roṭii khaa-nii] caah-ii
 FIROZ-ERG Shabnam-GEN bread.F eat-INF.F.SG want-PFV.F.SG

c. firoz-ne [roṭii khaa-nii] caah-ii
 FIROZ-ERG bread.F eat-INF.F.SG want-PFV.F.SG
 ‘Firoz wanted to eat bread.’

(Bhatt 2005: 784)

Gerunds are also opaque to A- and \bar{A} -extraction. This is demonstrated in (119), where *har kuttaa* ‘every dog’ is moved out of the gerund above the matrix subject. Regardless of whether it binds a pronoun from its landing site or not, the result is ill-formed. All extractions out of gerunds are thus impossible.

(119) *No A/ \bar{A} -movement out of gerunds*

*har kuttaa_i [us-kaa_{i/j} maalik] [siitaa-kaa t_i ghuuma-naa] caah-taa
 every dog 3.SG-GEN owner Sita-GEN walk-INF.M.SG want-IPFV.M.SG
 hai
 be.PRES.3SG

Intended: ‘Its_j owner wanted Sita’s walking every dog_i. / For every dog x , x ’s owner wanted Sita’s walking x .’

On the natural assumption that it is the nominal structure inside the gerund that blocks extraction out of it, we can conclude that regular nonfinite clauses lack such nominal structure, as they are transparent to both A- and \bar{A} -movement (with qualifications discussed in the next section).

Converging evidence comes from case-marked clauses. Nonfinite clauses embedded under certain predicates bear overt case marking in Hindi. In (120), for example, the nonfinite clause appears with the case marker *-ko*²⁴ because it is embedded under the matrix verb *kah* ‘say’. Interestingly, case-marked clauses once again exhibit a more restrictive locality profile than regular nonfinite clauses. In particular, they disallow A-movement out of them, as shown in (120a). This example demonstrates that extraction of *har kitaab* ‘every book’ is possible, but subject to weak crossover, hence obligatorily \bar{A} -movement. Furthermore, they block ϕ -agreement into them, as (120b) illustrates (Butt 1993: 77).

- (120) a. *Only \bar{A} -extraction out of case-marked infinitival clauses*
 har kitaab_i [us-ke_j/^{*}_i lekhak-ne] siitaa-se [t_i paṛh-ne]-ko kah-aa
 every book 3SG-GEN author-ERG Sita-INSTR read-INF -ACC say-PFV.M.SG
 ‘Its_j author told Sita to read every book_i.’ (bound reading impossible)
- b. *No ϕ -agreement into case-marked infinitival clauses*
 raam-ne siitaa-se [kitaab paṛh-ne]-ko kah-aa/^{*}-ii
 Ram-ERG Sita-INSTR book.F read-INF -ACC say-PFV.M.SG/^{*}-PFV.F.SG
 ‘Ram told Sita to read a book.’

Just like in the case of gerunds, the presence of overt reflexes of nominal structure has substantial effects on the locality properties of nonfinite clauses. The fact that regular nonfinite clauses are transparent to A-movement, \bar{A} -movement, and ϕ -agreement then provides evidence that they simply lack this nominal structure, a conclusion that I will adopt from here on.

In sum, I have argued for the structural distinction between finite and nonfinite clauses in (117). In the next section, I will provide evidence that there are in fact two types of nonfinite clauses, which differ in the amount of functional structure they contain. This size difference affects the locality properties of these two clauses.

²⁴ Due to the homophony between the accusative and dative case marker in Hindi, it is unclear exactly which case these clauses bear. I will gloss the case as accusative, but this commitment is irrelevant for my concerns here.

2.3.2 Two types of nonfinite clauses

Perhaps the most striking difference between local and long-distance agreement in Hindi is that local agreement is deterministic whereas long-distance agreement is optional.²⁵ To appreciate this contrast, recall from section 2.2.2 the ϕ -agreement algorithm in Hindi:

- (121) **Hindi ϕ -agreement algorithm** [= (84)]
If the subject does not bear a case marker \rightarrow agree with the subject
Otherwise: If object does not bear a case marker \rightarrow agree with the object
Otherwise: Use masculine singular default agreement.

In the case of local agreement – i.e., if a verb is agreeing with one of its arguments – agreement is obligatory if it is possible. In other words default agreement is a last resort in this case, admissible only if there is no DP that could control verbal agreement. This point is illustrated in (122). In (122a) the subject *larke* is not overtly case-marked and therefore has to control agreement on the verb and its auxiliary. Masculine singular default agreement is not an option here. (122b) demonstrates the same point for objects. Here the subject is overtly case-marked and hence invisible for verbal agreement, whereas the object does not bear an overt case marker. In this case, object agreement is obligatory. Finally, (122c), both the subject and the object bear overt case markers. In this case, default agreement arises.

- (122) a. *Local subject agreement is obligatory if possible*
- i. *larke* is *kitaab-ko* *parh-te* *hãĩ*
boys.M.PL this book.F.SG-ACC read-IPFV.M.PL be.PRES.3PL
'The boys are reading this book.'
- ii. **larke* is *kitaab-ko* *parh-taa* *hai*
boys.M.PL this book.F.SG-ACC read-IPFV.M.SG be.PRES.3SG

²⁵ This discrepancy between local and long-distance agreement is by no means restricted to Hindi. All long-distance agreement systems that I am aware of display a similar distinction. This includes, but is not limited to, Tsez (Polinsky & Potsdam 2001), Itelmen (Bobaljik & Wurmbrand 2005), Basque (Preminger 2009), and Innu-aimũn (Branigan & MacKenzie 2002). Whether the restructuring analysis proposed for Hindi here extends to these other systems remains to be evaluated.

b. *Local object agreement is obligatory if possible and subject agreement is impossible*

- i. laṛkō-ne yah kitaab paṛh-ii
 boys.M.PL-ERG this book.F.SG read-PFV.F.SG
 ‘The boys read this book.’
- ii. *laṛkō-ne yah kitaab paṛh-aa
 boys.M.PL-ERG this book.F.SG read-PFV.M.SG

c. *Last resort default agreement*

- laṛkō-ne is kitaab-ko paṛh-aa
 boys.M.PL-ERG this book.F.SG-ACC read-PFV.M.SG
 ‘The boys read this book.’

The ϕ -agreement pattern in local configurations is consonant with the logic of **OBLIGATORY OPERATIONS** (Preminger 2011, 2014): If a ϕ -probe can enter into an agreement relationship with a DP, it has to do so. If it cannot agree with any DP, default agreement arises, but the derivation does not crash.²⁶ This is the picture I will tentatively adopt from now on.

(123) **Agreement obligatoriness** (Preminger 2011, 2014)

If a verb can ϕ -agree with a DP, it has to.

Long-distance agreement presents an interesting challenge to this view of agreement in Hindi. As already mentioned in section 2.2.2, LDA is apparently optional. Thus, the LDA configuration in (124a) differs from (124b) only in that the former exhibits LDA and the latter default agreement. Both are well-formed.

(124) *Apparent optionality of long-distance agreement*

- a. laṛkō-ne [roṭii khaa-nii] caah-ii [= (87)]
 boys-ERG bread.F eat-INF.F.SG want-PFV.F.SG
 ‘The boys wanted to eat bread.’
- b. laṛkō-ne [roṭii khaa-aa] caah-aa [= (88)]
 boys-ERG bread.F eat-INF.M.SG want-PFV.M.SG
 ‘The boys wanted to eat bread.’

²⁶ In this sense, Preminger’s (2011, 2014) obligatory operations are similar to Chomsky’s (1957, 1965) obligatory transformations, though the conceptual and technical underpinnings are of course very different.

We have seen in section 2.2.2 evidence that LDA is not fed by (optional) movement of the agreement trigger or the result of (optional) verb cluster formation. Specifically, I have argued that LDA is genuinely cross-clausal in the sense that the matrix predicate agrees into the embedded clause. We are faced with the question, then, of why agreement is optional across a clause boundary, but obligatory otherwise.

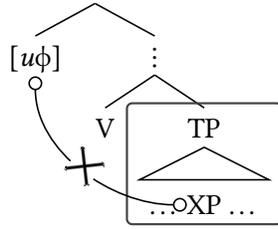
To address this problem, I propose a restructuring/clause pruning analysis in the sense of Chomsky (1981), Rochette (1988), Wurmbrand (2001) and others, according to which nonfinite clauses are not associated with a uniform clause structure, but can instead contain a varying degree of functional clause structure. That is, I adopt the common view that clause structures can be pruned at specific projections, which can then be embedded under a higher verb without the mediation of higher projections. For the case of Hindi, I propose that nonfinite clauses come in (at least) two structural sizes. The structurally larger clauses by assumption contain a barrier for agreement, which is absent in the structurally smaller ones. For the sake of concreteness, I will treat these two sizes as TP and vP , respectively, though little hinges on the labels. If the nonfinite clause is a TP, it is by assumption opaque to ϕ -agreement and default agreement will arise. On the other hand, if it is a vP , it will be transparent to ϕ -agreement. By (123), LDA will be obligatory in a vP structure because it is possible. Because TP and vP clauses are surface-identical, the apparent optionality of LDA follows as an instance of regular structural ambiguity. This allows us to retain the view in (123) that agreement in Hindi is always forced if it is possible.

(125) *Two types of nonfinite clauses*

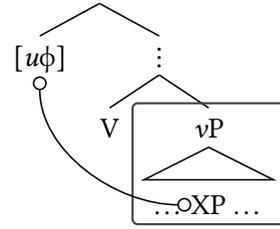
- a. TP clauses are opaque to ϕ -agreement.
- b. vP clauses are transparent to ϕ -agreement.

This yields the schematic structures in (126) and (127). I will remain agnostic for the time being about the precise location of the ϕ -probe in Hindi. See section 2.5.2 for arguments that it is located on T^0 .

(126) *TP nonfinite clauses*
 ~ ϕ -agreement impossible



(127) *vP nonfinite clauses*
 ~ ϕ -agreement obligatory



Restructuring analyses of Hindi LDA have also been proposed by Boeckx (2004) and Bhatt (2005), although the details differ substantially from the system proposed here. Some independent, though somewhat tentative, evidence that it is the presence or absence of a TP that underlies the distinction between LDA and default agreement structures comes from temporal adverbs. The adverb *kal* can mean either ‘yesterday’ or ‘tomorrow’, as illustrated in (128):

(128) *mujhe kal dillii jaa-naa thaa*
 me.DAT yesterday/tomorrow Delhi go-INF.M.SG be.PST.M.SG
 ‘I had to go to Delhi tomorrow (there was a plan, but it got canceled).’ Or:
 ‘I had to go to Delhi yesterday (but I didn’t go).’

Kal can in principle modify a nonfinite clause. This is shown in (129), in which *kal* is placed into two positions within the lower clause. To ensure that *kal* does indeed attach to the lower clause, the second temporal adverb *pichle hafte* ‘last week’ is placed at the beginning of the sentence, modifying the matrix predicate. Crucially, both sentences in (129) do not involve LDA. Instead, both involve default agreement.

(129) *Temporal adverbs and default agreement*

a. *pichle hafte raam-ne [kal yeh kitaab parh-naa]*
 last week Ram-ERG yesterday/tomorrow this book.F read-INF.M.SG
caah-aa thaa
 want-PFV.M.SG be.PST.M.SG
 ‘Last week, Ram had wanted to read the book yesterday/tomorrow.’

- b. *pichle hafte raam-ne* [*yeh kitaab kal* *parh-naa*]
 last week Ram-ERG this book.F yesterday/tomorrow read-INF.M.SG
caah-aa thaa
 want-PFV.M.SG be.PST.M.SG
 ‘Last week, Ram had wanted to read the book yesterday/tomorrow.’

As soon as default agreement in (129) is replaced with LDA, as in (130), anomaly arises. Both sentences in (130) only have an interpretation in which both adverbs jointly, and thus inconsistently, specify a single event.

(130) *Temporal adverbs and LDA*

- a. #*pichle hafte raam-ne* [*kal* *yeh kitaab parh-nii*]
 last week Ram-ERG yesterday/tomorrow this book.F read-INF.F.SG
caah-ii thii
 want-PFV.F.SG be.PST.F.SG
Intended: ‘Last week, Ram had wanted to read the book yesterday/tomorrow.’

- b. #*pichle hafte raam-ne* [*yeh kitaab kal* *parh-nii*]
 last week Ram-ERG this book.F yesterday/tomorrow read-INF.F.SG
caah-ii thii
 want-PFV.F.SG be.PST.F.SG
Intended: ‘Last week, Ram had wanted to read the book yesterday/tomorrow.’

The contrast between LDA and default agreement in (129) and (130) follows straightforwardly if (i) temporal adverbs require the presence of a TP layer in the nonfinite clause and (ii) LDA is impossible into TP clauses. An embedded construal of *kal* is then possible only in clauses that are too large to allow LDA into them.²⁷

Additional supporting evidence for a structural ambiguity of nonfinite clauses along the lines of (125) comes from coordination. As (131a) shows, infinitival clauses can be coordinated regardless of whether they show LDA or default agreement, as long as the agreement is the same in both. Furthermore, (131b) demonstrates that the combination of a nonfinite clause with agreement and another one with default agreement leads to ungrammaticality. It is irrelevant for this restriction whether the matrix verb shows LDA or default agreement.

²⁷ See Wurmbrand (2014) and references cited there for a discussion of tense in different types of nonfinite clauses.

(131) *Coordination*

a. *Matching agreement*

- i. raam-ne [roṭii khaa-nii] aur [kitaab paṛh-nii] caah-ii
Ram-ERG bread.F eat-INF.F.SG and book.F read-INF.F.SG want-PFV.F.SG
thii
be.PST.F.SG
'Ram wanted to eat bread and read a book.'
- ii. raam-ne [roṭii khaa-naa] aur [kitaab paṛh-naa] caah-aa
Ram-ERG bread.F eat-INF.M.SG and book.F read-INF.M.SG want-PFV.M.SG
thaa
be.PST.M.SG
'Ram wanted to eat bread and read a book.'

b. *Mismatching agreement*

- i. *raam-ne [roṭii khaa-nii] aur [kitaab paṛh-naa]
Ram-ERG bread.F eat-INF.M.SG and book.F read-INF.F.SG
caah-ii/-aa thii/thaa
want-PFV.F.SG/-PFV.M.SG be.PST.F.SG/be.PST.M.SG
- ii. *raam-ne [roṭii khaa-naa] aur [kitaab paṛh-nii]
Ram-ERG bread.F eat-INF.F.SG and book.F read-INF.M.SG
caah-ii/-aa thii/thaa
want-PFV.F.SG/-PFV.M.SG be.PST.F.SG/be.PST.M.SG

This restriction is accounted for if infinitives that are transparent for LDA and ones that are not differ syntactically and perhaps also semantically. According to (125), nonfinite clauses with LDA into them are *v*P_s, whereas nonfinite clauses without LDA are TP_s. Given a general parallelism constraint on coordination – e.g., Williams' (1978) Law of Coordination of Likes –, only matching coordinations will be allowed.²⁸

²⁸ A further expectation of a restructuring account like the one proposed here is that certain verbs should only select for one of the two clause sizes. For such verbs, LDA should not be optional. One candidate is the verb *lag* 'begin'. With this verb, LDA is obligatory, as (i) shows. Note that infinitives embedded under *lag* do not themselves show agreement. The contrast between LDA and default agreement therefore only manifests itself on the matrix verb.

(i) us-ko niind aa-ne lag-ii / *lag-aa
3SG-DAT sleep.F come-INF begin-PFV.F.SG *begin-PFV.M.SG
'S/he began to sleep.'

The fact that LDA is forced with *lag* 'begin' is accounted for if this verb only embeds *v*P_s. I am not of a verb that embeds bare nonfinite clauses but does not allow LDA. One verb that allows for a nonfinite complement clause but not LDA is *kah* 'say'. However, in this case the nonfinite clause bears an overt case marker, which may plausibly

2.3.3 The A-movement–agreement connection

In this section, I will provide converging evidence for the claim that nonfinite clauses in Hindi are ambiguous between two structures. I will show that A-movement and LDA interact in a curious way:

(132) **A-Movement–Agreement Generalization**

A-movement of *any* element out of a nonfinite clause makes LDA into this clause obligatory.
 \bar{A} -movement has no such effect.

A crucial aspect of this generalization is that it does not hold of individual items, but rather of the nonfinite clause as a whole. Specifically, the element that A-moves does not have to be the one that has to control LDA as a consequence. The *clause-based* nature of the implication in (132) strongly suggests that (i) nonfinite clauses are structurally ambiguous, (ii) the locality of A-movement is the same as that of ϕ -agreement, and (iii) both are distinct from the locality of \bar{A} -movement. The evidence for this generalization comes from three independent diagnostics: weak crossover, reciprocal binding, and quantifier scope, which I will discuss in turn.

2.3.3.1 Weak crossover

Initial evidence in favor of the generalization (132) comes from the paradigm in (133). In all three sentences, a nonfinite clause is embedded under a predicate that allows LDA. (133a) constitutes the baseline. The embedded object *har billii* ‘every cat’ remains in its base position and the pronoun *us-ke* ‘his/her’ inside the matrix subject is interpreted referentially.²⁹ LDA is optional in this configuration. In (133b), the embedded object *har billii* is moved into the matrix clause, to a position above the subject. Importantly, the pronoun retains a referential interpretation in (133b). This movement step does not have an effect on LDA, which remains optional. The crucial example is (133c). Here *har billii* is likewise moved above the matrix subject, just as in (133b), but in (133c) it binds the

be held responsible for the impossibility of LDA (see (120b) for an example). Whether the apparent absence of a predicate that takes a bare nonfinite clause but does not allow LDA into it is significant or not remains to be seen.

²⁹ A bound reading of the pronoun is unsurprisingly impossible in (133a), as *har billii* does not c-command it. This holds irrespective of the choice of agreement.

pronoun *us-ke* from its landing site. Under this interpretation, LDA becomes obligatory and default agreement is no longer an option.³⁰

(133) *Movement and LDA: Direct object*

- a. [us-ke_i maalik-ne] [har billii_j/^{*}_i ghumaa-nii/-naa]
 3SG-GEN owner-ERG every cat.F walk-INF.F.SG/-INF.M.SG
 caah-ii/-aa
 want-PFV.F.SG/-PFV.M.SG
 ‘His/her_i owner wanted to walk every cat_j.’
- b. har billii_i [us-ke_j maalik-ne] [t_i ghumaa-nii/-naa]
 every cat.F 3SG-GEN owner-ERG walk-INF.F.SG/-INF.M.SG
 caah-ii/-aa
 want-PFV.F.SG/-PFV.M.SG
 ‘Every cat_i, his/her_j owner wanted to walk (it).’
- c. har billii_i [us-ke_i maalik-ne] [t_i ghumaa-nii/^{*}-naa]
 every cat.F 3SG-GEN owner-ERG walk-INF.F.SG/^{*}-INF.M.SG
 caah-ii/^{*}-aa
 want-PFV.F.SG/^{*}-PFV.M.SG
 ‘For every cat *x*, *x*’s owner wanted to walk *x*.’

Thus, if the embedded object moves into the matrix clause *and* binds a pronoun from its landing site, LDA with this object becomes obligatory. Recall from section 2.2.1 that Hindi is similar to English in that \bar{A} -movement is subject to weak crossover, whereas A-movement is not. Put differently, A-movement can feed pronominal binding from its landing site, but \bar{A} -movement cannot. Because (133c) crucially involves binding of the pronoun *us-ke* by *har billii* from its landing site, we can infer that (133c) must involve A-movement of *har billii*. Pronominal binding thus acts as a diagnostic of A-movement in (133c). This A-movement step of the direct object has the effect that it renders agreement with this direct object obligatory. (133b), on the other hand, involves a mere change in word order without binding, and can hence be produced by either A- or \bar{A} -movement. These interactions are schematized in (134):

³⁰ A word of caution is in order regarding the examples in this section. Quite generally, speakers prefer LDA over the default agreement variant, with some variation in how strong this preference is for which verbs. The claims in (133) and what follows are to be understood as factoring out this baseline preference. Conversely, if a speaker has a strong baseline preference for LDA even in the absence of any movement, then the effect of A-movement will be correspondingly smaller for this speaker.

(134) Schematic structure of (133)

- a. [matrix clause ... [nonfinite clause ... DO ...]] → LDA with DO optional
- b. [matrix clause DO ... [nonfinite clause ... *t* ...]] → LDA with DO optional
- c. [matrix clause DO ... [nonfinite clause ... *t* ...]] → LDA with DO obligatory

A-movement of the embedded direct object into the matrix clause renders LDA with this direct object obligatory, whereas \bar{A} -movement has no such effect.³¹ There are two possible explanations for this interaction. First, the interaction could be *item-based* in the sense that A-movement of some element *x* renders LDA with *x* obligatory. This view is compatible with (133c) because both A-movement and LDA target the same element, namely the embedded direct object. An alternative view of the effect is that it is *clause-based* in the sense that the generalization is not stated over individual elements, but instead more abstractly over the embedded clause as a whole. On this latter view, A-movement out of the lower clause renders this clause penetrable to LDA, and the element that A-moves could in principle be distinct from the element controlling LDA.

To empirically differentiate between these two views, we need to consider cases in which an element other than the direct object is A-moved out of the lower clause. The paradigm in (135) differs from (135) in that the embedded predicate is ditransitive and takes the dative-marked indirect object *har bacce-ko* ‘every child-DAT’ and the unmarked direct object *film* ‘movie’. (135a) constitutes the baseline. No movement takes place and LDA with the embedded direct object *film* ‘movie’ is optional. In (135b), the embedded indirect object *har bacce-ko* ‘every child-DAT’ is moved into the

³¹ The following caveat should be noted: As we have seen in section 2.2.2 (see (90)), LDA is possible only if the matrix verb cannot locally agree with its subject. This restriction is preserved if A-movement takes place. In (i), the matrix predicate appears in the imperfective aspect and the matrix subject *us-kaa maalik* ‘its owner’ is not overtly case-marked as a result. Following the agreement algorithm in (84), the matrix verb *caah* ‘want’ has to agree with this subject and LDA is impossible. Furthermore, the embedded verb has to show masculine singular default agreement and can likewise not agree with *har billii* ‘every cat’.

- (i) har billii_i [us-kaa₁ maalik] [t_i ghumaa-naa/*-nii] caah-taa/tii
 every cat.F 3SG-GEN OWNER.M.SG walk-INF.M.SG/*-INF.F.SG want-IPFV.M.SG/*-IPFV.M.SG
 ‘For every cat *x*, *x*’s owner wants to walk *x*.’

Moreover, because *har billii* is A-moved in (i), but the embedded verb *ghumaa-naa* ‘walk’ cannot agree with it, we can infer that infinitival agreement cannot be a by-product of A-raising into the matrix clause. This provides further evidence that infinitival agreement is established as a side effect of LDA and not independently of it, in line with the conclusions reached in section 2.2.2.

matrix clause, but does not bind the pronoun. This movement does not affect the agreement options. In (135c), *har bacce-ko* ‘every child-DAT’ is likewise moved above the matrix subject, but crucially binds the pronoun inside this subject. In this case, LDA with *film* ‘movie’ is obligatory.

(135) *Movement and LDA: Indirect object*

- a. [us-kii_i mǎǎ-ne] [har bacce-ko_j **film** dikhaa-nii/-naa]
 3SG-GEN mother-ERG every child-DAT movie.F show-INF.F.SG/-INF.M.SG
 caah-ii/-aa
 want-PFV.F.SG/-PFV.M.SG
 ‘His_i mother wanted to show a movie to every child_j.’
- b. har bacce-ko_j [us-kii_i mǎǎ-ne] [*t_j* **film** dikhaa-nii/-naa]
 every child-DAT 3SG-GEN mother-ERG movie.F show-INF.F.SG/-INF.M.SG
 caah-ii/-aa
 want-PFV.F.SG/-PFV.M.SG
 ‘His_i mother wanted to show a movie to every child_j.’
- c. har bacce-ko_i [*us-kii_i* mǎǎ-ne] [*t_i* **film** dikhaa-nii/*?-naa]
 every child-DAT 3SG-GEN mother-ERG movie.F show-INF.F.SG/*?-INF.M.SG
 caah-ii/*?-aa
 want-PFV.F.SG/*?-PFV.M.SG
 ‘For every child *x*, *x*’s mother wanted to show *x* a movie.’

The remarkable feature of (135c) is that the element undergoing A-movement (*har bacce-ko* ‘every child-DAT’) is not the same as the element that has to control LDA as a result (*film* ‘movie’). Nonetheless, the same implication between A-movement and LDA holds as in (133c). This state of affairs is schematized in (136):

(136) *Schematic structure of (135)*

- a. [matrix clause ... [nonfinite clause ... IO DO ...]] → LDA with DO optional
- b. [matrix clause IO ... [nonfinite clause ... *t* DO ...]] → LDA with DO optional

$$\begin{array}{c} \xrightarrow{A/\bar{A}\text{-movement}} \\ \downarrow \end{array}$$
- c. [matrix clause IO ... [nonfinite clause ... *t* DO ...]] → LDA with DO obligatory

$$\begin{array}{c} \xrightarrow{A\text{-movement}} \\ \downarrow \end{array}$$

There is, moreover, no evidence that *film* itself has undergone any kind of movement in (135c). Notably, *film* can be interpreted as a weak indefinite (‘do film-showing’), indicating that it remains in its base position (Diesing 1992), at least as an option. Nonetheless, it has to obligatorily control LDA if the indirect object is A-extracted into the matrix clause. Analogous facts holds for existentially quantified direct objects:

(137) *Crossover and LDA: Indirect object 2*

- a. [uskii_i mää-ne] [har bacce-ko_j ek maalaa de-nii/-naa]
 his/her mother-ERG every child-DAT a necklace.F give-INF.F.SG/INF.M.SG
 caah-ii/-aa
 want-PFV.F.SG/-PFV.M.SG
 ‘His/her_i mother wanted to give a necklace to every child_j.’
- b. har bacce-ko_j [uskii_i mää-ne] [t_j ek maalaa de-nii/-naa]
 every child-DAT his/her mother-ERG a necklace.F give-INF.F.SG/-INF.M.SG
 caah-ii/-aa
 want-PFV.F.SG/-PFV.M.SG
 ‘His/her_i mother wanted to give a necklace to every child_j.’
- c. har bacco-ko_i [uskii_i mää-ne] [t_i ek maalaa de-nii/*-naa]
 every child-DAT his/her mother-ERG a necklace.F give-INF.F.SG/*-INF.M.SG
 caah-ii/*-aa
 want-PFV.F.SG/*-PFV.M.SG
 ‘For every child *x*, *x*’s mother wanted to give a necklace to *x*.’

Just as in (135c), A-movement of the indirect object *har bacce-ko* ‘every child-DAT’ makes LDA with the embedded direct object *ek maalaa* ‘a necklace’ obligatory in (137c). And as before, there is no indication whatsoever that *ek maalaa* has undergone any movement. In particular, it can be interpreted as a weak indefinite. (137c) does not entail that there is a specific necklace that every mother wanted to give to their respective child. Put differently, *ek maalaa* can take embedded scope in (137c), indicating that it does not need to undergo any form of movement.

The facts in (135) and (137) provide compelling support for a *clause-based* of the implication between A-movement and LDA: If the embedded clause is such that A-movement out of it takes place, then it is also necessarily transparent for LDA. Put differently, obligatory LDA with the direct

object can be conditioned by A-moving of another element. While A-movement and LDA may of course target the same element, as in (133c), they do not have to, as in (135c) and (137c).

The implicational relationship is not restricted to indirect objects either. In Hindi, possessor DPs may be moved out of their host DP. In (138), the embedded verb *parh* ‘read’ takes *har lekhak-kii kitabē* ‘every author’s books’ as an object. The possessor *har lekhak-kii* may then be subextracted. The paradigm in (138) is again parallel to the ones in (133), (135) and (137). In the case of LDA, agreement is controlled by *kitaabē* ‘books’. In the absence of any movement, LDA is optional in (138a), precisely as before. The same holds if the possessor *har lekhak-kii* is moved into the matrix clause without binding the pronoun (138c). If the possessor does bind the pronoun, as in (138c), LDA becomes obligatory.

(138) *Movement and LDA: Possessor*

- a. [us-kii_i patnii-ne] [[DP har lekhak-kii_j **kitaabē**] parh-nii/-naa]
 3SG-GEN wife-ERG every author-GEN books.F read-INF.F.PL/-INF.M.SG
 caah-īī/-aa
 want-PFV.F.PL/-PFV.M.SG
 ‘His_i wife wanted to read the books of every author_j.’
- b. har lekhak-kii_j [us-kii_i patnii-ne] [[DP t_j **kitaabē**] parh-nii/-naa]
 every author-GEN 3SG-GEN wife-ERG books.F read-INF.F.PL/-INF.M.SG
 caah-īī/-aa
 want-PFV.F.PL/-PFV.M.SG
 ‘His_i wife wanted to read the books of every author_j.’
- c. har lekhak-kii_i [us-kii_i patnii-ne] [[DP t_i **kitaabē**] parh-nii/*-naa]
 every author-GEN 3SG-GEN wife-ERG books.F read-INF.F.PL/*-INF.M.SG
 caah-īī/*-aa
 want-PFV.F.PL/*-PFV.M.SG
 ‘For every author *x*, *x*’s wife wanted to read *x*’s books.’

Just as in (135c) and (137c), A-movement of an element other than the direct object renders LDA with the direct object obligatory. Moreover, there is again no indication that the agreement trigger *kitabē* has itself undergone movement. Again, we see a case in which LDA with *one* element is triggered by A-movement of *another* element:

(139) *Schematic structure of (138)*

- a. [matrix clause ... [nonfinite clause ... [DO DP_{GEN} DO] ...]] → LDA with DO optional
- b. [matrix clause DP_{GEN} ... [nonfinite clause ... [DO *t* DO] ...]] → LDA with DO optional
- c. [matrix clause DP_{GEN} ... [nonfinite clause ... [DO *t* DO] ...]] → LDA with DO obligatory

Additional evidence that obligatory LDA does not require movement of the agreement trigger comes from the idiom *X-kii khuub marammat kar* ‘give X a good beating’ (lit. ‘do X’s many repairs’). As we saw in section 2.2.2 based on (98), repeated here, this idiom does not tolerate movement of *khuub marammat* ‘many repairs’:

(140) *No movement of ‘marammat’ when idiomatic*

- a. raam-ne prataap-kii khuub **marammat** kii [= (98)]
 Ram-ERG Pratap-GEN lot repair.F.SG do.PFV.F.SG
 ‘Ram gave Pratap a good beating.’ (lit. ‘Ram did Pratap’s many repairs.’)
- b. # [prataap-kii khuub **marammat**]_i raam-ne *t_i* kii
 Pratap-GEN lot repair.F.SG Ram-ERG do.PFV.F.SG
 (idiomatic reading deviant)
- c. # [khuub **marammat**]_i raam-ne prataap-kii *t_i* kii
 lot repair.F.SG Ram-ERG Pratap-GEN do.PFV.F.SG
 (idiomatic reading deviant)

We have also seen in (99) that this idiom may nonetheless control LDA. All else equal, such LDA is optional:

(141) *LDA with ‘marammat’ possible even if idiomatic*

- raam-ne [prataap-kii khuub **marammat** kar-nii/-naa] caah-ii/-aa
 Ram-ERG Pratap-GEN lot repair.F.SG do-INF.F.SG/-INF.M.SG want-PFV.F.SG/-PFV.M.SG
 ‘Ram wanted to give Pratap a good beating.’

As (142) demonstrates, A-movement of the possessor *har bacce-kii* ‘every child’ renders LDA with *marammat* ‘repair’ obligatory:

- (142) *A-movement renders LDA with ‘marammat’ obligatory*
- har bacce-kii_i [uskii_i mǎǎ-ne] [[DP t_i khuub marammat]
 every child-GEN 3.SG.GEN mother-ERG lot repair.F.SG
 kar-nii/*-naa] caah-ii/*-aa
 do-INF.F.SG/*-INF.M.SG want-PFV.F.SG/*-PFV.F.SG
 ‘For every child *x*, *x*’s mother wanted to give *x* a good beating.’

(142) falls under the same generalization as (133–138): A-movement out of the infinitival clause renders LDA into it obligatory. The fact that this implication also holds if the LDA controller demonstrably resists movement, as in (99), provides direct evidence that no movement of the element triggering LDA is required to establish agreement.

A final piece of converging evidence comes from instrumentals. In (143) the embedded predicate is *baat kar* ‘to talk’ (lit. ‘talk do’), which consists of the verb *kar* ‘do’, which takes the direct object *baat* ‘talk’ to form a complex predicate. Because *baat* ‘talk’ is feminine, it can control visible ϕ -agreement on the verb. Additionally, the predicate takes an instrumental argument. As before, if this argument does not move at all (143a) or undergoes movement that could be A- or \bar{A} -movement (143b), LDA with *baat* ‘talk’ is optional. If *har laṛke-se* ‘every child-INSTR’ A-moves into the matrix clause, LDA with *baat* becomes obligatory:

- (143) *Movement and LDA: Instrumental*
- a. [us-kii_i mǎǎ-ne] [har laṛke-se_j baat kar-nii/-naa]
 3SG-GEN mother-ERG every boy-INSTR talk.F do-INF.F.SG/-INF.M.SG
 caah-ii/-aa
 want-PFV.F.SG/-PFV.M.SG
 ‘His/her_i mother wanted to talk to every boy_j’
- b. har laṛke-se_j [us-kii_i mǎǎ-ne] [t_j baat kar-nii/-naa]
 every boy-INSTR 3SG-GEN mother-ERG talk.F do-INF.F.SG/-INF.M.SG
 caah-ii/-aa
 want-PFV.F.SG/-PFV.M.SG
 ‘His/her_i mother wanted to talk to every boy_j’

- c. har larke-se_i [us-kii_i maa-ne] [t_i baat kar-nii/*-naa]
 every boy-INSTR 3SG-GEN mother-ERG talk.F do-INF.F.SG/*-INF.M.SG
 caah-ii/*-aa
 want-PFV.F.SG/*-PFV.M.SG
 ‘For every boy *x*, *x*’s mother wanted to talk to *x*.’

The structure of the paradigm in (143) is provided in a schematic form in (144):

(144) *Schematic structure of (143)*

- a. [matrix clause ... [nonfinite clause ... DP_{INSTR} DO ...]] → LDA with DO optional
- b. [matrix clause DP_{INSTR} ... [nonfinite clause ... t DO ...]] → LDA with DO optional
- A/A-movement
- c. [matrix clause DP_{INSTR} ... [nonfinite clause ... t DO ...]] → LDA with DO obligatory
- A-movement

In sum, the range of data just presented all fall under the generalization in (145). If any element undergoes A-movement out of the nonfinite clause, then this clause is necessarily transparent for LDA and LDA emerges as obligatory. Obligatory LDA does not implicate or require movement of the agreement controller. Furthermore, no such interaction exists if the movement could be produced by either A- or \bar{A} -movement. This leads us to the conclusion that \bar{A} -movement does *not* interact with agreement in this way.³²

³² While I have used the embedding predicate *caah* ‘want’ in the examples above, the examples in (i) and (ii) demonstrate that the same implicational relationship also holds for *shuruu kar* ‘start’. (i) involves possessor extraction analogous to (138):

- (i) a. [us-kii_i patnii-ne] [[DP har lekhak-kii_j kitaabē] paṛh-nii/-naa] shuruu
 3SG-GEN wife-ERG every author-GEN books.F read-INF.F.SG/-INF.M.SG start
 kii/kiyaa
 do.PFV.F.PL/do.PFV.M.SG
 ‘His_i wife started to read every author’s_j books.’
- b. har lekhak-kii_j [us-kii_i patnii-ne] [[DP t_j kitaabē] paṛh-nii/-naa] shuruu
 every author-GEN 3SG-GEN wife-ERG books.F read-INF.F.SG/-INF.M.SG start
 kii/kiyaa
 do.PFV.F.PL/do.PFV.M.SG
 ‘His_i wife started to read every author’s_j books.’

- (145) **A-Movement–Agreement Generalization** [= (132)]
 A-movement of *any* element out of a nonfinite clause makes LDA into this clause obligatory.
 \bar{A} -movement has no such effect.

A crucial feature of this generalization is that it applies to the embedded clause rather than the individual syntactic items within it, because the implication holds even if A-movement and LDA target distinct elements. Put somewhat differently, (145) is an implication between properties of the embedded clause itself, not between properties of individual syntactic elements inside it.

The evidence for (145) discussed so far employed weak crossover as a diagnostic for A-movement. As discussed at length in section 2.2.1, crossover is not the only diagnostic distinguishing between A- and \bar{A} -movement in Hindi. If (145) is correct, we expect other diagnostics to give rise to the same implicational relationship with LDA. The next two sub-sections are devoted to demonstrating that this is indeed the case.

2.3.3.2 Reciprocal binding

Recall from section 2.2.1 above that A-movement can feed reciprocal binding, while \bar{A} -movement cannot. Using this as a diagnostic, (146) demonstrates that A-movement of the indirect object out of the nonfinite clause requires LDA by the direct object. In the baseline sentence (146a), no movement takes place and the reciprocal pronoun lacks an antecedent, rendering the sentence ungrammatical. In (146b), the indirect object *siitaa aur sangiitaa-ne* ‘Sita and Sangita-ERG’ is moved over the matrix subject, binding the reciprocal. LDA with *ek maalaa* ‘a necklace’ becomes obligatory as a result.

-
- c. har lekhak-kii_i [*us-kii_i* patnii-ne] [[DP *t_i* kitaabē] paṛh-nii/?*-naa] shuruu
 every author-GEN 3SG-GEN wife-ERG books.F read-INF.F.SG/?*-INF.M.SG start
 kīī/?*kiyaa
 do.PFV.F.PL/?*do.PFV.M.SG
 ‘For every author *x*, *x*’s wife started to read *x*’s books.

The example in (ii) makes use of the idiom *X-kii khuub marammat kar* ‘give X a good beating’ analogous to (142):

- (ii) har bacce-kii_i [*us-kii_i* maa-ne] [[DP *t_i* khuub marammat] kar-nii/?*-naa] shuruu
 every child-GEN 3SG-GEN mother-ERG lot repair.F.SG do-INF.F.SG/?*-INF.M.SG start
 kīī/?*kiyaa
 do.PFV.F.PL/?*do.PFV.M.SG
 ‘For every child *x*, *x*’s mother started to give *x* a good beating.’

For some speakers, LDA is strongly preferred with *shuruu kar*, in which case the effect of A-movement is diminished (see fn. 30 on p. 94).

Because only A-movement can feed reciprocal binding (see section 2.2.1), (146b) must involve A-movement.

(146) *Reciprocal binding and LDA: Indirect object*

- a. **ek-duusre-ke_i* bhaaiyō-ne [[siitaa aur sangiitaa]-ko_i ek maalaa
 each other's brothers-ERG Sita and Sangita -DAT a necklace.F
 de-nii/-naa] caah-ii/-aa
 give-INF.F.SG/-INF.M.SG want-PFV.F.SG/-PFV.M.SG
 'Sita and Sangita, each other's brothers wanted to give (them) a necklace.'
- b. [siitaa aur sangiitaa]-ko_i *ek-duusre-ke_i* bhaaiyō-ne [*t_i* ek maalaa
 Sita and Sangita -DAT each other's brothers-ERG a necklace.F
 de-nii/*-naa] caah-ii/*-aa
 give-INF.F.SG/*-INF.M.SG want-PFV.F.SG/*-PFV.M.SG
 'Sita and Sangita, each other's brothers wanted to give (them) a necklace.'

(146) thus provides additional support for the generalization (145). Further support comes from (147), in which the moving element is the possessor of the embedded object. Analogous to (146), A-extraction out of the lower clause makes LDA with the embedded direct object *kitaabē* 'books' obligatory.

(147) *Reciprocal binding and LDA: Possessor*

- a. **ek-duusre-ke_i* bhaaiyō-ne [[DP [siitaa aur sangiitaa]-kii_i kitaabē]
 each other's brothers-ERG Sita and Sangita -GEN books.F
 parh-nii/-naa] caah-īī/-aa
 read-INF.F.PL/-INF.M.SG want-PFV.F.PL/-PFV.M.SG
 'Sita and Sangita, each other's brothers wanted to read (their) books.'
- b. [siitaa aur sangiitaa]-kii_i *ek-duusre-ke_i* bhaaiyō-ne [[DP *t_i* kitaabē]
 Sita and Sangita -GEN each other's brothers-ERG books
 parh-nii/*-naa] caah-īī/*-aa
 read-INF.F.SG/*-INF.M.SG want-PFV.F.PL/*-PFV.M.SG
 'Sita and Sangita, each other's brothers wanted to read (their) books.'

Thus, utilizing reciprocal binding instead of weak crossover as an indicator for the presence of A-movement yields identical results, supporting (145).

2.3.3.3 Quantifier scope

The final, and more exploratory, test for A-movement out of the lower clause involves quantifier scope. As already mentioned in section 2.2.1, movement has a scope-widening effect in Hindi. Thus, the subject has to take scope over the object in the absence of movement in (148), whereas the reverse is possible if the object moves over the subject in (149b).

(148) *kisii laṛkii-ne har laṛke-ko ḍāāṭ-aa* [= (59)]
 some girl-ERG every boy-ACC scold-PFV.M.SG
 ‘Some girl scolded every boy.’ $(\exists > \forall; * \forall > \exists)$

(149) a. *sab tiin ciizē khariid-ēge* [= (60)]
 everyone three things buy-FUT.M.PL
 ‘Everyone will buy three things.’ $(\forall > 3)$

b. *tiin ciizē_i sab t_i khariid-ēge*
 three things everyone buy-FUT.M.PL
 ‘Everyone will buy three things.’ $(3 > \forall; \forall > 3)$

(Mahajan 1997: 199)

Importantly, only A-movement has a scope-widening effect. Movement out of finite clauses, which only allow \bar{A} -extraction, does not extend the domain of scope taking. Thus, *har kek* ‘every cake’ in (150) cannot scope over *kisii laṛke-ne* ‘some boy-ERG’, despite the fact that it occupies a structurally higher position.³³

(150) \bar{A} -movement does not extend scope
har kek_i kisii laṛke-ne soc-aa [*ki prataap-ne t_i khaa li-yaa*
 every cake some boy-ERG think-PFV.M.SG that Pratap-ERG eat take-PFV.M.SG
hai]
 be.3.SG
 ‘Every cake, some boy thought that Pratap has eaten (it).’ $(\exists > \forall; * \forall > \exists)$

³³ As mentioned in fn. 6 on p. 55, quantificational elements other than numerals exhibit a strong preference for taking surface scope with respect to other nominals even if they undergo movement. It is only if reconstruction is forced by some other means (e.g., anaphor binding) that scope reconstruction becomes possible. The scopal elements in this section are all of this type, because the scope isomorphism only arises with respect to other numerals. The crucial diagnostic in the examples here is scope with respect to a higher verb. Scope rigidity hence does not interfere in these cases. Moreover, the judgments reported here remain unchanged if the quantificational elements are replaced with numerals.

If only A-movement can extend scope domains, scope provides a third way of analytically separating A- from \bar{A} -movement. As Bhatt (2005) observes, scope and LDA do indeed interact. Consider the contrast in (151). Under LDA, the embedded object *har kitaab* ‘every book’ can take scope either above the matrix verb *caah* ‘want’ or under it, as (151a) shows. With default agreement, by contrast, only a low construal of *har kitaab* is possible, see (151b).

(151) *Scope and LDA: Direct object*

a. *LDA*

naim-ne har kitaab paṛh-nii caah-ii thii [= (101)]
 Naim-ERG every book.F read-INF.F.SG want-PFV.F.SG be.PST.F.SG
 ‘Naim wanted to read every book.’
 $\forall > want$: ‘For every book, Naim wanted to read it.’
 $want > \forall$: ‘Naim’s desire: to read every book’

b. *Default agreement*

naim-ne har kitaab paṛh-naa caah-aa thaa
 Naim-ERG every book.F read-INF.M.SG want-PFV.M.SG be.PST.M.SG
 ‘Naim wanted to read every book.’
 $*\forall > want$: ‘For every book, Naim wanted to read it.’
 $want > \forall$: ‘Naim’s desire: to read every book’ (Bhatt 2005: 799)

Because only A-movement may extend a nominal’s scope domain, high scope of *har kitaab* in (151) is possible only if this element undergoes (string-vacuous) A-movement into the matrix clause. Because such wide scope is possible only under LDA, the contrast in (151) falls under the broader generalization in (145): To take matrix scope, *har kitaab* ‘every book’ has to A-move into the matrix clause, producing obligatory LDA. At the same time, (151) is fully consistent with the claim defended here that LDA is possible in the absence of any movement. Thus, if *har kitaab* takes embedded scope in (151), either LDA or default agreement are possible.

We are now in the position to ask whether extending the scope of an element other than the direct object has an effect on LDA. If the generalization in (145) is correct, we predict the answer to be yes. As (152) demonstrates, this prediction is borne out. The indirect object *har laṛkii-ko* ‘every girl-DAT’ can take matrix scope only if the direct object *film* ‘movie’ controls LDA.

(152) *Scope and LDA: Indirect object*

a. *LDA*

naim-ne har lar̥kii-ko film dikhaa-nii caah-ii
Naim-ERG every girl-DAT movie.F show-INF.F.SG want-PFV.F.SG

‘Naim wanted to show a movie to every girl.’

$\forall > want$: ‘For every girl, Naim wanted to show a movie to that girl.’

$want > \forall$: ‘Naim’s desire: to show a movie to every girl.’

b. *Default agreement*

naim-ne har lar̥kii-ko film dikhaa-naa caah-aa
Naim-ERG every girl-DAT movie.F show-INF.M.SG want-PFV.M.SG

‘Naim wanted to show a movie to every girl.’

$??\forall > want$: ‘For every girl, Naim wanted to show a movie to that girl.’

$want > \forall$: ‘Naim’s desire: to show a movie to every girl.’

To take wide scope, *har lar̥kii-ko* ‘every boy-DAT’ has to A-move into the matrix clause. As a result of this raising, the embedded direct object *film* ‘movie’ has to control LDA.

2.3.3.4 Section summary

In this section, we have considered a range of configurations in which the optionality of LDA disappears and LDA becomes obligatory. All of these configurations fall under the unified generalization in (153):

(153) **A-Movement–Agreement Generalization** [= (145)]
A-movement of *any* element out of a nonfinite clause makes LDA into this clause obligatory.
 \bar{A} -movement has no such effect.

We have seen a wide range of evidence supporting (153): First, the application of A-movement can be diagnosed with either weak crossover, reciprocal binding or quantifier scope. Furthermore, it is insubstantial for (153) which element undergoes A-movement. We have seen that A-movement of embedded direct objects, indirect objects, possessors and instrumentals all have the effect of rendering LDA with the embedded direct object obligatory. The fact that such a range of tests and configurations converges on one and the same generalization provides strong support for (153).³⁴

³⁴ It is noteworthy that the generalization (153) also provides novel evidence that LDA in Hindi does not require movement of the agreement controller. To see this, consider what a movement-based analysis would have to say to capture (153). Because A-movement of, e.g., the indirect object renders LDA with the direct object obligatory, a movement analysis would have to stipulate that A-movement of the indirect object enforces raising of the direct

2.3.4 Implications for selective opacity

The most intriguing aspect of (153) is that the implicational relationship between A-movement and ϕ -agreement does not hold at the level of an individual DP, but rather at the level of the embedded clause as a whole. Thus, LDA into the embedded object becomes obligatory if there is any A-extraction out of this clause. As emphasized above, it is insubstantial for (153) which element undergoes this A-movement step. In other words, A-movement of one element can make LDA with some other element obligatory. In this sense, the generalization in (153) is fundamentally *clause-based*.

What does (153) tell us about selective opacity? I suggest that we can make sense of (153) if the following hold: (i) nonfinite clauses are structurally ambiguous, (ii) the locality of both ϕ -agreement and A-movement is sensitive to this ambiguity, and (iii) the locality of \bar{A} -movement is not sensitive to the structural distinction. In other words, we can understand (153) as resulting from a locality mismatch, hence selective opacity.

I have proposed in section 2.3.2 above that nonfinite clauses in Hindi are structurally ambiguous between a TP and a ν P structure. I also proposed that only ν P clause allow ϕ -agreement into them, while TP clauses are impermeable to ϕ -agreement. We have seen above that the two assumptions afford a simple account of the surface optionality of LDA. Independent evidence for this analysis came from the observation that LDA is strongly dispreferred if the embedded clause contains a temporal adverb and hence a TP structure (recall (130)).

The A-movement–agreement generalization (153) on the one hand provides further support for this account, and furthermore sheds light on the locality of A- and \bar{A} -movement with respect to TP and ν P clauses. What (153) reveals is that A-movement has the same locality profile as ϕ -agreement, while \bar{A} -movement does not. Thus, let us suppose that a TP is opaque to both A-movement and ϕ -agreement, but transparent to \bar{A} -movement. Bare ν P clauses, on the other hand, are transparent to all three operations. Assuming agreement obligatoriness (154), the requirement that ϕ -agreement is obligatory if it is possible, the consequences of this state of affairs are depicted in (155).

object into the matrix clause. Not only is it unclear how such *tandem movement* could be enforced technically, it is also, to the best of my knowledge, unattested in other syntactic domains and dubious for this reason. An analysis that allows direct agreement with the direct object in its base position does not face this problem, as we will see shortly.

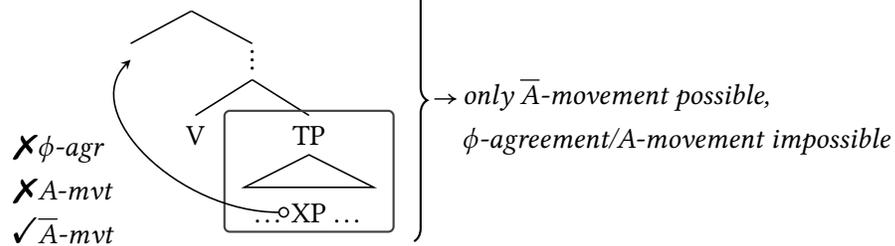
(154) **Agreement obligatoriness** (Preminger 2011, 2014)

[=(123)]

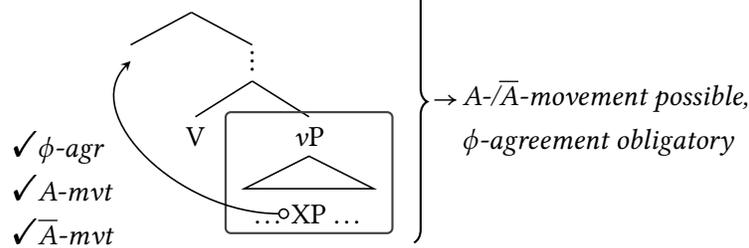
If a verb can ϕ -agree with a DP, it has to.

(155) *Two nonfinite-clause structures and their opacity profiles*

a. *TP nonfinite clauses*



b. *vP nonfinite clauses*



The A-movement–agreement generalization (153) then follows as a direct consequence of the situation in (155): Because A-movement is possible only out of vP clauses, A-extraction of any element entails the vP structure in (155b). By assumption, this vP structure is permeable to ϕ -agreement, which is hence obligatory (by (154)). \bar{A} -movement, by contrast, is possible out of vP and TP clauses and therefore does not disambiguate the structure of the nonfinite clause. As a consequence, LDA remains optional if \bar{A} -movement takes place. Because this line of reasoning is stated on the basis of the embedded clause, it effortlessly captures the core feature of (153), namely that the element that A-moves does not have to be the element that obligatorily triggers LDA as a result. Because A-extraction of *any* element out of the lower clause entails a vP structure, any A-extraction will render LDA obligatory. This straightforwardly accommodates the range of facts observed in section 2.3.3, all of which are subsumed under (153).

On the type of account just sketched, (153) is a particular instance of selective opacity, because the generalization emerges as a consequence of locality mismatches between operations. ϕ -Agreement has the same locality properties as A-movement, which are crucially different from \bar{A} -movement. As a result, ϕ -agreement interacts only with A-movement, not with \bar{A} -movement.

Adding these insights about the locality properties of TP and vP clauses to the properties of finite clauses discussed in section 2.2 yields the summary table in (156), an extension of (104) above.

(156) *Interim summary (to be extended)*

<i>Operation</i>	Finite clauses		Nonfinite clauses	
	CP	TP	vP	
\bar{A} -movement	✓	✓	✓	
A-movement	✗	✗	✓	
ϕ -agreement	✗	✗	✓	

(✓: transparent; ✗: opaque)

In addition to deepening our understanding of clausal structure in Hindi, the insights in (156) also shed light on the broader notion of selective opacity that is the primary concern of this thesis. Because (156) instantiates a locality mismatch between \bar{A} -movement on the one hand, and A-movement and ϕ -agreement on the other, it is an example of selective opacity. Crucially, this selective opacity pattern is not merely between different types of movement, but encompasses ϕ -agreement in an entirely parallel way.

Importantly, unlike in English, A-movement and ϕ -agreement are separate operations in Hindi and do not necessarily go together. While in English A-movement to subject position and ϕ -agreement entail each other (at least in the majority of cases), this is not the case in Hindi, where one may take place without the other. First, we have already seen that ϕ -agreement is possible in the absence of movement, because elements that resist movement can nonetheless control agreement (recall the discussion of the idiom *X-kii khuub marammat kar* ‘give X a good beating’ (lit. ‘do X’s many repairs’) in section 2.2.2). Second, we saw in section 2.2.2 that A-movement of the object over the subject does not affect agreement with the subject (see (86b)). A-movement of an element is hence independent of ϕ -agreement with that element. This characterization is also forced on us based on the evidence in section 2.3.3, which shows that A-movement and ϕ -agreement can target distinct elements. In sum, then, A-movement and ϕ -agreement are *operationally independent* in Hindi. The fact that they exhibit the same locality profile distinct from that of \bar{A} -movement must hence follow from a locality restriction that is abstract enough to restrict both A-movement and ϕ -agreement.

Furthermore, because LDA in Hindi does not require movement of the agreement controller. As a consequence, any notion of selective opacity that is solely based on movement dependencies will not extend to ϕ -agreement and hence leave the parallelism between A-movement and ϕ -agreement unaccounted for. To appreciate this point, consider the fact that \bar{A} -movement may leave a finite clause, while A-movement may not, a standard instance of selective opacity analogous to hyper-raising in English. No approach that derives this restriction from a constraint on movement proper is able to extend this restriction to the entirely parallel fact that ϕ -agreement is likewise impossible across a finite clause boundary, precisely because ϕ -agreement does not involve movement and constraints on movement therefore do not apply to it. As a result, any movement-based account of selective opacity is forced to invoke a second, unrelated constraint to restrict ϕ -agreement, which happens to mimic the locality of A-movement.³⁵ The evidence presented in this section further aggravates this problem because it demonstrates that the locality of ϕ -agreement matches precisely that of A-movement. The limitation of movement-based approaches is that this is nothing but a coincidence, as a matter of principle.

What these considerations highlight is the need for a theory of selective opacity that is more abstract than previous approaches in that it must encompass *both movement and ϕ -agreement*. Only an account that is abstract enough to constrain both types of operations in a uniform way allows us to capture the identical locality properties of A-movement and ϕ -agreement, and hence the A-movement–agreement generalization in (153), in a truly uniform way. At the same time, such an account has to preserve the key characteristic of selective opacity: The locality properties of \bar{A} -movement are evidently distinct from those of A-movement and ϕ -agreement.

An account of selective opacity that meets these requirements is developed in chapter 3. Before turning to this task, however, I will add a fourth operation to our considerations of Hindi: *wh*-licensing. As I will demonstrate in the next section, *wh*-licensing like ϕ -agreement does not

³⁵ Consider as an example May's (1979) classic account of improper movement, adopted in Chomsky (1981), according to which traces left by \bar{A} -movement are variables, subject to Principle C. Assuming furthermore that extraction out of a finite clause requires a prior \bar{A} -step to the edge of the clause, further A-movement is ruled out because the \bar{A} -trace would be bound from an A-position, violating Principle C. The limitation of such an account is that it has nothing to say about ϕ -agreement: There would be no expectation that the landing site of an \bar{A} -moved element should be invisible for ϕ -agreement. In other words, the fact that ϕ -agreement can never take place across a finite clause boundary in Hindi and that, consequently, elements which are \bar{A} -moved to the edge of a finite clause cannot control ϕ -agreement on a higher verb must be the result of a stipulation separate from the one barring them from undergoing A-movement.

involve movement. Furthermore, it displays a locality profile distinct from the ones considered so far.

2.4 *Wh*-licensing

This section provides an investigation into the properties of *wh*-licensing in Hindi, with a particular emphasis on its locality profile. First, I will show that *wh*-licensing does not involve movement, but a genuinely long-distance relationship between an interrogative DP and a C⁰ head. Second, we will see that *wh*-licensing is impossible across a CP boundary, but licit across TP and *v*P clauses. Its locality is hence distinct from the three operations considered so far (A-movement, \bar{A} -movement, and ϕ -agreement). These facts will thus demonstrate that selective opacity in Hindi is not a binary distinction, corroborating one of the core insights of the previous literature, reviewed in chapter 1. Moreover, they offer additional support for the conclusion reached in the preceding section, namely that selective opacity is not a distinctive property of movement, but rather encompasses movement and non-movement operations alike.

2.4.1 The absence of *wh*-movement: Evidence from focus intervention

It is well-known that *wh*-elements in Hindi do not undergo obligatory overt movement to SpecCP (Gurtu 1985, 1992, Mahajan 1990, Srivastav 1991b, Dayal 1994b, 1996, 2000, Manetta 2010, 2011). Instead, they have a weak tendency to appear in a preverbal focus position, but they may also undergo scrambling into other positions (see, e.g., Mahajan 1990, Dayal 1994b, 1996).

(157) *No overt wh-movement to SpecCP*

raam-ne kyaa khaa-yaa thaa
Ram-ERG what eat-PFV.M.SG be.PST.M.SG
'What did Ram eat?'

(Mahajan 1990: 125)

While it is thus clear that Hindi has no overt *wh*-movement to SpecCP, the question whether there is covert movement is much harder to answer. Traditional analyses of (157) answer in the affirmative (e.g., Srivastav 1991a,b, Mahajan 1990, Dayal 1994b, 1996), whereas, e.g., Simpson (2000)

and Manetta (2010, 2011) argue against it.³⁶ I will present here evidence from focus intervention against the existence of covert *wh*-movement in Hindi. Rather, Hindi *wh*-licensing is genuinely long-distance and does not require any overt or covert movement.

FOCUS INTERVENTION EFFECTS (sometimes called ‘**BECK EFFECTS**’) refer to the phenomenon that *wh*-elements cannot be associated with an interrogative C⁰ head if a quantificational or focus element intervenes between the two, as schematized in (159). Interveners in this sense include negation, negative quantifier, negative polarity items, focus operators, etc. See Hoji (1985), Takahashi (1990), Beck (1996, 2006), Beck & Kim (1997), Hagstrom (1998), Pesetsky (2000), Kim (2002), Cable (2007, 2010), Tomioka (2007), Miyagawa (2010), Kotek (2014), and references cited there for extensive discussion of focus intervention in a variety of languages and constructions.

(158) **Focus intervention** (Beck 2006: 5)

A quantificational or focusing element may not intervene between a *wh*-phrase and its licensing complementizer.

(159) * $[C_i [\dots [\text{intervener} [\dots \text{wh-phrase}_i \dots]]]]$

Focus intervention is illustrated on the basis of German examples in (160), from Beck (1996: 1,4,5). In (160a), the negative quantifier *niemand* ‘nobody’ intervenes between the interrogative C⁰ head and the *wh*-element *wo* ‘where’. The resulting sentence is ungrammatical. That this ungrammaticality is indeed caused by the negative quantifier is illustrated in (160b), in which the quantifier is replaced with a proper name and the resulting sentence is well-formed. Finally, (160c) demonstrates that no intervention arises if the *wh*-element is moved over the quantifier so that the latter no longer intervenes.

(160) *Focus intervention in German*

- a. *Wen hat niemand wo gesehen?
whom has nobody where seen
Intended: ‘Where did nobody see whom?’
- b. Wen hat Luise wo gesehen?
whom has Luise where seen
‘Where did Luise see whom?’

³⁶ Simpson (2000) argues that no *wh*-movement at all takes place, whereas Manetta (2010, 2011) proposes that *wh*-elements move to SpecvP but not higher.

- c. Wo_i hat **niemand** Karl t_i gesehen?
 where has nobody Karl seen
 ‘Where did nobody see Karl?’

Focus intervention effects have received considerable attention in the literature and it is not my goal to do justice to this literature here. Rather, my intention is to use our current understanding of focus intervention as a window into the syntax of *wh*-licensing in Hindi.

Beck (2006) develops an account in which focus intervention follows from the very semantics of *wh*-phrases and focus. The intuition underlying Beck’s (2006) account is that *wh*-elements and focus are interpreted using the same semantic mechanism. This causes interference in the configuration (159), which renders the entire structure uninterpretable. A crucial component of Beck’s (2006) analysis is that an in-situ *wh*-phrase does not covertly move to the CP that designates its scope. Rather, *wh*-licensing can apply long-distance, with the semantic association between the *wh*-element and its licensing C^0 being accomplished by projection of alternatives (Hamblin 1973, Rooth 1985, Kratzer & Shimoyama 2002). The requirement for one *wh*-element to move to the left periphery of the clause is then viewed as a purely syntactic property. I will briefly review the basics of Beck’s (2006) account and then turn to the relevance of focus intervention effects for the syntax of *wh*-licensing.

Borrowing from Rooth (1985, 1992), Kratzer (1991), and Wold (1996), Beck (2006) assumes that each structural node α is associated with two interpretations: an ordinary semantic value $\llbracket \alpha \rrbracket^g$ and focus semantic interpretation $\llbracket \alpha \rrbracket^{g,h}$, which contains a second assignment function h for distinguished variables. Focus semantics corresponds to the value assigned to such variables by h . In the absence of focus, $\llbracket \alpha \rrbracket^g = \llbracket \alpha \rrbracket^{g,h}$. Beck proposes that *wh*-elements lack an ordinary semantic value and only receive a focus semantic interpretation, as in (161).

- (161) a. $\llbracket who_1 \rrbracket^g$ is undefined
 b. $\llbracket who_1 \rrbracket^{g,h} = h(1)$

The undefinedness of the ordinary semantic value percolates up the tree from the *wh*-expression, with only the focus semantic value being interpreted.

- (162) a. $\llbracket \text{who}_1 \text{ left} \rrbracket^g$ is undefined
 b. $\llbracket \text{who}_1 \text{ left} \rrbracket^{g,h} = \lambda w. h(1) \text{ left in } w$

The semantic contribution of a C^0 head is to introduce question semantics in the form of a question operator Q , whereby questions denote a set of propositions (see Hamblin 1973, Karttunen 1977), and, importantly, to set the ordinary semantic value to the focus value:

- (163) If $X = [Q_i Y]$, then $\llbracket X \rrbracket^g = \lambda p \exists x [p = \llbracket Y \rrbracket^{g,h[x/i]}]$ (Beck 2006: 16)
 and $\llbracket X \rrbracket^{g,h} = \lambda p \exists x [p = \llbracket Y \rrbracket^{g,h[x/i]}]$

Applied to our example, the complementizer yields a question interpretation as the ordinary semantic value of the sentence:

- (164) $\llbracket [Q_1 [\text{who}_1 \text{ left}]] \rrbracket^g = \lambda p \exists x [p = \llbracket [\text{who}_1 \text{ left}] \rrbracket^{g,h[x/1]}]$
 $= \lambda p \exists x [p = \lambda w. x \text{ left in } w]$

To interpret focus, Beck (2006) employs Rooth's (1985, 1992)'s \sim operator, which is interpreted as in (165).³⁷ The details of (165) do not need to concern us here. What is crucial for our purposes is that one contribution of the operator is to set the focus semantic value to the ordinary semantic value (165b).

- (165) If $X = [\sim C Y]$, then
 a. $\llbracket X \rrbracket^g = \llbracket Y \rrbracket^g$ if $g(C) \subseteq \{\llbracket Y \rrbracket^{g,h'} : h' \in H \wedge h' \text{ is total}\}$,
 b. $\llbracket X \rrbracket^{g,h} = \llbracket X \rrbracket^g$ (Beck 2006: 15)

Because all focus computation involves the \sim operator, focus intervention in (159) is equivalent in Beck's (2006) system to a \sim intervening between a *wh*-expression and its licensing Q on a C^0 head. As (166) illustrates for a toy example corresponding to, e.g., a multiple question in German in which a *wh*-element remains below a focus intervener, such structures emerge as semantically undefined. Because TP_2 contains a *wh*-element, it has no ordinary semantic value and instead only receives a focus semantic interpretation (166b), analogous to (162). As stated in (165b), \sim defines the focus semantic value $\llbracket X \rrbracket^{g,h}$ as the ordinary semantic value $\llbracket X \rrbracket^g$. But because the ordinary

³⁷ I am using the calligraphic ' \mathcal{C} ' to refer to the focus anaphor to keep it notationally distinct from the complementizer, which is the location of the Q operator. H is the set of focus variable assignments (Wold 1996).

values of both TP₂ and TP₁ are undefined, their focus semantic values will likewise be undefined, as in (166c). With both values being undefined, even the subsequent addition of Q₁ will not recover a semantic interpretation for the sentence. Consequently, the entire sentence has neither an ordinary nor a focus-semantic value and it hence uninterpretable. In this way, Beck (2006) is able to derive focus intervention effects as semantic uninterpretability.

(166) *Only John saw who?

- a. [Q₁ [... [TP₁ ~C [TP₂ John saw who₁]]]]
- b. i. [[TP₂]]^g is undefined
 - ii. [[TP₂]]^{g,h} = λw. John saw h(1) in w
- c. i. [[TP₁]]^g = [[TP₂]]^g = undefined (by (165a))
 - ii. [[TP₁]]^{g,h} = [[TP₁]]^g = undefined (by (165b))

An important property of this account is that a focus element causes intervention only if it intervenes between the *wh*-element and its licensing C⁰ head. Movement of the *wh*-element over the focus element therefore obviates the intervention effect. This is because the semantic association between the landing site and the gap does not involve focus computation and is hence not disturbed by the presence of an intervening focus:

(167) *Movement obviates intervention*

[Q_i *wh*-phrase_i [... [intervener [... t_i]]]]

(168) *Wen_i hat niemand t_i gesehen?*

[German]

whom has nobody seen

‘Who did nobody see?’

The same facts hold for covert movement. Like overt movement, covert movement over a focus intervener should circumvent the intervention effect and produce an interpretable and grammatical structure. This is the view that has recently been argued for by Kotek (2014). Kotek extends work by Pesetsky (2000), who observed that some *wh*-in-situ configurations in English give rise to focus intervention, while others do not. Specifically, Pesetsky (2000) argues that focus intervention, which is normally absent in English, nonetheless arises in superiority-violating questions. This is illustrated in (169):

- (169) a. *Superiority-obeying questions*
 i. Which student did Fred introduce to which professor?
 ii. Which student did **only Fred** introduce to *which professor*?
- b. *Superiority-violating questions*
 i. Which professor did Fed introduce which student to?
 ii. *Which professor did **only Fred** introduce *which student* to? (Kotek 2014:18)

In (169a) as well as (169b), a negation appears between a *wh*-element and the interrogative C⁰ head hosting the Q operator. Curiously, this causes intervention only in superiority-violating questions. Simplifying somewhat, Kotek (2014) proposes that in-situ *wh*-expressions undergo a covert movement step over the intervener in superiority-obeying questions (169a), but not in superiority-violating questions (169b).³⁸

- (170) *Kotek's (2014) account for (169)*
- a. *Superiority-obeying questions*

$$\begin{array}{c} wh_1 \quad wh_2 \quad \dots \quad \text{intervener} \quad \dots \quad t_1 \quad \dots \quad t_2 \\ \uparrow \quad \uparrow \quad \underbrace{\hspace{10em}} \quad \uparrow \\ \text{---} \quad \text{---} \end{array} \quad \rightarrow \text{covert movement of } wh_2 \rightarrow \text{no intervention}$$
- b. *Superiority-violating questions*

$$\begin{array}{c} wh_1 \quad \dots \quad \text{intervener} \quad \dots \quad wh_2 \quad \dots \quad t_1 \\ \uparrow \quad \underbrace{\hspace{10em}} \quad \uparrow \end{array} \quad \rightarrow \text{no movement of } wh_2 \rightarrow \text{intervention}$$

As a result of the covert movement of *wh*₂ in (170a), no focus intervention effect arises, as the focus element does not intervene between the LF position of *which book* and its licenser. In the configuration for superiority-violating questions in (170b), on the other hand, *wh*₂ does *not* covertly move over the intervener, and focus intervention results, in a way that is parallel to Beck's (2006) original proposal.

³⁸ To implement the general picture in (170), Kotek (2014) assumes that *wh*-DPs can be of two types. They can either be QPs or DPs. QPs move to their licensing C head, although only the movement of the structurally highest QP is overt. DPs, on the other hand, cannot move and are instead interpreted via percolation of focus alternatives. Superiority is the requirement that the structurally highest QP undergo overt fronting. In superiority-obeying configurations, both *wh*-elements can be QPs and move (the first one overtly, the other covertly) over the focus intervener, thereby circumventing focus intervention. Superiority-violating questions result if the higher *wh*-element (*wh*₂ in (170b)) is a DP and therefore unable to *wh*-move. Its inability to move also entails that it has to be interpreted via percolation of alternatives and that it is hence susceptible to focus intervention.

An alternative view, pointed out to me by Kyle Johnson, is that the covert movement step in (170a) is QR. The pervasiveness of focus intervention in German and, as we will see, Hindi could then be attributed to the fact that these languages show scope rigidity in the base order and therefore lack QR.

For our concerns, the core conclusion of these considerations is that focus intervention arises if *neither overt nor covert movement take place*. In other words, focus intervention effects serve as a tool to diagnose the *absence of covert movement*. Applying this desideratum to Hindi, we observe that Hindi exhibits focus intervention effects, a fact already noted by Beck & Kim (1997) and Beck (2006).³⁹ This is illustrated with negative polarity items (NPIs) in (171). In (171a) the subject is an NPI licensed by the sentential negation *nahī̃*.⁴⁰ The object *kis-ko* ‘who-ACC’ appears in its base position, and focus intervention arises. Such intervention can be obviated by overtly moving the object over the NPI, as in (171b).⁴¹

³⁹ Beck & Kim (1997) and Beck (2006) give the contrast in (i):

- (i) a. ??koi nahī̃ kyaa paṛh-aa?
 anyone not what read-PFV.M.SG
Intended: ‘What did no one read?’
 b. kyaa koi nahī̃ paṛh-aa?
 what anyone not read
 ‘What did no one read?’

(Beck & Kim 1997: 377)

As they stand, however, both sentences are ungrammatical because they lack ergative case marking on the subject. Moreover, the contrast between (i.a) and (i.b) is confounded with an independent requirement for sentential negation in Hindi to occur in the immediately preverbal position. Thus, (i.a) is degraded even if the NPI is replaced with a proper name and the subject bears ergative case. The more suitable pair in (171) demonstrates that Beck & Kim’s (1997) generalization holds irrespective of this confound.

⁴⁰ See Bhatt & Homer (2014) for an investigation into the scope of negation in Hindi. They argue that *nahī̃* is not a direct realization of sentential negation. Instead, they propose that sentential negation is located higher in the tree but not itself realized.

⁴¹ Interestingly, Simpson & Bhattacharya (2003) note a restriction analogous to (171) in the related language Bangla, but take it as evidence in favor of overt movement of the *wh*-element.

- (i) a. *kew/sudhu meri ka-ke voṭ dæy-ni [Bangla]
 anyone/only Mary who-DAT vote gave-not
 b. ka-ke kew/sudhu meri voṭ dæy-ni
 who-DAT anyone/only Mary vote gave-not
 ‘Who did no one vote for?/Who did only Mary not vote for?’ (Simpson & Bhattacharya 2003: 140)

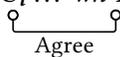
Instead of analyzing the ungrammaticality of (i.a) as a focus intervention effect as I have done here, Simpson & Bhattacharya (2003) propose that *wh*-elements in Bangla undergo obligatory overt movement to the left periphery, producing (i.b). The impossibility of (i.a) is attributed to the claim that NPIs and other quantified elements cannot topicalize and therefore cannot end up to the left of a fronted *wh*-element. While intriguing, there is evidence against this analysis at least in Hindi. As (ii) demonstrates, NPIs may freely move in Hindi, even out of a finite clause:

- (ii) koi-bhii_i mujhe nahī̃ lag-taa [ki t_i aa-yegaa]
 some-NPI me.DAT not seem-IPFV.M.SG that come-FUT.M.SG
 ‘It doesn’t seem to me that anyone will come.’

(171) *Focus intervention in Hindi*

- a. ??**kisi-bhii laṛke-ne** *kis-ko nahī̃ dekh-aa?*
 some-NPI boy-ERG who-ACC not see-PFV.M.SG
Intended: ‘Who did no boy see?’
- b. *kis-ko_i kisi-bhii laṛke-ne t_i nahī̃ dekh-aa?*
 who-ACC some-NPI boy-ERG not see-PFV.M.SG
 ‘Who did no boy see?’

Against the background of Beck’s (2006) and Kotek’s (2014) insights into focus intervention just discussed, the fact that (171a) results in a focus intervention effect provides evidence that Hindi lacks covert *wh*-movement. In Hindi, then, *wh*-elements do not obligatorily raise to CP, neither overtly nor covertly. I therefore conclude that *wh*-licensing in Hindi does not involve movement, but is instead accomplished by an Agree relation between an interrogative C⁰ and a *wh*-expression:

(172) [C_i ... *wh*-DP_i ...]


As a result of this Agree relation, the two elements are coindexed and interpreted via focus projection. As per Beck’s (2006) proposal, this projection is sensitive to intervention, which captures the degradedness of (171a).

2.4.2 The locality of *wh*-licensing

Having established that *wh*-licensing involves a long-distance relationship between a C⁰ head and an interrogative element, this section will consider the locality conditions on this relationship. One well-known restriction on *wh*-licensing in Hindi is that it cannot proceed past a finite clause boundary (see Mahajan 1990, 2000, Srivastav 1991a,b, Dayal 1994b, 1996, 2000, Lahiri 2002, Bhatt & Dayal 2007, Manetta 2010, 2011). Thus, a *wh*-element embedded inside a finite clause may take scope within this clause, but it may not take matrix scope. As an example, (173) can only be interpreted as an embedded question, not as a matrix question.

At least in Hindi, then, the impossibility of (171a) has nothing to do with the movement options of the NPI itself. Simpson & Bhattacharya’s (2003) analysis of Bangla therefore does not carry over to Hindi.

(173) *Finite clauses are islands for wh-scope*

tum jaan-te ho [ki us-ne *kyaa* ki-yaa]
 you know-IPFV.M.SG be.PRES.2SG that he-ERG what do-PFV.M.SG

‘You know what he did.’

Not: ‘What do you know he did.’

(Dayal 1996: 21)

If the embedding predicate does not allow for an interrogative complement clause, ungrammaticality results:

(174) *siitaa-ne soc-aa [ki ravii-ne *kis-ko* dekh-aa]
 Sita-ERG think-PFV.M.SG that Ram-ERG who-ACC see-PFV.M.SG

Intended: ‘Who did Sita think that Ravi saw?’

(Mahajan 2000: 319)

This restriction is fully general and holds irrespective of where the *wh*-element is located within the lower finite clause. As (175) shows, movement of *kis-ko* ‘who-ACC’ to the edge of the lower clause does not allow matrix scope either:

(175) *Finite clause edges are opaque to wh-licensing from matrix clause*

*siitaa-ne soc-aa [*kis-ko*_{*i*} ravii-ne *t*_{*i*} dekh-aa]
 Sita-ERG think-PFV.M.SG who-ACC Ram-ERG see-PFV.M.SG

Intended: ‘Who did Sita think that Ravi saw?’

It is worth noting that the inability for *wh*-licensing to pass a finite clause boundary is by no means a general feature of *wh*-in-situ languages. In Mandarin, for instance, such a relationship is perfectly possible (Huang 1982), and (176) is therefore ambiguous between an embedded and a matrix construal of the *wh*-phrase:

(176) Zhangsan zhidao [*shei* mai-le shu] (?)

[Mandarin]

Zhangsan know who buy-ASP book

‘Who does Zhangsan know bought books?’ *Or:*

‘Zhangsan knows who bought books.’

(Huang 1982: 254)

In contrast to finite clauses, nonfinite clauses do allow *wh*-licensing across them. This is illustrated in (177), in which *kyaa* ‘what’ inside the nonfinite clause takes matrix *wh*-scope.⁴²

⁴² Recall from section 2.3.1 that nonfinite clauses may not carry interrogative force in Hindi, a fact that I have taken to suggest that these clauses obligatorily lack a CP layer. The absence of an embedded *wh*-construal in (177) is a consequence of this general property of nonfinite clauses.

(177) *Nonfinite clauses are not wh-islands*

tum [*kyaa* kar-naa] jaan-te ho [= (108a)]
 you what do-INF.M.SG know-IPFV.M.PL be.PRES.3SG
 ‘What do you know to do?’
 Not: ‘You know what to do.’

I have argued in sections 2.3.2 and 2.3.3 that nonfinite clauses in Hindi are ambiguous between a TP and a *v*P structure. This raises the question of whether both types of clauses are transparent for *wh*-licensing or only one of them. Recall also that LDA allows us to distinguish between the two clauses, as it is possible into a *v*P clause, but not into a TP clause. As (178) shows, *wh*-association across a nonfinite clause boundary is possible with or without LDA.

(178) *Wh-licensing is independent of LDA*

a. *Default agreement*

raam-ne [*kaunse* baccō-ko film dikhaa-naa] caah-aa?
 Ram-ERG which child-DAT movie show-INF.M.SG want-PFV.M.SG
 ‘Which child did Ram want to show a movie to?’

b. *LDA*

raam-ne [*kaunse* baccō-ko **film** dikhaa-nii] caah-ii?
 Ram-ERG which child-DAT movie show-INF.F.SG want-PFV.F.SG
 ‘Which child did Ram want to show a movie to?’

Following (155), LDA diagnoses a *v*P structure, whereas default agreement diagnoses a TP structure. What (178) demonstrates is that *wh*-licensing is possible across both *v*P and TP structures.

There is independent evidence that *wh*-licensing of an element inside a nonfinite clause does not require movement of that element. This evidence comes from the *wh*-word *kyaa* ‘what’. A general property of this element is that it resists scrambling.

(179) *No movement of kyaa ‘what’*

a. sangiitaa-ne kyaa khaa-yaa
 Sangita-ERG what eat-PFV.M.SG
 ‘What did Sangita eat?’

b. ???kyaa_i sangiitaa t_i khaa-yaa
 what Sangita eat-PFV.M.SG

At the same time, *kyaa* can take matrix scope from inside an embedded nonfinite clause. This is illustrated in (180). (180a) shows that *kyaa* inside a nonfinite clause can be *wh*-licensed. (180b) shows that extraction of *kyaa* into the matrix clause is severely degraded. The inability of *kyaa* to undergo movement into higher clause then suggests that it is licensed in situ in (180). This in turn entails that nonfinite clauses must be transparent to *wh*-licensing.

(180) *Wh-licensing does not require movement*

- a. raam-ne [kyaa khaa-naa] caah-aa?
 Ram-ERG what eat-INF.M.SG want-PFV.M.SG
 ‘What did Ram want to eat?’
- b. ???kyaa_i raam-ne [_{t_i} khaa-naa] caah-aa?
 what Ram-ERG eat-INF.M.SG want-PFV.M.SG
Intended: ‘What did Ram want to eat?’

Converging evidence for the conclusion that nonfinite clauses are transparent to *wh*-licensing comes from temporal adverbs. Recall from the discussion at the end of section 2.3.2 that the temporal adverb *kal* ‘yesterday/tomorrow’ arguably requires a TP structure. The evidence for this claim came from the observation that LDA becomes dispreferred if *kal* modifies the embedded clause. (181) demonstrates that *wh*-licensing of *kyaa* inside a nonfinite clause is possible even if this nonfinite clause is modified by *kal* and hence a TP.

(181) *Wh-licensing into TP clauses*

- pichle hafte* raam-ne [*kal* kyaa khaa-naa] caah-aa
 last week Ram-ERG yesterday/tomorrow what eat-INF.M.SG want-PFV.M.SG
 thaa?
 be.PST.M.SG
 ‘What did Ram last week want to eat yesterday/tomorrow?’

Because *kyaa* is unable to move into the matrix clause, we can conclude from (181) that nonfinite TP clauses are also transparent for *wh*-licensing.

In sum, then, the picture that emerges from these considerations is that nonfinite TP and *vP* clauses are transparent to *wh*-licensing, but finite CP clauses are not:

(182) *Locality of wh-licensing*

Wh-licensing cannot cross a CP clause boundary in Hindi, but it may cross TP and *v*P clause boundaries.

Before moving on to integrate (182) into the broader picture developed in this chapter, I will briefly illustrate two strategies used in Hindi to form cross-clausal questions and argue that they do not constitute evidence against (182). Largely following the previous literature on these constructions, I will show that they allow the formation of cross-clausal questions precisely because they obviate the need for *wh*-licensing across a CP.

2.4.3 Detour 1: More *wh*-constructions in Hindi

As discussed in the preceding section, (183) is ungrammatical because matrix scope of *kis-ko* ‘who-ACC’ is blocked by (182) and embedded scope is incompatible with the semantics of the embedding predicate.

- (183) **siitaa-ne soc-aa [ki ravii-ne kis-ko dekh-aa]*
Sita-ERG think-PFV.M.SG that Ravi-ERG who-ACC see-PFV.M.SG
Intended: ‘Who did Sita think that Ravi saw?’ (Mahajan 2000: 319)

How does Hindi form questions corresponding to the English translation of (183)? There are two strategies. One strategy involves overt movement of the *wh*-element into the matrix clause:

- (184) *kis-ko_i siitaa-ne soc-aa [ki ravii-ne t_i dekh-aa]*
who-ACC Sita-ERG think-PFV.M.SG that Ravi-ERG see-PFV.M.SG
‘Who did Sita think that Ravi saw?’ (Mahajan 2000: 318)

Following Mahajan (1990) and Dayal (1994b, 1996, to appear), I will treat this construction as involving regular \bar{A} -movement into the higher clause, followed by *wh*-licensing from the landing site within that clause. That is, the movement in (184) is the same all-purpose \bar{A} -movement that was discussed in section 2.2.1 and that can target *wh* and non-*wh* elements alike. The claim that *wh*-elements can undergo regular \bar{A} -movement unrelated to their *wh*-construal is hardly controversial, as the example in (185) attests. The crucial difference between (184) and (185) is that the *wh*-element

takes embedded scope in (185), despite the fact that it is scrambled into the matrix clause, whereas in (184) the *wh*-element takes matrix scope.

- (185) *kaun_i* anu soc rahii hai [ki *t_i* aa-yegaa]
 who Anu wonder PROG.F.SG be.PRES.3SG that come-FUT.M.SG
 ‘Anu is wondering who will come.’ (Dayal to appear: (16))

(185) makes it clear that *wh*-elements can be moved into a clause higher than the one that they take scope in. Just like non-*wh* phrases, then, *wh*-elements can freely undergo \bar{A} -movement. (184) is thus entirely compatible with (182): The *wh*-item undergoes regular \bar{A} -movement across the CP and gets licensed by the matrix C^0 in its landing site. No CP is crossed by *wh*-licensing.

The second, and more common, strategy is what is often referred to as the **SCOPE-MARKING CONSTRUCTION**. In this construction, the *wh*-element remains in the lower clause, but the *wh*-word *kyaa* ‘what’ appears in the matrix clause:

- (186) *Scope-marking construction*
 siitaa-ne *kyaa* soc-aa [ki ravii-ne *kis-ko* dekh-aa]
 Sita-ERG what think-PFV.M.SG that Ravi-ERG who-ACC see-PFV.M.SG
 ‘Who did Sita think that Ravi saw?’ (Mahajan 2000: 317)

The scope-marking construction has received a significant amount of attention in the literature (see Srivastav 1991a,b, Mahajan 1990, 2000, Dayal 1994b, 1996, to appear, Lahiri 2002, Manetta 2010, 2011). It is not my goal here to provide an overview of this literature. Rather, I will briefly present the **INDIRECT-DEPENDENCY APPROACH** to this construction (developed in particular by Srivastav 1991b, Dayal 1994b, 1996, 2000, and Lahiri 2002) and demonstrate that this account is fully compatible with (182). For a more comprehensive presentation of the construction, its properties, and the range of analyses that have been explored, I refer the reader to these sources.

One of the distinguishing features of approaches to the scope-marking construction in (186) is whether the *kyaa* is a pleonastic or expletive argument (see Mahajan 1990, 2000, Manetta 2010, 2011), or whether it is the regular *wh*-word for ‘what’ (see Dayal 1994b, 1996, Lahiri 2002). The indirect-dependency approach treats it as the latter, justifiably so because the upper clause in (186) is itself a well-formed sentence in Hindi:

- (187) siitaa-ne kyaa soc-aa?
 Sita-ERG what think-PFV.M.SG
 ‘What did Sita think?’

Kyaa in (187) is a variable ranging over propositions. On the null assumption that the two instances of *kyaa* in (186) and (187) are the same element, Dayal (1994b, 1996) concludes that *kyaa* in (186) likewise ranges over propositions.

Against this background, the indirect-dependency approach claims that the embedded *wh*-element *kis-ko* ‘who-ACC’ in (186) is never associated with the matrix C^0 head. Rather, on this analysis there are two strictly local *wh*-dependencies, which are linked semantically. The embedded *wh*-element is licensed by the embedded C^0 and the matrix C^0 *wh*-licenses the *kyaa* in the matrix clause. The schematic structure of (186) on this account is given in (188).⁴³

- (188) $[_{CP} C^0 \dots kyaa \dots [_{CP} C^0 \dots kis-ko \dots]]$
-

On this structure, (186) contains two questions: *What did Sita think?* and *Who did Ravi see?* Working within Hamblin’s (1973) semantics for questions, Dayal (1994b, 1996) combines the two semantically by treating *kyaa* as an existential quantifier over a propositional variable. Crucially, she proposes that *kyaa* takes a potentially covert restrictor, which is coindexed with the lower clause. To illustrate, I will make use of a somewhat simplified version of Lahiri’s (2002) rendition of the indirect dependency account. As shown in (189), *kyaa* takes a propositional argument Q and also a restrictor argument T (for ‘topic’). The domain restriction T can be contextually provided, as in (187) or overtly by an embedded clause if this clause denotes a set of propositions (i.e., a question), as in (190).

- (189) $\llbracket kyaa \rrbracket = \lambda T \lambda Q \lambda p \exists q [T(q) \wedge Q(q)(p)]$

⁴³ Dayal (1994b, 1996) employs LF movement of both *wh*-elements. Because focus intervention arises in the scope-marking construction just as it does on local *wh*-dependencies, I will make use of a pure Agree relation instead, though nothing hinges on this choice as far as the indirect-dependency approach is concerned.

(190) *Denotation of embedded clause in (186)*
 $\lambda p \exists x [p = \wedge \text{saw}'(\text{Ravi}', x)]$

The intuition behind Dayal's (1994b, 1996) and Lahiri's (2002) accounts is thus that (186) is a question about Sita's thoughts, with the added restrictions that only her thoughts pertaining to the answer to the question who Ravi saw are relevant. Using the denotation of the lower clause in (190) as the restriction in (189) has the effect that (186) does not only range over propositions that Sita stands in the think-relation with, but that are also answer the question of who Ravi saw. Combining these pieces, the resulting interpretation of (186) is:

(191) a. $[[\text{(186)}]] = \lambda p \exists q [\exists x [q = \wedge \text{saw}'(\text{Ravi}', x)] \wedge [p = \wedge \text{think}'(\text{Sita}', q)]]$
 b. 'Which answer q to the question "Who did Ravi see?" is such that Ram thinks that q ?'

The most important aspect of this analysis is that the relationship between the embedded *wh*-element and content of the question is indirect: The embedded *wh*-element give rise to a question which restricts the domain of quantification of the matrix *wh*-element *kyaa*. For a more detailed development of this system, I refer the reader to Dayal (1994b, 1996) and Lahiri (2002).

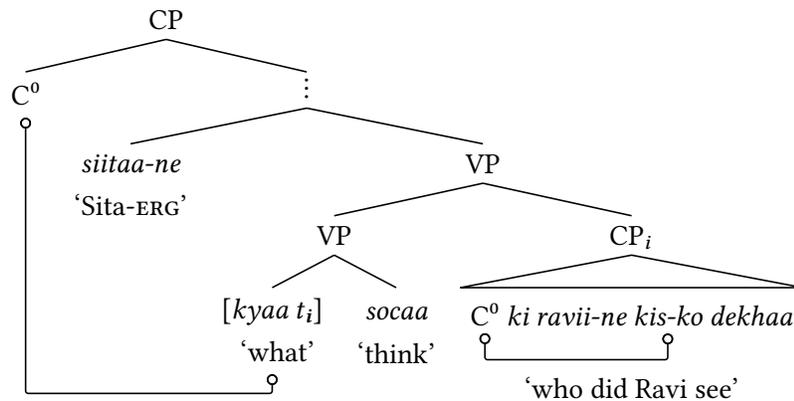
The syntactic connection between the two clauses on the indirect-dependency approach deserves some comments. Dayal (1994b, 1996) suggests a structure in which the lower clause is right-adjoined to the matrix IP. Problematic for this structure is Lahiri's (2002) discovery that quantificational elements in the matrix clause can bind pronouns in the lower clause, as in (192).

(192) *Embedded clause is c-commanded by matrix material*
 [har laṛkaa]_i kyaa soc-taa hai [ki kaun laṛkii use;
 every boy what think-IPFV.M.SG be.PRES.3SG that which girl him
 pasand kar-tii hai] ?
 like-IPFV.F.SG be.PRES.3SG
 'What does every boy_i think, which girl likes him_i?' (Lahiri 2002: 527)

Based on this observation, Manetta (2010, 2011) argues that Dayal's (1994b, 1996) adjunction structure is incompatible with such binding and that these facts provide evidence against an indirect-dependency account. However, as Dayal (2000) emphasizes, the indirect-dependency approach does not itself commit one to an adjunction structure (also see Dayal to appear). Other structural configurations are perfectly conceivable, such as the one in (193), which follows a proposal by

Lahiri (2002). On this structure, *kyaa* and the embedded clause form a constituent. The embedded clause is extraposed to the right but semantically reconstructs into its base-position, serving as the restrictor. Because matrix material c-commands the embedded CP in (193), the option of binding in (192) falls out naturally.⁴⁴

(193) *Schematic syntactic structure for (186)*



Most importantly for our present concerns, there is no cross-clausal *wh*-dependency (193) and it is this feature of the account that makes it fully compatible with (182). There are a number of arguments for this line of analysis, some of which I would like to mention here. First, any embedded question can occur in the scope-marking construction, regardless of whether this question contains a designated *wh*-word or not. Thus, in (194), the embedded clause is a yes/no question:

(194) *Embedded yes/no question in scope-marking construction*

ravii-ne kyaa kah-aa [ki anu aa-yegii yaa nahii] ?
 Ravi-ERG what say-PFV.M.SG that Anu come-FUT.F.SG or not
 'What do you think, will Any come or not?'

(Dayal 1996: 69)

The indirect-dependency approach extends to (194) without further ado. Following again the account in Dayal (1994b, 1996) for concreteness, the embedded clause of (194) receives the interpretation in (195a) and the entire sentence the semantics in (195b).

⁴⁴ Other structures are conceivable. Dayal (2000: 173) suggests a structure like (i), where the embedded clause is base-generated below *kyaa*, but coindexed with covert restrictor argument of *kyaa*. The embedded clause would then have to undergo covert raising into this position, leaving no trace behind.

(i) [CP siitaa-ne [VP [kyaa \emptyset_i] [VP socaa [CP_i ki ravii-ne kis-ko dekhaa]]]]
 Sita-ERG what think that Ravi-ERG who-ACC saw

- (195) a. $\lambda p [p = \wedge \text{will-come}'(\text{Anu}') \vee p = \wedge \neg \text{will-come}'(\text{Anu}')]]$
 b. $\lambda p \exists q [[q = \wedge \text{will-come}'(\text{Anu}') \vee q = \wedge \neg \text{will-come}'(\text{Anu}')] \wedge p = \wedge \text{think}'(\text{Ravi}', q)]$

A second desirable consequence of the indirect-dependency account is that it straightforwardly derives the fact that nonfinite embedded clauses cannot appear in the construction. Thus, a *wh*-word in a nonfinite clause cannot co-occur with *kyaa* in the matrix clause, in direct contrast to finite clauses:

(196) *No scope marking with nonfinite clauses*

- a. raam-ne [kis-ko dekh-naa] caah-aa thaa?
 Ram-ERG who-ACC see-INF.M.SG want-PFV.M.SG be.PST.M.SG
 ‘Who had Ram wanted to see?’
- b. *raam-ne *kyaa* [kis-ko dekh-naa] caah-aa thaa?
 Ram-ERG what who-ACC see-INF.M.SG want-PFV.M.SG be.PST.M.SG
- c. *raam-ne *kyaa* caah-aa thaa [kis-ko dekh-naa]?
 Ram-ERG what want-PFV.M.SG be.PST.M.SG who-ACC see-INF.M.SG

This contrast is entirely expected on an indirect-dependency approach because the embedded clause has to denote a question in order to act as a restrictor on the propositional variable contributed by *kyaa*. As we have seen in section 2.3.1, nonfinite clauses in Hindi cannot carry interrogative force (recall (108)). This general fact derives straightforwardly that they cannot occur in the scope-marking construction.

The third argument for an indirect-dependency approach comes from **SCOPE FREEZING** and has been pointed out by Lahiri (2002). As is well-known, cross-clausal amount questions are ambiguous in English (see, e.g., Cresti 1995):

- (197) How many people does John think I should talk to?
- a. For what *n*: there are *n*-many people *x* such that John thinks I should talk to *x*?
 (*wide scope*)
- b. For what *n*: John thinks that it should be the case that there be *n*-many people that I talk to?
 (*narrow scope*)

According to the wide scope reading in (197a), the question asks about specific people who John thinks that I should talk to. According to the narrow scope reading in (197b), on the other hand, the question asks about the number of people such that John thinks I should talk to that number of people, without having particular people in mind. Amount questions involve *wh*-quantification over degrees and existential quantification over individuals. The ambiguity in (197) arises because the existential quantification over individuals can be introduced either in the matrix or the embedded clause.

Lahiri (2002) observes that the corresponding scope-marking construction in Hindi is unambiguous. It lacks the wide scope reading:

(198) *Scope freezing*

ramesh *kyaa* soc-taa hai [ki raam-ne *kitnii kitabē*
 Ramesh what think-IPFV.M.SG be.PRES.3SG that Ram-ERG how many books
 paṛh-ī̃]?
 read-PFV.F.PL

*‘For what *n*: there are *n*-many books *x* such that Ramesh thinks that Ram read *x*’

(*wide scope*)

‘For what *n*: Ramesh thinks that there are *n*-many books *x* that Ram read?’

(*narrow scope*)

(Lahiri 2002: 520)

Lahiri (2002) points out that the absence of the wide scope reading in (198) follows from the indirect-dependency approach without further ado precisely because *kitnii kitabē* ‘how many books’ receives *wh*-scope in the embedded clause. Because the existential quantification has to be introduced lower than the *wh*-quantification for general semantic reasons, it has to likewise apply in the embedded clause. The unavailability of a wide reading is thus derived.

Covert movement of *kitnii kitaabē* ‘how many books’ into the matrix clause, on the other hand, would incorrectly make available the wide scope reading. This incorrect prediction also emerges on accounts that propose covert pied-piping of the complement clause (see Mahajan 1990, 2000), because overt clausal pied-piping gives rise to the ambiguity (Arregi 2003). A more recent alternative to the indirect dependency account has been proposed by Manetta (2010, 2011). On this alternative, the embedded *wh*-item remains in the lower clause, but enters into a *wh*-Agree relation with matrix material. The resulting structure is then interpreted via unselective binding of

choice functions à la Reinhart (1998). While Manetta (2010, 2011) claims that this account derives Lahiri’s (2002) scope freezing facts, it is far from clear that it does.⁴⁵ Barring ad hoc stipulations, the existential closure of the choice function argument of the DP could be introduced in either of the two clauses on Manetta’s (2010, 2011) account. This would yield the two interpretations in (199), where ‘ f_{CH} ’ denotes a choice function:

- (199) a. For what n : Ramesh thinks that $\exists f_{CH}$ Ram read $f_{CH}(\text{book}')$ and $|f_{CH}(\text{book}')| = n$
(*narrow scope*)
- b. For what n : $\exists f_{CH}$ Ramesh thinks that Ram read $f_{CH}(\text{book}')$ and $|f_{CH}(\text{book}')| = n$
(*wide scope*)

In (199a), existential closure of the choice function takes place inside the lower clause and a narrow reading results. In (199b), on the other hand, closure applies in the matrix clause, yielding a wide scope reading. The choice function account thus predicts a wide scope reading to be available, contrary to fact (also see Dayal to appear for related discussion).⁴⁶ Therefore, scope freezing is problematic for the choice function account, but it follows straightforwardly from the indirect-dependency account, precisely because this account only involves two separate, and strictly local, *wh*-dependencies.

In sum, there is compelling evidence for the view that the scope-marking construction in Hindi does not contain a *wh*-dependency that crosses a CP boundary. I have adopted here an indirect-dependency account of this construction, following the lead of Srivastav (1991b), Dayal (1994b, 1996, 2000) and Lahiri (2002). The analytical details of this construction are irrelevant for what is to come. The main point here is that on the independently motivated indirect-dependency account, the conclusion reached in the previous section stands: CPs are barriers for *wh*-licensing:

- (200) *Locality of wh-licensing* [=(182)]
Wh-licensing cannot cross a CP clause boundary in Hindi, but it may cross TP and ν P clause boundaries.

⁴⁵ I am indebted to Rajesh Bhatt, who pointed this problem of the choice function account out to me.

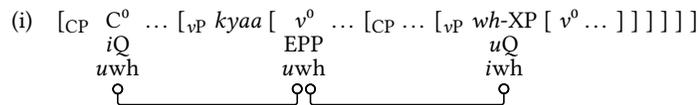
⁴⁶ As Dayal (to appear) emphasizes, this is a desirable effect for other *wh*-in-situ languages like Japanese, which exhibits precisely the ambiguity that is absent in Hindi.

The two strategies available in Hindi to form cross-clausal questions are possible precisely because they respect (200) by obviating the need for a cross-clausal *wh*-dependency. First, it is possible to use regular \bar{A} -movement of the *wh*-element to leave the finite clause and thus circumvent the CP boundary (recall that \bar{A} -movement is possible out of CPs). Second, one may have two strictly local, clause-internal *wh*-dependencies, neither of which crosses a CP. As a result, a consideration of these two strategies are fully in line with the generalization in (200).⁴⁷

2.4.4 Detour 2: Structure vs. positions – The case of Bangla

The characterization of the locality of *wh*-licensing in Hindi in this section has taken as its point of departure differences in the internal structure of finite and nonfinite clauses in the language. The fact that nonfinite clauses are transparent to *wh*-licensing but finite clauses are not correlates with the presence or absence of a CP layer within those clauses and I have proposed that *wh*-licensing as an operation is sensitive to such a CP layer. However, as is well-known (e.g., Dayal 1996), finite and nonfinite clauses in Hindi not only differ in their internal structure, but also with respect to their linear position relative to the matrix verb. In particular, finite complement clauses always and invariably appear to the *right* of their embedding verb, whereas nonfinite clauses typically appear to the *left*. All else equal, then, one might entertain a different empirical generalization: that

⁴⁷ A recent alternative account of the scope-marking construction has been proposed by Manetta (2010, 2011). Simplifying somewhat, Manetta proposes the structure in (i). On her account, the *wh*-element resides in SpecvP of the lower clause. The element *kyaa* is not a *wh*-word on her account, but a contentless expletive that is inserted into the matrix SpecvP for purely syntactic reasons. Manetta proposes that embedded *wh*-element agrees with the matrix v^0 head, which subsequently agrees with matrix C^0 , triggering unselective binding of choice functions.



While I will not adopt this account for the reasons given in the text, it highlights in a rather dramatic fashion the need for a theory of selective opacity. A crucial component of Manetta's (2010) account is that CP and vP are phases. Adopting a proposal by Rackowski & Richards (2005), she suggests that the matrix v^0 agrees with the embedded CP, a process that voids CP's phasehood and allows the matrix v^0 to continue probing and agree with the embedded v^0 . As Manetta (2010: 21n20) herself notes, removing the phasal status of the embedded C would allow *all* matrix probe to enter the embedded clause. For example, matrix ϕ -agreement into the embedded clause should then be possible, contrary to fact. To address this point, Manetta suggests that matrix T^0 cannot probe past an embedded T^0 , thus preventing ϕ -agreement into the embedded vP. I would like to note that this is, of course, an instance of selective opacity, because the embedded T^0 has to be a barrier for the ϕ -probe, but not other probes (like v 's *uwh*). Moreover, this analysis would predict that movement of an embedded argument out of the embedded TP would allow for agreement with the matrix verb. This again is incorrect. Hence, Manetta's (2010, 2011) is not an alternative to selective opacity because it presupposes it.

it is the linear position of the embedded clause that determines opacity for *wh*-licensing, not the presence of a CP layer. This is in fact the position taken by Mahajan (1990, 1997) and Dayal (1996) (also see Bhatt & Dayal 2007), and it has also been argued for by Simpson & Bhattacharya (2003) for the related language Bangla, although the technical implementations differ substantially.

The matter is not an easy one to decide. *Prima facie* evidence for a position-based characterization comes from the fact that nonfinite clauses may also appear to the right of the matrix verb. In this case, matrix scope is ruled out. Because nonfinite clauses also do not allow *wh*-elements to take scope inside them, as discussed in section 2.3.1, a local *wh*-construal is also impossible. This is illustrated in (201), where extraposition makes a regular question interpretation unavailable. Only an echo reading is possible.

- (201) a. tum [kyaa paṛh-naa] caah-te ho
 you what read-INF.M.SG want-IPFV.M.PL be.2PL
 ‘What do you want to read?’
- b. #tum t_i caah-te ho [kyaa paṛh-naa]_{*i*}
 you want-IPFV.M.PL be.2PL what read-INF.M.SG (*only echo interpretation*)

At first glance, this fact appears to strongly support a position-based characterization. The matter is rendered more complicated by the fact that it is not just *wh*-elements contained inside extraposed clauses that are barred from taking matrix scope. In fact, even unembedded *wh*-elements that follow the matrix predicate are barred from matrix scope (Mahajan 1997, Bhatt & Dayal 2007, Manetta 2012).⁴⁸

- (202) a. siitaa-ne dhyaan-se kis-ko dekh-aa thaa
 Sita-ERG care-with who-ACC see-PFV.M.SG be.PST.M.SG
 ‘Who had Sita looked at carefully?’
- b. #siitaa-ne dhyaan-se t_i dekh-aa thaa kis-ko_{*i*}
 Sita-ERG care-with see-PFV.M.SG be.PST.M.SG who-ACC
 (*only echo interpretation*)
 (based on Bhatt & Dayal 2007: 290–291)

⁴⁸ Bhatt & Dayal (2007) in fact argue that (202b) does not involve extraposition of the *wh*-DP itself, but rather of a remnant VP that contains the *wh*-DP. See Manetta (2012) for discussion of this proposal.

One valid characterization of these facts is extraposed material is invisible to *wh*-licensing, however this is technically enforced. On the assumption that finite clauses have to obligatorily extrapose, their opacity for *wh*-licensing could then be derived without appeal to their CP status. This line of approach hence constitutes an alternative characterization of the facts to the one advocated in the preceding sections and therefore merits consideration.

The first thing to note is that this alternative characterization does not affect the key conclusion of this chapter that Hindi instantiates multiple layers of selective opacity. The reason is that nonfinite clauses that are extraposed remain transparent to A- and \bar{A} -movement (see, e.g., (63b) and (80b)) and to ϕ -agreement (see (95)). Extraposed clauses are hence selectively opaque even on a position-based characterization of *wh*-opacity.

Second, while the data just discussed complicate somewhat a size-based characterization of *wh*-opacity, they do not render it invalid. One straightforward way of reconciling the extraposition facts with the characterization arrived at in the previous sections is to adopt Cable's (2007, 2010) account of movement and require that extraposed clauses are encapsulated inside an additional projection that undergoes extraposition. The transparency or opacity of extraposed constituents for various operations can then be stated with respect to this additional projection, rather than the linear position of the clause.

An alternative, and somewhat more radical characterization of the interaction between extraposition and *wh*-licensing has been suggested by Manetta (2012). She proposes that the ban on non-echo questions with post-verbal *wh*-material is the result of the information structure constraints imposed by extraposition. Specifically, she claims that extraposed material in Hindi is discourse-old. If *wh*-elements appear in this position, the requirement that they have to be interpreted as discourse-old precludes an interpretation as a regular information-seeking question. Only an echo reading is available. Interestingly, Manetta's (2012) proposal leaves open *wh*-elements inside finite clauses. Because finite clauses cannot appear anywhere else other than in an extraposed position, no specific information structure constraint is imposed on them in this position. Furthermore, Manetta (2012) argues that finite clauses never move, but are simply linearized to the right of the main verb in their base position. The question then arises as to why they are islands

for *wh*-licensing.⁴⁹ The generalization I have argued for here provides an answer to this question. It is their CP status.

The empirical landscape in Hindi is thus somewhat ambiguous, with no account emerging as clearly superior. Deciding evidence would come from the behavior of intraposed finite clauses. If it is the internal clause structure that determines *wh*-opacity (as argued for here), intraposed clauses should remain islands for *wh*-licensing. If it is the linear position that determines *wh*-opacity, they should be transparent. Unfortunately, these predictions simply cannot be tested in Hindi because finite clauses categorically resist intraposition.

Bangla, an Indo-Aryan language related to Hindi, does not share this restriction on the placement of finite clauses. Evidence from Bangla has indeed been used to argue in favor of a position-based generalization by Simpson & Bhattacharya (2003). I will briefly review the key evidence and then argue that, upon closer scrutiny, Bangla in fact supports a size-based characterization over a position-based one, providing indirect evidence for the line of approach taken for Hindi here.

The crucial empirical fact of interest to us is that in Bangla, finite clauses may appear in either an intraposed or an extraposed position. Importantly, a *wh*-element embedded inside such a clause can receive matrix scope only if the clause is intraposed, as in (203a), but not if it is extraposed, as in (203b) (Bayer 1996). In the latter case, if the matrix predicate allows a *wh*-complement, embedded *wh*-scope is possible. Otherwise, the structure is simply ungrammatical.

- (203) a. *Intraposition: Matrix scope possible* [Bangla]
 ora [CP ke aṣbe] ṣuneche
 they who come.will heard
 ‘Who have they heard will come?’ Or:
 ‘They have heard who will come.’

⁴⁹ Following an account by Simpson & Bhattacharya (2003) for Bangla (to be discussed immediately below), Manetta (2012) proposes that Hindi *wh*-elements undergo covert movement to SpecvP. She argues that to take scope outside of a finite clause, *wh*-elements in Hindi have to either overtly move to SpecvP of the higher clause or employ a *wh* expletive in the higher clause that acts as a linker between the interrogative C⁰ and the *wh*-phrase.

- b. *Extraposition: Matrix scope impossible*
 ora şuneche [CP ke aşbe]
 they heard who come.will
 #‘Who have they heard will come?’
 ‘They have heard who will come.’ (Simpson & Bhattacharya 2003: 128)

At first glance, the contrast in (203) seems to support the position-based view, and this is indeed what Simpson & Bhattacharya (2003) argue. Yet closer scrutiny shows that this conclusion is premature. The reason is that pre-verbal and post-verbal clauses differ in their internal structure in Bangla. Two central contrasts involve complementizers. Both pre-verbal and post-verbal clauses can contain the complementizer *je*, but the linear position of *je* within the clause differs (Bayer 1996, Bhattacharya 2002, Hsu 2015). In pre-verbal clauses, *je* has to be preceded by at least one constituent and possibly more. In post-verbal clauses, on the other hand, *je* has to be clause-initial.

- (204) a. *Intraposition: ‘je’ non-initial* [Bangla]
 John [ma je kal rate oSudh kheyechē] jane
 John mother COMP last night.LOC medicine ate knows
 ‘John knows that mother took medicine last night.’
 b. *Extraposition: ‘je’ initial*
 John jane [je ma kal rate oSudh kheyechē]
 John knows COMP mother last night.LOC medicine ate
 (Bhattacharya 2002: 101)

Reversing the position of *je* in pre- and post-verbal clauses leads to ungrammaticality:

- (205) a. *Intraposition* [Bangla]
 *amra [je ma aSbe] jantam
 we COMP mother come.will knew
Intended: ‘We knew that mother will come.’
 b. *Extraposition*
 *amra jantam [ma je aSbe]
 we knew mother COMP come.will
Intended: ‘We knew that mother will come.’ (Bhattacharya 2002: 101–102)

Moreover, pre-verbal clauses may contain the complementizer *bole*, which appears in a clause-final position. This complementizer is blocked from post-verbal clauses.

- (206) a. *Intraposition: 'bole' possible* [Bangla]
 chele-Ta [or baba aSbe **bole**] Suneche
 boy-CF his father come.will COMP heard
 'The boy has heard that his father will come.'
- b. *Extraposition: 'bole' impossible*
 ?* chele-Ta Suneche [or baba aSbe **bole**]
 boy-CF heard his father come.will COMP (Bayer 1996: 255)

Against this background, the crucial observation is that the choice of complementizer affects a clause's transparency for *wh*-licensing. Recall from (203a) that intraposed clauses are transparent to *wh*-licensing. However, if this clause contains the complementizer *je*, it becomes opaque to *wh*-licensing, as shown in (207). It is irrelevant whether the *wh*-element precedes or follows *je* in its clause.⁵⁰

- (207) *Intraposition + 'je': Scope island* [Bangla]
 a. *tumi [ke je baʒi korbe] bhabcho
 you who COMP house make.will think
 b. *tumi [baʒi je ke korbe] bhabcho
 you house COMP who make.will think (Saurov Syed, p.c.)

If the intraposed clause contains the complementizer *bole*, it remains transparent for *wh*-licensing:

- (208) *Intraposition + 'bole': No scope island* [Bangla]
 tumi [ke baʒi korbe **bole**] bhabcho
 you who house make.will COMP think
 'Who do you think will build the house?' (Saurov Syed, p.c.)

The contrast between (207) and (208) makes it clear that it is the internal structure of the finite clause, rather than its linear position, that determines whether it is a scope island or not.

As shown in (203b) above, post-verbal clauses are opaque to *wh*-licensing:

⁵⁰ The sentences in (207) are ungrammatical because the matrix verb *bhabcho* 'think' does not allow an embedded *wh*-construal (Simpson & Bhattacharya 2003: 129).

- (209) *tumi bhabcho [_{CP} ke baṭi korbe] [Bangla]
 you think who house make.will (Simpson & Bhattacharya 2003: 129)

Adding *je* to the clause does not change this fact:

- (210) *tumi bhabcho [*je* ke baṭi korbe] [Bangla]
 you think COMP who house make.will (Saurov Syed, p.c.)

The transparency or opacity of a clause for *wh*-licensing can thus be predicted from its internal structure: If the clause takes *bole*, it is transparent, if it takes *je*, it is opaque. On the assumption that the overt realization of *bole* is optional, the fact that complementizer-less pre-verbal clauses are transparent to *wh*-licensing (see (203a)) likewise follows. Moreover, the fact that post-verbal clauses are opaque to *wh*-licensing is attributed to independent fact that they do not allow a *bole*-structure (206b).

A straightforward account of these facts is that *bole*-clauses are structurally smaller than *je*-clauses. In other words, *je* is the realization of a higher head in the left periphery than *bole*. Let us refer to these projections mnemonically as *jeP* and *boleP*, respectively. On this view, *jePs* are opaque to *wh*-licensing and *bolePs* are transparent. An independently needed constraint that requires post-verbal clauses to be *jePs* (recall (206b)) ensures that post-verbal clauses are always too large to allow *wh*-licensing into them.

Needless to say, this account is merely a sketch and the empirical facts and generalization deserve more careful study than I have been able to give them here. A more thorough investigation into the constructions at hand would be necessary to develop the details of an account along these lines. I will not embark on such an investigation here. Analytical details aside, the generalizations uncovered here, though tentative, appear to provide rather striking support for a size-based account of the locality of *wh*-licensing in Bangla and indirectly for Hindi. It seems clear that the internal structure of a clause is a crucial conditioning factor for its opacity to *wh*-licensing. The linear position of a clause appears to be relevant only to the extent that it is confounded with its internal structure, i.e., if clauses in a certain position are required to have a certain internal structure. The basic contrast in (203), then, is misleading, as it obscures differences in the internal structure of intraposed and extraposed clauses. I conclude, then, (i) that any satisfactory account of *wh*-licensing in Bangla has to refer to the internal structure of the embedded clause, and (ii) that there is no clear

evidence that linear position plays a direct role. To the extent that the syntax of *wh*-licensing in Bangla is similar to that of Hindi, the Bangla facts considered here thus provide evidence in favor of the basic approach to the locality *wh*-licensing in Hindi.

2.4.5 Section summary

This section has added *wh*-licensing to the array of operations in Hindi investigated here. I have argued for two central claims: First, I have provided evidence from focus intervention effects that *wh*-licensing in Hindi does not involve movement, but is instead a long-distance operation. In this respect, *wh*-licensing is similar to ϕ -agreement. Conversely, movement is possible without *wh*-licensing in the landing site (see (185)). As a consequence, \bar{A} -movement and *wh*-licensing are independent operations. Second, I have investigated the locality of *wh*-licensing. We have seen that it may proceed across TP and *v*P clauses, but that it is blocked by CP clauses. Adding these insights the summary table in (156) yields the summary table in (211):

(211) *Interim summary (to be extended)*

<i>Operation</i>	Finite clauses		Nonfinite clauses	
	CP	TP	<i>v</i> P	
\bar{A} -movement	✓	✓	✓	
<i>wh</i> -licensing	✗	✓	✓	
A-movement	✗	✗	✓	
ϕ -agreement	✗	✗	✓	

(✓: transparent; ✗: opaque)

The addition of *wh*-licensing to the domain of investigation leads us to a number of conclusions: First, the Hindi pattern in (211) instantiates selective opacity. One and the same clause structure is transparent to some operations, but opaque to others. Second, this selective opacity pattern is *not binary*, because there are three types of locality profiles. Third, selective opacity effects span both movement and non-movement operations, because two of the four operations of interest here (namely, ϕ -agreement and *wh*-licensing) do not involve movement, but are instead established long-distance. Fourth, there is no indication of a systematic divide between movement and non-movement operations. For one, locality mismatches between different movement types are mirrored

in locality mismatches between different types of non-movement operations: Just like the locality of A-movement differs from that of \bar{A} -movement, so the locality of ϕ -agreement differs from that of *wh*-licensing. In other words, we observe selective opacity effects in operations that do not involve movement. Further evidence for this conclusion comes from the observation that the locality of A-movement is identical to that of ϕ -agreement. To capture this convergence, the two operations have to be subject to a unified constraint. Such a constraint must hence not be confined to movement. Instead, it must be general enough to apply to movement and non-movement alike. Lastly, the constraint has to be selective, because other movement and non-movement operations are not subject to it.

A fairly general conclusion emerges from these considerations: Selective opacity is not limited to movement, but is instead also observable for syntactic operations that do not involve movement. Thus, while virtually all previous investigations into selective opacity have focused on movement dependencies, the phenomenon is revealed to be much more general. Finally, it should also be noted that the coexistence of three locality profiles in (211) corroborates the conclusion reached in chapter 1 that selective opacity is not a binary distinction between A- and \bar{A} -operations, but more fine-grained and nuanced.

2.5 Regularities of selective opacity

Before developing an analysis of the Hindi facts presented in this section and selective opacity more generally, I will inspect a prediction that the claims laid out so far lead to. Recall from chapter 1 that the previous literature on selective opacity has identified two meta-generalizations: Upward Entailment and the Height–Locality Connection. If both are indeed hallmark properties of selective opacity and if the Hindi pattern in (211) exemplifies selective opacity, we expect the two meta-generalizations to hold in Hindi as well. This section will demonstrate that this expectation is indeed borne out.

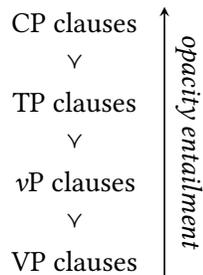
2.5.1 Upward entailment

The first generalization, Upward Entailment, is repeated from chapter 1 in (212) and illustrated in (213):

(212) **Upward Entailment**

If a clause of a certain structural size is opaque to an operation, then clauses that are structurally larger are also opaque to this operation.

(213) *Illustration of Upward Entailment:*



According to (212), selective opacity effects are not distributed at random. Although syntactic operations differ in their locality properties, there exist implicational relationships between the opacity profiles of clauses of different structural sizes. If a clause of a particular structural size (i.e., a clause that is pruned at a particular projection) is opaque to some movement, then larger clauses (i.e., clauses pruned at a higher projection) are likewise opaque. To repeat just one example from chapter 1, TP clauses in English are opaque to extraposition. Upward Entailment then states that CP clauses, in virtue of being structurally ‘bigger’, will likewise be opaque to extraposition. Put differently, (212) precludes the existence of a movement type that can leave a finite clause, but not a structurally smaller nonfinite clause. While such a locality profile is perfectly conceivable, it does not seem to exist empirically.

The Hindi pattern motivated in this chapter provides clear further support for Upward Entailment. This is represented in (214):

(214) *Upward Entailment in Hindi selective opacity*

	CP	TP	vP
\bar{A} -movement	✓	✓	✓
<i>wh</i> -licensing	✗	✓	✓
A-movement	✗	✗	✓
ϕ -agreement	✗	✗	✓

← *opacity entailment*

(214) highlights that the transparency/opacity profiles in Hindi are not random. Specifically, the addition of structural layers to an embedded clause may render a clause opaque to an operation, but the inverse never happens: removal of a structural layer never results in added opacity. In this sense, opacity is *monotonic*. This monotonicity is nothing other than Upward Entailment (212). For example, if a TP clause is opaque to an operation (A-movement and ϕ -agreement), then a CP clause is likewise opaque. Conversely, if a TP clause is transparent for an operation, then a structurally smaller vP clause is also transparent to this operation, and so on.

The finding that selective opacity in Hindi conforms to Upward Entailment is important for two reasons. First, it provides additional evidence for the validity of Upward Entailment as a generalization that holds across a range of languages and constructions. Second, it is noteworthy that Upward Entailment holds of all four operations in (214), regardless of whether they involve movement or not. This provides further support for the conclusion reached in the previous section: Selective opacity is not a property of only movement, but instead also comprises non-movement operations like ϕ -agreement and *wh*-licensing. Upward Entailment supports this view because the empirical generalizations that govern selective opacity in the domain of movement also hold for selective opacity beyond movement.

2.5.2 The Height–Locality Connection

In addition to Upward Entailment, chapter 1 identified a second meta-generalization across selective opacity effects (Williams 2003, 2011, 2013, Abels 2007, 2009, 2012a, Müller 2014a,b):

(215) **Height–Locality Connection**

Movement types differ in their landing sites. The higher the landing site of a movement type is in the clausal structure, the more kinds of structures are transparent to this movement type.

According to (215), the locality profile of a movement type is at least partially a function of the height of that movement type's landing site in the clausal spine. Specifically, the higher the landing site of a movement type, the more types of clauses are transparent to this movement type. Conversely, movement types that land low in the clausal spine tend to be subject to stricter locality constraints and can access a smaller set of embedded clauses. To repeat just one example supporting (215) from chapter 1, A-movement in English lands in a position lower than \bar{A} -movement (SpecTP vs. SpecCP, respectively), and this relative difference in their landing site corresponds with their locality properties: A-movement cannot leave a finite clause, whereas \bar{A} -movement can. According to (215), this convergence between the height of the landing site and the locality profiles of the two movement types is not a coincidence, but a manifestation of a systematic correspondence between the two.

The Height–Locality Connection (215) is particularly intriguing because it does not follow from any standard theory of movement locality. As such, any connection between a movement's landing site and its locality remains unaccounted for on standard locality principles. In this section, I will show that (215) also holds in Hindi. This finding not only provides further support for its validity, it will also demonstrate that the connection between height and locality holds for non-movement operations as well. In order to ascertain the validity of (215) for Hindi, we need to determine the height of the landing site for A- and \bar{A} -movement and of the triggers for ϕ -agreement and *wh*-licensing.

2.5.2.1 The landing site of A-movement

Due to the head-final phrase structure and the optionality of movement in Hindi, determining the landing sites of A- and \bar{A} -movement is a difficult task and I am not aware of previous attempts to discern the landing sites of the two movements.⁵¹ To narrow down the landing site possibilities of

⁵¹ Mahajan (1990) conjectures that A-movement lands in an L-related specifier position, while \bar{A} -movement targets a non-L-related adjoined position, but he introduces this difference solely in order to account for the differences

A-movement in Hindi, it is instructive to consider whether A-movement can land inside a nonfinite clause.

It is possible to A-move an element inside a nonfinite clause over another element inside the same clause. Thus, in (216), the embedded direct object *har kuttaa* ‘every dog’ is A-moved over the indirect object *us-ke baccō-ko* ‘its children-DAT’. A-movement is diagnosed by the bound interpretation of the pronoun, analogous to the reasoning in section 2.2.1 above.

- (216) raam-ne har kuttaa_i us-ke_i baccō-ko t_i dikhaa-naa caah-aa
 Ram-ERG every dog 3SG-GEN children-DAT show-INF.M.SG want-PFV.M.SG
 ‘Ram wanted to show every dog *x* to *x*’s children.’

The limitation of (216) is that it does not allow us to determine whether *har kuttaa* lands inside the nonfinite clause or in the matrix clause. Therefore, to determine whether A-movement of *har kuttaa* can remain solely within the nonfinite clause, we need to consider a structure that demarcates in one way or another the left edge of the nonfinite clause. I will present two such structures.

The first structure is similar to (216), but additionally involves extraposition of the embedded clause. In (217), the A-moved element *har kuttaa* appears to the right of the matrix verb and hence remains inside the nonfinite clause. The grammaticality of this sentence then indicates that A-movement of *har kuttaa* can take place solely within the nonfinite clause. Therefore, A-movement must be able to land within a nonfinite clause.

- (217) *A-movement can land inside nonfinite clause*
 raam-ne t_j caah-aa [har kuttaa_i us-ke_i baccō-ko t_i dikhaa-naa]_j
 Ram-ERG want-PFV.M.SG every dog 3SG-GEN children-DAT show-INF.M.SG
 ‘Ram wanted to show every dog *x* to *x*’s children.’

The same conclusion can be reached with binding into an adverb:

with respect to binding etc. between these two movement types. In other words, Mahajan’s (1990) reasons for postulating a difference in landing site are purely theory-internal.

- (218) *A-movement can land inside nonfinite clause*
- siitaa-ne t_j caah-aa [har laṛkii-ko_i us-kii_i shaadii ke dauraan t_i
 Sita-ERG want-PFV.M.SG every boy-ACC 3SG-GEN wedding during
 dekh-naa]_j
 see-INF.M.SG
 ‘Sita wanted to see every girl x at x ’s wedding.’

A second configuration that allows us to unambiguously identify A-movement within a nonfinite clause involves subject clauses. This argument is based on the independent observation that subject clauses are islands in Hindi. The contrast in (219) establishes the islandhood of subject clauses. In (219a), a clause in subject position is preceded by an adverb that modifies the matrix predicate. As (219b) shows, it is impossible for *phal* ‘fruit’ to move out of the subject clause over this adverb. This follows if subject clauses are islands for extraction.⁵²

- (219) *Subject clauses are islands*
- a. sehat ke-liye [phal khaa-naa] acchaa ho-taa hai
 health for fruit eat-INF.M.SG good be-IPFV.M.SG be.PRES.3.SG
 ‘Eating fruits is good for health.’
- b. *phal_i sehat ke-liye [t_i khaa-naa] acchaa ho-taa hai
 fruit health for eat-INF.M.SG good be-IPFV.M.SG be.PRES.3.SG
Intended: ‘Eating fruits is good for health.’

Subject clauses, then, are islands to both A- and \bar{A} -movement. Against this background, consider the example (220):

- (220) *A-movement inside nonfinite subject clause*
- [har kuttaa_i us-ke_i baccō-ko t_i dikhaa-naa] ajiib baat hai
 every dog 3SG-GEN children-DAT show-INF.M.SG weird thing be.PRES.3SG
 ‘Showing every dog x to x ’s children is weird.’

In (220), the direct object *har kuttaa* is A-moved over the indirect object *us-ke baccō-ko*, indicated by pronominal binding from the landing site. In light of the complete islandhood of subject clauses,

⁵² The sentence in (219b) is grammatical under a reading in which *sehat ke-liye* ‘for health’ modifies the subject clause (‘Eating fruits for health is good.’) Under this reading no extraction has taken place and the grammaticality is expected.

har kuttaa has to remain within the subject clause. As a result, A-movement must be able to land within a nonfinite clause.

As was the case with extraposed clauses, this result may be replicated with binding into an adjunct:

(221) *A-movement inside nonfinite subject clause*

[har larḱii-ko_i us-kii_i shaadii ke dauraan dekh-naa] acchii baat hai
 every girl-DAT 3SG-GEN wedding during see-INF.M.SG good thing be.PRES.3SG
 ‘Seeing every girl *x* during *x*’s wedding is a good thing.’

To summarize, we have seen converging evidence that A-movement can take place within a nonfinite clause in Hindi. In light of this observation, recall from section 2.3.1 that nonfinite clauses are maximally TPs in Hindi and lack a CP layer (see (117)). Because A-movement is nonetheless possible in these clauses, A-movement must land in a position lower than SpecCP:

(222) A-movement lands in SpecTP or lower.

(222) thus states an upper bound on the landing site of A-movement.

2.5.2.2 The landing site of \bar{A} -movement

In direct contrast to A-movement, \bar{A} -movement cannot land inside a nonfinite clause in Hindi. This is demonstrated by the paradigm in (223).⁵³ All sentences in (223) involve a double embedding structure in which a finite clause is embedded inside a nonfinite clause, which is in turn embedded inside a finite matrix clause. The baseline structure without any movement is given in (223a). (223b) is then derived from (223a) by moving the DP *kitaab* ‘book’ from the lowermost clause into the intermediate nonfinite clause. As shown, the result is ungrammatical. In (223c), on the other hand, the same DP *kitaab* is moved all the way into the matrix clause and the resulting sentence is grammatical.

⁵³ I am grateful to Klaus Abels for suggesting this paradigm to me.

- (223) a. *Base configuration:* \checkmark [matrix clause [non-finite clause [finite clause DP]]]
- māĩ caah-taa hũũ [kah-naa [ki māĩ-ne *kitaab* paṛh
 I want-IPFV.M.SG be.PRES.1SG say-INF.M.SG that I-ERG book.F read
 l-ii hai]]
 take-PFV.F.SG be.PRES.3SG
 ‘I want to say that I read the book.’
- b. *No \bar{A} -mvt into non-finite clauses:* * [matrix clause [non-finite clause DP [finite clause *t*]]]
- *māĩ caah-taa hũũ [*kitaab*_{*i*} kah-naa [ki māĩ-ne *t*_{*i*} paṛh
 I want-IPFV.M.SG be.PRES.1SG book say-INF.M.SG that I-ERG read
 l-ii hai]]
 take-PFV.F.SG be.PRES.3SG
- c. *\bar{A} -mvt into finite clauses:* \checkmark [matrix clause DP [non-finite clause [finite clause *t*]]]
- kitaab*_{*i*} māĩ caah-taa hũũ [kah-naa [ki māĩ-ne *t*_{*i*} paṛh
 book I want-IPFV.M.SG be.PRES.1SG say-INF.M.SG that I-ERG read
 l-ii hai]]
 take-PFV.F.SG be.PRES.3SG

Movement of *kitaab* in both (223b) and (223c) proceeds out of a finite clause. We have seen above in section 2.2.1 that finite clauses allow \bar{A} -movement out of them, but block A-movement (see (83)). As a consequence, the movement of *kitaab* in (223b,c) must be \bar{A} -movement. The ungrammaticality of (223b) then demonstrates that \bar{A} -movement *cannot* land inside a nonfinite clause. Crucially, if the same element is moved into the highest finite clause, as in (223c), the result is grammatical. This makes it clear that it is not \bar{A} -movement of *kitaab* itself that constitutes the problem in (223b), but rather the fact that this \bar{A} -movement lands in a nonfinite clause.

The constraint that \bar{A} -movement cannot land inside a nonfinite clause follows straightforwardly if \bar{A} -movement targets SpecCP. Because nonfinite clauses obligatorily lack a CP layer in Hindi (recall (117) in section 2.3.1), they simply lack the functional structure necessary to provide a landing site for \bar{A} -movement and the paradigm in (223) is derived.

(224) \bar{A} -movement lands in SpecCP.

Despite superficial appearance, then, A- and \bar{A} -movement differ in the height of their landing sites.

An interesting consequence of (224) concerns the landing site of \bar{A} -movement out of a finite clause with respect to the matrix subject. In many cases, such movement can land either above or below the subject:

- (225) a. *is kitaab-ko_i siitaa-ne soc-aa* [ki *sangiitaa-ne t_i paṛh-aa*
 this book-ACC Sita-ERG think-PFV.M.SG that Sangita-ERG read-PFV.M.SG
thaa]
 be.PST.M.SG
 ‘This book, Sita thought that Sangita had read.’
- b. *siitaa-ne is kitaab-ko_i soc-aa* [ki *sangiitaa-ne t_i paṛh-aa*
 Sita-ERG this book-ACC think-PFV.M.SG that Sangita-ERG read-PFV.M.SG
thaa]
 be.PST.M.SG
 ‘This book, Sita thought that Sangita had read.’

The claim that \bar{A} -movement targets SpecCP in Hindi has the consequence that (225b) is syntactically more complex than (225a). In (225a), the \bar{A} -moved DP lands above the matrix subject (in SpecCP according to (224)), but the syntax of (225b) must involve \bar{A} -movement of *is kitaab-ko* ‘this book’, followed by a second \bar{A} -movement of the matrix subject *siitaa-ne* above it. We might ask whether there is evidence supporting this discrepancy.

As it turns out, a clear difference between the two structures emerges if the moving element has the same case as the matrix subject, i.e., ergative in (225). If the embedded subject *sangiitaa-ne* ‘Sangita-ERG’ is moved into the matrix clause, it has to land above the matrix subject:

- (226) a. *sangiitaa-ne_i siitaa-ne soc-aa* [ki *t_i is kitaab-ko paṛh-aa*
 Sangita-ERG Sita-ERG think-PFV.M.SG that this book-ACC read-PFV.M.SG
thaa]
 be.PST.M.SG
 ‘Sangita, Sita thought had read the book.’

- b. *siitaa-ne sangiitaa-ne_i soc-aa [ki t_i is kitaab-ko parh-aa
 Sita-ERG Sangita-ERG think-PFV.M.SG that this book-ACC read-PFV.M.SG
 thaa]
 be.PST.M.SG
Intended: ‘Sangita, Sita thought had read the book.’

(226b) is ungrammatical on the structure given. It is grammatical only if *sangiitaa-ne* is the matrix subject and *siitaa-ne* is \bar{A} -extracted out of the lower clause, hence a structure analogous to that of (226a). It is clear that the ungrammaticality of (226b) is not due to a general ban against adjacent elements with identical cases. The relevant restriction specifically targets the landing site of \bar{A} -movement relative to the matrix subject. This restriction follows from (224) if any string is assigned the structure compatible with it that has the fewest possible steps of optional movement. As is well-known since Ross (1967), such a constraint is necessary under any theory with optional movement. To illustrate, we have seen in section 2.2.1, soV word order exhibits scope rigidity, whereas the oSV order is ambiguous. To ensure the scope rigidity of the soV order, one has rule out a Duke-of-York derivation (see Pullum 1979) in which movement first produces an oSV order, followed by a second movement of the subject over the object, returning soV again and scope flexibility. One possibility of ruling out such a derivation is to require that any string is assigned the structure compatible with it that has the fewest instances of optional movements. An soV order is compatible with no application of movement and therefore a structure that assigns it a Duke-of-York derivation is blocked. This principle might plausibly be a parsing principle, but its proper characterization does not need to concern us here.

Notably, this blocking principle offers a rationale for the contrast in (226) only if the structure in (226b) is derivationally more complex than that of (226a). If \bar{A} -movement has to land in SpecCP in Hindi, the relative complexity of these two structures is directly predicted. Consequently, in a string that comprises two ergative DPs next to each other and a gap inside an embedded clause, the structural analysis with the fewest possible movement steps is as in (227a), where the moved element is the initial DP. Analyzing the linearly second DP as the moved element requires the structure in (227b), which involves two movement dependencies and is hence blocked by the principle just outlined. (226b) is therefore correctly blocked.

- (227) a. [CP DP-ERG_i [DP-ERG ... [CP t_i ...]]]
 b. * [CP DP-ERG_j DP-ERG_i [t_j ... [CP t_i ...]]]

This analysis also correctly predicts no contrast in (225) because here the case marking unambiguously encodes the base position of the two elements.

In sum, this account crucially requires that an element moved out of a finite clause cannot directly land below the matrix subject, but that such configurations are derived indirectly, by movement over the subject followed by movement of the subject itself. If \bar{A} -movement lands in a very high position (SpecCP), this fact follows without further ado. The discussion thus supports (224).

2.5.2.3 The location of the ϕ -probe

Determining the location of the ϕ -probe is more difficult, for general reasons. Often, morphological perturbations lead to the morphological expression of agreement in a position distinct from its syntactic source (tense/agreement lowering in English is an example). Therefore, the locus of agreement morphology is an unreliable indicator of the syntactic position of a ϕ -probe.

It is possible, however, to draw inferences about the structural location of the ϕ -probe by considering the algorithm for local verb agreement, repeated in (228).

- (228) **Hindi ϕ -agreement algorithm** [= (84)]
If the subject does not bear a case marker \rightarrow agree with the subject
Otherwise: If object does not bear a case marker \rightarrow agree with the object
Otherwise: Use masculine singular default agreement.

If the subject is not overtly case-marked, the verb agrees with it, as in (229a). If it is case-marked, then the verb agrees with the object if the object is not case-marked, as in (229b).

- (229) a. *Subject zero-marked* \rightarrow *subject agreement*
 laṛke is kitaab-ko paṛh-te hãĩ [= (85)]
 boys.M.PL this book.F.SG-ACC read-IPFV.M.PL be.PRES.3PL
 ‘The boys are reading this book.’

- b. *Subject overtly marked, object zero-marked* → *object agreement*

laṛkō-ne yah kitaab paṛh-ii
 boys.M.PL-ERG this book.F.SG read-PFV.F.SG
 ‘The boys read this book.’

Consequently, both subjects and objects can in principle control verb agreement if they do not bear a case marker. Crucially, if both the subject and the object are *a priori* viable agreement controllers, subject agreement must take place, in line with (228) and exemplified in (230).

- (230) *Subject agreement preempts object agreement*

a. laṛke kitaab paṛh-te hāī
 boys.M.PL book.F.SG read-IPFV.M.PL be.PRES.3PL
 ‘The boys are reading a book.’

b. *laṛke kitaab paṛh-ti hai
 boys.M.PL book.F.SG read-IPFV.F.SG be.PRES.3SG
Intended: ‘The boys are reading a book.’

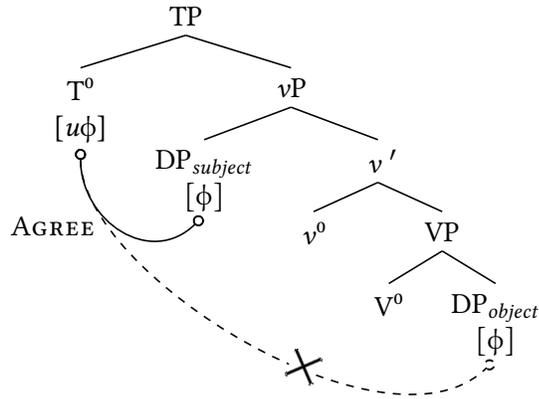
Subject agreement thus preempts object agreement. This pattern can be straightforwardly modeled if the ϕ -probe resides on a head higher than v^0 . On this view, the preference for subject agreement then emerges as a consequence of locality. Assuming the definition of **AGREE** in (231), a probe agrees with the closest eligible goal in its c-command domain.

- (231) **Agree**

An unvalued feature [uX] (the ‘probe’) serially searches through its c-command domain for a valued counterpart (a ‘goal’). It agrees with the closest goal.

To illustrate, if the ϕ -probe is located on T^0 in Hindi, subject agreement will take priority over object agreement simply because the subject is closer to [$u\phi$], as schematized in (232). Object agreement correctly arises only if the subject is not a legible goal for T ’s ϕ -probe, i.e., if it case-marked.

(232) *Subject agreement preempts object agreement*



This view is consistent with Béjar & Rezac’s (2009) analysis of agreement displacement in Basque (also see Carstens 2016). They show that in Basque, the situation is reversed: object agreement takes priority over subject agreement. Béjar & Rezac (2009) analyze this pattern by placing the verbal ϕ -probe on v^0 . To the extent that their account is on the right track, it provides indirect support for the view that the ϕ -probe in Hindi is located higher than v^0 .

There is furthermore reason to believe that the ϕ -probe is indeed located on T^0 , as opposed to, say, C^0 . The argument to this effect is based on the conclusion in (222) above that A-movement in Hindi must land lower than C^0 , hence in SpecTP or lower, and the general fact, already noted in section 2.2.2 (cf. (86)), that A-movement does not affect verbal agreement in Hindi. Thus, as shown in (233), movement of the object over the subject does not make object agreement possible. In other words, the verb in (233) has to agree with the subject *laṛke* ‘boys’, just as in (230).

(233) *Object movement does not override subject agreement*

- a. *kitaab_i laṛke t_i paṛh-te hāĩ*
 book.F.SG boys.M.PL read-IPFV.M.PL be.PRES.3PL
 ‘The boys are reading a book.’
- b. **kitaab_i laṛke t_i paṛh-ti hai*
 book.F.SG boys.M.PL read-IPFV.F.SG be.PRES.3SG
Intended: ‘The boys are reading a book.’

Unsurprisingly, the same is true for movement that is unambiguously A-movement as diagnosed by the absence of weak crossover:

(234) *Object A-movement does not override subject agreement*

- a. har gaar_{ii}_i [us-kaa_i maalik (=hii)] t_i saaf kar-egaa
 every car.F 3SG-GEN owner.M(=only) clean do-FUT.M.SG
 ‘Every car *x* will be cleaned by *x*’s owner (not by anybody else).’
- b. *har gaar_{ii}_i [us-kaa_i maalik (=hii)] t_i saaf kar-egii
 every car.F 3SG-GEN owner.M(=only) clean do-FUT.F.SG
Intended: ‘Every car *x* will be cleaned by *x*’s owner (not by anybody else).’

If A-movement were to land below the head that hosts the ϕ -probe, and hence in its c-command domain, it should be able to feed verbal ϕ -agreement, contrary to fact. I thus propose based on (233) that the landing site of A-movement must lie outside the c-command domain of the head hosting the ϕ -probe.

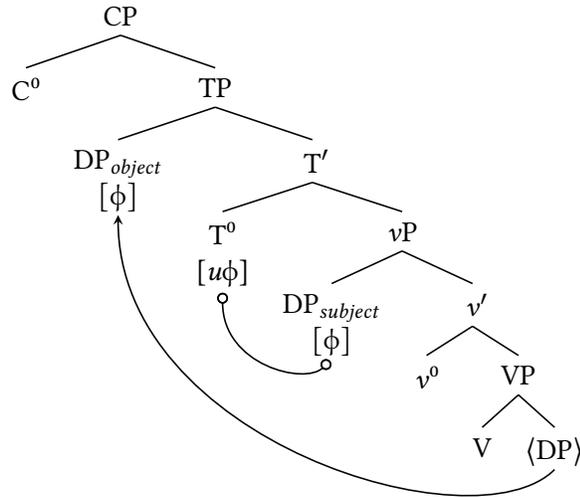
Combining these various facts allows us to pinpoint the locus of ϕ -agreement and A-movement.

(i) The ϕ -probe must reside higher than v^0 in order to capture the preference for subject agreement (230). (ii) we have seen that A-movement must land in SpecTP or lower (see (222)). (iii) The landing site of A-movement must be outside the c-command domain of the ϕ -probe (see (234)). Assuming as before a C–T– v –V clausal spine, this constellation of facts requires that A-movement lands in SpecTP and ϕ -probe resides in T^0 . (235) thus refines the generalizations in (222) and (224).

- (235) a. A-movement lands in SpecTP.
 b. The ϕ -probe [$u\phi$] is located on T^0 .

(235) correctly derives the various facts observed in this section: (i) Subject agreement takes preference over object agreement because the ϕ -probe is located on T^0 (see (232)); (ii) A-movement is possible within a nonfinite clause because it lands in SpecTP and nonfinite clauses may be TPs (see (216)–(221)); and (iii) A-movement of the object does not obviate subject agreement because it lands outside the c-command domain of the ϕ -probe (see (234)). The structure of (234) that results from this view is given in (236). It is irrelevant for the derivation in (236) whether ϕ -agreement or A-movement applies first, but considerations of cyclicity would favor the former.

(236) *The non-interaction between A-movement and ϕ -agreement*



In (236), A-movement of the object over the subject does not affect verbal agreement because this movement does not place the object in a position between the ϕ -probe and the subject. The landing site is too high for $[u\phi]$ and hence agreement remains with the subject. Notice, incidentally, that this line of reasoning entails that an A-moving object directly moves from its base position to SpecTP, without an intermediate touchdown in SpecvP. In other words, it requires that vP not be a phase. This is indeed a conclusion I will reach on independent grounds in chapters 6 and 7.⁵⁴

If \bar{A} -movement targets SpecCP, as concluded in (224) above, then we expect \bar{A} -movement to likewise not be able to feed ϕ -agreement. This expectation is correct, as (237) attests, where the

⁵⁴ Although the facts just discussed are sharp, there is an interesting complication, which suggests that A-movement is able to feed ϕ -agreement in a narrowly confined set of circumstances. The observation comes from extraposed clauses. While LDA into extraposed clauses is possible in some cases (e.g., (95)), it is usually somewhat degraded compared to the default agreement variant. This fact holds irrespective of whether the agreement trigger remains in its base position or is (A-)moved:

- (i) raam-ne t_k caah-aa/??-ii [har billii_i us-ke_{i/j} baccō-ko dikhaa-naa/??-nii]_k
 Ram-ERG want-PFV.M.SG/??-PFV.F.SG every cat.F its children-DAT show-INF.M.SG/??-INF.F.SG
 ‘Ram wanted to show every cat_i to its_{i/j} children.’

This restriction is arguably a manifestation of a freezing effect:

- (ii) ϕ -agreement into extraposed infinitival clauses is degraded.

I will not investigate the syntactic source of (ii) here. What is relevant for our concerns is that if a DP is moved out of an extraposed clause, it readily controls LDA. In (iii.a), the nonfinite clause is extraposed and agreement with *har billii* ‘every cat’ is degraded, just as in (i) above. In (iii.b), *har billii* is moved into the matrix clause without binding the pronoun and LDA improves considerably. In (iii.c), *har billii* is likewise moved into the matrix clause but binds the pronoun from its landing site. In this case, LDA becomes obligatory.

- (iii) a. [us-ke_j malik-ne] caah-aa/??-ii [har billii_i ghumaa-naa/??-nii]
 3SG-GEN owner-ERG want-PFV.M.SG/??-PFV.F.SG every cat.F walk-INF.M.SG/-??-INF.F.SG
 ‘Its_j owner wanted to walk every cat_i.’

\bar{A} -moved element *ghazal* cannot trigger agreement in the matrix clause in spite of the fact that it is located in the same clause. Like A-movement, \bar{A} -movement lands outside the c-command domain of the ϕ -probe and therefore does not feed agreement.

(237) \bar{A} -movement cannot result in ϕ -agreement

ghazal_i firoz-ne soc-aa/*-ii [ki monaa-ne t_i gaa-yii
 ghazal.F Firoz-ERG think-PFV.M.SG/*-PFV.F.SG that Mona-ERG sing-PFV.F.SG
 thii]
 be.PST.F.SG
 ‘Firoz thought that Mona had sung a ghazal.’

In sum, the facts considered in this section allow us to pinpoint the structural loci of A- and \bar{A} -movement and ϕ -agreement in Hindi. We have seen that (i) A-movement lands in SpecTP, (ii) \bar{A} -movement lands in SpecCP, and (iii) the ϕ -probe is located on T⁰. This structural assignment captures a wide range of distributional properties of these operations.

-
- b. har billii_i [us-ke_j malik-ne] caah-aa/-ii [t_i ghumaa-naa/-nii]
 every cat 3SG-GEN owner-ERG want-PFV.M.SG/-PFV.F.SG walk-INF.M.SG/-INF.F.SG
 ‘Its_j owner wanted to walk every cat_i.’
- c. har billii_i [us-ke_i malik-ne] caah-?*aa/-ii [t_i ghumaa-?*naa/-nii]
 every cat 3SG-GEN owner-ERG want-?*PFV.M.SG/-PFV.F.SG walk-?*INF.M.SG/-INF.F.SG
 ‘For every cat x, x’s owner wanted to walk x.’

If agreement in (iii.b–c) were established with the trace of *har billii* in the extraposed clause, LDA should be degraded, just as in (iii.a). That this is not the case suggests that LDA in (iii.b–c) is in fact established with the DP in its landing site in the matrix clause. Furthermore, the fact that A-movement of that DP renders LDA obligatory would seem to show that the A-landing site of *har billii* in (iii.c) feeds ϕ -agreement, in apparent violation of the conclusion reached in the main text.

Despite initial appearance, the two generalizations just motivated are not contradictory and in fact fully compatible with the structural assumptions argued for here. In the configurations considered in the main text, agreement with the subject is not affected by scrambling of the object. In (iii), on the other hand, it is default agreement, hence the lack of agreement, that is overridden by object scrambling. An overarching generalization that reconciles these facts is thus that A-movement feeds agreement *only if there is no alternative agreement controller*. This intriguing interaction in fact follows from (235) on the assumption that a head that has unsuccessfully probes its c-command domain subsequently probes its specifier (e.g., Béjar & Rezac 2009, Carstens 2016), hence a second-cycle Agree effect. On this view, the ϕ -probe on T⁰ first searches through its complement and agrees with the closest goal if there is one. Due to cyclicity, this probing will always take place before A-movement (i.e., movement to SpecTP takes place) and it will hence be unaffected by such movement. As a result, if T⁰ can agree with the subject, it will do so, regardless of whether the object is subsequently moved to SpecTP. This derives (233)–(234). If the ϕ -probe does not find a suitable goal in its c-command domain, then it re-probes into its specifier once A-movement has taken place. In (iii), A-movement is hence able to override default agreement. \bar{A} -movement, i.e., movement to SpecCP has no such effect, because it lands in SpecCP, hence not in the specifier of T⁰ (see (224)). This derives the fact that LDA is obligatory in (iii.c), but optional in (iii.b).

2.5.2.4 The location of the *wh*-licensor

The final operation whose structural height remains to be determined is *wh*-licensing. I will adopt here the standard view that *wh*-licensing is a property of C^0 and that the corresponding probe feature [*uwh*] is located on C^0 . Notice that this view immediately captures the fact that nonfinite clauses cannot carry interrogative force or license *wh*-elements (see the discussion in section 2.3.1). Lacking a CP layer, these clauses cannot bear the relevant [*uwh*] feature as a matter of necessity.

(238) The *wh*-probe [*uwh*] is located on C^0 .

As discussed in the next section, the desiderata in (224), (235), and (238) corroborate the Height–Locality Connection.

2.5.2.5 Consequences for the Height–Locality Connection

The conclusions about the structural location of the various Hindi operations in (224), (235), and (238), in addition to the locality profiles uncovered in sections 2.2–2.4, results in the summary table in (239). Recall that all four operations in (239) are principally independent of each other. In particular, ϕ -agreement and A-movement are distinct operations that may target different elements. The same is true for *wh*-licensing and \bar{A} -movement, as \bar{A} -movement is possible in the absence of *wh*-licensing and as *wh*-licensing does not require \bar{A} -movement.

(239) *Summary: Selective opacity across Hindi clauses*

<i>Operation</i>	<i>Landing site/ probe location</i>	<i>Size of embedded clause</i>		
		CP	TP	vP
\bar{A} -movement	C	✓	✓	✓
<i>wh</i> -licensing	C	✗	✓	✓
A-movement	T	✗	✗	✓
ϕ -agreement	T	✗	✗	✓

(✓ : transparent; ✗ : opaque)

The Height–Locality Connection is repeated in (240):

- (240) **Height–Locality Connection** [= (215)]
Movement types differ in their landing sites. The higher the landing site of a movement type is in the clausal structure, the more kinds of structures are transparent to this movement type.

The distribution in (239) conforms to the Height–Locality Connection (240): The two operations that are associated with the C domain – \bar{A} -movement and *wh*-licensing – have access to more types of embedded clauses than the two operations associated with the T domain. The locality profile of an operations is hence at least partially a function of the structural height of its landing site or probe location. The Hindi evidence thus provides independent support for the Height–Locality Connection.

At the same time, (239) makes it clear that the Height–Locality Connection is not a one-to-one mapping between location and locality. Both *wh*-movement and \bar{A} -movement are associated with the CP layer, yet they differ slightly in their locality: CPs are transparent to \bar{A} -movement, but they are opaque to *wh*-licensing. Yet at the same time, both have access to more types of embedded clauses than the two operations associated with the TP layer (*A*-movement and ϕ -agreement). The Height–Locality Connection is thus precisely that – a “connection,” not a direct “correlation,” a conclusion that will play a major role in the account developed in the next chapter.

Furthermore, this connection between height and locality is not limited to movement, but also constrains ϕ -agreement and *wh*-licensing, which, as we have seen, do not involve movement. This finding lends further support to the conclusion that selective opacity is not a property of only movement, but extends to other syntactic operations. The fact that the meta-generalizations that govern selective opacity in the domain of movement – like the Height–Locality Connection – can also be observed in non-movement operations strongly suggests that selective opacity transcends both movement and non-movement operations.

To preview the account developed in the next chapter, I will assume, following much work since Chomsky (2000, 2001) that both movement and non-movement operations involve the operations Agree. In particular, I will assume that movement is parasitic on the successful establishment of an Agree relation with a movement-inducing feature. On this view, difference in the landing sites of movement types translate into difference in the location of the probe triggering this movement.

This perspective allows us to formulate a suitably generalized version of the Height–Locality Connection that accommodates both movement and non-movement operations:

(241) **Height–Locality Connection** (*probe-based formulation*)

The higher the location of a probe is in the clausal structure, the more kinds of structures are transparent to this probe.

Deriving this link between a probe’s locality properties and its syntactic locations is one of the more difficult tasks that a theory of selective opacity faces. I will take on this task in section 3.5 of the next chapter. I will develop an account of the Height–Locality Connection that derives a link between a probe’s structural position and its locality profile, while still allowing some amount of variation between the two. This account not only derives the Height–Locality Connection, but is also flexible enough to handle cases in which height does not completely determine locality.

2.6 Chapter summary

Based on an in-depth study of selective opacity effects in Hindi, this chapter has argued for the empirical picture in (239). This arrangement of facts supports the following conclusions:

1. Selective opacity effects beyond movement

The phenomenon that a clause of given size is transparent to some operations but opaque to others can likewise be observed for non-movement operations like ϕ -agreement and *wh*-licensing. Selective opacity is thus not restricted to movement dependencies, but considerably more general. Not only does Hindi show selective opacity between A- and \bar{A} -movement, i.e., within the realm of movement dependencies, but also between movement and non-movement operations. *Wh*-licensing differs in its locality from both A- and \bar{A} -movement and ϕ -agreement differs from \bar{A} -movement. Finally, there are locality differences within the class of in-situ dependencies between *wh*-licensing exhibits a different locality profile than ϕ -agreement: TP clauses allow *wh*-licensing into them, but not ϕ -agreement. Thus, locality mismatches can be observed (i) within the domain of movement, (ii) within the domain of in-situ relations, and (iii) between movement and in-situ dependencies. There is therefore no reason to believe that locality differences are in any way confined to movement or to a categorical split between

movement and non-movement operations. Rather, selective opacity is a property of syntactic operations more generally.

2. Domain-based interactions as selective opacity

We have observed in section 2.3.3 a domain-based interaction between ϕ -agreement and A-movement: If A-movement out of an embedded nonfinite clause takes place, then this clause is necessarily transparent for ϕ -agreement into it, and LDA becomes forced. A remarkable feature of this interaction is that it is altogether irrelevant which element undergoes the A-movement step. The implicational relationship holds at the level of the clause, not at the level of individual syntactic items. I have proposed that this otherwise peculiar generalization can be understood as a consequence of selective opacity. By assumption, ϕ -agreement has the same locality profile as A-movement, but is critically distinct from \bar{A} -movement and *wh*-licensing. I have shown how this derives the fact that A-movement interacts with ϕ -agreement, but \bar{A} -movement and *wh*-licensing does not.

3. Movement–agreement mismatches as selective opacity

As mentioned in chapter 1, it has sometimes been argued that movement is subject to at least partly different locality constraints than agreement (see, e.g., Bošković 2003, Bobaljik & Wurmbrand 2005, Chomsky 2012, 2013). The pattern in (239) raises the questions of whether the locality mismatches between movement and agreement operations are any different in nature than the locality mismatches between different types of movement or agreement. I suggest that they are not. The Hindi data do not indicate a systematic binary distinction between the locality of movement and non-movement operations, precisely because neither movement dependencies nor non-movement ones pattern as a natural class. Moreover, the one pair of operations that does exhibit the same locality profile is A-movement and ϕ -agreement. While locality differences between movement and non-movement operations are clearly attested, there is no compelling reason to believe that they are the result of a mechanism above and beyond that underlying locality differences in the classes of movement and non-movement dependencies. I therefore conclude that selective opacity is significantly more widespread than it is standardly taken to be.

4. The non-binarity of selective opacity

The selective opacity pattern in (239) is not a binary pattern.⁵⁵ There are not just two types of locality profiles (corresponding to A- vs. \bar{A} -operations), but at least three (\bar{A} -movement vs. *wh*-licensing vs. A-movement/ ϕ -agreement). Any strictly binary approach to selective opacity is thus under-equipped to capture the full range of the facts. This provides further support for the conclusion in chapter 1 that locality mismatches between types of movement are not binary either.

5. Meta-generalizations of selective opacity

In reviewing the previous literature on selective opacity, I highlighted in chapter 1 two very general meta-generalizations of selective opacity noted in the literature. One is Upward Entailment, the other the Height–Locality Connection.

(242) **Upward Entailment** [=(212)]
If a clause of a certain structural size is opaque to an operation, then clauses that are structurally larger are also opaque to this operation.

(243) **Height–Locality Connection** (*probe-based formulation*) [=(241)]
The higher the location of a probe is in the clausal structure, the more kinds of structures are transparent to this probe.

The Hindi case study has demonstrated that the Hindi pattern adheres to these generalizations, thus lending further support to them. In addition, the Hindi facts also make it clear that both generalizations hold for movement and non-movement dependencies alike. This finding in turn supports the view argued for here that selective opacity transcends both types of operations.

In the next chapter, I will present my account of selective opacity and apply it to the Hindi facts presented in this section. I will then demonstrate how the account derives Upward Entailment and the Height–Locality Connection.

⁵⁵ Also see Williams (2003), who generalizes the binary A/ \bar{A} -distinction to what he calls, somewhat ironically, the A/ \bar{A} / $\bar{\bar{A}}$ / $\bar{\bar{\bar{A}}}$... distinction.

CHAPTER 3

HORIZONS

3.1 Introduction

The previous chapter has conducted an in-depth case study of selective opacity in Hindi. I have argued based on this case study that selective opacity effects are not restricted to movement, but also encompass non-movement operations like ϕ -agreement and *wh*-licensing. I have furthermore argued that selective opacity is not a binary phenomenon, in that there can be several layers of selective opacity. Finally, I have provided additional evidence for the two generalizations discussed in chapter 1: Upward Entailment and Height–Locality Connection. In this chapter, I will propose and develop a general account of selective opacity effects that is abstract enough to comprise both movement and non-movement operations and that derives Upward Entailment and the Height–Locality Connection.

The account is based on the intuition that a probe has characteristic *horizons* that define the limits of the domain that is visible to it. The central novelty of horizons is that they can differ between probes, ultimately resulting in locality differences between various processes. More concretely, I propose that a probe’s search can terminate if this probe encounters projections of a certain type. Such projections delimit the search space of a probe. Upon encountering the relevant projection, a probe terminates its search. Material embedded inside such projections will then be necessarily out of reach for this probe. No syntactic interactions with such material and the probe is hence possible and the domain will emerge as opaque for this probe. Selective opacity

of a domain follows because a domain may allow search by one probe into it but block search by another. Horizons are thus not absolute locality domains but probe-specific. Just like horizons in the real world, it is not possible for an observer to look beyond their respective horizon and at the same time horizons are inherently relative to the observer.

We will see that the concept of horizons not only affords a systematic and comprehensive account of selective opacity, it also offers a major reassessment of the underlying geometry of ‘improper movement’ because it differs from previous approaches in a number of fundamental ways. The most radical of these differences is that virtually all previous approaches to improper movement and selective opacity more generally have pursued an account in terms of the moving element, its properties or its syntactic position. The horizon account, by contrast, does not attribute any role to the moving element. On the view presented here, the empirical restrictions arise solely from the interplay of probes and their horizons. I will argue that this reassessment leads to an account that is empirically and conceptually superior.

I will assume here a standard model of clause structure comprising the projections CP, TP, ν P, and VP. All discussion in this section will also apply to models in which clause structure contains significantly more projections. One aspect of clause structure that will be crucial for the study of horizons will be whether a projection is higher or lower than another projection in the clausal spine. Thus, CP is higher than TP, which is in turn higher than ν P and so on. These relations remain constant, of course, if clause structure is expanded with additional projections. Consequently, my adopting a ‘CP > TP > ν P > VP’ clause structure here is simply for concreteness and represents neither a requirement nor an entailment of the account proposed here. The results obtained here generalize to a wide range of views on the details of clause structure.

This chapter will proceed as follows: In section 3.2, I will begin by laying out some of the problems that selective opacity in Hindi presents for standard locality principles like phases, the A-over-A Principle, and case-based accounts. I will show that these problems arise as consequences of fundamental properties of these constraints and I will conclude from these considerations that standard locality principles simply do not have the right structure to handle selective opacity effect. This conclusion motivates the need for a theory of locality that captures selective opacity effects. I propose and develop such a theory in section 3.3, which introduces and motivates the concept of a ‘horizon’ for a probe. An analysis based on horizons is then applied to a variety of empirical patterns

in section 3.4, including the various instances of selective opacity in Hindi, superraising in English, instances of movement type interactions between distinct elements, and locality differences between movement and agreement. Section 3.5 then discusses how the horizon-based account imposes constraints on possible selective opacity patterns by deriving the Height–Locality Connection (section 3.5.1). I furthermore lay out the extent to which crosslinguistic variation in locality can be captured in the proposed system (section 3.5.2). It then considers one instance of a pattern that appears to challenge the link between height and locality that horizons predict, namely topicalization and topic islands in English (section 3.5.3). I will show that this challenge is only apparent. Finally, I will introduce and motivate the concept of ‘default-horizons’ (section 3.5.4), which establishes a default locality setting for probes based on their structural position. Section 3.6 concludes.

3.2 The puzzle for standard principles

I have argued in chapter 1 that standard principles of syntactic locality like phases, subjacency, Relativized Minimality or an appeal to case fail to capture selective opacity effects for very general reasons. This section illustrates that the same conclusion emerges when we consider the Hindi pattern. Thus, I will show that a successful account of selective opacity requires a novel *type* of locality principle.

3.2.1 Phases

Let us first consider what problems an account of selective opacity in terms of phases would encounter. According to phase theory (Chomsky 2000, 2001, *et seq.*), a syntactic derivation proceeds in small chunks of structure, which are periodically removed from the workspace, a process called **SPELL-OUT** (or *Transfer*). Elements that have undergone Spell-Out are then inaccessible to further computations. There are a variety of specific proposals that all follow this basic conception, but the specific details of each account do not need to concern us here as the challenge that selective opacity poses for these accounts is quite general.

For the sake of the exposition, I will assume here Chomsky's (2000) classical version of phase theory, according to which C^0 and v^0 are the phase heads. Once the next higher head is merged into the structure, the complement of a phase undergoes Spell-Out. This results in the **PHASE IMPENETRABILITY CONDITION** (PIC) in (244), where H's *domain* is its complement and its *edge* its specifier(s):¹

(244) **Phase Impenetrability Condition** (Chomsky 2000: 108)

In phase α with head H, the domain of H is not accessible to operations outside of α , only H and its edge are accessible to such operations.

The PIC thus prohibits syntactic operations between material in a phase head complement and material outside the highest projection of the phase head. How does the PIC relate to selective opacity?

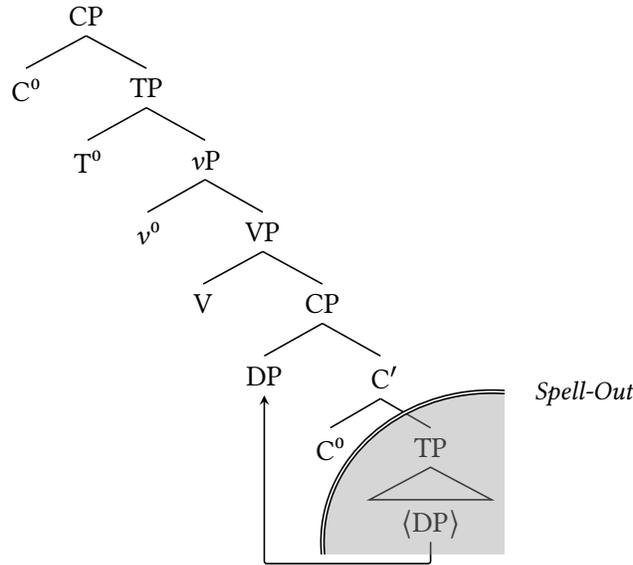
The core intuition behind phase theory is that certain syntactic material is derivationally removed from the workspace. Removal of material has the consequence that this material becomes inaccessible to *all* syntactic operations. In other words, the PIC imposes an *absolute* locality boundary, in the sense that this boundary, as a matter of principle, holds without exception for all dependencies. On phase theory, then, locality is *binary*. Either a given part of the structure is already spelled-out or it is not yet spelled-out. If it is spelled-out, it is inaccessible to all operations; if it is not yet spelled-out, it is accessible to all operations.

This binary nature of phase-based locality puts selective opacity effects beyond the reach of phases. Phase locality does not have the appropriate structure to handle patterns in which a given element is accessible to some operations, but inaccessible to others. Finite clauses in English and Hindi serve to illustrate the problem. CP being a phase (Chomsky 2000, 2001), nothing but the specifier of CP is accessible for operations in a higher clause. This is illustrated schematically in (245).²

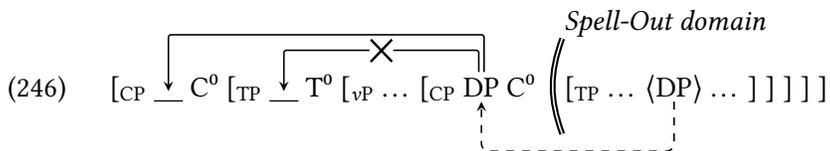
¹ Chomsky (2001) proposes a second version of the PIC (see section 6.3.2 in chapter 6 for discussion). The problem discussed here applies to both.

² I am ignoring vP phases in (245). The analytical problems discussed in the text arise regardless of whether vP is a phase or not. The relation between selective opacity and vP phases will be taken up in section 6.3 of chapter 6.

(245) *Edge accessibility in CP phases*



In (245), the DP moved to SpecCP of the embedded clause remains accessible to material in the matrix clause, simply because subsequent \bar{A} -movement of DP into the matrix clause is possible. At the same time, A-movement from the same position is ruled out, a familiar case of superraising:



An element at the edge of the phase is hence eligible for \bar{A} -movement, but not for A-movement. Because the PIC (244) merely states that the edge of a phase is accessible, the asymmetry between A- and \bar{A} -movement cannot possibly be the result of the PIC, which is the reason for why additional principles unrelated to phases (like the ban on improper movement) have been invoked.

The same point can be made on the basis of the Hindi evidence presented in chapter 2. We know independently that a DP can escape CP's Spell-Out domain by movement to SpecCP in Hindi because a DP can undergo \bar{A} -movement out of a CP clause, just like in English. A relevant example is given in (247), repeated from (62) in chapter 2.

(247) \bar{A} -movement from SpecCP

mohan-ko_i raam-ne soc-aa [*t_i* ki siitaa-ne *t_i* dekh-aa thaa]
 Mohan-ACC Ram-ERG think-PFV.M.SG that Sita-ERG see-PFV.M.SG be.PST.M.SG
 'Ram thought that Sita had seen Mohan.'

But if the DP in (245) can remain accessible to the matrix clause by moving to SpecCP, the PIC entails that it is accessible to *all* operations. Yet as I have argued at length in the previous chapter, this is incorrect. A-movement, ϕ -agreement, and *wh*-licensing cannot operate on the embedded DP even if it is at the edge of the lower clause. This is shown in (248)–(250), repeated from (76b), (103) and (175), respectively:

(248) *No A-movement from SpecCP*

*har larke-ko_i [us-kii_i bahin] soc-tii hai [t_i ki raam-ne t_i
 every boy-ACC 3SG-GEN sister think-IPFV.M.SG be.PRES.3SG that Ram-ERG
 dekh-aa]
 see-PFV.M.SG

Intended: ‘For every boy *x*, *x*’s sister thinks that Ram saw *x*.’

(249) *No ϕ -agreement with SpecCP*

larḱõ-ne soc-aa/*-ii [ghazal_i monaa-ne t_i gaa-yii thii]
 boys-ERG think-PFV.M.SG/*-PFV.F.SG ghazal.F Mona-ERG sing-PFV.F.SG be.PST.F.SG
 ‘The boys thought that Mona had sung ghazal.’

(250) *No wh-licensing in lower SpecCP*

*siitaa-ne soc-aa [kis-ko_i ravii-ne t_i dekh-aa]
 Sita-ERG think-PFV.M.SG who-ACC Ram-ERG see-PFV.M.SG

Intended: ‘Who did Sita think that Ravi saw?’

We may call this the **RELATIVITY PROBLEM**. Whether a given position is accessible or inaccessible is relative to the operation involved. Such relativity cannot be attributed to phases *per se* because the locality that phases give rise to is binary and absolute.

Furthermore, the direction of the relativity problem is quite curious: After all, the DP has to remain accessible for the \bar{A} -probe on the higher C⁰ head, but inaccessible to the ϕ - and A-probes on T⁰. Consequently, no constraint that requires shorter dependencies will be able to capture this contrast.

A second problem related to relativity is what one may call the **POROSITY PROBLEM**. Phases are by design porous in that they allow extraction out of them as long as this extraction proceeds through the edge of the phase. A general consequence of this property is that phases do not produce opaque domains, at least unless additional and unrelated stipulations are invoked, a point that has

been raised by Boeckx & Grohmann (2007) and Abels (2012b), among others. Consequently, phases in and of themselves do not offer an account of configurations in which a domain is opaque to extraction. All else equal, all extraction should be licit as long as it proceeds through the phase edge.

In sum, phases do not have the right analytical structure to provide an account of selective opacity effects. Because phase locality is binary in the sense that a given domain is either entirely accessible to all operations or entirely opaque to all operations, phenomena in which a domain's opacity is relative to operation involved fall outside the purview of phases.

This conclusion is independent of the specific version of phases employed. Thus, the same problem arises on accounts in which every phrase constitutes a phase (see, e.g., Manzini 1994, Bošković 2002, Epstein & Seely 2002, Boeckx 2003, Fox & Lasnik 2003, Lahne 2009, Müller 2010b, 2011), or on accounts on which the phasehood of a node is determined contextually (e.g. Bošković 2005, 2014, den Dikken 2007, Gallego & Uriagereka 2007, Takahashi 2011, 2012). While these proposals modify the distribution of phases, they retain the fundamentally binary nature of phase locality that was shown to underlie the relativity problem. Similar remarks apply to historical predecessors to phases, namely *subjacency* (Chomsky 1973, 1977, 1981) and *barriers* (Chomsky 1986), because, like phases, these principles are binary in nature.

I hasten to add that the existence of selective opacity effects does not constitute an argument *against* the existence of phases. In fact, I will argue in chapters 6 and 7 that at least CP phases are independently required. My point here is that selective opacity *cannot* be a consequence of phases, however modeled, because phases do not have the right theoretical structure to handle cases in which different operations have different locality profiles. Selective opacity must hence be traced back to a locality principle with a character quite different from phases.

3.2.2 A-over-A Principle

A second locality principle worth considering in the context of selective opacity is the A-over-A Principle. The A-over-A Principle prevents an operation from targeting an element that is embedded inside a constituent that is itself a licit target for that operation. While originally formulated on the basis of categorial features by Chomsky (1964, 1973) and Bresnan (1976), the A-over-A Principle

has been further revised and extended in subsequent work with particular attention to differences between movement types. Müller (1996, 1998), for instance, argues that extraction out of a remnant is impossible if this remnant itself undergoes the same extraction type and attributes this restriction to a revised version of the A-over-A Principle (also see Takano 1994, Koizumi 1995, Kitahara 1997, Sauerland 1999, Fitzpatrick 2002, van Urk & Richards 2015). Müller (2011) calls this the *F-over-F Principle*, formulated in (251), where [uF] is a movement-inducing probe feature.

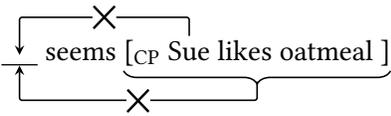
(251) **F-over-F Principle** (Müller 2011: 42)

In a structure $\alpha_{[uF]} \dots [\beta_{[F]} \dots [\gamma_{[F]} \dots] \dots] \dots$, movement to [uF] can only affect the category bearing the feature [F] that is closer to [uF].

According to (251), an element γ cannot embedded inside β cannot undergo a movement if β can also undergo this movement. What is crucial for our concerns is that the F-over-F Principle is stated on the basis of features and hence discriminates between different movement types: If different movement types are triggered by different features, a constituent that undergoes scrambling is opaque for scrambling out of it but is transparent for topicalization, and so on. In other words, the F-over-F Principle gives rise to *relativized* locality effects. Whether or not a given movement out of a domain is possible depends on whether the domain can itself undergo this movement. There are indeed a number of accounts that attribute one instance of selective opacity, the ban on superraising, to this principle (McFadden 2004, Halpert 2015, van Urk 2015, and also Nunes 2008, 2010).

Yet while there are specific instances in which a selective opacity pattern might plausibly be attributed to the F-over-F Principle (see Nunes 2008, 2010 for Brazilian Portuguese and Halpert 2015 for Zulu), it is doubtful whether the F-over-F Principle provides a general account of selective opacity effects. To repeat the argument advanced in chapter 1, the ban on A-movement out of a finite clause (superraising) in English cannot be attributed to the F-over-F Principle because the finite clause is not itself a closer target for A-movement. Thus, the structure of the ungrammatical sentence in (252) before A-movement *Sue* is given in (253). Importantly, not only is A-movement of *Sue* impossible, but A-movement of the embedded clause is illicit as well.

(252) *Sue_i seems [CP t_i likes oatmeal]

(253) 

a. CP movement → *[CP Sue likes oatmeal] seems.

b. DP movement → *Sue seems [CP t likes oatmeal].

It is an empirical fact of English that the embedded CP in (253) cannot undergo A-movement into the matrix subject position. As such, it does not constitute a closer target for T's EPP-probe and A-movement of the embedded subject DP is therefore allowed by both the classical A-over-A Principle and the more recent F-over-F Principle. Analogous considerations apply to ϕ -agreement on Iatridou & Embick's (1997) proposal that CPs lack ϕ -features. Because CPs are thus not themselves viable targets to A-movement and ϕ -agreement, the A-over-A Principle and the F-over-F Principle do not lend themselves to an account of why CPs in English block these operations across them, At least unless further stipulations unrelated to these principles are added.

Selective opacity in Hindi poses exactly the same problem. Despite the general freedom of word order in Hindi, finite clauses resist movement. As (254) demonstrates, it is altogether impossible to move a finite complement clause (also see Bhatt & Dayal 2007 and Manetta 2010 for related recent discussion):

(254) *No movement of finite clauses*

a. raam-ne kah-aa [ki paris sundar hai]
 Ram-ERG say-PFV.M.SG that Paris beautiful be.PRES.3SG
 'Ram said that Paris is beautiful.'

b. *raam-ne [ki paris sundar hai]_i kah-aa t_i
 Ram that Paris beautiful be.PRES.3SG say-PFV.M.SG

c. *[ki paris sundar hai]_i raam-ne kah-aa t_i
 that Paris beautiful be.PRES.3SG Ram say-PFV.M.SG

To correctly rule out (254b,c), finite clauses must be blocked from undergoing A- as well as \bar{A} -movement (and any other movement type there may be). The source of this restriction does not need to concern us here. What is crucial is that despite their inability to undergo movement themselves,

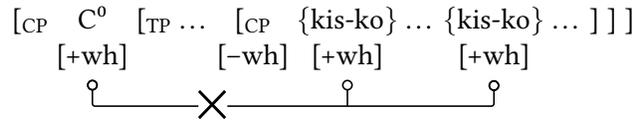
they block A-movement out of them. As in the case of English superraising just discussed, the F-over-F Principle has nothing to say about this constraint. To make the problem even worse, recall from section 2.2.1 in chapter 2 that finite clauses in Hindi only block A-movement, but not \bar{A} -movement. The fact that they can themselves undergo neither casts serious doubts on any approach that attempts to analyze selective opacity in terms of the F-over-F Principle, simply because a clause's opacity for a process seems unrelated to its own properties with respect to this process.

Syntactic relations that do not involve movement provide additional illustrations of this general problem. Consider *wh*-licensing in Hindi. As discussed in detail in section 2.4 of chapter 2, there is evidence that Hindi *wh*-licensing is a genuine long-distance relationship between a C^0 head and a *wh*-element that is not mediated by overt or covert movement. We have also seen there that *wh*-licensing partakes in selective opacity: It cannot take place across finite clauses boundaries even if it is at the edge of that clause, as illustrated in (255) (a domain that is transparent to \bar{A} -movement). Nonfinite clauses, on the other hand, are transparent to *wh*-licensing, regardless of whether they are TPs or *v*Ps (see section 2.4.2). It is hence the CP projection in the lower clause that induces the opacity for *wh*-licensing.

- (255) *No wh-licensing across finite clause boundary* [= (174), (175)]
 *siitaa-ne soc-aa [{kis-ko} ravii-ne {kis-ko} dekh-aa]
 Sita-ERG think-PFV.M.SG {who-ACC} Ram-ERG {who-ACC} see-PFV.M.SG
Intended: 'Who did Sita think that Ravi saw?'

The fundamental problem that an F-over-F account of this restriction faces is that *all* CPs are opaque to *wh*-licensing, irrespective of whether these CPs are [+wh] or [-wh]. In the structure in (255), a matrix question is intended and the embedded clause is therefore non-interrogative. The structure of (255) is thus as given in (256). In this structure, it is not possible for the matrix C^0 to license the *wh*-element inside the lower clause. Note that this locality restriction does not reduce to phasehood because it holds even if the *wh*-element is at the edge of the lower clause.

(256) *Schematic structure of (255)*



Because the F-over-F Principle rules out syntactic relations only if these relations cross an element that bears the same feature, the structure in (256) does not violate it. More generally, from the point of view of Relativized Minimality, it is entirely mysterious why a [-wh] CP should intervene for *wh*-licensing. The general shape of the problem is the same as with the movement cases above: Domains may be opaque to a relation even if they could not themselves enter into this relation. Appeals to minimality of dependencies do not capture such situations.

In sum, selective opacity arises even in configurations in which the opaque domain is not itself a licit target for an operation. In such configurations, the F-over-F Principle has nothing to say about the opacity of this domain. While there are plausibly instances where movement out of a lower clause is impossible because this clause constitutes a closer target for the movement (see, e.g., Nunes 2008, 2010, Halpert 2015), the F-over-F Principle itself does not provide a comprehensive account of selective opacity effects.

3.2.3 Case and the Activity Condition

A third type of approach quite distinct in nature from phases and minimality is based on case. Case has traditionally played a large role in constraining the distribution of nominal elements (e.g., Chomsky's 1981 Case Filter). More recently, Chomsky (2000, 2001) has proposed that a DP whose case is valued is rendered invisible to subsequent A-processes, the so-called **ACTIVITY CONDITION**. This account rules out superraising in English because an element inside a finite clause will receive case within that clause. As a consequence, it cannot then undergo A-movement out of this finite clause. The gist of case-based approaches to the distribution of DPs is thus that it is the internal properties of a DP (valued vs. unvalued for case) that determine its movement options.

A thorough discussion and evaluation of case-based approaches in the context of selective opacity will be carried out in section 5.2 of chapter 5. To summarize the key conclusions reached there, case-based approaches are narrowly geared towards the properties of A- and \bar{A} -movement in English and therefore they do not extend to selective opacity effects more generally. For example,

we saw that in Hindi *wh*-licensing and \bar{A} -movement, which are distinct operations, have different locality properties. Because both are \bar{A} -operations and unrelated to case, a case-based approach has nothing to say about them. This point generalizes. A cursory inspection of the various examples of selective opacity presented in chapter 1 makes it clear that selective opacity is much more widespread and diverse than a simple split between A- and \bar{A} -operations. The same conclusion is reached in the detailed investigation of selective opacity in German in chapter 5. In sum, there is no indication that selective opacity in its variety of forms is in any way connected to properties of case, although locality differences between movement types may of course happen to coincide with differences in case.

Furthermore, I will argue in section 5.2 that A- and \bar{A} -movement in Hindi differ from their English counterparts in that both can apply to DPs whose case is already valued. In other words, neither feeds case assignment. A case-based account will therefore be unable to distinguish between them and their locality difference cannot be attributed to case. As a result, even locality differences between A- and \bar{A} -movement in languages other than English fall outside the scope of a case-based account.

Finally, once a theory of selective opacity is available that is general enough to cover a wide array of locality mismatches – like the account proposed here – superraising in English will fall under this more general theory, making any appeal to case redundant and hence superfluous.

3.2.4 Section summary

In this section, I have argued that standard locality principles like phases, A-over-A minimality, and case do not have the right structure to capture selective opacity effects. On the one hand, this demonstrates that selective opacity effects make it necessary to explore a new *type* of locality, distinct in structure and scope from more familiar constraints. On the other hand, this section has also served to sharpen our understanding of what properties a theory of selective opacity needs to have. First, selective opacity requires *relativized* locality, in the sense that one and the same domain can be transparent and opaque at the same time, albeit to different operations. It is this requirement that makes implausible a phase-based account of selective opacity. Second, a domain can be selectively opaque to an operation even if this domain cannot itself participate in

this operation. This property poses a severe problem for an A-over-A/F-over-F account of selective opacity. Third, because selective opacity is not confined to the binary divide between A- and \bar{A} -processes, the account has to be fine-grained enough to capture this non-binarity. This excludes a simple case-based account. I thus conclude that selective opacity effects require an altogether different approach, which I will embark on in the next section.

3.3 The proposal

3.3.1 Horizons

The account of selective opacity to be developed in this chapter takes locality differences between operations at face value. At the most general level, then, I propose that syntactic operations can be subject to distinct locality boundaries, and that at least some locality boundaries are *relative* in that they discriminate between different operations. This is a key difference to classical locality conditions like subadjacency, barriers or phases, all of which are absolute and strictly binary.

I have argued at length in chapter 2 that selective opacity is not proprietary to movement, but rather a property of syntactic operations more generally, regardless of whether they involve movement or not. This parallelism suggests that both types of operations share some component which the relevant constraint applies to. While several ways of developing this unification are possible, I will follow here the view in Chomsky (2000, 2001) that movement is parasitic on the operation Agree. Because on this view Agree is a component of both movement and non-movement operations, constraints on Agree will restrict both in a uniform fashion.³

I have argued in chapter 2 that A-movement, \bar{A} -movement, ϕ -agreement, and *wh*-licensing are all distinct operations in Hindi. Specifically, A-movement can take place without ϕ -agreement and vice versa, and *wh*-licensing can apply without \bar{A} -movement and vice versa. All four processes are

³ Alternative approaches to the unification of movement and non-movement operations are of course conceivable and worth exploring. One alternative option is to view ϕ -agreement and *wh*-licensing as involving covert movement and that all agreement relationships are Spec-Head relationships (see Koopman 2006 for a recent defense of this view). Such an account would have to address the various arguments advanced in chapter 2 to the effect that no overt or covert movement is involved in the establishment of ϕ -agreement and *wh*-licensing. Another alternative would be to postulate that in-situ relationships necessarily involve feature movement (e.g., Chomsky 1995b, Pesetsky 2000). On such a view, a constraint on movement would also constrain ϕ -agreement and *wh*-licensing. In light of the conceptual proximity between Agree and feature movement, I do not know of a way to empirically distinguish between these two views and the two views strike me as notational variants. I adopt the Agree-based view to be consonant with the current literature, but would like to emphasize that this is primarily an expository choice.

hence operationally independent. I will thus assume that these four operations are each triggered by a corresponding probe feature, which I will notate as $[uF]$, where ‘F’ is a placeholder for the feature involved. A- and \bar{A} -movement are parasitic on the prior establishment of a successful Agree relationship with their respective features, which I will simply notate mnemonically as $[uA]$ and $[u\bar{A}]$. Movement-inducing features like $[uA]$ and $[u\bar{A}]$ thus require a local relationship between the probe and the goal that is brought about by movement into the specifier of the head hosting the features.

I take optional movement of this sort to be triggered by information-structural reasons and hence features like [topic] and [focus], a view that is by no means novel, and has been independently argued for by, e.g., Kidwai (2000) for Hindi, by Frey (2004) and Grewendorf (2005) for German, by Horvath (1986, 1995) for Hungarian, and by Miyagawa (2010) and references cited there for Japanese. As my primary concern here is not the information-structural impact of A- and \bar{A} -movement in Hindi, but rather the locality conditions they are subject to, I will use the labels $[uA]$ and $[u\bar{A}]$ to refer to their respective probe features. The view that apparently optional movement is triggered by syntactic features has been independently argued for by Müller (1996, 2010b), Grewendorf & Sabel (1999), and Sauerland (1999) in the context of scrambling. ϕ -Agreement and *wh*-licensing are triggered by the features $[u\phi]$ and $[uwh]$, respectively. As for the structural location of these four probe features, we saw in section 2.5.2 of chapter 2 that ϕ -agreement and A-movement are associated with the T domain, whereas *wh*-licensing and \bar{A} -movement are associated with the C domain. We can thus deduce from the evidence presented there that the ϕ -probe and the A-probe are located on T^0 , while the *wh*-probe and the \bar{A} -probe reside on C^0 . This is summarized in (257), where a subscript indicates the location of a probe. All four probes in (257) are separate syntactic entities and probe separately from each other, but they may of course find the same goal in certain cases.

(257) *Hindi probes and their locations*

- a. ϕ -agreement: $[u\phi]_{T^0}$
- b. A-movement: $[uA]_{T^0}$
- c. *wh*-licensing: $[uwh]_{C^0}$
- d. \bar{A} -movement: $[u\bar{A}]_{C^0}$

The features $[uA]$ and $[u\bar{A}]$ trigger movement of the agreed-with element to the specifier of the head hosting the probe. $[u\phi]$ and $[uwh]$, on the other hand, can be agreed with non-locally and they do not induce movement. I will have nothing to say about the source or structural implementation of this distinction. For concreteness, we may assume that $[uA]$ and $[u\bar{A}]$ have an EPP-property or, equivalently, are ‘strong’ features. Because the account developed here focuses solely on the Agree relation and not the movement component, I will not notationally distinguish between the two types of probes.

I will follow the standard assumption that probes are features of heads. That is, syntactic trees consist of a sequence of projections headed by lexical items (V^0 , v^0 , T^0 , etc.) and probes are part of the featural makeup of heads. Crucially, any given head can host more than one probe feature: T^0 in Hindi hosts $[uA]$ and $[u\phi]$ and C^0 contains $[u\bar{A}]$ and $[uwh]$. I will also assume, though not crucially so, that each probe is inherently located on a specific head. That is, every probe has as part of its specification exactly one head that it is part of (though see fn. 18 on p. 226 for a possible alternative). For example, a language may have one ϕ -probe on one head and a second one on another head. In such a case, I will treat both ϕ -probes as distinct syntactic objects. That is, I will conceive of probes as tokens, not as types.

These features probe into their c-command domain for a matching counterpart (on a DP, for all the operations discussed here). This operation ‘Agree’ is stated in (258):

(258) **Agree**

An unvalued feature $[uX]$ (the ‘probe’) serially searches through its c-command domain for a valued counterpart (a ‘goal’). It agrees with the closest goal.

Following the *obligatory operations* framework of Preminger (2011, 2014), Agree is obligatory if it is possible, i.e., if there is a matching goal in a probe’s search space. If a probe does not find a matching goal to agree with, no ungrammaticality arises.

(259) **Agree obligatoriness**

If a probe finds a matching goal, the two have to agree. If a probe does not find a matching goal, no ungrammaticality results and the unvalued probe is spelled out.

Against this background, I propose that search by a probe is subject to **HORIZONS**, characteristic nodes that define the outer bounds of the domain visible to that probe. When a probe’s search

through its c-command domain reaches that probe's horizon, the search terminates. Due to the serial nature of the search process in (258), any element beyond a probe's horizon is inaccessible to that probe as its search terminates before the probe encounters the element.

An immediate question is what defines such horizons. For reasons to become clear momentarily, I will explore the view that it is *category features* that can lead to search termination. That is, a probe can lead to search termination if it comprises a certain category specification. Crucially, probes may differ in which category features terminate their search. This in turn will lead to locality differences between operations. Termination is *defective* in the sense that the element that leads to termination is not itself a viable goal for the probe.⁴ I will refer to category labels that block a probe's search as **PROBE-HORIZONS** or just **HORIZONS**:

(260) **HORIZONS**

If a category label X is a horizon for probe [*uF*] (notated as '[*uF*] –|| X'), then a [*uF*]-initiated search terminates at XP. All elements dominated by XP are therefore outside [*uF*]'s search space.

According to (260), particular category features can lead to termination of a search process and thus delimit a probe's search space. If, for example, a probe has the category feature C as its horizon, a CP node will terminate that probe's search in virtue of being of the category C. Recall that a probe serially searches through its c-command domain. As soon as such a probe encounters a node that bears the relevant category feature, the search terminates and no Agree or valuation takes place. As a result, the probe will fail to enter into a syntactic relationship with *any* element. As I will discuss in detail momentarily, the crucial hypothesis behind horizons is that they are probe-specific and can hence differ between probes.

The fact that horizons are fundamentally based on category features places them into a broader range of cases in which category features play a crucial role. These include selection/subcategorization, EPP-probing in English (which requires a DP), and perhaps case features as uninterpretable

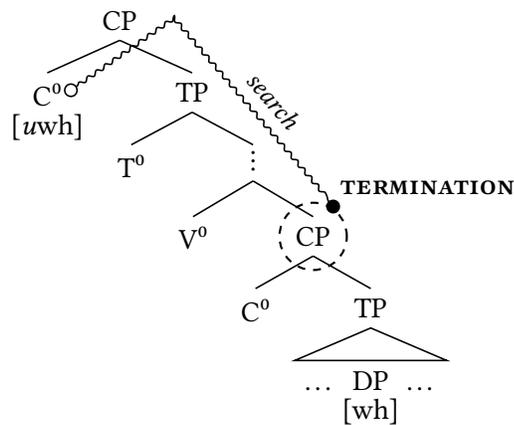
⁴ In a sense, my proposal amounts to a defective version of the F-over-F Principle discussed in section 3.2.2. This defective character of my proposal ensures that the cases that were problematic for the F-over-F Principle do not pose problems for my proposal. I would also like to note that the recognition of a defective version of the F-over-F Principle is strongly reminiscent of the existence of defective cases of c-command-based intervention effects, like Chomsky's (2001) defective intervention effects. Put differently, (260) is a dominance-based counterpart to defective intervention. While I take this analytical symmetry to be desirable, a unification of the two principles is not straightforward as it is not clear that defective intervention is crucially related to the category of the intervener.

category features on nominals (Pesetsky & Torrego 2001). Horizons thus constitute one more domain in which syntactic category features play a central role. Moreover, in some sense, horizons are an “inverse EPP”. While the EPP limits a probe’s search to elements of a certain category, horizons terminate (and in a way repel) a probe’s search.

A note on notation: For the sake of clarity, I will distinguish between heads and syntactic categories. For example, I will refer to the syntactic head that projects a TP in the standard way as ‘T⁰’. The syntactic category/label will be referred to as ‘T’. Thus a TP node has the syntactic category T, but of course it is not a T⁰. Crucially, horizons are sensitive to syntactic categories, not heads.

To illustrate the basic workings of (260), consider the structure in (261), which contains a probe feature [*uwh*] on C⁰, which, by assumption, has C as its horizon. This probe searches serially through its c-command domain. Upon encountering the topmost CP node of an embedded nonfinite clause, [*uwh*]’s search terminates. As a consequence, [*uwh*] will be unable to enter into an Agree relationship with any material inside the CP clause. In the structure in (261), the embedded clause does contain a DP with a [*wh*] feature matching the probe, but Agree is ruled out because [*uwh*]’s search terminates before it reaches this DP.

(261) [*uwh*] \neq C \rightarrow no Agree into a CP



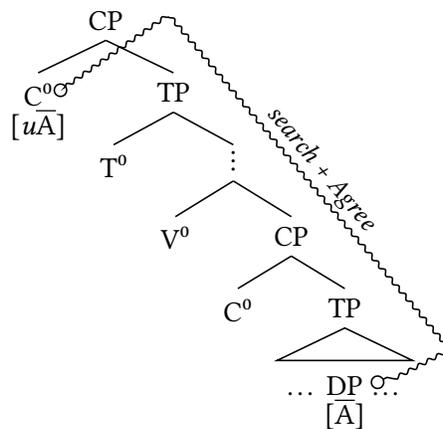
The key consequence of (261) is that *wh*-licensing will be impossible across a CP node and hence into an embedded CP clause, because C constitutes a horizon for [*uwh*]. As the example of [*uwh*] and [*uA*] makes clear, horizons are features of probes. Two probes that are situated on the same syntactic head may therefore differ in their locality. There is hence no direct analytical link between

heads and the horizons of the probes located on these heads. However, we will see in section 3.5 below that the account pursued here derives an indirect link between heads and locality, which will capture the connection between height and locality motivated in chapters 1 and 2.

It is important to note that it is the topmost node of the embedded clause that bears the C specification and that as a consequence it is this node that leads to search termination. Consequently, *no* element inside that CP will be visible to $[uwh]$, including elements at the edge of the embedded clause. A crucial consequence is thus that horizons impose absolute opacity. Just like horizons in the real world, anything that lies beyond them is invisible. There are no edges with special properties as far as horizons are concerned. In this regard, they are fundamentally different from phases.

Probes may differ in their horizons. A category specification that constitutes a horizon for one probe may not be another probe's horizon. Again, just like in the real world, horizons are relative to the observer. Contrast the structure in (261) to the one in (262). Here a second probe, $[u\bar{A}]$, which also resides on C^0 does not have CP as its horizon. Consequently, $[u\bar{A}]$'s search is not delimited by a CP node and is therefore able to probe past it into the embedded clause. It may then enter into an Agree relation with a DP embedded inside this CP clause. As a consequence, an embedded CP clause is opaque to $[uwh]$, but transparent to $[u\bar{A}]$.⁵

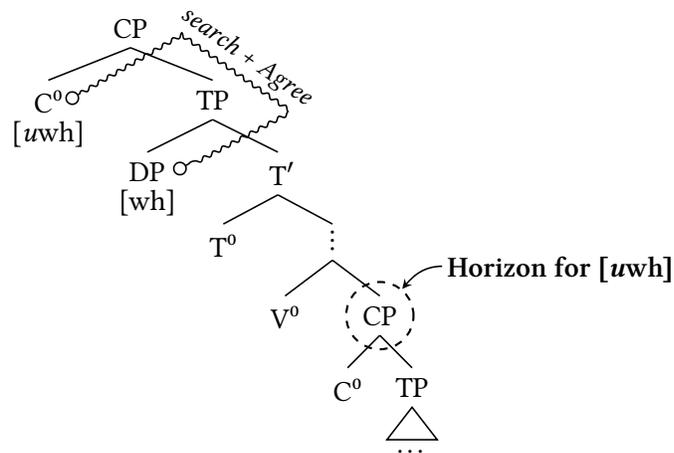
(262) $[u\bar{A}] \neq C \rightarrow$ Agree into CP possible



⁵ I sketch in (262) the search space of $[u\bar{A}]$ relative to horizons. There may, of course, be additional constraints that impose additional limitations. In fact, I will argue in chapter 6 that CP phases constitute a second constraint on Agree relations, which will limit $[u\bar{A}]$'s search space in (262) to the edge of the lower clause. The key difference between (261) and (262) remains: The edge of the lower clause is visible only to $[u\bar{A}]$, not to $[uwh]$.

Termination induced by horizons is a derivational process. That is, termination occurs if a probe encounters a horizon on its serial search through its c-command domain. As just discussed, this prevents a probe from agreeing with anything dominated by its horizon. At the same time, elements that are located between a probe and its horizon are visible to this probe, simply because the probe makes contact with these elements before it does with its horizon and search terminates. This is sketched in (263). No termination takes place in (263) because $[uwh]$ finds a goal before its CP horizon would induce termination.

(263) $[uwh] \dashv\vdash C \rightarrow$ *matrix material remains visible*



The notions of horizons and search termination provide a means to approach selective opacity effects that circumvents the problems that selective opacity poses for more traditional concepts of locality. With phases, horizons share the defectivity of the intervention effect: A node induces search termination even if it does not itself constitute a goal for the probe. Unlike phases, however, horizons are fundamentally relative to the probe initiating the search process. One and the same node can lead one probe to terminate but not another, effectively rendering the domain it dominates opaque to the former probe but transparent to the latter. A second key difference to phases is that the edge of the domain does not play any privileged role. If a phrasal node incurs search termination, all material dominated by this node will be invisible to the probe, regardless of whether this material is located at the edge of this domain or not.

With the F-over-F Principle, horizons share the relativity of the boundary. On the other hand, horizons diverge from the F-over-F Principle in that a node can block probing into it regardless of whether this node constitutes a goal to the probe. This characteristic ensures that an account in

terms of horizons does not suffer from the problems that plague an F-over-F account of selective opacity discussed in section 3.2.2.

It is clear that selective opacity follows directly from horizons: A given domain Δ is opaque to an operation α but transparent to the operation β if Δ is dominated by a node δ , whose category is a horizon for the α -probe but not for the β -probe. A question that immediately emerges is, of course, how horizons are distributed. Specifically, we seen in chapters 1 and 2 that selective opacity exhibits two meta-generalizations, i.e., Upward Entailment and the Height–Locality Connection. Selective opacity is consequently not distributed randomly, a fact that any account of selective opacity has to capture. I will next consider Upward Entailment and demonstrate how Upward Entailment can be derived within the broader hypothesis of horizons. I will defer a discussion of the Height–Locality Connection until section 3.5.

3.3.2 Deriving Upward Entailment

Recall from the discussion in chapters 1 and 2 the meta-generalization of Upward Entailment, suitably generalized in (264) to comprise both movement and non-movement dependencies.

(264) **Upward Entailment**

If a clause of a certain structural size is opaque to an operation, then clauses that are structurally larger are also opaque to this operation.

Upward Entailment rules out an operation that, for instance, can cross a finite clause boundary but not a nonfinite one, while the reverse is allowed. It is clear that there are operations that can access a nonfinite clause but not a structurally larger finite one. To give an illustration, A-movement and ϕ -agreement are impossible across a TP clause boundary. Upward Entailment states that as a consequence, they are also impossible out of CP clauses, as they are structurally larger than TP clauses. Extraposition in English provides an analogous example. To the best of my knowledge, the inverse – cases in which only the structurally larger clause is transparent to an operations – are unattested.⁶ Addition of clause structure leads to increased opacity, never to increased transparency.

⁶ Interestingly, Halpert (2015) argues that in Zulu finite clauses, but not nonfinite ones, allow A-movement out of them, in apparent violation of Upward Entailment. As already mentioned in fn. 7 on p. 22, she also provides evidence that nonfinite clauses are nominalized and clauses are not. If this is correct, A-movement in Zulu is

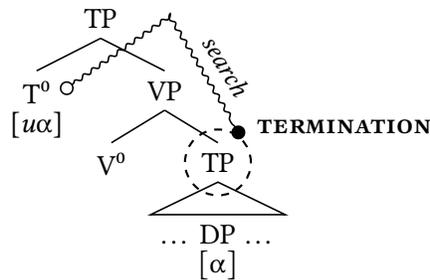
3.3.2.1 The puzzle

The notion of horizons in (260) by itself does not derive (264). To appreciate this fact, consider the exemplary probes in (265). The probe $[u\alpha]$ is located on T^0 and has T as its horizon, hence not being able to search into a node of category T.⁷ It is irrelevant whether $[u\alpha]$ is a movement-inducing probe or not.

(265) *Hypothetical horizon:*
 $[u\alpha]_{T^0} \dashv\!\! \dashv T$

All else equal, such a probe would give rise to a pattern that violates Upward Entailment. Let us consider first a configuration in which a bare TP clause is embedded, as schematized in (266), where νP is not shown for simplicity. In this structure, $[u\alpha]$'s search terminates at the embedded TP node. Consequently the embedded clause is entirely opaque to $[u\alpha]$: If $[u\alpha]$ is a movement-inducing probe, no α -extraction is possible out of TPs, and if $[u\alpha]$ is an agreement probe, no agreement is possible into TPs. The DP in (266) is shown with a feature $[\alpha]$ matching the probe.

(266) *Opacity of TP clauses under (265)* ($[u\alpha] \dashv\!\! \dashv T$)

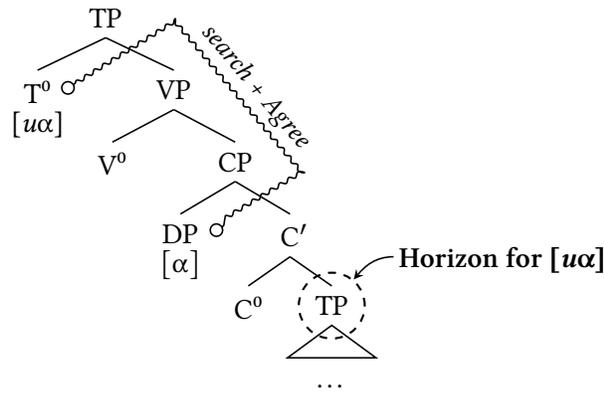


Compare this state of affairs to one in which the embedded clause is a CP instead of a TP, as sketched in (267). The account developed so far will produce an incorrect result for such a case. Because the embedded clause contains a C^0 head and CPs are not horizons for $[u\alpha]$, an element base-generated in SpecCP would be accessible to $[u\alpha]$ and Agree could be established. In this situation, then, $[u\alpha]$ could agree into the edge of an embedded CP clause, allowing extraction out of it if $[u\alpha]$ induces movement and agreement into it if it does not.

blocked by the presence of nominal structure. Because only nonfinite clauses contain such nominal structure, the Zulu pattern is fully consistent with (264).

⁷ Recall that I notate the syntactic category feature as T and corresponding head as X^0 .

(267) *Transparency of CP clauses under (265) (to be revised)*



The contrast between (266) and (267) would violate Upward Entailment because CP clauses would be descriptively transparent to the process α – be it movement or agreement –, while structurally smaller TP clauses would not be. This would constitute the inverse of the locality pattern of English A-movement. As mentioned above, such cases do not seem to exist.

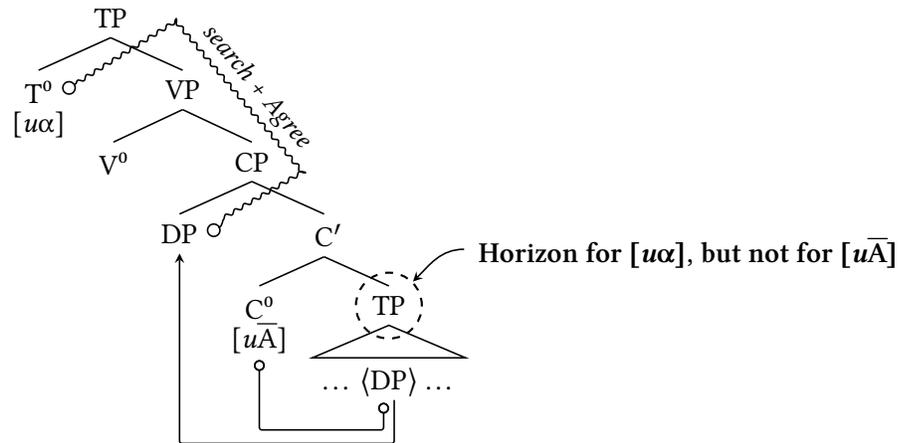
The DP in the structure in (267) is base-generated in SpecCP and accessible there. The resulting problem for Upward Entailment is not limited to base-generated material, but extends to cases of movement to the embedded SpecCP as well. To see this, consider a slightly more complex situation that contains the two probes in (268). In addition to the probe $[u\alpha]$, (268) contains a second probe $[u\bar{A}]$, which does induce movement and is located on C^0 . For the sake of the argument, suppose that $[u\bar{A}]$ has no horizon.

(268) *Hypothetical horizons*

- a. $[u\alpha]_{T^0} \dashv\!\! \dashv T$
- b. $[u\bar{A}]_{C^0} \dashv\!\! \dashv \emptyset$

TP clauses are opaque to $[u\alpha]$, as in (266) above. In the case of CP clauses, the probe $[u\bar{A}]$ on C^0 is able to move a DP inside this lower clause to the specifier of the embedded CP, as TP is not a horizon for this probe. This DP will then be visible to the matrix $[u\alpha]$ -probe, because it is not until the embedded TP is reached that $[u\alpha]$'s search terminates. Elements in the embedded SpecCP are therefore visible to this probe and Agree can be established.

(269) *Transparency of CP clauses under (265) (to be revised)*



3.3.2.2 Proposal: Category inheritance

There are a number of ways of addressing this pathological prediction of the horizons proposal in its current form. The most direct, and least interesting, way of doing so is to impose a designated stipulation on horizons. This line of attack would state, as a matter of stipulation, that if a given category is a horizon for a probe, then higher projections in the clausal spine are also horizons for this probe. This account would hence impose an implicational restriction on the distribution of horizons across categories. Applied to the example just discussed, this principle would require that $[u\alpha] \dashv\vdash C$, because $[u\alpha] \dashv\vdash T$ and because C is higher in the clausal spine than T . In other words, this principle would make sure that every probe that has T as its horizon also has C as a horizon. With this in place, the pathological patterns in (267) and (269) would disappear because $[u\alpha]$'s search would terminate at the embedded CP node in (267) and (269). A DP at the edge of this CP would hence be invisible and Agree between $[u\alpha]$ and DP would be ruled out in (266), (267), and (269), as desired. This would ensure that Upward Entailment is never violated.

While this line of approach is entirely feasible and derives the correct empirical results, I would like to explore a somewhat different option here. Rather than imposing a designated implicational constraint on the distribution of horizons, I will pursue the hypothesis that horizons are themselves unconstrained, that is, that in principle any probe can be associated with any horizon. All generalizations over selective opacity patterns should, I submit, *emerge* from the interplay of unconstrained horizons and independently motivated properties of the syntactic structure. How,

then, can Upward Entailment be captured in a system in which horizons are freely distributed and ruling out the problematic situation in (265) by invoking a designated constraint on horizons is not an option?

To develop an account of Upward Entailment, it is instructive to consider the notion of **EXTENDED PROJECTION**, the unit that consists of a lexical projection plus the cloud of functional projections it is embedded under. The term ‘extended projection’ is due to Grimshaw (1991, 2000), also see van Riemsdijk (1998). The underlying idea behind extended projections is that syntactic projections may form a structure that is larger than the immediate projection of a head. Appeal to ‘nominal’ and ‘verbal’ domains in syntax is a standard example as these domains typically comprise a sequence of projections. While the term ‘extended projections’ is commonly used (see, e.g., Bošković 2014 for a recent example), alternative names for tightly related concepts include *functional sequence* or *functional hierarchy* (Cinque 1999, Starke 2001, Ramchand & Svenonius 2014, Müller 2014a), *hierarchy of projections* (Adger 2003, 2013) or *F-structure* (Williams 2009, 2011, 2013). I will use the traditional term ‘extended projections’ here, though this is merely a terminological choice.

Van Riemsdijk (1988, 1998) and Grimshaw (1991, 2000), among others, have drawn attention to the fact that extended projections form a unit in a way that standard X-bar theory or bare phrase structure does not readily capture. Van Riemsdijk (1988, 1998) in particular argues that extended projections are *endocentric*. Endocentricity is a well-established fact about syntactic projections, enshrined in both X-bar theory and bare phrase structure. Endocentricity amounts to the claim that an NP needs to contain an N, a VP needs to contain a V, and so on. With the proliferation of functional projections, nominal and verbal structures have become increasingly ramified and broken up into a multitude of endocentric projections, but the nominal and verbal domains that these projections form are not endocentric as far as their phrase-structural representation is concerned. For the purposes of X-bar theory or base phrase structure, there is no substantive difference between a verb taking a DP as an argument, and a *v* taking a VP as an argument. In both cases, the category of the resulting constituent is exclusively determined by the head (*v*⁰ or V⁰). The information about the complement is overwritten. But as van Riemsdijk (1988, 1998) and also Williams (2009) emphasize, this view misses a fundamental asymmetry between these two cases. Embedding a VP under a *v* yields another ‘verbal’ projection (a *v*P), whereas embedding a DP under a verb does not

yield a ‘nominal’ projection. That is, combining a v^0 and a VP preserves the verbiness of VP, whereas combining a V^0 and a DP does not preserve the DP’s nouniness. VPs, v Ps, TPs, CP, and so on are all part of the ‘verbal’ domain, whereas NPs, n Ps, DPs, and so on are all part of the ‘nominal’ domain. In this sense, all of these projections are endocentric: They preserve the ‘verbiness’ or ‘nouniness’ of the lexical head whose spine they are in. This endocentricity of extended projections is quite analogous to endocentricity in an immediate projection. Just like an NP entails the presence of an N inside it, so a CP or TP entails the presence of a V inside it. Within standard X-bar theory or bare phrase structure, there is no systematic way of expressing this relationship. It is impossible to state that a v P retains the ‘verbiness’ of its complement because endocentricity holds in these models only between the v P and its v^0 head. Similar remarks apply to the other functional projections just mentioned. It is this endocentricity that underlies the widespread intuition that there is a nominal and a verbal spine, that there are clauses, that there is restructuring, etc. Extended projections are thus projections beyond the immediate projection of a head that belong to the same domain in this sense.

I will explore here the view that Upward Entailment arises from the general fact that extended projections are endocentric. Grimshaw (1991, 2000) and Shlonsky (2006) have presented evidence that features can be percolated up through an extended projection. The arguments come from selection. Grimshaw (1991, 2000) observes that selection requires access to information that is not necessarily present on the highest member of an extended projection. She gives the example of subjunctive selection. The verb *request* requires a subjunctive complement clause, while a verb like *think* is incompatible with it.

- (270) a. We requested that he leave/?left at 6.
 b. We thought that he left/*leave at 6. (Grimshaw 2000: 130)

Crucially, both types of complement clauses are headed by the complementizer *that*. Grimshaw thus concludes that the relevant distinction is not present on the C^0 head itself because this C^0 is the same in both cases. This raises the question of how the matrix verb can select for the property of an embedded clause that is encoded more deeply than the highest projection of that clause. Grimshaw suggests that the features of a head percolate up within an extended projection. Because the TP forms an extended projection with the CP, information about the T^0 head will percolate

to the CP level and can there be selected for by the matrix predicate. Shlonsky (2006) provides analogous observations from French, Italian, and Hebrew and furthermore emphasizes the nature of this problem for cartographic approaches to the left periphery.

- (271) a. Jean pense que Marie dort. [French]
 Jean thinks that Maria sleeps.IND
- b. Jean veut que Marie dorme.
 Jean wants that Marie sleeps.SUBJ (Shlonsky 2006: (1))

Considerations of endocentricity and selection thus converge on the conclusion that information can percolate up through an extended projection. I propose that Upward Entailment is a consequence of this general percolation of information within extended projections. Following the spirit of Grimshaw's (1991, 2000) and Shlonsky's (2006) proposal, I propose that category information percolates up as well. This is formulated in the recursive formulation in (272). It states that within an extended projection (e.g., $\langle CP > TP > vP > VP \rangle$), at least the categorial features percolate up. We will momentarily turn to a labeling algorithm that has (272) as its result.⁸

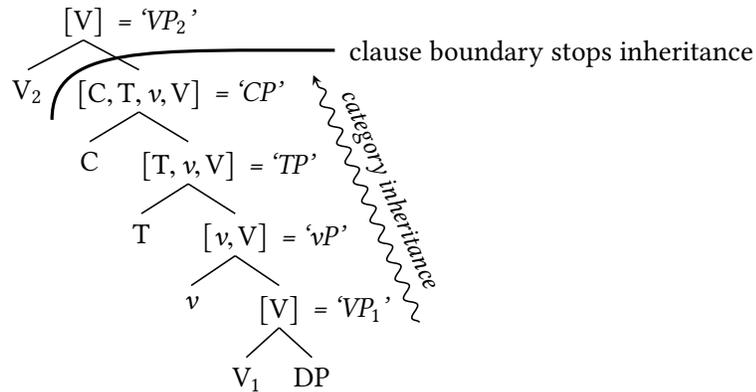
(272) **CATEGORY INHERITANCE**

Given an extended projection $\Phi = \langle \Pi_n > \Pi_{n-1} > \dots > \Pi_1 \rangle$, where Π_x 's are all phrases, the categorial features of Π_m are inherited up to Π_{m+1} .

According to (272), if a head takes a complement that is part of the same extended projection, the category of the resulting constituent is a function of both the category of the head and its complement. The category of complex expression is hence determine *bilaterally* in this case. Across extended projections, on the other hand, only the category of the head projects. This is illustrated in (273).

⁸ Alternatives to (272) are of course conceivable. One alternative line of account more in keeping with van Riemsdijk's (1988, 1998) and Grimshaw's (1991, 2000) original proposals is as follows. Grimshaw (1991, 2000) proposes that all elements of an extended projection share the category feature of the lexical head. In current terminology, *v*, *T* and *C* would all be of the lexical category [verb]. They are distinguished by a functional feature ($\{F0\}$, $\{F1\}$, $\{F2\}$ and so on), which specifies the level of this head in the extended projection. *T*, for instance, would have the dual specification of being a [verb] and $\{F2\}$ (see van Riemsdijk 1988, 1998 for a related proposal). It is then possible to state inheritance in terms of level features. If level features are cumulative rather than substitutive, category inheritance can be redefined as level inheritance. If these levels are then taken to constitute horizons, Upward Entailment may be captured in a way analogous to the account in the main text, but without recourse to category inheritance.

(273) Schematic example of (272)



Thus, within the extended projection spanning V_1 up to the CP, category is cumulative: Category features percolate up beyond the immediate projection of the head by (272), following and extending Grimshaw's (1991, 2000) proposal. Importantly, category inheritance applies only to heads within a single extended projection. Returning to the schematized structure in (273), merging v and VP_1 ($[V]$) creates a complex mother label $[v, V]$, which reflects the categories of both of its daughters, as v and V are part of the same extended projection. The same holds for Merge of T and vP , which creates the complex label $[T, v, V]$, and so forth. On the other hand, V_2 and CP are not part of the same extended projection. Thus, at this juncture point between extended projections, the category of the mother node is 'reset' and unilaterally determined by only V_2 , thus creating VP_2 ($[V]$), another VP , setting into motion another cycle of bilateral labeling.

We may ask how category inheritance relates to label projection in more familiar settings. I will adopt here the null assumption that they are not meaningfully different. This is, the features that percolate up within an extended projection are all and only the features projected up in an immediate projection of a head. On the standard view that the category feature of a head is projected onto the phrase headed by this head, category percolation within extended projections follows immediately.

In Chomsky's (1995a) classical conception of bare phrase structure, the label of the output of a Merge operation is defined as the label of either of the two elements that have been merged.

(274) $\text{Merge}(\alpha, \beta) = \{k, \{\alpha, \beta\}\}$, where k = the label of either α or β (Chomsky 1995a)

Questions emerge on how to determine the projecting element, but those are irrelevant here.⁹ We may conceptualize category inheritance as nothing more than a variation in the definition of the label in (274). Rather than picking out the label of either α or β , category inheritance amounts to *bilateral labeling*. Category inheritance follows if both elements being merged contribute their respective label. The label of the resulting node is the union of the labels of the elements being merged. As defined in (272) and highlighted in (273), such bilateral labeling must be restricted to Merge within an extended projection, with unilateral labeling being drawn into service across extended projections. This is defined in (275):¹⁰

(275) **LABELING**

Let α be a head, whose label is λ_α , and let β a possibly complex syntactic object, whose label is λ_β .

a. If α is not part of the same extended projection as β , then

$$\text{Merge}(\alpha, \beta) = \{[\lambda_\alpha], \{\alpha, \beta\}\}$$

b. If α is part of the same extended projection as β , then

$$\text{Merge}(\alpha, \beta) = \{[\lambda_\alpha, \lambda_\beta], \{\alpha, \beta\}\}$$

The definition of Merge in (275) has as a consequence that labeling is bilateral and cumulative within an extended projection. It is thus one conceivable way of implementing category inheritance (272).

The usual questions as to the nature and content of the label that is projected arises. It is standardly assumed that the label of a complex syntactic object comprises at least the syntactic category of the head (a *verb* projects a *verb* phrase). I take it that this is likewise the case for bilateral labeling, though the label may of course contain more fine-grained information as well. As mentioned above, I will thus tentatively assume that the features that make up a label are the same

⁹ See Chomsky (1995a) for discussion and Moro (2000), Ott (2012, 2015), and Chomsky (2013) for recent proposals addressing this question.

¹⁰ Interestingly, Chomsky (1995a: 397) briefly considers and rejects bilateral labeling, presumably on the basis of examples like *cold water*, which behaves like a noun, not like an adjective. This observation is compatible with (275) as adjectives and nouns head distinct extended projections and labeling will hence be unilateral (275a).

for bilateral labeling as they are for unilateral labeling. For the goal of deriving Upward Entailment, it will be sufficient that the label contains the category features and I will hence put the question aside here.

On the view just laid out, the endocentricity of an extended projection is achieved in exactly the same way as endocentricity within an immediate projection: by projection of a label. The ‘verbi-ness’ of CPs and the ‘nouniness’ of DPs then follows straightforwardly on this view. CPs quite literally contain a V specification as part of their complex label [C, T, ν , V].

How does the proposal of category inheritance in (272) relate to Grimshaw’s (1991, 2000) and Shlonsky’s (2006) original observation? On the face of it, the selection facts would appear to require more than just the category feature to percolate because two flavors of the same head (indicative vs. subjunctive T) would have to be distinguished. There are at least two options of incorporating this subjunctive/indicative distinction into the notion of category inheritance. First, percolation within an extended projection could simply not be confined to category features, but also comprise features that adequately distinguish indicative and subjunctive T. Note that this does not entail that all features percolate. There is, for instance, no selection of a clause based on whether that clause contains an agent-introducing ν or not, so this information must not be percolated. While I have no principled proposal to offer, assumptions about inheritance are not arbitrary. If the features that are inherited through an extended projection are the same as those projected within the immediate projection of a head, a claim about one makes testable predictions about the other. For example, the claim that the distinction between agent-introducing and defective ν is not inherited through an extended projection and available to a higher verb for selection immediately predicts that a T^0 head is not able to locally select for one flavor of ν or another. That is, there should be no tenses that exclusively occur with or without external arguments. This seems entirely correct.

An alternative path to pursue is that subjunctives contain less functional structure than indicatives, but that the missing structure is not removed from the top of the clausal spine, but from the middle, creating what Lisa Travis (p.c.) has fittingly called a ‘donut.’ On this view, indicatives contain a layer of clause structure lower than C^0 that indicatives lack. Call this projection IndP. The greater morphological paucity of the subjunctive is certainly consistent with this purported structural asymmetry. Inheritance of category features would then be sufficient to encode the

presence or absence of IndP on the topmost CP node of the clause. Verbs could then locally selective for CPs that either contain or lack the category feature Ind.

For the analysis in this dissertation, it is inconsequential which of these two views is adopted, as I will be concerned here with horizons, which are solely based on category features, and clause structures created by pruning at the top. What is crucial for the purposes of selective opacity is that only category features are inherited, but of course this view precludes neither the inheritance of additional features nor the existence of clauses with a ‘donut’ structure. All that is crucial is that the label of a clause contains the label of the projections that make up its spine.

3.3.2.3 The Horizon Inheritance Theorem

Assuming this general view of extended projection and their endocentric properties, we are now in a position to address Upward Entailment (264), the generalization that structurally larger clauses are opaque to an operation if structurally smaller ones are. Recall from the discussion surrounding (267) and (269) that horizons in and of themselves would allow violations of Upward Entailment. Once category inheritance within extended projections is acknowledged, however, this pathology disappears and Upward Entailment emerges as a consequence of extended projections.

The crucial consequence of category inheritance (272) is that labels are cumulative within an extended projection. A TP for instance has the complex label $[T, v, V]$. The label of a CP ($[C, T, v, V]$) properly includes the label of a TP. Because category labels are horizons, projection of category labels also projects horizons. This consequence is stated in (276).

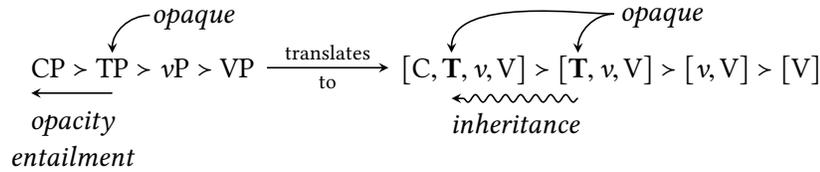
(276) **HORIZON INHERITANCE THEOREM**

Given a probe $[uF]$ and an extended projection $\Phi = \langle \Pi_n > \Pi_{n-1} > \dots > \Pi_1 \rangle$, if $\Pi_m \in \Phi$ is a horizon for $[uF]$, then all projection Π_{m+1}, \dots, Π_n are likewise horizons for $[uF]$ (due to category inheritance (272)).

Crucially, (276) is not a stipulation of the system, but a theorem that emerges from it. To understand (276), it is helpful to consider a concrete example. Consider a probe that has T as its horizon and will therefore not be able to probe into a TP (as schematized in (266)). Upward Entailment states the empirical generalization that such a probe will also be unable to probe into a CP clause. This entailment now follows from category inheritance. By category inheritance, the label of CP clause is $[C, T, v, V]$ and therefore contains a T specification. If a probe has T as a horizon, this T specification

will trigger termination of this probe's search. As a result, CPs will induce probe termination just like TPs do and hence be opaque to that probe. This is schematized in (277).

(277) *Upward Entailment follows from category inheritance*



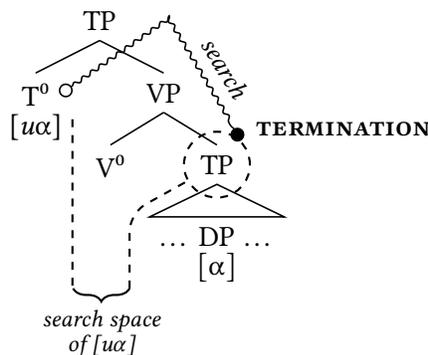
Against the background of category inheritance and the Horizon Inheritance Theorem (276) it gives rise to, consider again the schematic example considered in (265) above, repeated here for convenience:

(278) *Hypothetical horizons* [=(265)]

- a. $[u\alpha]_{T^0} \dashv\vdash T$
- b. $[u\bar{A}]_{C^0} \dashv\vdash \emptyset$

In a structure in which the embedded clause is a bare TP, nothing changes from above: The TP node induces search termination. As a result, the embedded clause is opaque to the probe $[u\alpha]$.

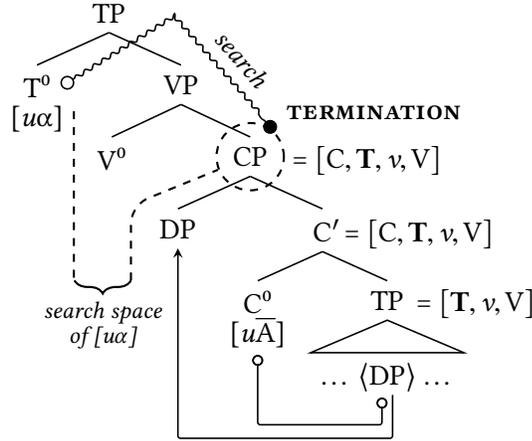
(279) *Opacity of TP clauses under (278)*



The situation differs, however, with respect to CP clauses. We saw in (267) and (269) above that in the absence of category inheritance, the edge of a CP clause would be visible to $[u\alpha]$, an undesirable result. With category inheritance in place, however, CP contains a T as part of its label. It will hence induce termination of an $[u\alpha]$ -induced search just like a TP would. As a result, DP within CP will

be out of $[u\alpha]$'s reach regardless of whether it undergoes movement to the edge of the lower clause or not.

(280) *Opacity of CP clauses under (278) with category inheritance*



If T is a horizon for a probe, this probe will be blocked not only by TPs, but also by CPs, as a matter of principle. v P clauses, on the other hand, would be transparent to this probe as they do not bear a T feature. This derives Upward Entailment. Addition of functional structure never leads to transparency.

This outcome is completely general, of course. If, for example, a probe has v as its horizon, then TP clauses as well as CP clauses will be inaccessible to that probe because both TP and CP nodes contain a v category feature as part of their label. Consequently, if v P clauses are inaccessible to a probe, then TP clauses and CP clauses will be as well.

Upward Entailment can thus be derived from independently motivated properties of extended projections. Importantly, Upward Entailment has now been derived without imposing any constraint on the distribution of horizons. As far as the system just laid out is concerned, any probe can be associated with any horizon. Due to category inheritance, however, no assignment of horizons will end up violating Upward Entailment: Even a probe like $[u\alpha]$ in (278), for which T is a horizon but C is not, will be terminated by both CPs and TPs. We have thus derived a descriptive constraint on selective opacity without a stipulated constraint on the theoretical distribution of horizons. In this sense, Upward Entailment is an emergent property.

3.4 Accounting for selective opacity

In this section, I will apply the concept of horizons to the selective opacity facts of Hindi motivated in chapter 2, to familiar superraising configurations, and to instances of locality differences between movement and agreement.

3.4.1 Selective opacity in Hindi

The distribution of the four operations considered there over the three clause sizes is repeated in (281).

(281) *Selective opacity across Hindi clauses* [=(239)]

<i>Operation</i>	<i>Landing site/ probe location</i>	<i>Size of embedded clause</i>		
		CP	TP	vP
\bar{A} -movement	C	✓	✓	✓
<i>wh</i> -licensing	C	✗	✓	✓
A-movement	T	✗	✗	✓
ϕ -agreement	T	✗	✗	✓

(✓: transparent; ✗: opaque)

One of the key conclusions of chapter 2 is that selective opacity is not limited to movement, but also encompasses in-situ relations like ϕ -agreement and *wh*-licensing. Moreover, the selective locality profile of a movement dependency can be identical to that of an in-situ dependency, as is the case for A-movement and ϕ -agreement in Hindi. I have argued that this provides strong evidence that selective opacity is not a property of movement, but of syntactic operations more generally and that any account that states selective opacity solely in terms of movement is unable to capture the full pattern in (281).

Horizons offer a systematic account of (281) because the locality restriction is stated at the level of the Agree operation. Assuming, as is standard, that both movement and in-situ relations are dependent on the successful establishment of an Agree dependency, horizons provide a unified account of the complex pattern in (281).

The four probes of interest and their respective horizons are listed in (282). The subscript next to the probe feature indicates the location of the probe and is taken directly from the table in (281). Independent evidence for these probe locations was provided in section 2.5.2 of chapter 2.

(282) *Hindi probes and their horizons*

- | | | |
|----|--|--------------------------|
| a. | $[uA]_{T^0} \dashv\!\! \dashv T$ | → terminates at TP or CP |
| b. | $[u\phi]_{T^0} \dashv\!\! \dashv T$ | → terminates at TP or CP |
| c. | $[uwh]_{C^0} \dashv\!\! \dashv C$ | → terminates at CP |
| d. | $[u\bar{A}]_{C^0} \dashv\!\! \dashv \emptyset$ | → no horizons |

The A-movement probe $[uA]$ is located on T^0 and has T as its horizon. This entails that its search terminates at TP and CP nodes, because both contain a T specification as part of their label (by category inheritance (272)). Note that a single specification of $[uA]$'s horizons ($[uA] \dashv\!\! \dashv T$) has the effect that two nodes (TP and CP) induce search termination, a consequence of the Horizon Inheritance Theorem (276) that is derived from category inheritance, as discussed in the previous section.

The ϕ -probe $[u\phi]$ is likewise located on T^0 and has T as its horizon. The *wh*-probe $[uwh]$ and the \bar{A} -probe $[u\bar{A}]$ are both located on C^0 , but differ in their horizons. $[uwh]$ has C as its Horizon, thus terminating at CPs, whereas $[u\bar{A}]$ has no horizon whatsoever.

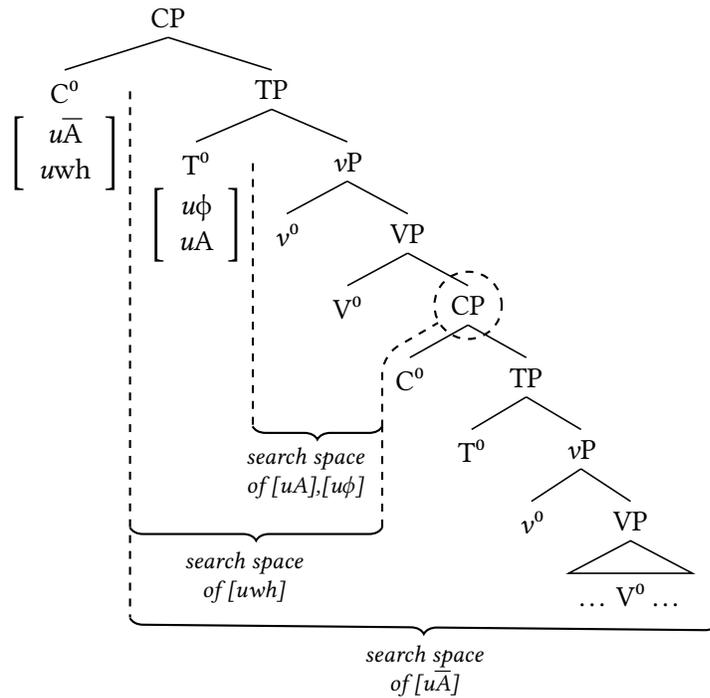
I will continue to use the terms ‘A-movement’ and ‘ \bar{A} -movement’ as convenient labels for movement triggered by $[uA]$ and $[u\bar{A}]$, respectively, due to their differential behavior with respect to weak crossover, reciprocal binding, etc. (see chapter 2). I will have nothing to say about why movement triggered by $[uA]$ can feed binding of pronouns and reciprocals, whereas movement triggered by $[u\bar{A}]$ cannot. I will be purely concerned with their landing sites, probe features, and locality properties here. As discussed in chapter 1, I will adopt the widely held assumption that their interpretive properties have a source separate from their locality.

With (282) in place, let us now consider the emerging opacity profiles of the three clause types of interest.

3.4.1.1 Finite (CP) clauses

(283) gives the schematic structure for an embedded CP clause with respect to horizons.¹¹ Only the clausal structure is shown, specifiers and other elements are suppressed for readability.

(283) *Finite clause (CP) embedding in Hindi*



As shown, a search initiated by $[uA]$, $[u\phi]$ or $[uwh]$ terminates at the embedded CP level. Consequently, anything embedded inside this CP is invisible to these probes, including elements at the CP edge. Therefore, CP clauses block A-movement out of them and ϕ -agreement and *wh*-licensing into them. Only $[u\bar{A}]$ is not terminated by the embedded CP node, in virtue of having no horizon. It is therefore able to search past this node and agree with elements inside the CP clause. \bar{A} -movement out of CP clauses is hence possible.¹²

¹¹ Of course, there may be other locality constraints in addition to horizons that impose further restrictions on the search spaces in (283). I will indeed argue in chapter 6 that horizons coexist with CP phases. Incorporating CP phases into (283) would amount to limiting $[u\bar{A}]$'s search space to the specifier of the lower CP but it would not affect the search spaces of $[uA]$, $[u\phi]$, and $[uwh]$. Thus, an element in the embedded SpecCP will be accessible only to $[u\bar{A}]$, but not to any of the other three probes due to their differing horizons. Everything I am going to say here is compatible with this narrower search space as well. The crucial contrast is whether a probe is able to access the edge of the lower clause, as determined by horizons. Because I focus on horizons here, I will not indicate the effects of phases and return to them in chapter 6.

¹² See section 2.4.4 in chapter 2 for some remarks about finite clauses in Bangla. There I present evidence that finite clauses in Bangla come in different types, a distinction that manifests itself in the choice of complementizer and that affects the possibility of *wh*-licensing across them. If this distinction maps unto differences in clause size, these

This derives the facts pertaining to finite clauses in chapter 2. First, finite clauses allow \bar{A} -movement out of them, but block A-extraction:

- (284) a. *Finite clauses allow \bar{A} -extraction* [= (76)]
 har laṛke-ko_i [us-kii_j bahin] soc-tii hai [ki raam-ne t_i
 every boy-ACC 3SG-GEN sister think-IPFV.F.SG be.PRES.3SG that Ram-ERG
 dekh-aa]
 see-PFV.M.SG
 ‘His/her_j sister thinks that Ram saw every boy_i.’
- b. *Finite clauses block A-extraction*
 *har laṛke-ko_i [us-kii_i bahin] soc-tii hai [ki raam-ne t_i
 every boy-ACC 3SG-GEN sister think-IPFV.M.SG be.PRES.3SG that Ram-ERG
 dekh-aa]
 see-PFV.M.SG
Intended: ‘For every boy x , x ’s sister thinks that Ram saw x .’

Furthermore, finite clauses are opaque to ϕ -agreement and *wh*-licensing, regardless of whether the goal DP is located at the edge of the clause or is more deeply embedded:

- (285) *Finite clauses block ϕ -agreement into them* [= (102), (103)]
- a. laṛkō-ne soc-aa/*-ii [ki monaa-ne ghazal gaa-yii
 boys-ERG think-PFV.M.SG/*-PFV.F.SG that Mona-ERG ghazal.F sing-PFV.F.SG
 thii]
 be.PST.F.SG
- b. laṛkō-ne soc-aa/*-ii [ghazal_i monaa-ne t_i gaa-yii
 boys-ERG think-PFV.M.SG/*-PFV.F.SG ghazal.F Mona-ERG sing-PFV.F.SG
 thii]
 be.PST.F.SG
 ‘The boys thought that Mona had sung ghazal.’

observations are compatible with the horizons account developed here, but require more fine-grained structural distinctions in the CP domain à la Rizzi (1997).

(286) *Finite clauses block wh-licensing into them* [= (174), (175)]

- a. *siitaa-ne soc-aa [ki ravii-ne kis-ko dekh-aa]
 Sita-ERG think-PFV.M.SG that Ram-ERG who-ACC see-PFV.M.SG
Intended: ‘Who did Sita think that Ravi saw?’
- b. *siitaa-ne soc-aa [kis-ko_i ravii-ne t_i dekh-aa]
 Sita-ERG think-PFV.M.SG who-ACC Ram-ERG see-PFV.M.SG
Intended: ‘Who did Sita think that Ravi saw?’

As discussed above, the fact that it is irrelevant whether an element is located at the edge of the clause or not in (285) and (286) follows because the edge of a clause does not have a special role as far as horizons are concerned. Because the search of $[uwh]$, $[uA]$, and $[u\phi]$ terminates at the CP node of the embedded clause, everything dominated by that node is invisible to these probes, including elements in SpecCP.

The locality profiles of these four operations thus follows straightforwardly. Possible and impossible feeding relations between operations likewise follow from the distribution in (282). Consider first the relation between A-movement and \bar{A} -movement. Because \bar{A} -movement lands in SpecCP and A-movement is triggered by a probe on T^0 , as demonstrated in section 2.5.2 of chapter 2.2, it immediately follows that \bar{A} -movement into the matrix clause cannot be followed by A-movement within the matrix clause, as the higher SpecCP falls outside the c-command domain of the T^0 head and therefore \bar{A} -movement lands too high to be followed by A-movement in the same clause. As a general consequence, \bar{A} -movement of an element can *never* be followed by A-movement because an \bar{A} -movement element in SpecCP is beyond the horizon of an A-probe in a higher clause and too high for an A-probe in the same clause. This derives the ban on superraising in Hindi. No designated stipulation that renders \bar{A} -positions invisible to the A-movement probe is required. Rather, the impossibility of \bar{A} -movement feeding A-movement is an indirect consequence of the distribution of landing sites and horizons. The inverse – A-movement followed by \bar{A} -movement – is of course possible.

Identical facts hold for ϕ -agreement, as $[u\phi]$ is also located on T^0 . ϕ -Agreement into a finite clause is impossible because C constitutes this probe’s horizon (see (283)). Furthermore, an element \bar{A} -moved into the matrix clause lands in SpecCP and hence outside of $[u\phi]$ ’s search space, and agreement with it is likewise impossible. As we saw in section 2.5.2 of chapter 2, this is correct:

(287) \bar{A} -movement does not feed ϕ -agreement [=(237)]

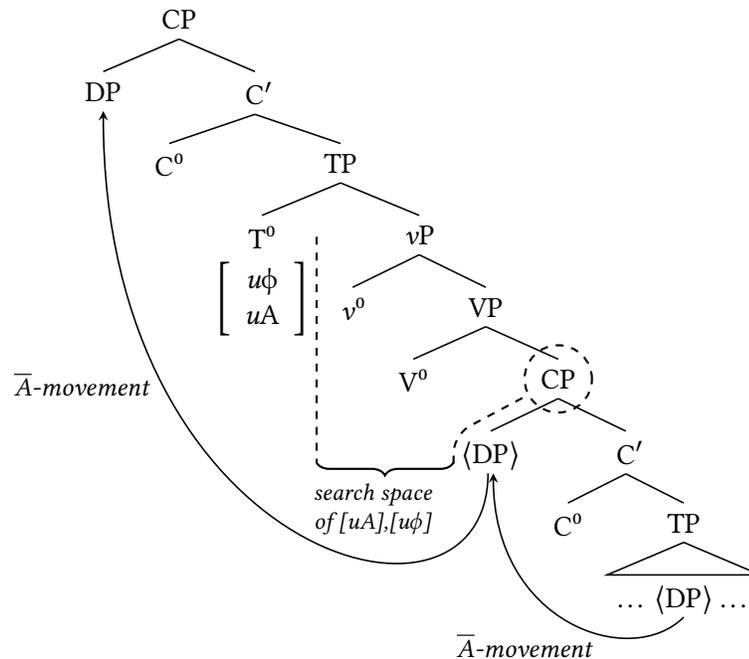
ghazal_i firoz-ne soc-aa/*-ii [ki monaa-ne t_i gaa-yii
 ghazal.F Firoz-ERG think-PFV.M.SG/*-PFV.F.SG that Mona-ERG sing-PFV.F.SG
 thii]
 be.PST.F.SG

‘Firoz thought that Mona had sung ghazal.’

In (287), *ghazal* is moved out of a finite clause. Because only $[u\bar{A}]$ is able to search into a CP clause, *ghazal* must land in SpecCP of the matrix clause. This position is not c-commanded by the matrix ϕ -probe and agreement with it is consequently ruled out.

ϕ -Agreement with an \bar{A} -position as well as A-movement from an \bar{A} -position are therefore ruled out without recourse to \bar{A} -positions per se. Rather, because \bar{A} -movement lands in SpecCP, its application entails the presence of a CP structure in that clause. This CP structure will shield any element inside from agreeing with $[u\phi]$ or $[uA]$ in a higher clause. Moreover, the landing site of \bar{A} -movement falls outside the c-command-based search space of probes on T^0 in the same clause. The situation is schematized in (288).

(288) *The invisibility of \bar{A} -positions to A-movement and ϕ -agreement*



The \bar{A} -position in the lower clause is beyond the horizon of $[uA]$ and $[u\phi]$ and hence out of their reach. The \bar{A} -position in the higher clause is outside the c-command domain of these two probes

and hence likewise invisible to them. Notice that it is crucial for this account that an element \bar{A} -moved out of a lower CP clause does not land in Spec ν P of the higher clause, as this position would fall into $[uA]$'s and $[u\phi]$'s search domain and hence be visible to them. I do indeed take this to be an argument against the view that ν is a phase head. See chapter 6 for extensive discussion.

We have thus derived the fact that \bar{A} -movement can never feed ϕ -agreement or A-movement. Importantly, this restriction is not attributed to a designated constraint on admissible derivational sequences of operations. Instead, it follows from $[u\phi]$'s and $[uA]$ horizon and their location relative to the landing site of \bar{A} -movement.¹³

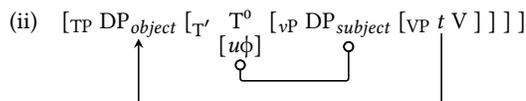
An important consequence of the horizons account is that it blocks two different derivations of A-movement out of a finite clause like (284b) in a uniform way: The first derivation involves one-fell-swoop movement, where the element moves directly from its base position into the matrix clause, shown in (289a). The second derivation is makes use of successive-cyclic successive-cyclic movement, such that the element first undergoes \bar{A} -movement to the edge of the lower clause, followed by A-movement from there, schematized in (289b). While the two derivations are standardly excluded by separate principles (subjacency/phases and the ban on improper movement),

¹³ The account for why \bar{A} -movement cannot feed A-movement and ϕ -agreement in Hindi is similar to the explanation in section 2.5.2 of chapter 2 for why A-movement does not affect ϕ -agreement. We observed there that in a simple transitive clause, A-movement of the object over the subject does not override agreement with the subject. The relevant example (234) is repeated here as (i). Only subject agreement is possible.

(i) *Object A-movement does not override subject agreement*

- a. har gaari_i [us-kaa_i maalik (=hii)] t_i saaf kar-egaa
 every car.F 3SG-GEN owner.M(=only) clean do-FUT.M.SG
 'Every car x will be cleaned by x 's owner (not by anybody else).'
- b. *har gaari_i [us-kaa_i maalik (=hii)] t_i saaf kar-egii
 every car.F 3SG-GEN owner.M(=only) clean do-FUT.F.SG
 Intended: 'Every car x will be cleaned by x 's owner (not by anybody else).'

In chapter 2, I take this fact as evidence for the view that A-movement lands in SpecTP, whereas the ϕ -probe is located on T⁰. A-movement of the object therefore lands too high to be visible to the ϕ -probe. Only the launching site of the object is visible to $[u\phi]$ and subject agreement arises as a minimality effect.



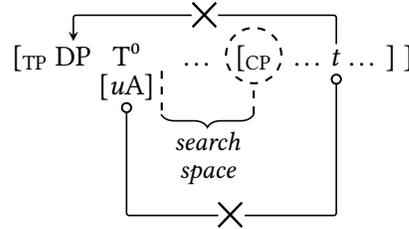
See fn. 54 on p. 152 for some qualifications compatible with this account.

Just like in the structure (288), a crucial component of this explanation is that the moving element does not and cannot stop over in Spec ν P but directly moves to SpecTP or SpecCP, respectively. In other words, ν P cannot be a phase. This is indeed the conclusion I reach on independent grounds in chapters 6 and 7.

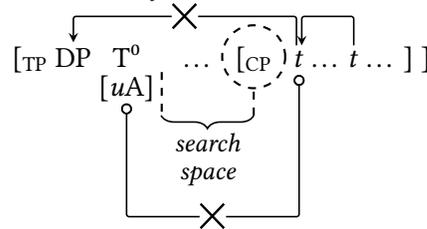
the horizons account provides a unified account that applies to both because the $[uA]$ -probe cannot search in a CP clause at all. It is entirely irrelevant where the element is located inside that CP. I will return to this point in section 3.4.2, where I argue more generally that the horizons account captures generalizations that the standard account misses.

(289) *Two excluded derivations for superraising*

a. *One-fell-swoop movement:*



b. *Successive-cyclic movement:*



While the analysis thus rules out derivations in which \bar{A} -movement of an element is followed by A-movement of the same element, it allows the inverse ordering, i.e., A-movement feeding \bar{A} -movement. In (290), the embedded object *har laṛke-ko* ‘every boy-ACC’ first undergoes A-movement inside the embedded clause to produce the bound reading of the pronoun *uskii*, followed by \bar{A} -extraction out of the finite clause. Unsurprisingly, the result is grammatical.

(290) har laṛke-ko_i raam-ne soc-aa [t_i us-kii_i bahin-ne t_i dekh-aa]
 every boy-ACC Ram-ERG think-PFV.M.SG 3SG-GEN sister-ERG see-PFV.M.SG
 ‘Ram thinks that for every boy *x*, *x*’s sister saw *x*.’

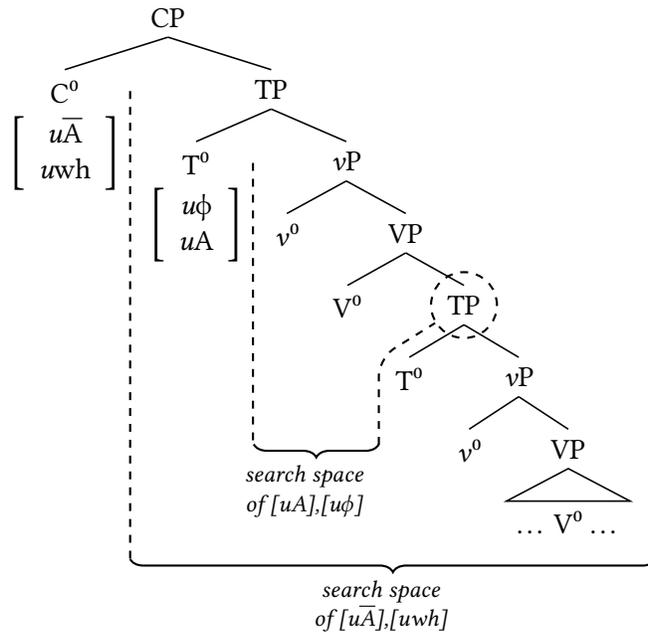
(290) is successfully generated by the horizons system. The analysis developed here thus affords a new view on \bar{A} -invisibility and on the differential ability of movement types to bleed and feed one another. Crucially, both emerge as epiphenomena on this account as there are no direct interactions between either A- and \bar{A} -movement or A- and \bar{A} -positions. \bar{A} -positions in and of themselves are no more or less visible to a probe than A-positions. But because they require the presence of a CP layer

that encapsulates them (at least in English and Hindi), \bar{A} -positions invariably end up outside the search domain of the A-probe, ruling out a feeding relationship. It follows, then, that \bar{A} -movement bleeds A-movement. The inverse does not hold simply because A-movement does not entail the presence of a horizon for the \bar{A} -probe.

3.4.1.2 Large nonfinite (TP) clauses

The locality profile of TP clauses is schematized in (291). Due to the lack of a CP projection, $[uwh]$ is able to probe into the embedded clause as it does not encounter a horizon. $[uA]$ and $[u\phi]$, on the other hand, still terminate at the topmost TP node of the embedded clause. As was the case in (283), the search spaces in (291) is shown only with respect to horizons.

(291) *Large nonfinite clause (TP) embedding in Hindi*

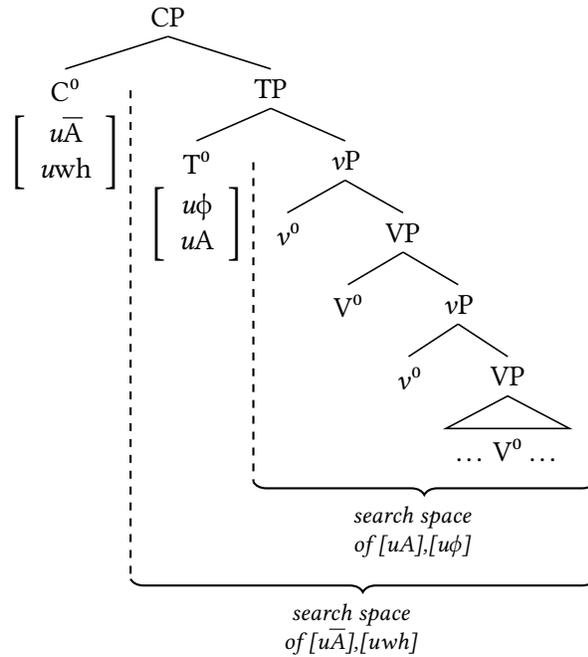


TP clauses differ from CP clauses in that *wh*-licensing into them is possible. Thus, an embedded *wh*-element may take matrix clause. TP clauses are still opaque to A-movement and ϕ -agreement, and LDA into them is therefore impossible.

3.4.1.3 Small nonfinite (vP) clauses

Lastly, the opacity profile of *vP* clauses is given in (292). With the T layer removed as well, the embedded clause is transparent to search from all four probes.

(292) *Small nonfinite clause (vP) embedding in Hindi*



In this structure, the embedded clause allows A- and \bar{A} -movement out of it and ϕ -agreement and *wh*-licensing into it. Furthermore, we saw evidence in section 2.3.2 of chapter 2 that ϕ -agreement is obligatory in Hindi if it is possible, following Preminger’s (2011, 2014) *obligatory operations*.

(293) **Agreement obligatoriness** (Preminger 2011, 2014) [= (123)]
 If a verb can ϕ -agree with a DP, it has to.

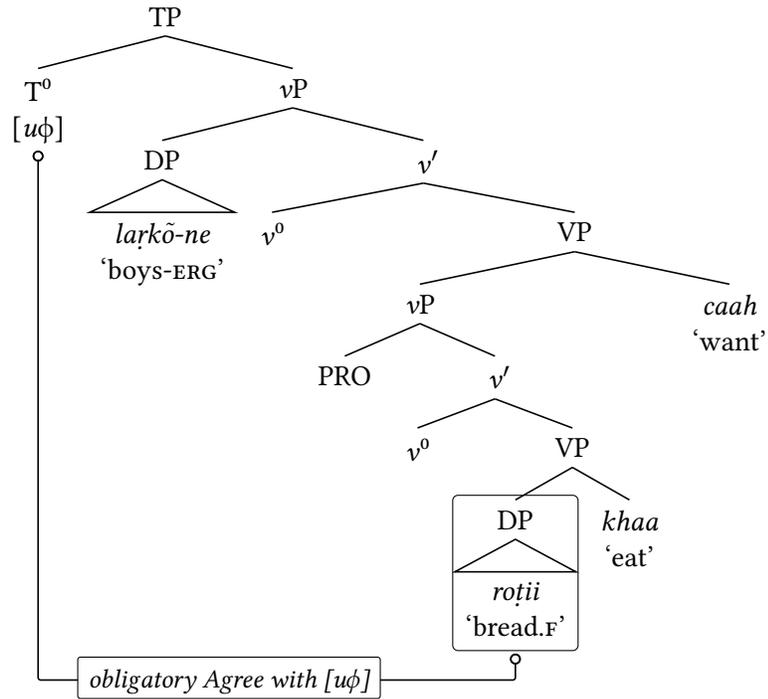
Because an embedded *vP* clause is transparent to the matrix ϕ -probe, long-distance agreement (LDA) is in fact obligatory in (292). Following the reasoning in section 2.3, the surface optionality of LDA arises because TP and *vP* clauses are surface-identical. Thus, the sentence in (294a) exemplifies *vP* embedding as in (295), whereas (294b) instantiates TP embedding (296). In (295), ϕ -agreement is obligatory, where in (296), the TP clause constitutes a horizon for the matrix ϕ -probe and LDA is hence impossible. The result is default agreement.

(294) a. *Long-distance agreement*
 laṛkō-ne [_{vP} roṭii khaa-nii] caah-ii [= (87), (88)]
 boys-ERG bread.F eat-INF.F.SG want-PFV.F.SG
 ‘The boys wanted to eat bread.’

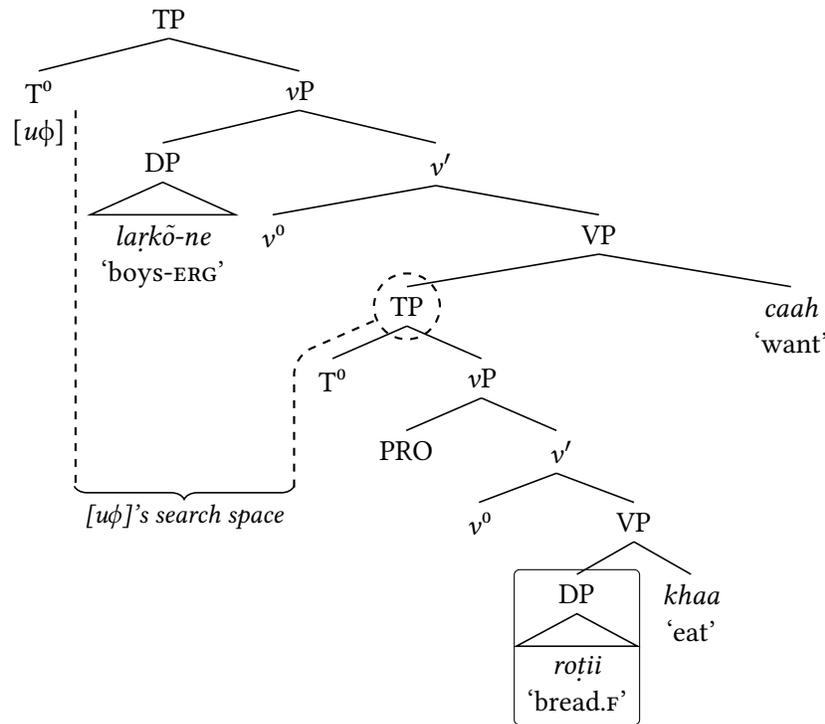
b. *Default agreement*

laṛkō-ne [TP roṭii khaa-aa] caah-aa
 boys-ERG bread.F eat-INF.M.SG want-PFV.M.SG
 ‘The boys wanted to eat bread.’

(295) *Structure of (294a): Long-distance agreement*



(296) Structure of (294b): Default agreement



Because no movement of *rotii* ‘bread’ is necessary to establish LDA in (295), the evidence presented in section 2.2.2 of chapter 2 that no movement of the element controlling LDA is necessary falls out immediately.

The status of the PRO inside the embedded clause in (295) and (296) deserves some comments. Section 2.3.1 of chapter 2 has presented evidence for the view that nonfinite clauses in Hindi contain a vP projection and a PRO subject, irrespective of whether LDA takes place or not. One might wonder then why this PRO subject does not constitute a closer goal for the ϕ -probe in (295) and block agreement across it. One possible explanation for this lack of intervention is that PRO in Hindi simply lacks ϕ -features. On this account, it would simply not constitute a closer goal for the ϕ -probe and hence not interfere with LDA. Evidence from secondary predicates supports this view. Some secondary predicates, like *nangaa* ‘nude’, have to obligatorily agree in ϕ -features with the subject, as in (297a). Crucially, if these predicates modify a PRO subject, no such agreement is possible, as shown in (297b). In this example, *nangaa* has to appear in its masculine singular form, despite the fact that the controller of PRO is feminine.

- (297) a. miinaa nang-ii/*-aa naac-tii hai
 Mina.F nude-F.SG/*-M.SG dance-IPFV.F.SG be.PRES.3SG
 ‘Mina dances nude.’
- b. miinaa [PRO nang-aa/*-ii naac-naa] caah-tii hai
 Mina.F nude-M.SG/*-F.SG dance-INF.M.SG want-IPFV.F.SG be.PRES.3SG
 ‘Mina wants to dance nude.’

The contrast between (297a) and (297b) is accounted for if PRO simply does not contain any ϕ -features. This conclusion solves the intervention problem in (296) above. Lacking ϕ -content, PRO will be invisible for [$u\phi$] and hence not block Agree over it. I therefore take the presence of PRO inside the embedded clause to be unproblematic for the account.

The horizon account also captures without further ado the A-movement–Agreement Generalization established in section 2.3.3 of chapter 2 and repeated here as (298).

- (298) **A-Movement–Agreement Generalization** [=(153)]
 A-movement of *any* element out of a nonfinite clause makes LDA into this clause obligatory.
 \bar{A} -movement has no such effect.

An example of a sentence falling under (298) is repeated in (299). Here the embedded indirect object *har bacce-ko* ‘every child-DAT’ is A-moved out of the embedded clause (diagnosed by the weak crossover configuration). As a consequence of this movement, the embedded direct object *film* ‘movie’ has to obligatorily control LDA.

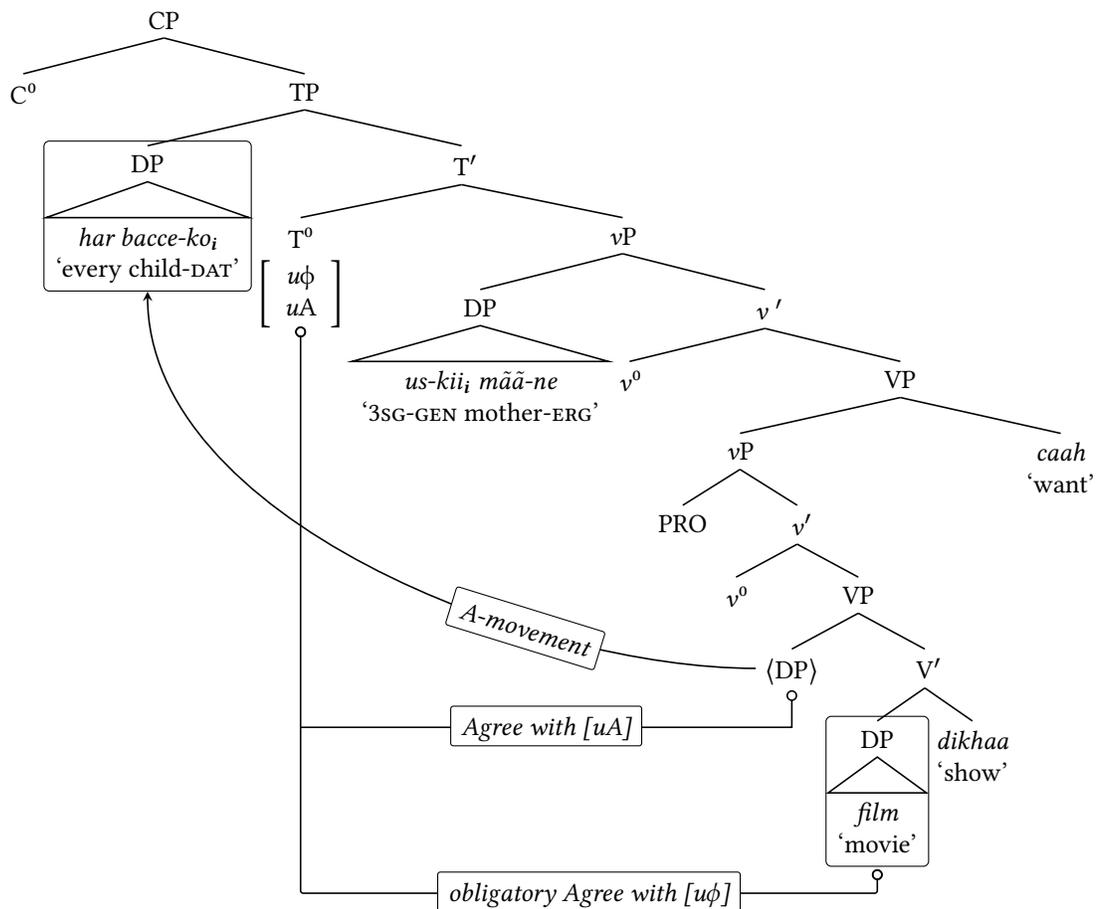
- (299) *A-movement out of nonfinite clause renders this clause obligatorily transparent to LDA*
 har bacce-ko_i [*us-kii_i* m \tilde{a} \tilde{a} -ne] [*t_i* **film** dikhaa-nii/*?-naa]
 every child-DAT 3SG-GEN mother-ERG movie.F show-INF.F.SG/*?-INF.M.SG
 caah-ii/*?-aa [=(135c)]
 want-PFV.F.SG/*?-PFV.M.SG
 ‘For every child x , x ’s mother wanted to show x a movie.’

As discussed in detail in chapter 2, the remarkable feature of (298) is that it is a generalization over the embedded clause. The element that undergoes the A-movement and the one controlling LDA as a result do not have to be the same.

This interaction between LDA and A-movement follows from the system developed here. Because [uA] –|| T, TP clauses are opaque to the A-movement probe. Only vP clauses are transparent.

Consequently, A-extraction out of the lower clause disambiguates this clause towards a vP structure. Because vPs are transparent to $[u\phi]$ and Agree is obligatory by (293), LDA is obligatory in this case. This state of affairs is represented in (300).¹⁴ Notice that T^0 triggers two operations in (300): (i) A-movement of *har bacce-ko* ‘every child-DAT’ and (ii) ϕ -agreement with *film* ‘movie’. Because the A-probe and the ϕ -probe are separate probes, the fact they agree with distinct elements in (300) is unremarkable. Note also that there is no particular order in which the two operations have to apply, simply because they are independent of each other.

(300) Structure of (299)



¹⁴ I have placed the ergative matrix subject in (300) in Spec vP . Because ergative subjects are invisible to verbal ϕ -agreement (recall the algorithm in (84) in chapter 2), it is irrelevant for the purposes of LDA whether this subject resides in SpecTP or Spec vP . In general, ergative subjects exhibit the same behavior as non-overtly case-marked subjects with respect to every syntactic test apart from verbal agreement (Pandharipande & Kachru 1977, Mohanan 1994, Anand & Nevins 2006). There is, in other words, no evidence that ergative DPs occupy a different structural position than non-case-marked ones. Some authors have proposed that this position is SpecTP for both kinds of subjects (Davison 2004a,b, Anand & Nevins 2006), and this view is fully compatible with the account presented here.

The same account applies to the various other instances of the A-movement–agreement generalization (298) presented in section 2.3.3 of chapter 2.

Notice that in order to capture the fact that LDA is obligatory in (299) and its structure (300), ϕ -agreement has to be able to proceed into a ν P. Consequently, ν P cannot be a phase. This is indeed the conclusion I draw in chapters 6 and 7 on the basis of a range of facts.

In contrast to [uA], [$u\bar{A}$] is not blocked by an embedded TP and can hence agree into one. Therefore, \bar{A} -extraction out of nonfinite clause does not disambiguate this clause towards either structure and LDA remains optional on the surface. The crucial fact that A-movement, but not \bar{A} -movement, interacts with LDA in (298) thus follows as a consequence of the horizons of the two probes.

The same is true for *wh*-licensing. As we saw in (178), repeated here, *wh*-licensing into a clause is possible with both LDA and default agreement:

(301) *Wh-licensing is independent of LDA* [=(178)]

a. *Default agreement*

raam-ne [*kaunse baccō-ko film dikhaa-naa*] caah-aa?
 Ram-ERG which child-DAT movie show-INF.M.SG want-PFV.M.SG
 ‘Which child did Ram want to show a movie to?’

b. *LDA*

raam-ne [*kaunse baccō-ko film dikhaa-nii*] caah-ii?
 Ram-ERG which child-DAT movie show-INF.F.SG want-PFV.F.SG
 ‘Which child did Ram want to show a movie to?’

This fact follows because [uwh] is able to probe into both TP and ν P clauses. Unlike A-movement, it does not disambiguate the structure of the nonfinite clause.

Horizons are therefore flexible enough to capture both interactions between operations (such as between A-movement and LDA in the form of the A-movement–agreement generalization) and non-interactions between operations (such as between LDA on the one hand and \bar{A} -movement or *wh*-licensing on the other). Interactions as well as non-interactions crucially are the result of the fact that different probes have differential horizons and therefore locality profiles. That is, interactions and non-interactions are instances of selective opacity and therefore accounted for by horizons.

Because of category inheritance, and hence Upward Entailment, the horizon account makes the immediate prediction that the inverse of the A-movement–agreement generalization should be impossible. That is, while extraction out of a clause can make this clause transparent to another operation like ϕ -agreement, there should be no cases in which extraction out of a clause that normally allows a process like ϕ -agreement into it renders this clause opaque to this process. For example, there should be no language Hindi' in which A-movement out of a nonfinite clause enforces default agreement instead of LDA. The reason is the following: Because of Upward Entailment, movement out of a clause can disambiguate this clause towards a structurally smaller variant, but not towards a structurally larger one. Moreover, if two clause structures differ in their opacity for an operation like ϕ -agreement, it is necessarily the smaller variant that is transparent, once again because of Upward Entailment. Consequently, the hypothetical language Hindi' could exist only if at least one of the two operations violated Upward Entailment. Because Upward Entailment is a consequence of the horizons account, Hindi' would be incompatible with horizons. While interactions of this sort are perfectly conceivable logically, they are unattested to the best of my knowledge. I take this to be additional evidence for Upward Entailment and the horizons account that derives it.

3.4.1.4 Summary

The notion of horizons affords a uniform account of the various layers of selective opacity in the Hindi paradigm in (281). First, because they apply to the operation Agree, the intricate locality profiles of all four operations in (281) can be expressed in one and the same system.

Second, the system allows us to capture interactions between the various operations. Because horizons apply to movement and non-movement dependencies alike, they allow us to express the fact that both A-movement and ϕ -agreement partake in selective opacity in the same way, i.e., with the same locality profile. This overlap in their locality then provides a straightforward account of the A-movement–Agreement Generalization (298). Domain-based interactions like (298) can therefore be captured in a systematic way that is out of the reach of any account of selective opacity that is confined to movement alone.

Third, in addition to interactions between operations, the system also accounts for the lack of interaction between other operations. We have seen, for example, that A-movement interacts

with LDA, but that \bar{A} -movement or *wh*-licensing do not. This discrepancy now falls out as a direct consequence of selective opacity. Because both $[u\bar{A}]$ and $[uwh]$ have different horizons than $[u\phi]$ and $[uA]$, their application does not disambiguate the embedded clause structure in a way that A-movement does. Hence no effect on LDA is observed.

3.4.2 The syntax of superraising and other movement asymmetries

The ban on superraising in English follows from horizons in exactly the same way as it does in Hindi. C is a Horizon for the A-movement probe, but not for \bar{A} -movement probes. For the sake of concreteness we may add the extraposition probe to this collection, whose horizon is (at least) as low as T:

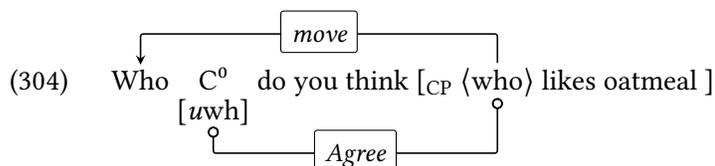
(302) *Horizons for English probes*

- a. $[uwh] \dashv\vdash \emptyset$
- b. $[uA] \dashv\vdash C$
- c. $[u\text{Extr}] \dashv\vdash T$

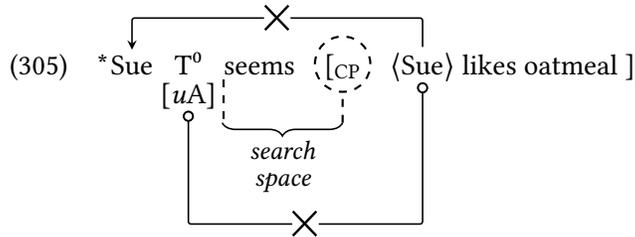
The ban on superraising – A-movement out of a finite clause –, illustrated in (303), then follows as a horizon effect.

- (303) a. Who_i do you think [*t_i* likes oatmeal]?
- b. *Sue_i seems [*t_i* likes oatmeal].

In the case of \bar{A} -movement, the $[uwh]$ -probe can search into the embedded finite clause, enter into an Agree-relation with an embedded *wh*-element and thereby cause it to move to the matrix SpecCP position, allowing \bar{A} -movement out of a finite clause:

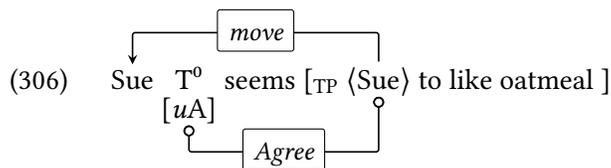


T's EPP-features, here notated as [*uA*], on the other hand, is unable to search into a CP due to its horizon. Movement being parasitic on an Agree-relationship, A-movement out of a CP is therefore ruled out.



As in the discussion of Hindi above, there is no designated constraint against A-moving an \bar{A} -moved element. It is simply the case that \bar{A} -movement of an element inside the lower clause implicates the presence of a CP layer in that clause. This CP layer then constitutes a horizon for the A-probe, blocking A-movement out of it. A consequence of this line of analysis is that A-movement out of a CP is impossible regardless of the launching site of this movement. I will return to this point immediately below.

Once the CP layer of the embedded clause is removed, it becomes transparent to probing by [*uA*] and hence A-movement. A-movement out of nonfinite clauses is therefore correctly allowed.



TP clauses are still opaque to extraposition, as T is [*uExtr*]'s horizon. Extraposition out of nonfinite clauses is therefore impossible, as illustrated in (307), where extraposition of *that Fred is crazy* cannot leave the embedded nonfinite clause.

(307) *No extraposition out of nonfinite clauses:*

*[John_{*i*} is believed [*t_i* to be certain *t_j*] by everybody [that Fred is crazy]_{*j*}].

By category inheritance, CP nodes contains as part of their label a T specification and therefore likewise terminate search by [*uExtr*]. This derives the fact that finite clauses are likewise are opaque to extraposition, an example of Upward Entailment.

(308) *No extraposition out of finite clauses:*

*[It is believed [that John is certain t_j] by everybody [that Fred is crazy] $_j$].

This general account of superraising and other instances of selective opacity has a number of distinctive features that set it apart from previous approaches. First, the ban on A-movement out of finite clauses has nothing to do with movement on this account. Rather, it is the Agree-relation that such movement is parasitic on that cannot be established in the first place. In this sense, improper movement can be characterized as ‘improper agreement’ on this account. A side effect is that the account straightforwardly extends to the fact that ϕ -agreement into a finite clause is likewise impossible in English.

Second, as mentioned above, there is no direct interaction between movement types or types of positions on the horizons account. Just as in the case of Hindi in the previous section, the inability of \bar{A} -movement to feed A-movement follows indirectly: \bar{A} -movement entails the presence of a CP structure and hence CP node that immediately dominates them. This CP node constitutes a horizon for the matrix [uA]-probe and thereby places the \bar{A} -moved element beyond [uA]’s horizon. Just as in Hindi the inverse feeding relationship is, of course, allowed. Unlike on the traditional approach to superraising, no designated stipulation that A-movement cannot target \bar{A} -positions is necessary. Similarly, no designated role is assigned to A- vs. \bar{A} -positions, case, etc. Interactions between operations and positions are entirely epiphenomenal.

Third, the account extends beyond the binary A/ \bar{A} -distinction, a desirable result given the discussion in chapter 1. As noted there, the previous literature has observed selective opacity effects for movements that fall outside the A/ \bar{A} -divide. Extraposition provides an example from English, as just discussed. Ample further illustration of selective opacity outside the A/ \bar{A} -divide has been given in chapters 1 and 2 and the horizons account is comprehensive and versatile enough to extend to these cases. A particularly striking example of a system that comprises at least four levels of selective opacity is arguably German, which is investigated in detail in chapter 4.

Fourth, horizons radically diverge from previous approaches to selective opacity in that horizons do not attribute any role to the moving element whatsoever. While selective opacity has been investigated from a variety of analytical perspectives and while previous accounts differ substantially, virtually all of them crucially focus on the moving element, be it its syntactic position

(Chomsky 1973, 1977, 1981, May 1979), properties of its trace (Müller & Sternefeld 1993, Müller 1995), its internal properties like case (Chomsky 2000, 2001) or a record of its derivational history (Müller 2014a,b) or whether this element is affected by other movements (Abels 2007, 2009, Neeleman & van de Koot 2010). These properties of the moving item are then held responsible for that item's inability to target certain positions. In this sense, these previous accounts are *item-based*. For example, an item may be unable to move to SpecTP in virtue of being located in an \bar{A} -position or because it has a valued case features (see chapter 5 for a thorough comparison of horizons to previous approaches).

Horizons constitute a major departure from this line of analysis. On the horizon account, it is the relation between the probe and the embedded clause that determines whether a given extraction out of that clause is possible or not. The moving item itself is entirely irrelevant. To illustrate using superraising, A-movement out of CP is impossible because search by the [*uA*]-probe terminates at the CP node. Any and all elements dominated by this CP node are thus out of reach and [*uA*] cannot trigger any DP movement across this CP. It is insubstantial what position this DP occupies within the CP or what its internal features are. Consequently, selective opacity has nothing to do with the moving element on the horizons account. It arises solely from the interplay of probes and their horizons. In this sense, the horizons account is *domain-based* because it amounts to the statement that an entire domain (i.e., a clause) is impenetrable to a probe. The internal structure of this domain is irrelevant. In their domain-based nature, horizons are much more similar to the line of account pursued by Williams (2003, 2011, 2013) than they are to the other accounts just cited.

This shift in perspective has direct ramifications, both conceptual and empirical, which I will discuss in turn. On the conceptual side, the horizons account captures a generalization that traditional item-based approaches miss. Consider the empirical generalization that A-movement out of a CP clause is impossible in English. Of course, there are multiple logically possible derivations that would end up moving an element out of a CP into an A-position. To capture the empirical generalization, each and every one of them must be ruled out. A traditional item-based account has to appeal to a conspiracy of several unrelated constraints to achieve this goal. The horizon account, by contrast, manages to exclude *all* offending derivations in a uniform way, precisely because of its domain-based nature.

To develop this claim in greater detail, it is well-known that there are at least two distinct derivations that would result in A-movement out of a CP clause. One involves one-fell-swoop

movement of the element from an A- to an A-position, as shown in (309a). The other derivation involves movement to SpecCP of the lower clause, followed by movement to the A-position in the higher clause, as shown in (309b). The two derivations are usually ruled out by two distinct principles. The one-fell-swoop derivation (309a) is excluded by a constraint that requires movement to proceed through the edge of a finite clause (subjacency or phases). The successive-cyclic derivation (309b) is ruled out by a ban on movement from SpecCP (an \bar{A} -position) to an A-position (the ban on improper movement).

- (309) a. *One-fell-swoop derivation:*

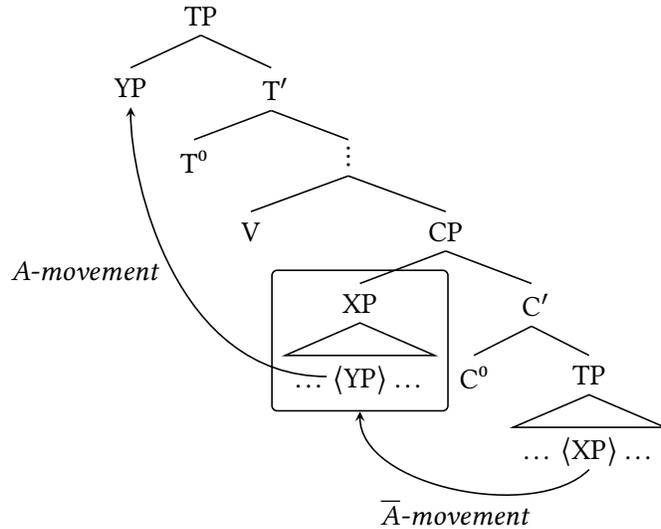
$$*Sue_i \text{ seems } [_{CP} C^0 [_{TP} t_i [_{vP} \dots]]]$$
 b. *Successive-cyclic derivation:*

$$*Sue_i \text{ seems } [_{CP} t_i C^0 [_{TP} t_i [_{vP} \dots]]]$$

Thus, on traditional item-based analyses the ban on superraising arises as a conspiracy of two separate constraints. These accounts lack the ability to express the empirical generalization that A-movement out of a CP is impossible directly. The horizons account, on the other hand, rules out both derivations in (309) in a uniform way. On both cases, movement would require the [uA]-probe to agree past a CP, violating horizons. It is altogether irrelevant where inside the CP *Sue* is located. The domain-based nature of horizons thus allows it to capture a generalization that the standard item-based account misses.

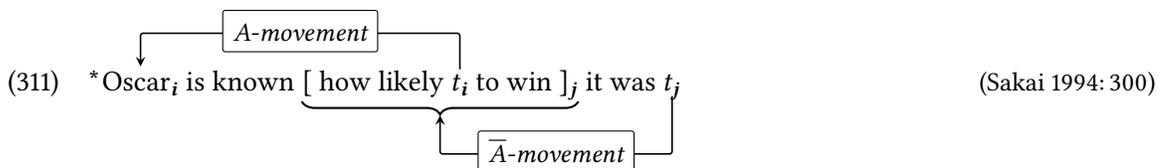
Closer inspection reveals that the conspiracy problem for item-based accounts is even worse. In addition to one-fell-swoop movement and successive-cyclic movement in (309), there is a third conceivable derivation that would result in A-movement out of a CP clause that is not ruled out by either of the two constraints invoked for (309). This derivation involves \bar{A} -movement of an XP to SpecCP, followed by A-movement of a second constituent YP out of XP, as schematized in (310). Collins (2005a,b) refers to derivations of this type as ‘smuggling’; Sauerland (1999) uses the term ‘surfing’.

(310) *Illicit smuggling derivation*

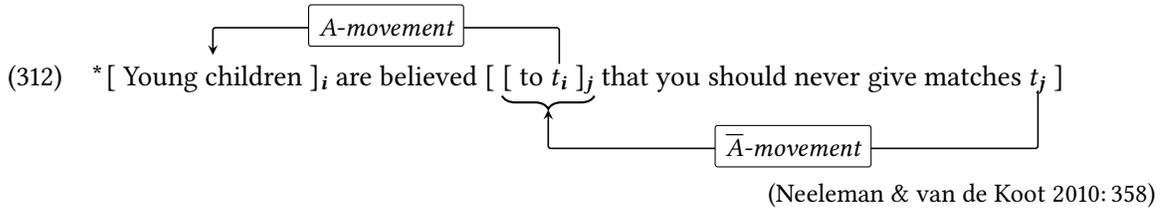


In (310), YP is smuggled to the edge of the lower clause inside XP and then A-subextracted into the higher clause. This derivation violates neither of the two constraints used to rule out (309): Subjacency and phases are not violated because YP is part of the edge of the lower clause. Furthermore, the ban on improper movement is respected because no element undergoes movement from an \bar{A} - to an A-position.¹⁵

Interestingly, as Sakai (1994), Collins (2005a) (who notes it as a problem), Abels (2007, 2009), and Neeleman & van de Koot (2010) observe, the derivation in (310) is nonetheless impossible. Relevant examples are provided in (311) and (312). In (311), A-movement is from a raising configuration, in (312) A-movement is an instance of pseudo passivization from inside a PP. In both cases, the result is completely ungrammatical.



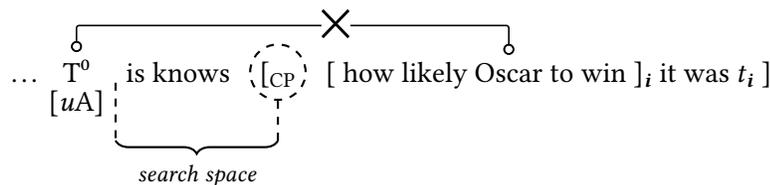
¹⁵ In fact, Hicks' (2009) account of *tough*-constructions crucially relies on the fact that structures like (310) are not ruled out by the ban on improper movement. However, the fact that such smuggling derivations are demonstrably ungrammatical elsewhere (see (311) and (312)) may suggest that the smuggling account of *tough*-constructions only restates the problem, rather than solves it.



It is clear that that (311) and (312) fall under the overarching empirical generalization that A-movement out of a finite CP clause is impossible. But because neither subjacency/phases nor the ban on improper movement succeeds in ruling them out, a traditional item-based account has to invoke a third unrelated constraint specifically targeted towards smuggling derivation. The most obvious candidate for this constraint is *freezing* (Wexler & Culicover 1980), which prohibits movement out of moved constituent. While the addition of freezing provides an accounts for the ungrammaticality of (311) and (312), it is noteworthy that item-based constraints now need a conspiracy of three distinct constraints to rule out A-movement from a finite clause: (i) a constraint that rules out one-fell-swoop movement into the matrix subject position, (ii) a constraint that rules out movement from an \bar{A} - to an A-position, and (iii) a constraint that blocks extraction out of moved constituents. The simplicity and elegance of the empirical generalization – no A-movement out of a finite clause – is lost.

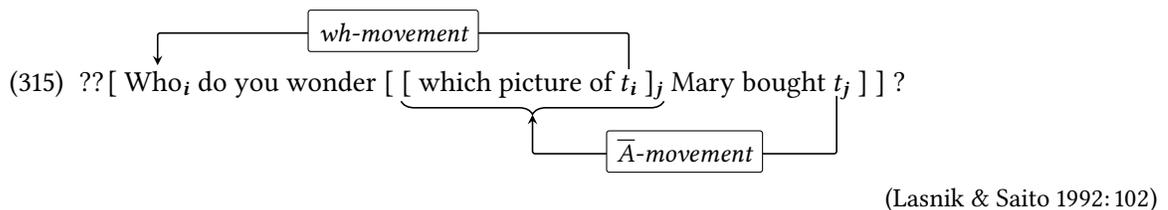
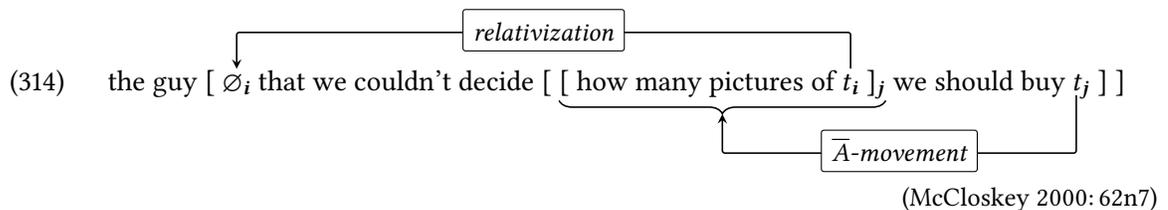
The horizon analysis, on the other hand, extends to (311) and (312) without further ado. Because anything inside the CP clause is invisible to the matrix A-probe, no Agree-relation can be formed with t_i in (311) and (312), as both are encapsulated within a CP. A-movement from these positions is hence impossible for exactly the same reason it is impossible in (303b): Any element encapsulated in a CP is inaccessible to the A-probe, regardless of precisely where this element is located within the CP. In (313), the search space of the A-movement/EPP-probe is delimited by the embedded CP node. This probe hence cannot contact *Oscar* and A-movement of *Oscar* is impossible. This derives the ungrammaticality of (311). The account of (312) is analogous.

(313) *Impossible A-movement out of \bar{A} -moved constituent (311)*



The empirical generalization that finite clauses are opaque for A-movement thus receives a systematic account on the horizons analysis, but not on traditional item-based accounts. The domain-based nature of horizons thus renders them conceptually superior.

In addition to this conceptual point, smuggling configurations like the one schematized in (310) also provide an empirical argument in favor of horizons. As I have just mentioned, an item-based account might appeal to freezing in order to rule out the sentences in (311) and (312). However, a simple freezing constraint that rules out all extraction from moved constituents (and hence all smuggling derivations) is too strong, because such extraction itself exhibits selective opacity effects. For example, \bar{A} -movement out of an \bar{A} -moved element, while not perfect, is considerably better than A-movement out of an \bar{A} -moved element (Chomsky 1986: 25–27, Lasnik & Saito 1992: 102, Sakai 1994: 300n9, McCloskey 2000: 62n7, Rizzi 2006: 114), as (314) illustrates for relativization and (315) shows for *wh*-movement:

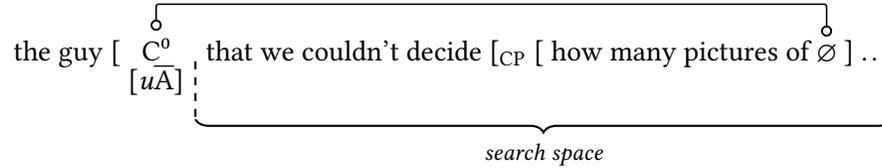


While these structural are somewhat degraded to some speakers, they are vastly better than A-movement out of an \bar{A} -moved element in (311) and (312). Any account that simply bans all movement out of moved constituents predicts the sentences in (311)–(315) to all be equally ungrammatical, contrary to fact.

The horizons account, by contrast, successfully captures the contrast between the two smuggling derivations. A-movement out an \bar{A} -moved element as in (311) and (312) violates [*uA*]'s CP horizon. By contrast, \bar{A} -movement out of an \bar{A} -moved element as in (314) and (315) is considerably improved, because \bar{A} -movement probes in English do not have CP horizons. This is schematized

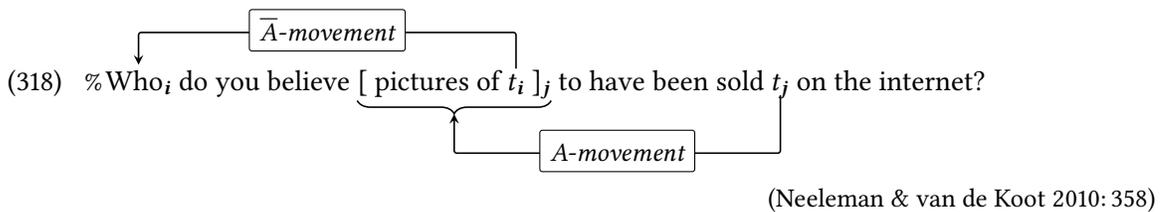
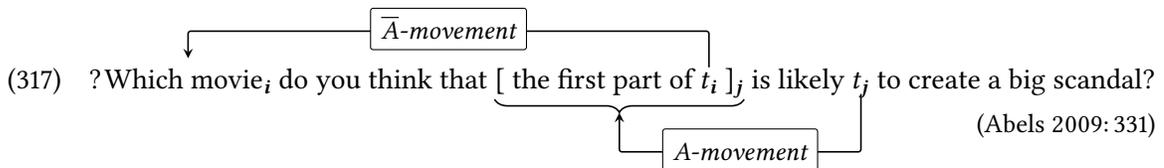
for (314) in (316), where the relativization probe can agree with the null relative pronoun and hence attract it to its specifier.

(316) *Possible relativization out of \bar{A} -moved constituent (314)*



The fact that smuggling derivations are subject to selective opacity follows from the horizons account, but it is beyond the reach of traditional item-based accounts.

Furthermore, Abels (2007, 2009) and Neeleman & van de Koot (2010) have drawn attention to the fact that \bar{A} -movement out of an A-movement constituent is likewise possible, as (317) and (318) demonstrate:



As in the cases just discussed, a simple appeal to omnibus freezing for the purposes of ruling out (311) and (312) predicts (317) and (318) to be equally impossible, again incorrectly so.

There is again an ordering asymmetry here. Just like A-movement of an element may be followed by \bar{A} -movement of the same element while the reverse is ungrammatical, so A-movement of a constituent may be followed by \bar{A} -subextraction out of that constituent, while the inverse is impossible. The horizon account captures this parallelism. (317) and (318) are possible because they do not involve A-movement over a CP. By contrast, (311) and (312) do and therefore constitute violations of [uA]'s horizons.

In sum, smuggling derivations themselves exhibit selective opacity effects. The horizons account thus provides an account when smuggling derivations are possible and when they are

impossible. It is thus empirically superior to a blanket freezing constraint, which simply bans all smuggling derivations and is hence too restrictive, as well as to a system that allows all smuggling and is hence too permissive.

The observation that movement type asymmetries extend to configurations in which movement of *one* item affects movement of *another* extends to configurations other than A- and \bar{A} -movement in English. For additional discussion of the smuggling cases just discussed, see Abels (2007, 2009) and section 4.4 of chapter 4 of this dissertation. Asymmetries in smuggling configurations provide strong support for the horizons approach over traditional item-based accounts, which are inherently unable to express this relationship.

A final point worth noting is that the horizon account also derives for free a constraint on *remnant movement* (what Sauerland 1999 calls *diving paths*) that have been independently argued for in the literature. It is uncontroversial that any theory of movement must countenance remnant movement as in (319).

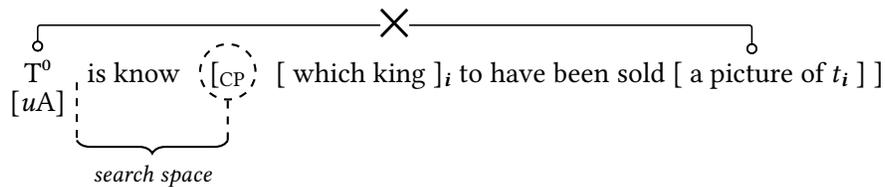
- (319) \bar{A} -movement of remnant created by A-movement
 [How likely t_i to win]_j is Sue_i t_j ?

In (319), A-movement of *Sue* is followed by \bar{A} -movement of the remnant. Interestingly, there are constraints on remnant movement, which bear an uncanny resemblance to the constraints on moving a single constituent, as specifically argued for by Grewendorf (2003, 2015), Williams (2003), and Abels (2007, 2009). Thus, the inverse of (319), i.e., A-movement of a remnant created by \bar{A} -movement, is impossible, as the example in (320c), from (Abels 2009: 331), shows. (320a,b) are provided as baselines. In (320a), *which king* is \bar{A} -extracted out of the object DP. (320b) demonstrates that such \bar{A} -movement is also possible within a nonfinite clause. In (320c), the remnant object DP is then A-moved into the matrix clause. The embedded clause in (320c) is passivized and nonfinite in order to remove a case assigner for the remnant DP in the lower clause. The remnant therefore receives case only in the matrix subject position. Nonetheless, the resulting structure is ungrammatical because a remnant created by \bar{A} -movement is itself A-moved.

- (320) a. *Baseline: \bar{A} -extraction*
 It is known [which king_i they sold [a picture of t_i]]
- b. *\bar{A} -extraction into nonfinite clause*
 It isn't known [which king_i to sell [a picture of t_i]]
- c. *A-movement of remnant created by \bar{A} -movement*
 * [A picture of t_i]_j is known [which king_i to have been sold t_j]

The contrast between (319) and (320) follows from the horizon account without further ado because (320c) would require Agree between the [uA] probe and something past its CP horizon (see (321), while (319) does not.

- (321) *Impossible A-movement out of a remnant created by \bar{A} -movement (320)*



As in the case of smuggling derivations, constraints on remnant movement are interesting because movement type interactions arise between movements of distinct elements. In virtue of its domain-based nature, the horizons account captures these non-identity cases in a completely general and unified way. Horizons thus share the advantages of other completely or partially domain-based approaches to selective opacity, namely Abels (2007, 2009) and Williams (2003, 2011, 2013). I will delay a comparison between these accounts until chapter 5.

3.4.3 Accounting for movement–agreement mismatches

The horizons account also provides an account of instances of movement–agreement mismatches, already discussed in chapter 1. As noted there, a number of authors have observed differences in the locality of movement and ϕ -agreement dependencies, particularly in the context of long-distance agreement (LDA) crosslinguistically.

One example of a movement–agreement mismatch comes from Itelmen, as analyzed by Bobaljik & Wurmbrand (2005). As illustrated in section 1.4 of chapter 1, Itelmen allows optional LDA into a nonfinite clause. Crucially, Bobaljik & Wurmbrand (2005) observe that LDA with an embedded

object forces this object to take scope over the matrix predicate, as exemplified in (322), where the agreement controller *miʃ okno-ʔn* ‘all windows’ has to take matrix scope.

- (322) t'-əntxa-čeʔn [miʃ okno-ʔn sop-es] [Itelmen]
 1SG-forget-3PL.OBJ all window-PL close-INF
 ‘I forgot to close all the windows.’ (∀ > forget; *forget > ∀)
 (Bobaljik & Wurmbrand 2005: 849)

Bobaljik & Wurmbrand’s (2005) analysis of this restriction crucially rests on a locality differences between movement and agreement. They propose that the highest projection of a clause is an *agreement domain*, which is impenetrable to agreement but transparent to movement. Consequently, LDA with the embedded object can only be established if this object moves into the matrix clause. This movement manifests itself in obligatory high scope in (322). This account thus requires that there be proprietary locality constraints on agreement which do not constrain movement.

The inverse locality mismatch between movement and agreement is likewise attested. In their investigation into LDA in Tsez, Polinsky & Potsdam (2001) note that Tsez does not allow any crossclausal movement:

- (323) a. kid-bā [uʒ-ā hibore-d bikori žāk'-ru-li] esis [Tsez]
 girl-ERG boy-ERG stick-INSTR snake hit-PASTPART-NMLZ said
 ‘The girl said that the boy had hit the snake with a stick.’
 b. *bikori kidbā [užā hibored žāk'ruʃi] esis
 snake girl boy stick hit said (Polinsky & Potsdam 2001: 590)

At the same time, Tsez allows LDA into an embedded clause, as in (324), where the embedded object *magalu* ‘bread’ triggers gender agreement in class III on the matrix verb.

- (324) enir [užā magalu bāc'ruʃi] b-iyxo [Tsez]
 mother boy bread.III.ABS ate III-know
 ‘The mother knows the boy ate the bread.’ (Polinsky & Potsdam 2001: 584)

To be precise, Polinsky & Potsdam (2001) argue that only the specifier of the TopicP of the lower clause is accessible to matrix agreement. But it is clear from the contrast between (323) and (324) that this SpecTopicP must be accessible to matrix ϕ -agreement but not to matrix movement. As in

Itelmen, there is a locality difference between movement and agreement. In contrast to Itelmen, however, in Tsez it is movement that is subject to stricter locality constraints than agreement.

It is clear from these two examples that movement and agreement may differ in their locality and there does not appear to be any directionality in this difference. In Itelmen, agreement is more tightly restricted than movement; in Tsez, it is the other way around. Horizons provide a framework flexible enough to capture these locality differences and their crosslinguistic variation. For Itelmen, we may assume that the ϕ -probe has T as its horizon, whereas the movement probe has no horizon. This is stated in (325), where I simply use $[u\mu]$ as a convenient label for the movement probe.¹⁶

(325) *Itelmen horizons*

- | | | | |
|----|------------------------------------|---|--|
| a. | $[u\phi] \dashv\parallel T$ | } | $\rightarrow \phi$ -agreement more local than movement |
| b. | $[u\mu] \dashv\parallel \emptyset$ | | |

For the Tsez pattern, we may assume the horizons in (326). For a full translation of Polinsky & Potsdam's (2001) account, an additional probe that moves a DP to a left-peripheral topic position and thereby enables agreement with $[u\phi]$ in the matrix clause would be necessary. Because topicalization is likewise clausebounded in Tsez, this probe too would have to have C as its horizon.¹⁷

(326) *Tsez horizons*

- | | | | |
|----|-------------------------------------|---|--|
| a. | $[u\phi] \dashv\parallel \emptyset$ | } | \rightarrow movement more local than ϕ -agreement |
| b. | $[u\mu] \dashv\parallel C$ | | |

On the horizons account, locality differences between movement and agreement are now unified with locality differences between types of movement and with locality differences between types

¹⁶ Needless to say, the sparse evidence just presented underdetermines the precise choice of horizons. Part of the evidence necessary for a full account revolves around the agreement and movement options for finite embedded clauses, a domain where there are gaps in the empirical data (see Bobaljik & Wurmbrand 2005: 846n26). (325) is sufficient for the point made in the main text. Determining the location of the probes in (325) is likewise difficult in light of the sparsity of available evidence. The precise locations do not matter for the point at hand, but the horizons account makes clear predictions about possible and impossible probe locations, in the form of the Height–Locality Connection to be discussed in section 3.5.

¹⁷ As was the case for Itelmen, a broader range of data would have to be considered to determine the horizons more precisely. One important qualification of (326) is that, all else equal, it would allow a ϕ -probe to search arbitrarily deep into an embedded clause. This is not in line with the analysis put forth by Polinsky & Potsdam (2001), who, as mentioned above, argue that only SpecTopicP is visible to the ϕ -probe. The question of edge agreement is taken up in detail in chapter 6, where I argue that CP phases coexist with horizons. The complete Tsez pattern then emerges as the combined effects of horizons and CP phases (also see fn. 4 on p. 386). As for the location of the probes in (326) remarks analogous to those in fn. 16 for Itelmen apply here as well.

of agreement. All of these mismatches are instantiations of the overarching pattern of selective opacity.

3.5 Constraining selective opacity: The Height–Locality Connection

Having thus developed the concept of horizons and demonstrated its utility in accounting for a wide variety of selective opacity effects in a range of apparently unrelated empirical domains, this section will consider the question of which limits horizons impose on possible selective opacity patterns. We saw in section 1.3 of chapter 1 that selective opacity effects are not distributed randomly across operations and domains but subject to overarching generalizations that emerge once selective opacity is treated as a unified phenomenon. Two broad generalizations discussed in section 1.3 of chapter 1 and also demonstrated for Hindi in section 2.5 of chapter 2 are (i) Upward Entailment and (ii) the Height–Locality Connection. The account of Upward Entailment within the horizons framework was discussed in section 3.3.2, where I proposed that it follows as a consequence of category inheritance, a property of extended projections. In this section, I will turn to the Height–Locality Connection. I demonstrate that horizons give rise to a systematic link between a probe’s location and its possible horizons that has immediate consequences for which clauses can and cannot be transparent to this probe. This property of the system derives a version of the Height–Locality Connection. Horizons therefore not only offer an account of selective opacity but also derive the its limits.

3.5.1 Deriving the Height–Locality Connection

The Height–Locality Connection in its probe-based form is repeated in (327). As before, I am going to use the term “height” to refer to the location of a probe in the functional structure.

- (327) **Height–Locality Connection** (*probe-based formulation*) [=(241)]
The higher the location of a probe is in the clausal structure, the more kinds of structures are transparent to this probe.

What (327) states is that there is a connection between the locality profile of an operation and the structural height of the position it targets. Probes located high in the clausal spine have access to

more kinds of structures (i.e., more types of embedded clauses) than probes that are located in lower positions. The claim that there is an empirical link between height and locality has been argued for in one form or another in a sequence of recent papers by Williams (2003, 2011, 2013), Abels (2007, 2009, 2012a), and Müller (2014a,b). See section 1.3.1 of chapter 1 for examples. I have also argued in depth in chapter 2 that it holds for Hindi as well and chapter 4 will provide additional evidence from German.

As discussed in section 1.3.1 of chapter 1 and in section 2.5.2 of chapter 2, the Height–Locality Connection (HLC) is an empirical generalization over selective opacity effects across a variety of different languages and constructions. A straightforward example is A- vs. \bar{A} -movement in English. It is evident that \bar{A} -movement lands in a position higher than that targeted by A-movement (SpecCP vs. SpecTP) and that \bar{A} -movement is triggered by a probe higher than that responsible for A-movement. This difference in the relative height of the two operations correlates with a difference in their locality: \bar{A} -movement can leave a finite clause, whereas A-movement cannot, and both can leave nonfinite clauses. The Height–Locality Connection states that the height of a probe is more generally connected to its locality: Probes located on structurally high heads, like \bar{A} -movement in English, can access more types of embedded clauses than probes located on lower heads.

I have argued in chapter 2 that Hindi likewise bears out the Height–Locality Connection, because the two probes located on C^0 (\bar{A} -movement and *wh*-licensing) are able to search into more types of embedded clause than the two probes located on T^0 (A-movement and ϕ -agreement).

The HLC is particularly intriguing because standard principles of syntactic locality like phases do not establish a link between the structural height of an operation and its locality properties. From the perspective of traditional locality principles, the empirical link between the two is therefore surprising. In this section, I will argue that horizons provide us with the means to derive a version of (327).

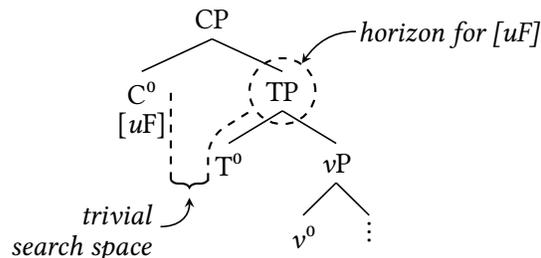
Because on the view adopted here, all selective opacity effects are the result of restrictions on Agree, the HLC translates into the observation that there is a connection between a probe's location and its horizon: Probes that are located on a structurally high head have structurally higher horizons. It is, of course, possible to simply stipulate such a connection in one way or another. But

I will argue here that horizons in fact *derive* a version of the HLC and hence allow us to understand *why* the HLC should hold in the first place.

No modification of the horizons system proposed so far is necessary to derive a connection between a probe's location and its locality profile, and hence a version of the HLC (327). I will thus retain the view that the distribution of horizons is entirely unconstrained as far as the analysis is concerned, but that distributional constraints of selective opacity emerge from the interplay between horizons and other aspects of the system, in particular Upward Entailment, which is itself a consequence of extended projections (see section 3.3.2). As I will now show, despite the fact that in principle any pairing of location and horizon of a probe is defined and licit, certain pairings will result in probes that are inherently unable to trigger movement or agreement operations. Put differently, movement and agreement can only arise from probes for which location and horizon stand in a particular relationship to each other. The HLC follows as an emergent property of the horizons system.

To better see the reasoning behind the account of the HLC that I propose, it is instructive to investigate a specific counterfactual example. Consider a probe located on C^0 . We have already seen in Hindi an example of a probe on C^0 with C as its horizon (namely, $[uwh]$) and of a probe on C^0 with no horizon ($[u\bar{A}]$). Consider a counterfactual situation in which a hypothetical probe $[uF]$ is located on C^0 and has T as its horizon. This probe's search would terminate at a TP node, rendering embedded TP clauses opaque to $[uF]$. But by the same token, *any* TP node would lead to termination of $[uF]$'s search. Because $[uF]$ is located on C^0 , its very sister TP would terminate its search. As a result, the very first element $[uF]$ encounters on its search constitutes its horizon. Consequently, $[uF]$'s search space would only comprise its sister node TP and nothing else. This counterfactual state of affairs is depicted in (328). For reasons to be discussed immediately, I will call such probes *vacuous*.

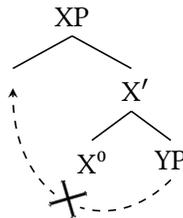
(328) Example of a vacuous probe: $[uF]$ on C^0 with $[uF] \dashv\vdash T$



What are the possible syntactic manifestations of probes with trivial search spaces? Suppose first that [*uF*] in (328) is a movement-inducing probe. Abels (2003, 2012b) has argued that no element may move from the complement position of a head to the specifier position of the same head, as such movement would be too local. He deduces this prohibition from *Last Resort*, a constraint that prohibits movement unless it enables the satisfaction of a feature that cannot otherwise be satisfied. He reasons that because a head stands in a mutual c-command relationship with its complement, moving that complement into the specifier position of the same head would not result in a c-command relationship that does not already exist in the base configuration. As a completely general consequence, there is no need for a complement to ever move into the specifier position of the same head. Since such movement is unnecessary, it is prohibited by Last Resort. The **ANTI-LOCALITY** constraint that emerges as a consequence is stated in (329).

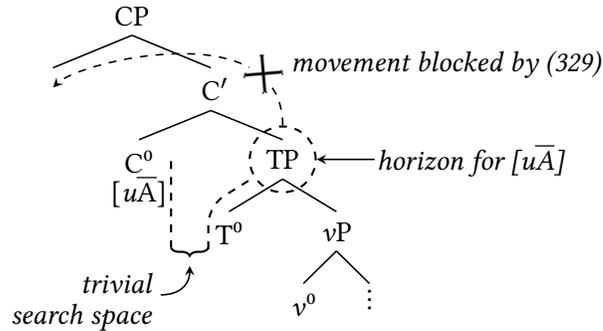
(329) **ANTI-LOCALITY CONSTRAINT** (Abels 2003)

Movement from the complement to the specifier position of the same projection is impossible.



The consequence of anti-locality (329) in the present context is clear: If a movement-inducing probe has its sister as its horizon, this probe will be unable to trigger any movement operation, as schematized (330) for our example probe [*uF*]. The one and only element in [*uF*]'s search space is its complement TP, which [*uF*] cannot move due to anti-locality. Furthermore, all elements that would be far enough away from [*uF*] to move to SpecCP, i.e., all elements contained inside that TP, are outside of [*uF*]'s search space.

(330) Another example of a vacuous probe: $[u\bar{A}]$ on C^0 with $[u\bar{A}] \dashv\vdash T$



As a result, $[uF]$ is unable to trigger any movement whatsoever in (330). $[uF]$'s existence would therefore be undetectable as it would have no discernible effects on the syntactic structure or derivation. In this sense, $[uF]$ is vacuous.

The second case to consider is one where the probe $[uF]$ does not trigger movement, but involves 'pure' Agree. As before, the only element within $[uF]$'s search space is its sister TP node. Agreement between $[uF]$ and this TP node is possible, but effectively indistinguishable from subcategorization under sisterhood. Genuine agreement, i.e., agreement between elements in an asymmetric c-command relationship, is ruled out because all elements that $[uF]$ asymmetrically c-commands are outside of its search domain. As in the case of a movement-inducing probe, we arrive at the conclusion that $[uF]$ would be unable to give rise to agreement or any other long-distance dependencies in (328).

In sum, probes like $[uF]$ in (328) whose sister is their horizon are systematically unable to trigger movement, agreement or any other dependencies that do not involve immediate sisterhood. For this reason, I will refer to such probes as **VACUOUS**. Their only possible impact on the syntactic derivation is indistinguishable from subcategorization/selection. In this sense, vacuous probes are undetectable in the output of the system.

The reasoning has so far focused on probes whose horizon is their sister node. As it turns out, it automatically generalizes beyond this case. Consider a slightly different counterfactual scenario in which the probe $[uF]$ on C^0 has v as its horizon (thus, $[uF]_{C^0} \dashv\vdash v$). In this case, vP s are opaque to the probe. Recall, however, that horizons are inherited up an extended projection, a consequence of category inheritance (272), as proposed in section 3.3.2 above.

- (331) **HORIZON INHERITANCE THEOREM** [=(276)]
 Given a probe $[uF]$ and an extended projection $\Phi = \langle \Pi_n > \Pi_{n-1} > \dots > \Pi_1 \rangle$, if $\Pi_m \in \Phi$ is a horizon for $[uF]$, then all projection Π_{m+1}, \dots, Π_n are likewise horizons for $[uF]$ (due to category inheritance (272)).

As a result, if vP is opaque to $[uF]$, then TP is as well, and we again arrive at the state of affairs in (328). Therefore, $[uF]$ in this case is vacuous even if it has v as its horizon. By parity of reasoning, $[uF]$ would also be vacuous if it had V as its horizon. As a consequence, a probe $[uF]$ located on C^0 could not have either T , v or V as its horizon without being vacuous. Continuing to use the term ‘agreement’ to refer to a dependency between elements in an asymmetric c-command relationship, we arrive at (332):

- (332) *Examples of vacuous height–locality pairings*
- | | | | |
|----|----------------------------------|---|---|
| a. | $[uF]_{C^0} \dashv\!\! \dashv T$ | } | <i>vacuous</i> \rightarrow <i>could not trigger any movement or agreement</i> |
| b. | $[uF]_{C^0} \dashv\!\! \dashv v$ | | |
| c. | $[uF]_{C^0} \dashv\!\! \dashv V$ | | |

We have now established that there is a connection between the location of a probe and its possible horizons. If a probe located on C^0 has as its horizon a projection lower than C in the extended projection, this probe will necessarily be vacuous and hence unable to establish any movement or agreement dependencies.

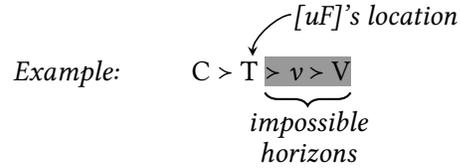
The fact that the horizon settings in (332) are impossible for a non-vacuous probe on C^0 immediately entails that embedded TP clauses, vP clauses, and VP clauses are necessarily transparent to such a probe, because they could never constitute a horizon for such a probe, given (332).

The relationship between location and horizons generalizes to probes on heads other than C^0 as well. The relationship is expressed in its most general form in the Height–Locality Theorem in (333). (333a) restricts what may constitute a horizon for a probe given the syntactic position of that probe. Conversely, (333b) states a restriction on the syntactic position a probe may occupy given its horizon specification. All probes that violate (333) are vacuous.

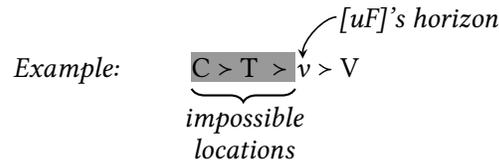
(333) **HEIGHT-LOCALITY THEOREM**

Given an extended projection $\Phi = \langle \Pi_n > \Pi_{n-1} > \dots > \Pi_1 \rangle$, for any non-vacuous probe $[uF]$,

- a. If $[uF]$ is located on Π_m , then a projection $\in \{\Pi_{m-1}, \dots, \Pi_1\}$ cannot be a horizon for $[uF]$.



- b. If $[uF]$ has Π_m as a horizon, $[uF]$ cannot be located on a projection $\in \{\Pi_n, \dots, \Pi_{m+1}\}$.



(333) describes the relationship that must hold between a probe's location and its horizon for this probe to be non-vacuous and hence be able to trigger movement, agreement or any other long-distance dependency. Knowing the location of a non-vacuous probe imposes restrictions on its possible locality properties. Conversely, knowing the locality properties of a probe imposes restrictions on its possible locations. There is thus an abstract but entirely general connection between a probe's location and its possible horizon settings. The consequence that certain horizon settings are thus ruled out given the location of a probe immediately entails that some clauses are necessarily transparent to this probe.¹⁸

It is important to note that neither (333a) nor (333b) are stipulations of the system. Nothing in the account precludes the existence of probes that do not conform to (333), hence vacuous probes. As far as the axioms of the account are concerned, all pairings of location and horizons are allowed. But as I have just shown, certain pairings of location and horizon result in a vacuous probe, which are unable to trigger movement or agreement dependencies. It follows that all attested movement

¹⁸ As mentioned in section 3.3.1 above, I presuppose here that every probe is inherently linked to exactly one syntactic head. That is, while a single head may contain more than one probe feature, every probe feature cannot variably appear on one head or another within a language (thus, $[u\phi]$ always appears on T^0 in Hindi, $[u\bar{A}]$ always appears on C^0 in Hindi, and so on). As Kyle Johnson (p.c.) has pointed out to me, a more flexible approach compatible with (333) would be to not specify the location of a probe and let the choice be free but subject to the Height-Locality Theorem. One case in which this line of account may prove useful is QR, which may target variable positions but exhibits a uniform locality pattern. On Kyle Johnson's suggestion, the QR probe would have a specified horizon, and this horizon would impose limitations on the possible locations of the probe, by (333). Within the set of admissible locations, the choice is free and variable. I will not explore this possibility further here.

and agreement dependencies must be triggered by probes for which height and locality stand in the relationship described by (333), with immediate consequences for what types of clauses are necessarily transparent to each dependency. This derives the fact that all movement and agreement dependencies exhibit an empirical link between height and locality, and hence a version of the Height–Locality Connection (327).¹⁹

On this account of the Height–Locality Connection, the link between height and locality is purely a property of the *output* of the system. As far as the theoretical axioms are concerned, every conceivable pairing of location and horizon is possible, but only probes for which the relation between the two satisfies (333) are able to trigger movement or agreement operations. It follows that in the output of the system, all movement and agreement dependencies exhibit a link between height and locality. In other words, despite the fact that the system itself does not impose any connection between height and locality, such a connection arises in the output of the system. The Height–Locality Connection is thus an *emergent* feature of the horizons analysis and derived from the more basic architecture of horizons.

The Height–Locality Theorem (333) automatically translates into predictions about the transparency or opacity of embedded clauses for particular probes given the location of these probes. For example, if a probe is located on, e.g., C^0 , then it follows that from (333) that neither T nor v nor V could constitute a horizon for it, which in turn entails that TP clauses, vP clauses, and VP clauses are necessarily transparent to this probe. Similarly, for a probe located on T^0 , embedded vP clauses and VP clauses are necessarily transparent, and so on. Knowing the location of a probe has immediate consequences for which clauses can and cannot be opaque to that probe and vice versa.

To illustrate the effects of (333) using a concrete example, (334) describes its effects for the Hindi facts presented in chapter 2. (334a) illustrates entailments from location/height to locality. (334b) exemplifies an entailment from locality to location/height. All entailments in (334) are factually correct in Hindi.²⁰

¹⁹ Considerations of language acquisition may entail that vacuous probes do not actually exist, as there would be no motivation for a language learner to postulate the existence of a probe that never has any effect on output. But as far as the axioms of the analysis are concerned, there is nothing wrong with vacuous probes besides the fact that their effects are invisible.

²⁰ Recall that I am using the terms ‘A-movement’ and ‘ \bar{A} -movement’ to refer to movement triggered by $[uA]$ and $[u\bar{A}]$, respectively, but that I will have nothing to say here about why movement triggered by $[uA]$ is able to obviate

(334) *Empirical effects of the Height–Locality Theorem (333) in Hindi*

a. **Height → Locality (333a):**

i. *Wh-licensing*

As motivated in section 2.5.2, the *wh*-probe [*uwh*] is located on C^0 in Hindi. By (333a), it follows that neither T nor *v* nor V can be horizons for [*uwh*], or else [*uwh*] would be vacuous. This entails that TP clauses and *v*P clauses are *necessarily* transparent to [*uwh*]. Thus, nonfinite clauses *could not* be opaque to *wh*-licensing. Moreover, *wh*-licensing *could not* interact with long-distance agreement in the sense that *wh*-licensing into an embedded clause renders LDA obligatory. Both entailments are correct, as shown in section 2.4.2.

ii. \bar{A} -movement

We also saw evidence in section 2.5.2 that the \bar{A} -probe [$u\bar{A}$] is located on C^0 as well. By analogous reasoning, neither T nor *v* nor V can be horizons for [$u\bar{A}$] and TP clauses as well as *v*P clause are *necessarily* transparent (by (333a)). Consequently, nonfinite *could not* be islands for movement triggered by [$u\bar{A}$]. Furthermore, such movement *could not* interact with long-distance agreement in that \bar{A} -movement out of the lower clause makes LDA into that clause obligatory (as in the case of A-movement). This entailment is again correct, as shown in section 2.3.3.

iii. *A-movement*

Section 2.5.2 in chapter 2 has provided distributional evidence for the claim that the A-probe [*uA*] is situated on T^0 . By (333a), this entails that neither *v* nor V can be horizons for [*uA*]. *v*P and VP clauses are hence *necessarily* transparent for movement triggered by [*uA*]. Against the background of the Hindi structures argued for in chapter 2, this entails that [*uA*] *must* be able to move an element out of nonfinite clauses (i.e., if they are *v*Ps). This entailment is again correct.

iv. ϕ -Agreement

Section 2.5.2 in chapter 2 has argued that the ϕ -probe [*u ϕ*] in Hindi is located on T^0 . By reasoning analogous to that for A-movement, (333a) entails that *v*Ps clauses are *necessarily* transparent to ϕ -agreement. This predicts LDA in Hindi to be possible into nonfinite clauses precisely because Hindi allows embedded *v*P clauses.

b. **Locality → Height (333b):**

A-movement

In section 2.3.3, we saw evidence that in Hindi, A-movement out of TP clause is

crossover and feed reciprocal binding, whereas movement triggered by [$u\bar{A}$] is not. That is, I focus here on the landing site of a movement, its probe feature, and its locality.

impossible. This means that $[uA]$ has T as its horizon. By (333b), it follows that $[uA]$ cannot be located on C^0 . Consequently, A-movement must be able to land inside nonfinite clauses. This is correct, as shown in section 2.5.2.

The horizons framework hence links a probe's location to its locality profile and vice versa. Therefore, horizons not only provide us with the means of accounting for the various selective opacity facts observed in Hindi and beyond, it also contributes to our understanding of *why* selective opacity in Hindi is distributed in the way it is. Several aspects of the Hindi system now emerge as necessary aspects of this system.

While the Height–Locality Theorem thus captures a connection between a probe's location and its locality properties, it is important to note that this relation is not a one-to-one mapping. For instance, we have seen in chapter 2 that the \bar{A} -movement probe and the *wh*-licensing probe are both located on C^0 . Nonetheless, they exhibit a subtle difference in their locality profiles: Finite CP clauses are transparent for $[u\bar{A}]$, but opaque for $[uwh]$. This fact is compatible with the Height–Locality Theorem because neither choice of horizon violates it. What is ruled out by (333a) are probes located on C^0 with any projection lower than CP as their horizons. Both $[u\bar{A}]$ and $[uwh]$ conform to this requirement and are hence well-formed. An important feature of the Height–Locality Theorem is thus that it restricts the space of licit pairings of locations and horizons, but it does not completely reduce one to the other. In other words, knowing a probe's location or horizon narrows down the potential choices for the other, but it leaves some residual variability. This is a desirable outcome because it allows us to capture empirically attested variability relations between height and locality. This issue is discussed in greater depth in section 3.5.2 below and in section 5.8 of chapter 5, where I compare the Height–Locality Theorem to previous approaches.

It is instructive to further illustrate the Height–Locality Theorem on the basis of an empirical generalization that Abels (2012a) has argued for on independent grounds. Based on an intriguing investigation of the Italian left periphery, Abels (2012a) argues that there exists an empirical link between the height of the landing site of a movement type in the left periphery and its ability to leave various kinds of clauses. This pattern is an instantiation of the Height–Locality Connection and Abels concludes that an account that does not relate a movement type's height and locality misses a generalization. Specifically, Abels (2012a) argues that at least some cartographic statements about the order of elements in the left periphery are superfluous once the locality properties of

various movement types are taken into consideration. I will illustrate his reasoning on the basis of Italian evidence and in particular the complementizer *se* (Abels 2012a: 244–245). As (335) illustrates, focus fronting can take place in a clause that contains *se* but if it does, *se* has to precede the focused element:

(335) *Italian: Focus fronting lands to the right of se*

- a. Mi domando **se QUESTO** gli volessero dire (non qualcos' altro).
 I wonder if **THIS** they wanted to say to him not something else
- b. *Mi domando **QUESTO se** gli volessero dire (non qualcos' altro).
 I wonder **THIS** if they wanted to say to him not something else

(Rizzi 2001: 289)

Abels (2012a: 244) observes that the inability of a focused element to cross *se* extends to cross-clausal dependencies. (336) shows that clauses containing *se* are islands for focus fronting.

(336) *Italian: Focus fronting cannot cross se*

- ***QUESTO** mi domando **se** gli volessero dire (non qualcos' altro).
THIS I wonder if they wanted to say to him not something else

(Ilaria Frana, p.c.)

Abels (2012a) notes that a locality constraint according to which focus movement cannot cross *se* automatically captures the local ordering restriction in (335) as well. (335a) would require focus movement over *se*, yet (336) shows that this is impossible. Abels (2012a) concludes that a cartographic statement according to which *se* occupies a position higher in the left periphery than the landing site of focus movement is redundant because this positional constraint is already entailed by the locality of focus fronting.

The Height–Locality Theorem that follows from horizons preserves Abels' (2012a) key insight. For the sake of simplicity, let us refer to the head that *se* realizes as ' C_{se}^0 ' and the probe that triggers focus movement [*uFoc*]. (336) shows that C_{se} is a horizon for [*uFoc*], as stated in (337a). It follows then from (333b) that [*uFoc*] *cannot* be located higher than C_{se}^0 . Instead, it must be located on a lower head, which I will simply call Focus^0 . The Height–Locality Theorem thus immediately entails from

(337a) that Focus^0 is located lower than C_{se}^0 (337b) and consequently that focus movement must land to the right of *se*. The facts in (335) thus follow. A probe's locality has immediate implications for its syntactic location. Horizons thus provide a straightforward way of capturing the links between height and locality that Abels (2012a) argues for. That said, horizons are markedly different from Abels' (2012a) theoretical proposal in a variety of other respects. These differences are discussed in section 5.8.2 in chapter 5.

- (337) a. $[\text{uFoc}]_{\text{Focus}^0} \dashv\vdash C_{se}$ entailment (333b)
 b. $\dots > C_{se} > \dots > \text{Focus} > \dots$ ←

In sum, the horizons account not only affords a systematic and general account of locality differences for movement and non-movement operations alike, it also imposes a general limit on such locality differences. In particular, the present account derives a version of the Height–Locality Connection without the need to make any additional assumptions regarding height and locality. As far as the core principles of the account are concerned, all pairings of height and locality are viable. But it follows from the Horizon Inheritance Theorem (276) that certain pairings of height and locality produce vacuous probes, which are unable to trigger long-distance operations like movement or agreement. Only probes for which height and locality stand in a particular relationship to each other will be non-vacuous. This relationship is what the Height–Locality Theorem (333) describes. Because limitations on the possible horizon settings of non-vacuous probes immediately entail which types of embedded clauses are necessarily transparent to such a probe, a version of the Height–Locality Connection follows: The higher the probe is located, the more types of clauses are necessarily transparent to this probe.

It is worth emphasizing that the curious connection between height and locality emerges here from a system that lacks any designated link between height and locality. It is merely a property of the output of this system and hence emergent. The empirical link between height and locality is thus derived, providing an explanation for why it should hold in the first place.

3.5.2 Variation in horizons

The previous section has demonstrated how horizons manage to derive the link between a probe's location and its locality properties that we observe empirically. In this section, I draw attention to the fact that this link is not predicted to be a one-to-one relationship. That is, while a probe's height restricts its possible locality profiles and vice versa, neither completely determines the other. To appreciate this fact, consider once more a hypothetical probe $[uF]$ located on C^0 . Due to (333a), no projection lower than C^0 can be a horizon for this probe. This does not completely determine a choice, however. $[uF]$ could either have C as its horizon or alternatively have no horizon at all.

(338) *Possible and impossible height–locality pairings for a probe located on C^0*

- | | | | |
|----|---|---|------------------------------------|
| a. | $[uF]_{C^0} \dashv \parallel \emptyset$ | } | \checkmark (non-vacuous) |
| b. | $[uF]_{C^0} \dashv \parallel C$ | | |
| c. | $[uF]_{C^0} \dashv \parallel T$ | } | vacuous (do not conform to (333a)) |
| d. | $[uF]_{C^0} \dashv \parallel v$ | | |
| e. | $[uF]_{C^0} \dashv \parallel V$ | | |

Both (338a) and (338b) give rise to non-vacuous probes and should hence both be possible and attested. This is indeed precisely what we find. Recall that there are two probes in Hindi that are located on C^0 : $[uwh]$ and $[u\bar{A}]$, repeated from (282) in (339). $[u\bar{A}]$ exemplifies (338a), while $[uwh]$ instantiates (338b).

- (339) a. $[u\bar{A}]_{C^0} \dashv \parallel \emptyset$
 b. $[uwh]_{C^0} \dashv \parallel C$

In spite of their identical location, $[u\bar{A}]$ and $[uwh]$ have different horizons. This makes it clear that a probe's syntactic location does not completely determine its locality profile. This partial indeterminacy is captured in the horizons analysis, because a probe's location restricts the choice of horizons by ruling out all projections lower than that location, but it does not force a particular choice. Staying with the example at hand, no probe located on C^0 can be blocked by VP, vP or TP clauses, but it may or may not be blocked by a CP clause. This is precisely the situation in Hindi.

Horizons, then, allow for some variability in the relation between height and locality, but at the same time impose a strict limit on the amount of this variability. As I will discuss in greater detail in chapter 5.8, this property places horizons in between accounts that do not derive any connection between height and locality on the one hand (e.g., Müller & Sternefeld 1993, Abels 2007, 2009), and accounts that treat the Height–Locality Connection as a correlation with no potential for variability (Williams 2003, 2011, 2013, Abels 2012a).

Due to the indirect nature of the relationship between height and locality, horizons afford an account of crosslinguistic variation in, e.g., the locality of A-movement. It is well-known that not all languages pattern like English or Hindi in prohibiting A-movement out of finite clauses. Possible superraising has perhaps been most thoroughly investigated for Bantu languages (e.g., Carstens 2010, 2011, Diercks 2012, Halpert 2012), but it is also attested for Greek (Alexiadou & Anagnostopoulou 2002), Brazilian Portuguese (Nunes 2008, though see Williams 2011 for critical discussion), Rumanian (Grosu & Horvath 1984), and several other languages (see Ura 1994 for extensive discussion). The example in (340) from the Bantu language Zulu provides an illustrative example.

- (340) *A-movement out of finite clause in Zulu*
 uZinhle u-bonakala [ukuthi u-zo-xova ujeqe]
 AUG.1Zinhle 1SG-seem that 1SG-FUT-make AUG.1steamed.bread
 ‘It seems that Zinhle will make steamed bread.’ (Halpert 2012: 246)

The embedded subject *uZinhle* ‘Zinhle’ moved out of the finite clause controls verb agreement on the matrix predicate in (340).

The evident crosslinguistic variation in whether or not a language allows A-movement out of a finite clause provides another argument that the location of a probe does not completely determine its locality profile and falls wholly within the range of variation permitted by the Height–Locality Theorem in (333). In English, the A-probe is specified as in (341b), whereas the Zulu A-probe lacks a horizon (341a).²¹ A third pattern predicted by the present account is that the A-movement probe has T as its horizon. This is the case in Hindi. It is also arguably the case in Russian, where neither

²¹ Work on superraising in Bantu has additionally unearthed more fine-grained distinctions than the ones given in (341). Carstens & Diercks (2013) point out interesting constraints on superraising in various Bantu languages. For example, Lubukusu superraising is possible across the complementizer *mbo*, as in (i.a) but not across the agreeing complementizer *-li*, as in (i.b):

finite nor nonfinite clauses allow A-movement out of them, as illustrated in (342) and discussed in detail by Stepanov (2007).²²

(341) *Crosslinguistic variability in the locality of A-movement*

- a. *Zulu:*
[uA]_{T⁰} -|| ∅
- b. *English:*
[uA]_{T⁰} -|| C
- c. *Hindi/Russian:*
[uA]_{T⁰} -|| T

(342) *No A-movement out of finite and nonfinite clauses in Russian*

- a. *Èti studenty_i kažutsja [_{t_i} znat' tri jazyka].
these students seem.3PL know.INF three languages
- b. *Èti studenty_i kažutsja [(čto) t_i znajut tri jazyka].
these students seem.3PL (that) know.3PL three languages
- c. Kažetsja [čto èti studenty znajut tri jazyka].
seem.3SG that these students know.3PL three languages
'It seems that these students know three languages.' (Ekaterina Vostrikova, p.c.)

-
- (i) a. Babaandu ba-lolekhana (mbo) ba-kwa [Lubukusu]
2.people 2SA-seem (that) 2SA.PST-fall
'The people seem like they fell/The people seem to have fallen.'
 - b. *Mikaeli a-lolekhana a-li a-si-kona
Michael 1SA-seem that 1SA-PERS-sleep
'Michael seems to be still sleeping.' (Carstens & Diercks 2013: 100,109)

Carstens & Diercks (2013) propose that *-li* is the realization of a head that is structurally higher in the left periphery than *mbo*. On this assumption, the constraint on superraising in Lubukusu is plausibly yet another effect of horizons, yielding the setting in (ii), where I use the label 'C^{-li}' to the head that hosts the complementizer *-li*, following Carstens & Diercks' (2013) structural proposal.

- (ii) *Lubukusu:* [uA] -|| C^{-li}

It furthermore makes it clear that the existence of superraising does not necessarily entail that an A-probe has no horizon. It may merely mean that its horizon is shifted upward and hence a broader class of clauses is transparent to this probe. Carstens & Diercks (2013) also discuss superraising Lusaamia and argue that it is possible only if the embedded clause lacks a complementizer, a constraint that also seems amenable to a horizon-based account.

²² The horizons in (341) are based on the assumption that in Russian nonfinite clauses can be TPs, like in English and Hindi. Stepanov (2007), by contrast, attributes their opacity to A-movement to the assumption that they are invariably CPs in Russian.

Crosslinguistic variation in the domain in which scrambling can take place points to a similar picture. In German, for instance, scrambling cannot leave a finite clause, an observation due to Bierwisch (1963) and Ross (1967):

- (343) *Maria hat den Mann_i gesagt, [_{CP} dass Elke t_i getroffen hat]. [German]
 Maria has the man said that Elke met has
Intended: ‘Maria said that Elke met the man.’

In Russian, by contrast, scrambling out of a finite clause is possible even if it lands in a relatively low position (Müller & Sternefeld 1993), as shown in (344a). Such scrambling may also apply across multiple finite clause boundaries, as in (344b), which is degraded but possible, at least in oral speech:

- (344) a. On ètu knigu_i dumajet [_{CP} što Pëtr pročital t_i]. [Russian]
 he this book thinks that Pjotr read
 ‘He thinks that Pjotr read this book.’
 b. ?Marina ètu knigu_i skazala [_{CP} što ona думаet [_{CP} što Pëtr pročital t_i]]
 Marina this book said that she thinks that Pjotr read
 ‘Marina said that she thinks that Pjotr read this book.’ (Ekaterina Vostrikova, p.c.)

This crosslinguistic difference between German and Russian may be captured in the horizons framework in a fashion analogous to superraising above. Assuming for concreteness that the scrambling probe, notated here as [_{uScr}], resides on v^0 , we have the following contrast (see chapter 4 for an in-depth discussion of German):

- (345) *Crosslinguistic variability in the locality of scrambling*
 a. *German:*
 [_{uScr}]_{v⁰} -|| C
 b. *Russian:*
 [_{uScr}]_{v⁰} -|| ∅

The limited amount of locality variation we observe both within a single language and across languages is thus consistent with the Height–Locality Theorem. All cases of variation discussed revolve around structurally low probes, which may or may not be blocked by structurally higher projections. This is precisely the configuration that the Height–Locality Theorem does *not* restrict, and so we predict variation in this domain. What the Height–Locality Theorem restricts is that a

structurally high probe cannot be blocked by a structurally lower projection. This is the case in all the cases just considered. An apparent counter-example is discussed in section 3.5.3. See section 5.8 of chapter 5 for a comparison between the Height–Locality Theorem and previous approaches to the Height–Locality Connection that do not share this flexibility. A core aspect of the horizon account is thus that horizons setting are arbitrary within the space of options permitted by the Height–Locality Theorem.

The relation between patterns that are ruled out by the Height–Locality Theorem and ones that are admitted may be precisely quantified. It is a function of the number of projections that an extended projection is assumed to comprise. The greater the number of projections, the more pairings of location and horizon are ruled out. The number of excluded pairings can be calculated as in (346):

(346) *Height–locality pairings ruled out by the Height–Locality Theorem (333)*

Let n be the number of projections in an extended projection. The number of height–locality pairings of probes excluded by (333) equals:

$$\sum_{k=1}^{n-1} k = \frac{n^2 - n}{2}$$

It is moreover possible to calculate the proportion of possible pairings that are ruled out with the equation in (347):

(347) *Proportion of height–locality pairings ruled out by (333)*

Let n be the number of projections in an extended projection. The proportion of height–locality pairings of probes excluded by (333) equals:

$$\frac{\sum_{k=1}^{n-1} k}{n(n+1)} = \frac{\frac{n^2-n}{2}}{n^2+n} = \frac{n-1}{2n+2}$$

The proportion increases with a greater number of projections. The limit of (347) is 0.5, meaning that in the limit 50% of all a priori possible pairings are ruled out by (333).²³

²³ The proportions of pairings ruled out by (333) is greater the more projections an extended projection is assumed to contain and hence most notable in cartographic approaches. The effect can be increased if it is assumed that probes are confined to functional projections and hence absent on lexical projections. In this case, the number of excluded pairings would be the same as in (346), but because the number of logically possible horizon settings has decreased from $n(n+1)$ in the denominator of (347) to $(n-1)(n+1)$, the proportion of excluded settings increases to

$$\frac{\sum_{k=1}^{n-1} k}{(n-1)(n+1)} = \frac{\frac{n^2-n}{2}}{n^2-1} = \frac{n^2-n}{2n^2-2}$$

Before moving on, it is worth pointing out that the Height–Locality Theorem (333) gives rise to an asymmetry: The restrictions that hold for structurally high probes are stricter than those for structurally low probes. To see this point more clearly, consider a simple extended projection consisting of V, v , T and C and let us compare a probe $[uF]$ located on v^0 with one residing on C^0 . If $[uF]$ is on v^0 , only V as an Agree-barrier will render the probe vacuous and hence be ruled out (see (348a)). By contrast, if $[uF]$ is located on C^0 , then V, v and T will all be excluded (see (348b)).

- (348) a. *Non-vacuous horizon options for $[uF]$ on v^0 :*
 $[uF] \dashv\!\! \dashv v; [uF] \dashv\!\! \dashv T; [uF] \dashv\!\! \dashv C; [uF] \dashv\!\! \dashv \emptyset$
- b. *Non-vacuous horizon options for $[uF]$ on C^0 :*
 $[uF] \dashv\!\! \dashv C; [uF] \dashv\!\! \dashv \emptyset$

Structurally low probes thus enjoy a greater variability in the horizons than high probes. Comparing languages that do not allow superraising – like English – and languages that do – like Zulu – demonstrates that probes can differ crosslinguistically in whether a structurally higher projection counts as a horizon for this probe or not. An important prediction of horizons is that no such indeterminacy should exist with respect to projections that are structurally lower: these are ruled out as horizons on systematic grounds. All mismatches that we have considered so far indeed fall into the former group and are hence unproblematic. In the next section, I will discuss an apparent example of the latter group: topicalization in English. The topicalization probe appears to require a horizon lower than its location. If such a situation is indeed attested, it would provide evidence against horizons. The next section is devoted to assessing this situation.

3.5.3 Topic islands in English:

A problem for the Height–Locality Theorem?

Assuming that the complementizer *that* is a realization of C^0 in English, topicalization evidently lands below C^0 :

- (349) a. Bill says that John, Mary doesn't like.
 b. *Bill says John, that Mary doesn't like.

The limit still equals 0.5 for $n \rightarrow \infty$, but the function approaches this limit faster.

For the sake of concreteness, I will tentatively assume here Lasnik & Saito's (1992) analysis, according to which topicalization targets IP/TP, but this assumption is not critical.

Moreover, it is standardly assumed that *wh*-movement and relativization target SpecCP in English. Within the probe-centric view adopted here, they must thus be triggered by probes on C^0 .

Given this state of affairs, the Height–Locality Theorem (333) makes a clear prediction: Because *wh*-movement and relativization are triggered by probes located higher than the landing site of topicalization, they should not be blocked by it. Concretely, neither [*uwh*] nor [*urel*] could have T as their horizons and therefore they should not be affected by the presence of topicalization. Interestingly, this prediction is standardly taken to not be borne out. Many authors, including Rochemont (1989), Lasnik & Saito (1992), Müller & Sternefeld (1993), and Williams (2013), claim that topicalization induces islandhood for *wh*-movement and/or relativization, based on facts like the following:

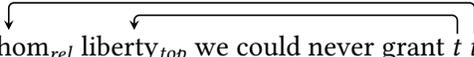
(350) *Topic islands*

- a. *What does John think that Bill, Mary gave to?
- b. *This is the man who that book, Mary gave to. (Rochemont 1989: 147)

Sensitivity of a probe to the presence of a movement that lands in a position lower than that probe is precisely what horizons rule out as an illicit pairing of height and locality. The examples in (350) thus present an apparent problem for the horizons account.

Upon closer scrutiny, however, there is good evidence that both relativization and *wh*-movement can take place over a topicalized element and out of clause in which topicalization has taken place, as predicted by horizons. Consider, for instance, the sentence in (351) from Baltin (1982), in which relativization has taken place in the same clause in which topicalization has applied. Evidently, topicalization does not bleed relativization over it in (351).

(351) He's a man [to whom_{rel} liberty_{top} we could never grant *t t*] (Baltin 1982: 17)



It is also possible for relativization to leave a clause in which topicalization has taken place:²⁴

²⁴ The judgments in this section are due to Ethan Poole.

(352) He's a man [to whom_{rel} I believe [liberty_{top} we could never grant *t t*]]

In addition, Culicover (1996) provides the example in (353), in which *wh*-movement over a topicalized element has taken place and concludes from this that topicalization does not create islands in English:

(353) I was wondering [to what kinds of people]_{wh} [books like these]_{top} you would actually have given *t t* if you had had the chance. (Culicover 1996: 460)

As was the case for relativization, *wh*-movement is likewise possible out of a clause in which topicalization has applied:

(354) [To what kinds of people]_{wh} did she say [(that) [books like these]_{top} she would actually have given *t t* if she had had the chance]?

With Culicover (1996), I conclude that topicalization does *not* create an island for subsequent *wh*-movement or relativization in English, be it for movement within the same clause or into a higher clause. An interesting question that arises is what causes the degradedness of the examples in (350). Culicover (1996) argues that it is related to the fact that the two moving elements in (350) are both DPs, whereas in the wellformed examples in (351)–(354) one moving element is a DP and the other one is a PP. He suggests that identification of the gap position is significantly harder in the case of category identity, and that this is responsible for the perceived ungrammaticality of (350). Once this obstacle is removed, judgments clearly improve.

Moreover, even a category mismatch between the two moving items is not strictly necessary. As Lyn Frazier has pointed out to me, even structures that involve two moving DPs can be quite good. An example is given in (355). (355a) gives that baseline in which only relativization and no topicalization takes place. (355b) adds topicalization. If an appropriate prosody is assigned to the sentences, it is quite acceptable despite the fact that both moving elements are DPs.

(355) a. This is the inquisitive woman [who]_{rel} Zora gave a chapter of her manuscript about quantum physics to *t*.
 b. This is the inquisitive woman [who]_{rel} [her manuscript about quantum physics]_{top} Zora gave a chapter of *t* to *t*.

In (355), relativization and topicalization both apply within the same clause. As (356) shows, it is also possible for relativization to leave a clause in which topicalization has taken place. Both moving elements are again DPs.

(356) This is the inquisitive woman [[who]_{rel} Peter thought [(that) [her manuscript about quantum physics]_{top} Zora had given a chapter of *t* to *t*].

It appears, then, that both *wh*-movement and relativization over a topic is in principle possible even if both elements are DPs. This further supports the conclusion that topicalization does not induce islandhood for further extraction. The severe degradedness of (350) must therefore be attributed to some other source.²⁵ In general, embedded topicalization in and of itself is marked and often dispreferred to begin with. Moreover, embedded topicalization induces a garden path effect because the topicalized element is initially construed as the subject of the lower clause, an analysis that has to be revised when the actual subject is encountered. Furthermore, two unassigned DP fillers are difficult to keep track of quite generally. This is consonant with the existence of *similarity-based interference effects* in language processing (see, e.g., McElree, Foraker & Dyer 2003, Van Dyke & Lewis 2003, Lewis & Vasishth 2005, Lewis, Vasishth & Van Dyke 2006, Bartek, Lewis, Vasishth & Smith 2011). The fact that the type of the DP has an effect on the status of the structures (compare (350) and (356)) is consistent with the role of parsing principles in the account of these structures. In sum, the bad cases of movement over a topic are bad due to a range of factors, including principles of sentence parsing, and once these factors are adequately controlled for, the examples become acceptable. The issue merits a systematic study, which I will not undertake here.

The prediction of the horizons account is thus borne out, despite initial appearance to the contrary: Topicalization does not create islands for relativization and *wh*-movement. This provides additional evidence for the Height–Locality Theorem and horizons, which give rise to it. It also justifies the asymmetry noted at the end of the previous section. While there exists variability in whether projections higher than a probe’s location are opaque to this probe or not, there is no variability with respect to lower projections: These will invariably be transparent. The discussion in this section has demonstrated this to be the case.

²⁵ I am indebted to Lyn Frazier for discussions of the various issues surrounding movement over an embedded topic.

3.5.4 Default-horizons

The discussion in sections 3.5.1–3.5.3 has reached two conclusions. First, horizons impose a restriction on possible height–locality pairings of probes. A probe’s location constrains its possible locality profiles, and a probe’s locality properties constrain its possible structural locations, a consequence formulated as the Height–Locality Theorem in (333). We have seen that this aspect of the account derives on a systematic basis the Height–Locality Connection and hence a core empirical property of selective opacity effects. Second, horizons do not place height and locality in a one-to-one relationship with each other. While one restricts the other, they do not completely determine each other. As discussed in section 5.8 of chapter 5, this is the core difference to previous approaches to the Height–Locality Connection. We have also seen empirical evidence for the limited variability in height–locality pairings allowed by horizons, both within a single language and across different languages. In this section, I will propose that within the range of possible height–locality pairings compatible with horizons, there are distinguished or ‘default’ settings, which can be overridden based on purely positive evidence.

To approach the notion of default horizons, recall from the discussion in section 3.3.2 that the locality effects of horizons stand in an entailment relationship. For any projection Π , if Π is a horizon for a probe $[uF]$, then all projections higher than Π within the same extended projection will also block $[uF]$. This is the Horizon Inheritance Theorem (276), repeated here:

(357) **HORIZON INHERITANCE THEOREM** [=(276)]
Given a probe $[uF]$ and an extended projection $\Phi = \langle \Pi_n > \Pi_{n-1} > \dots > \Pi_1 \rangle$, if $\Pi_m \in \Phi$ is a horizon for $[uF]$, then all projection Π_{m+1}, \dots, Π_n are likewise horizons for $[uF]$ (due to category inheritance (272)).

The Horizon Inheritance Theorem is itself derived from category inheritance (272). One important consequence of Horizon Inheritance discussed at length in section 3.3.2 is that it derives Upward Entailment, one core empirical property of selective opacity.

A second consequence of (357) is that different choices of horizons yield locality profiles that stand in an entailment relationship to each other. To illustrate, (358) states for various choices of horizons which types of embedded clauses are transparent under this choice.

(358) *Locality profiles induced by various horizon settings*

	Horizon			
	v	T	C	∅
CP clauses	-----			} transparent
TP clauses	-----			
vP clauses	} transparent		} transparent	
VP clauses	} transparent		} transparent	

For example, if a probe has C as its horizon, CP clauses will be opaque to it, whereas TP, vP, and VP clauses will be transparent. If the probe has T as its horizon, then CP and TP clauses will be opaque to it. The choice of T as a horizon hence yields a strictly stronger locality restriction than choosing C, a direct consequence of the Horizon Inheritance Theorem. This relationship generalizes: The lower the horizon of a probe, the stricter the locality constraint of that probe.

At the same time, the Height–Locality Theorem (333) imposes a lower limit on the choice of horizons. As detailed in section 3.5.1, no probe can have a projection lower than the one it is located on as its horizon, or else this probe will be vacuous. This part of the Height–Locality Theorem is repeated from (333a) in (359).

(359) *Height–locality implication* [=(333a)]
 Given an extended projection $\Phi = \langle \Pi_n > \Pi_{n-1} > \dots > \Pi_1 \rangle$, for any non-vacuous probe $[uF]$, if $[uF]$ is located on Π_m , then a projection $\in \{ \Pi_{m-1}, \dots, \Pi_1 \}$ cannot be a horizon for $[uF]$.

(359) thus entails the existence of a lower bound on possible horizon settings for a given probe. In general, for a probe located on a projection Π_i , this projection Π_i is also the lowest possible horizon for this probe. The facts that (i) lower horizons give rise to more restrictive locality patterns, and (ii) there exists a lower bound on the choice of horizon gives rise to the consequence that for each probe there is a choice of horizon that is maximally restrictive without rendering this probe vacuous. This choice is the projection that the probe is itself located on. For reasons to be made clear immediately, I will refer to this choice of Horizon as a default horizon:

(360) **DEFAULT HORIZON**
 For any probe $[uF]$ located on head X^0 , the default horizon is $[uF]_{X^0} \dashv \dashv X$.

The setting $[uF]_{X^0} \dashv\vdash X$ represents the strictest possible locality constraint on $[uF]$. Any projection higher than X would result in a more permissive locality profile (because then XPs would be transparent to $[uF]$), and any projection lower than X would render $[uF]$ vacuous.

It has been common since at least Chomsky (1995b, 2000, 2001, 2005) to view grammar as structured in a way that minimizes computational burden, presumably for the purposes of language processing. To give just one prominent example of this reasoning, Chomsky (2000, 2001, 2005) argues that the cyclic Spell-Out model of phase theory has its underlying motivation in computational considerations. On this view, cyclic removal of built syntactic structure is utilized by the grammar because it minimizes the amount of structure in the workspace and the length of syntactic dependencies. If we adopt this general outlook, then the default horizons in (360) receive a privileged role precisely because they minimize to the greatest extent possible the search space of a probe. Thus, consider two probes $[uF_1]$ and $[uF_2]$ with $[uF_1] \dashv\vdash T$ and $[uF_2] \dashv\vdash C$. In the case of an embedded TP clause, the structure that $[uF_1]$ will have to search through comprises only the higher clause, as the embedded TP is opaque to this probe. $[uF_2]$'s search space, on the other hand, encompasses the embedded clause as well and is hence strictly larger than that of $[uF_1]$. Choice of T as an horizon thus minimizes the amount of structure that a probe needs to traverse. In this sense, the default horizons in (360) are the most economical setting, as they have the greatest impact on restricting probing.

The view that the setting $[uF]_{X^0} \dashv\vdash X$ has a privileged meta-grammatical status gains plausibility from the fact that three out of the four probes in Hindi instantiate the default setting. The only probe that does not adhere to (360) is $[u\bar{A}]$. Direct evidence for this horizon setting of $[u\bar{A}]$ comes from the possibility of \bar{A} -movement out of finite clauses.

- (361) *Hindi probes and their horizons* [=(282)]
- a. $[uA]_{T^0} \dashv\vdash T$
 - b. $[u\phi]_{T^0} \dashv\vdash T$
 - c. $[uwh]_{C^0} \dashv\vdash C$
 - d. $[u\bar{A}]_{C^0} \dashv\vdash \emptyset$

It is plausible, then, to adopt the following picture: Certain height–locality pairings are categorically ruled out by the Height–Locality Theorem (333). Within the range of options allowed by the Height–Locality Theorem, the pairing in (360) receives a privileged meta-grammatical status.

One might now speculate that the default setting in (360) also plays a role in the acquisition of selective opacity. Suppose that the default setting is what children assume in the absence of evidence to the contrary. Because, as just discussed, this setting is the most restrictive possible setting, any divergence from it may then be acquired on the basis of strictly positive evidence, i.e., by observing a well-formed dependency that this setting would rule out. If a probe is located in T^0 , for example, the default assumption would be that T is a horizon for this probe, meaning that this probe cannot enter into a dependency with material inside TPs and CPs. As we have seen, this is the case of Hindi A-movement and ϕ -agreement. Alternatively, a child may encounter in its input dependencies between this probe and elements inside a TP. In the face of such evidence, the child would then minimally adjust the Horizon upward, to C . This would produce the locality of, e.g., English A-movement. Finally, a child exposed to wellformed dependencies into CP clauses. In this case, the child would again upward adjust the horizon to accommodate the input and remove all horizons from this probe. This would produce superraising in languages that exhibit it (see the discussion in section 3.5.2).²⁶

On this view, the absence of superraising would constitute the default setting and thus not require explicit stipulation. Because this default setting is restrictive, departures from it, and thus crosslinguistic differences in horizons, could be acquired on the basis of strictly positive evidence. This proposed acquisition procedure thus has the additional benefit that it would explain how locality differences between operations and hence selective opacity are acquired in the absence of negative evidence, thus obviating a poverty of the stimulus problem.

²⁶ An interesting consequence arising from default horizons that Rajesh Bhatt (p.c.) has made me aware of is that on the assumption that intermediate movement steps are also triggered by probes (see the discussion in chapter 6), there can be no unbounded successive-cyclic movement if all probes have their default horizon. This is because a probe that has a default horizon is only able to search into clauses properly smaller than the clause that the probe is located in. Each intermediate landing site then entails a reduction in clause size and therefore a finite number of embeddings. The desirability of unbounded successive-cyclic dependencies may thus provide another rationale for observed departures from default horizons.

3.6 Chapter summary

This chapter has proposed and developed an account of selective opacity effects and its crosslinguistic and crossconstructional properties. I have taken as my point of departure the conclusion of chapter 2 that selective opacity effects are not limited to the domain of movement, but in fact encompass in-situ relations like ϕ -agreement as well. I have taken this fact to show that selective opacity is more abstract than previously thought in that it instantiates a constraint on the operation Agree. Because Agree is involved in the establishment of both movement and genuine long-distance relations, this view enables a unified account of selective opacity across these two domains.

The account proposed in this section pursues the view that locality differences in syntax are just that – differences in the locality profiles of different processes. One of the main conclusions of this thesis is thus that locality is not uniform across different processes in all cases. Rather, the locality profile of various operations may be distinct, a conclusion with far-reaching consequences for the study of syntactic locality quite generally.

I have explored the intuition that certain syntactic nodes can lead a probe to terminate its search, effectively delimiting that probe's search space. As a result, the probe will be unable to enter into an Agree relation with any element dominated by that node. Opacity is the result. A crucial aspect of this proposal is that probes differ in what category features they are sensitive to. This induces locality differences between probes and hence selective opacity. Developing this intuition in greater detail, I have introduced the notion of an 'horizon', a category feature that induces probe termination. Because of its generality, this system is able to capture not only locality differences between a wide range of operations, but also interactions between various operations. I have illustrated the system by developing an account of the various movement and non-movement operations and their locality profiles and interactions in Hindi. I have also applied the system to superraising in English, focusing also on a number of generalizations that traditional accounts are ill-equipped to accommodate due to their item-based nature. Finally, I proposed that differences between the locality of movement and agreement that have been occasionally observed in the previous literature are not indicative of a deep syntactic difference between the two, but are manifestations of the same principle that underlies locality differences within the class of movement or agreement dependencies. In other words, they are nothing but general selective opacity effects and I have

demonstrated that the horizons account developed here straightforwardly extends to these cases as well, in virtue of being abstract enough to constrain movement and agreement. While I have developed these accounts assuming a standard clause structure comprising C, T, *v*, and V, the results reached here apply without modification to expanded clause structures.

The key distinctive features of the horizon account of selective opacity discussed in this chapter are as follows. For a comprehensive comparison of horizons to previous approaches to selective opacity, see chapter 5.

- (362)
- a. Horizons involve a constraint on Agree, not movement.
 - b. There is no direct interactions between movement types.
 - c. Horizons are domain-based in that they exclusively invoke probes and their horizons. The moving item itself is irrelevant.
 - d. Horizons extend beyond the binary A/\bar{A} -distinction.
 - e. Horizons extend to selective opacity in non-identity cases (smuggling and remnant movement).
 - f. Horizons extend to locality differences between movement and agreement.
 - g. Horizons derive Upward Entailment and a version of the Height–Locality Connection.

Notably, in addition to providing an account of an intricate range of facts that poses substantial puzzles for traditional accounts, horizons also manage to derive the meta-generalizations of selective opacity discussed in chapters 1 and 2, namely Upward Entailment and the Height–Locality Connection. I have proposed that Upward Entailment follows from independently motivated properties of extended projections, according to which syntactic labels are inherited through an extended projection. Moreover, I have shown how this view also derives a version of the Height–Locality Connection essentially for free, because it imposes a *de facto* constraint on possible pairings of a probe’s location and its horizon, hence its locality. The gist of this line of explanation is that only probes for which height and locality stand in a particular relationship to each other are able to trigger any movement and agreement operation. As a consequence, all movement and agreement dependencies exhibit an empirical link between height and locality. I have shown how this account succeeds in deriving various aspects of the Hindi system as necessary consequences of horizons.

The analysis, then, not only offers an account of selective opacity effects of various types, it also imposes systematic and non-stipulatory constraints on possible selective opacity patterns.

It bears reiterating that the account does not impose stipulations on possible and impossible horizon settings *per se*. That is, as far as narrow-syntactic principles are concerned, any probe can have any category feature as its horizon. The observation that selective opacity is nonetheless not distributed randomly, but subject to Upward Entailment and the Height–Locality Connection are purely *emergent* properties of this system. Upward Entailment is derived from a distributional constraint on category features, which form horizons. The Height–Locality Connection follows because certain pairings of height and locality, while admissible, give rise to a vacuous probe. Vacuous probes will never enter into movement or agreement dependencies and hence be invisible in the output of the system. Thus, Upward Entailment and the Height–Locality Connection are not axioms of the system but descriptions of the output that the system produces. Both are hence derived.

CHAPTER 4

SELECTIVE OPACITY IN GERMAN

4.1 Introduction

In the previous chapter, I have proposed and developed a general account of selective opacity effects that is based on the notion of horizons. I have applied this system to the selective opacity patterns of Hindi as well as improper movement more generally and locality differences between movement and agreement. I have also shown how horizons derive Upward Entailment and the Height–Locality Connection, two pervasive meta-generalizations of selective opacity effects across constructions and languages.

In this chapter, I will present and analyze selective opacity effects in German. Like Hindi, German does not just exhibit a simple binary distinction in locality, but opacity *layering*. I will discuss four movement types, each with their own locality profile. Intriguingly, however, these four locality profiles are not random, but conform to Upward Entailment and the Height–Locality Connection. They are thus subject to the entailment relationships and constraints observed in the preceding three chapters.

Investigating movement locality in German provides further support for several of the conclusions reached in the preceding chapters. First, it provides evidence for opacity layering, Upward Entailment, and the Height–Locality Connection as systematic properties of selective opacity effects. Second, it demonstrates how the horizon-based account proposed in chapter 3 carries over to another complex set of locality patterns. While the individual selective opacity patterns discussed

in this chapter have been noted and analyzed in the previous literature on German, virtually all previous accounts have focused on only a subpart of the pertinent generalizations and accounts devised for one instance of selective opacity typically fail to extend to other ones. They fail to capture the overarching patterns that underlie the various locality mismatches. The horizon-based approach, on the other hand, offers a comprehensive and uniform account to selective opacity in German that is flexible and general enough to accommodate the full range of data discussed here and to capture the overarching generalizations of selective opacity. Finally, the horizon-based account is a restrictive account of selective opacity because it imposes limits on possible selective opacity patterns (see section 3.5) and therefore derives various properties of the selective opacity patterns in German that remain elusive on traditional approaches.

I will focus here on four movement operations, all of which differ in their locality profiles. The first operation of interest is **SCRAMBLING**, which reorders element in the middle field ('Mittelfeld') of a clause, as in (363). The middle field is the region of a German clause between the complementizer and the finite verb in subordinate clauses and between the finite verb and nonfinite verbal elements in main clauses. See, e.g., Grewendorf & Sternefeld (1990) and Corver & van Riemsdijk (1994) for overviews of the properties of scrambling in German and analytical approaches to it.

- (363) a. dass keiner das Buch gelesen hat
 that nobody the book read has
 'that no one read the book'
- b. dass das Buch_i keiner t_i gelesen hat
 that the book nobody read has
 'that no one read the book'

The second movement type of interest is **RELATIVIZATION**, illustrated in (364):

- (364) das Buch [das_i keiner t_i gelesen hat]
 the book that nobody read has
 'the book that no one read'

Third, I will consider **WH-MOVEMENT IN VERB-FINAL CLAUSES**, that is, *wh*-movement in embedded clauses in which the finite verb appears at the end of the sentence:

- (365) Ich weiß nicht [welches Buch_i keiner t_i gelesen hat].
 I know now which book nobody read has
 ‘I don’t know which book no one read.’

The fourth and final descriptive movement type of interest here is movement to the preverbal position in a V2 clause. I will follow common practice and refer to this movement as **TOPICALIZATION**. Topicalization can target a variety of constituents, like the direct object in (366a). Moreover, *wh*-elements can also appear in this position, as shown in (366b).

- (366) a. Dieses Buch_i hat keiner t_i gelesen.
 this book has no one read
 ‘No one read this book.’
 b. Welches Buch_i hat keiner t_i gelesen?
 which book has no one read
 ‘Which book did no one read?’

Following common practice, I will use the term ‘topicalization’ as a convenient label for movement to the preverbal position in a V2 clause. The term is somewhat misleading as this movement does not actually induce a topic interpretation (but it is of course compatible with one). No commitment with respect to the information-structural effects of this movement are intended.

Although the movement to the preverbal position in (366a,b) plausibly instantiates the same movement, I will descriptively refer to structures like (366b) as **WH-MOVEMENT INTO V2 CLAUSES** or as **WH-TOPICALIZATION**. I employ this terminology mainly for expository reasons. As we will see, *wh*-movement in a V-final clause as in (365) exhibits locality properties that are somewhat different from *wh*-movement in a V2 clause such as (366b). Thus, there is no uniform locality of *wh*-movement in German. Rather, the type of the clause (V-final vs. V2) makes a difference. To highlight this curious empirical fact, I will terminologically distinguish between the two types of *wh*-movement. At this point, this terminological choice is not to be taken as an analytical commitment.

This chapter will proceed as follows: Section 4.2 presents the relevant empirical evidence, focusing on the transparency and opacity properties of various types of embedded clauses for the four movement types just introduced. I will focus on coherent and non-coherent infinitives

(section 4.2.1), V-final finite clauses (section 4.2.2), and V2 clauses (section 4.2.3). This section will also provide evidence that the internal clause structure differs between these four clause types. Section 4.3 then demonstrates that the empirical patterns motivated in section 4.2 conform to opacity layering, Upward Entailment, and the Height–Locality Connection. Section 4.4 then presents the core account in terms of horizons. Section 4.5 then provides evidence that the notion of horizons coexists with that of phases. The two have distinct syntactic effects, with horizons determining whether a syntactic relation across a given clause boundary is possible or not and phases mandating that possible extraction be successive-cyclic. As we will see, both concepts are necessary. Section 4.6 then demonstrates the effects of the Height–Locality Theorem in German and shows how the horizon-based account derives various properties of the selective opacity patterns, imposing limits on the relation between a probe’s structural location and its locality properties. Finally, the appendix in section 4.7 briefly discusses an additional instance of selective opacity argued for in some of the previous literature, namely between topicalization and *wh*-movement into a V2 clause. I will argue against such a distinction.

4.2 Clause types and their locality profiles

In this section, I will lay out the locality properties of the four movement types just mentioned. I will do so by considering four types of embedded clauses: (i) coherent infinitives, (ii) non-coherent infinitives, (iii) V-final finite clauses, and (iv) V2 clauses. I proceed by demonstrating for each type of embedded clause whether this clause is transparent or opaque for each movement type. An overview of the empirical generalizations that will be reached is provided in (367).

(367) *Preview: Selective opacity in German*

Operation	probe location	Size of embedded clause			
		nonfinite		finite	
		coherent (vP)	non-coherent (TP)	V-final (CP)	V2 (ForceP)
scrambling	T ⁰	✓	✗	✗	✗
relativization	C ⁰	✓	✓	✗	✗
<i>wh</i> -movement into V-final clause	C ⁰	✓	✓	✓	✗
topicalization/ <i>wh</i> -movement into V2 clause	Force ⁰	✓	✓	✓	✓

(✓ : transparent; ✗ : opaque)

4.2.1 Coherent and non-coherent infinitives

In his seminal work on German infinitival constructions, Bech (1955/1957) identified two classes of infinitives in German, which are standard labeled **COHERENT** (“*kohärent*”) and **NON-COHERENT** (“*inkohärent*”). Coherent infinitives pattern like monoclausal structures in that they are transparent for a number of otherwise clausebounded processes, whereas non-coherent infinitives are opaque to such processes. The two infinitival constructions are superficially identical, so that the difference between the two classes has no surface morphological manifestation. Verbs differ in which class of infinitives they select for, with some verbs occurring with both classes. Verbs that at least optionally embed coherent infinitives are *versuchen* ‘try’ and *vergessen* ‘forget’. Verbs whose infinitival complement clause is invariably non-coherent include *ablehnen* ‘refuse’, *anbieten* ‘offer’, *behaupten* ‘claim’, and *ankündigen* ‘announce’. See Wurmbrand (2001: 327–340) for a list of verbs and illustrations. Two examples of are provided in (368) and (369).

(368) *Coherent infinitives*

- a. weil der Mechaniker [den Wagen zu reparieren] vergessen hat
 because the mechanic.NOM the car.ACC to repair forgotten has
 ‘because the mechanic forgot to repair the car’

- b. weil sie [der Sache auf den Grund zu gehen] versucht hat
 because she the matter.DAT to the bottom to go tried has
 ‘because she tried to get to the bottom of things’

(369) *Non-coherent infinitives*

- a. weil der Mechaniker [den Wagen zu reparieren] abgelehnt hat
 because the mechanic.NOM the car.ACC to repair refused has
 ‘because the mechanic refused to repair the car’
- b. weil sie [der Sache auf den Grund zu gehen] angekündigt hat
 because she the matter.DAT to the bottom to go announced has
 ‘because she announced that she would get to the bottom of things’

Despite their superficial similarity, the two classes differ with respect to a number of syntactic properties. Following Bech’s (1955/1957) study, the two classes of infinitivals have received a great deal of attention in the literature and a wide variety of theoretical approaches has been explored (see, e.g., Haegeman & van Riemsdijk 1986, von Stechow & Sternefeld 1988, Fanselow 1989, Haider 1993, 2003, 2010, Bayer & Kornfilt 1990, von Stechow 1990, Grewendorf & Sabel 1994, Wurmbrand 2001, Hinterhölzl 2006, Lee-Schoenfeld 2007 among many others). Also see Schmid, Bader & Bayer (2005) for an experimental investigation.

It is impossible to do justice to this rich body of literature here. Because my primary concern here is the locality constraints governing the various movement types in German, I will confine my attention to the locality properties of the two classes of infinitives with regards to movement types.

4.2.1.1 Scrambling

A well-known difference between the two classes of infinitives is that coherent infinitives allow scrambling out of them, while non-coherent infinitives do not. This is illustrated in (370) and (371), where scrambling of the embedded object over the matrix subject takes place:¹

(370) *Coherent infinitives: Scrambling possible*

- a. weil den Wagen_i jemand [_{t_i} zu reparieren] vergessen hat
 because the car.ACC someone.NOM to repair forgotten has
 ‘because someone forgot to repair the car’

¹ The matrix subject is *jemand* ‘someone’ in (370) and (371) because scrambling of an indefinite object over of a definite subject is generally much better than scrambling of a definite object over a definite subject.

- b. weil der Sache_i jemand [t_i auf den Grund zu gehen] versucht
 because the matter.DAT someone.NOM to the bottom to go tried
 hat
 has
 ‘because someone tried to get to the bottom of things’

(371) *Non-coherent infinitives: Scrambling impossible*

- a. ?*weil den Wagen_i jemand [t_i zu reparieren] abgelehnt hat
 because the car.ACC someone.NOM to repair refused has
Intended: ‘because someone refused to repair the car’
- b. ?*weil der Sache_i jemand [t_i auf den Grund zu gehen]
 because the matter.DAT someone.NOM to the bottom to go
 angekündigt hat
 announced has
Intended: ‘because someone announced that they would go to the bottom of things’

This contrast between coherent and non-coherent infinitives with respect to scrambling may also be illustrated through remnant movement of the embedded clause to the clause-initial position. In (372) and (373), the embedded object is not fronted within the embedded clause and therefore must have been scrambled out of it. Such scrambling is possible only for coherent infinitives:²

(372) *Coherent infinitives: Scrambling possible*

- a. [t_i Zu reparieren]_j hat der Mechaniker den Wagen_i t_j vergessen.
 to repair has the mechanic.NOM the car.ACC forgotten
 ‘The mechanic forgot to repair the car.’

² Unsurprisingly, if the embedded object remains inside the lower clause that is being fronted, the sentence in (373) become grammatical:

- (i) a. [Den Wagen zu reparieren]_j hat der Mechaniker t_j abgelehnt.
 the car to repair has the mechanic refused
 ‘The mechanic refused to repair the car.’
- b. [Der Sache auf den Grund zu gehen]_j hat sie t_j angekündigt.
 the matter.DAT to the bottom to go has she announced
 ‘She announced that she would get to the bottom of things.’

This makes it clear that it is not the fronting of the embedded clause that causes the ungrammaticality of (373), but rather the scrambling out of it.

- b. [t_i Auf den Grund zu gehen] $_j$ hat sie der Sache $_i$ t_j versucht.
 to the bottom to go has she the matter.DAT tried
 ‘She tried to get to the bottom of things.’

(373) *Non-coherent infinitives: Scrambling impossible*

- a. ?*[t_i Zu reparieren] $_j$ hat der Mechaniker den Wagen $_i$ t_j abgelehnt.
 to repair has the mechanic.NOM the car.ACC refused
Intended: ‘The mechanic refused to repair the car.’
- b. ?*[t_i Auf den Grund zu gehen] $_j$ hat sie der Sache $_i$ t_j angekündigt.
 to the bottom to go has she the matter.DAT announced
 ‘She announced that she would get to the bottom of things.’

A parallel demonstration comes from extraposition of the embedded clause. Extraposition of a remnant embedded clause created by scrambling of the embedded object out of it is possible with coherent infinitives but impossible with non-coherent ones.³

(374) *Coherent infinitives: Scrambling possible*

- a. weil der Mechaniker den Wagen $_i$ t_j vergessen hat [t_i zu reparieren] $_j$
 because the mechanic.NOM the car.ACC forgotten has to repair
 ‘because the mechanic forgot to repair the car’
- b. weil sie der Sache $_i$ t_j versucht hat [t_i auf den Grund zu gehen] $_j$
 because she the matter.DAT tried has to the bottom to go
 ‘because she announced that she would get to the bottom of things’

³ This configuration is often referred to as the *third construction* (“*dritte Konstruktion*”) in the literature on German, a term originating with den Besten & Rutten (1989). It is discussed and analyzed in great detail by Wöllstein-Leisten (2001). As before (see fn. 2), extraposition is possible in (375) in the absence of scrambling:

- (i) a. weil der Mechaniker t_j abgelehnt hat [den Wagen zu reparieren] $_j$
 because the mechanic refused has the car to repair
 ‘because the mechanic refused to repair the car’
- b. weil sie t_j angekündigt hat [der Sache auf den Grund zu gehen] $_j$
 because she announced has the matter.DAT to the bottom to go
 ‘because she announced that she would get to the bottom of things’

(375) *Non-coherent infinitives: Scrambling impossible*

- a. ?*weil der Mechaniker den Wagen_i t_j abgelehnt hat [t_i zu reparieren]_j
because the mechanic.NOM the car.ACC refused has to repair
Intended: ‘because the mechanic refused to repair the car’
- b. ?*weil sie der Sache_i t_j angekündigt hat [t_i auf den Grund zu gehen]_j
because she the matter.DAT announced has to the bottom to go
Intended: ‘because she announced that she would get to the bottom of things’

A final distinction between coherent and non-coherent infinitives to be discussed here involves scope. As (376) demonstrates, a quantificational element in the nonfinite clause can take either embedded or matrix scope in a coherent infinitive. This contrasts with the non-coherent infinitives in (377), in which such an element is confined to embedded scope.

(376) *Coherent infinitive: Matrix or embedded scope*

- a. weil der Mechaniker keinen Wagen zu reparieren vergessen hat
because the mechanic no car to repair forgotten has
‘because the mechanic forgot to repair no car’ (forget > $\bar{\#}$; $\bar{\#}$ > forget)
- b. weil sie nur einer einzigen Sache auf den Grund zu gehen versucht
because she only a single matter.DAT to the bottom to go tried
hat
has
‘because she tried to get to the bottom of only one affair’
(try > only one; only one > try)

(377) *Non-coherent infinitive: Only embedded scope*

- a. weil der Mechaniker keinen Wagen zu reparieren abgelehnt hat
because the mechanic no car to repair refused has
‘because the mechanic refused to repair no car’ (refuse > $\bar{\#}$; ?* $\bar{\#}$ > refuse)
- b. because sie nur einer einzigen Sache auf den Grund zu gehen
because she only a single matter.DAT to the bottom to go
angekündigt hat
announced has
‘because she announced that she would get to the bottom of only one affair’
(announce > only one; ?*only one > announce)

The scope contrast between (376) and (377) is arguably just another manifestation of the disparity with respect to scrambling just observed. Frey (1993) and Krifka (1998) argue that scope in German is governed by the principle in (378). This principle only applies to overt movement, hence applies at S-structure in GB terminology:

- (378) If α, β are operators occurring in a sentence S, then S has reading in which α has scope over β if and only if:
- a. α c-commands β , or
 - b. α c-commands a trace of β . (Krifka 1998: 76)

If coherent infinitives are transparent for scrambling, the scope ambiguity in (376) follows from (378) and chain-vacuous movement of the negative element into the matrix clause, hence above the matrix predicate. No such scrambling is possible out of the non-coherent infinitives in (377), and only narrow scope with respect to the matrix verb is possible as a result.

In sum, coherent infinitives allow scrambling out of them, whereas non-coherent ones do not. The literature on infinitival constructions in German has identified additional domains in which the two constructions differ, but I will not discuss these here. Comprehensive overviews from different theoretical perspectives are provided by Wurmbrand (2001), Lee-Schoenfeld (2007), Haider (2010) and the references cited there.

4.2.1.2 Relativization

While scrambling is sensitive to the difference between the two types of infinitives, relativization is not. Thus, relativization is possible out of a non-coherent infinitive, just like it is out of a coherent one (Lee-Schoenfeld 2007: 14, Bayer & Salzmann 2013: 311, Müller 2014b: 130–131).⁴

⁴ Bayer & Salzmann (2013) provide the example of a non-coherent infinitive (i), in which the embedded clause is extraposed:

(i) das Buch [das_i ich erwartet habe [t_i geschenkt zu bekommen]]
 the book which I expected have given to get
 ‘the book I expected to be given as a present’ (Bayer & Salzmann 2013: 311)

(379) *Coherent infinitives: Relativization possible*

- a. der Wagen [den_i der Mechaniker [_{t_i} zu reparieren] vergessen hat]
the car which the mechanic to repair forgotten has
'the car which the mechanic forgot to repair'
- b. die Sache [der_i sie [_{t_i} auf den Grund zu gehen] versucht hat]
the matter which.DAT she to the bottom to go tried has
'the matter that she tried to get to the bottom of'

(380) *Non-coherent infinitives: Relativization possible*

- a. der Wagen [den_i der Mechaniker [_{t_i} zu reparieren] abgelehnt hat]
the car which the mechanic to repair refused has
'the car which the mechanic refused to repair'
- b. die Sache [der_i sie [_{t_i} auf den Grund zu gehen] angekündigt hat]
the matter which.DAT she to the bottom to got announced has
'the matter that she announced that she would get to the bottom of'

Relativization is also possible if the embedded clause is extraposed, again in contrast to scrambling.

(381) *Extraposed coherent infinitives: Relativization possible*

- a. der Wagen [den_i der Mechaniker _{t_j} vergessen hat [_{t_i} zu reparieren]_j]
the car which the mechanic forgotten has to repair
'the car which the mechanic forgot to repair'
- b. die Sache [der_i sie _{t_j} versucht hat [_{t_i} auf den Grund zu gehen]_j]
the matter which.DAT she tried has to the bottom to go
'the matter that she tried to get to the bottom of'

(382) *Extraposed non-coherent infinitives: Relativization possible*

- a. der Wagen [den_i der Mechaniker _{t_j} abgelehnt hat [_{t_i} zu reparieren]_j]
the car which the mechanic refused has to repair
'the car which the mechanic refused to repair'
- b. die Sache [der_i sie _{t_j} angekündigt hat [_{t_i} auf den Grund zu
the matter which.DAT she announced has to the bottom to
gehen]_j]
go
'the matter that she announced that she would get to the bottom of'

4.2.1.3 *Wh*-movement into V-final clause

Wh-movement into a V-final clause is likewise possible out of both classes of infinitives (e.g., Müller 1998).

(383) *Coherent infinitives: Wh-movement possible*

- a. Ich weiß nicht [welchen Wagen_i der Mechaniker [t_i zu reparieren] vergessen
I know not which car the mechanic to repair forgotten
hat].
has
'I don't know which car the mechanic forgot to repair.'
- b. Ich weiß nicht [welcher Sache_i sie [t_i auf den Grund zu gehen]
I know not which matter.DAT she to the bottom to go
versucht hat].
tried has
'I don't know which matter she tried to get to the bottom of.'

(384) *Non-coherent infinitives: Wh-movement possible*

- a. Ich weiß nicht [welchen Wagen_i der Mechaniker [t_i zu reparieren] abgelehnt
I know not which car the mechanic to repair refused
hat].
has
'I don't know which the mechanic refused to repair.'
- b. Ich weiß nicht [welcher Sache_i sie [t_i auf den Grund zu gehen]
I know not which matter.DAT she to the bottom to go
angekündigt hat].
announced has
'I don't know which matter she announced that she would get to the bottom of.'

4.2.1.4 Movement into V2 clause

Finally, the distinction between coherent and non-coherent infinitives is also irrelevant for movement into a V2 clause. The examples in (385) and (386) illustrate this for movement of a *wh*-element.⁵

⁵ Müller (1998) demonstrates the difference between scrambling and the two kinds of *wh*-movement for non-coherent infinitives using the following triplet:

- (385) *Coherent infinitives: Wh-movement into V2 clause possible*
- a. Welchen Wagen_i hat der Mechaniker [_{t_i} zu reparieren] vergessen?
 which car has the mechanic to repair forgotten
 ‘Which car did the mechanic forget to repair?’
- b. Welcher Sache_i hat sie [_{t_i} auf den Grund zu gehen] versucht?
 which matter.DAT has she to the bottom to go tried
 ‘Which matter did she try to get to the bottom of?’
- (386) *Non-coherent infinitive: Wh-movement into V2 clause possible*
- a. Welchen Wagen_i hat der Mechaniker [_{t_i} zu reparieren] abgelehnt?
 which car has the mechanic to repair refused
 ‘Which car did the mechanic refuse to repair?’
- b. Welcher Sache_i hat sie [_{t_i} auf den Grund zu gehen] angekündigt?
 which matter.DAT has she to the bottom to go announced
 ‘Which matter did she announced that she would go to the bottom of?’

Furthermore, the examples in (387) and (388) show that the same holds for fronting of non-*wh* elements:

-
- (i) a. *Scrambling*
 ?*dass das Buch_i keiner [_{t_i} zu lesen] abgelehnt hat
 that the book someone to read refused has
Intended: ‘that nobody refused to read the book’
- b. *Wh-movement into V2 clause*
 Was_i hat keiner [_{t_i} zu lesen] abgelehnt?
 what has someone to read refused
 ‘What has someone refused to read?’
- c. *Wh-movement into V-final clause*
 Ich weiß nicht [wen_i er [_{t_i} zu küssen] abgelehnt hat].
 I know not who he to kiss refused has
 ‘I don’t know how he refused to kiss.’ (Müller 1998: 17,299,304)

In addition, Müller (2014b) gives the example in (ii) for relativization out of a non-coherent infinitive:

- (ii) Das ist ein Buch [das_i ich abgelehnt habe [_{t_i} zu kaufen]]
 this is a book that I refused have to buy (Müller 2014b: 131)

(387) *Coherent infinitives: Topicalization possible*

- a. Diesen Wagen_i hat der Mechaniker [_{t_i} zu reparieren] vergessen.
this car has the mechanic to repair forgotten
'This tractor, the mechanic forgot to repair.'
- b. Dieser Sache_i hat sie [_{t_i} auf den Grund zu gehen] versucht.
this matter.DAT has she to the bottom to go tried
'This matter, she tried to get to the bottom of.'

(388) *Non-coherent infinitive: Topicalization possible*

- a. Diesen Wagen_i hat der Mechaniker [_{t_i} zu reparieren] abgelehnt.
this car has the mechanic to repair refused
'This car, the mechanic refused to repair.'
- b. Dieser Sache_i hat sie [_{t_i} auf den Grund zu gehen] angekündigt.
this matter has she to the bottom to go announced
'This matter, she announced that she would get to the bottom of.'

In sum, the distinction between coherent and non-coherent infinitives affects the possibility of scrambling, but has no impact on the other three movement types considered here.

(389) *Locality of infinitival clauses*

Coherent infinitives allow scrambling out of them, non-coherent infinitives do not. Both types of infinitives are transparent to relativization, *wh*-movement into a V-final clause, and movement into a V2 clause.

4.2.1.5 The structure of infinitival clauses

As for the structure of the two classes of infinitives, I will follow here the lead of Tappe (1984) and Wurmbrand (2001), among others, in treating them as structurally deficient. There is, for instance, no evidence that either type of infinitive contains a CP layer, because both lack properties associated with C⁰. First, they do not license embedded *wh*-scope. Infinitival questions are altogether impossible in German, regardless of the embedding predicate. This is illustrated in (390a) the verb *vergessen* 'forget', which embeds a coherent infinitive. (390b) illustrate the same fact for the verb *entscheiden* 'decide', which embeds a non-coherent infinitive.⁶

⁶ The two sentences in (390) are grammatical on a non-*wh* interpretation of *wen* 'who', in which case they mean *He forgot/decided to call someone*. This interpretation is irrelevant for our concerns.

- (390) a. *Er hat vergessen [wen anzurufen].
 he has forgotten who to.call
Intended: ‘He forgot who to call.’
- b. *Er hat entschieden [wen anzurufen].
 he has decided who to.call
Intended: ‘He decided who to call.’

Note that the absence of an embedded *wh*-scope in (390) is not due to a general incompatibility between the embedded predicates and an embedded question. If the embedded clause is finite, embedded question interpretations are readily available:

- (391) a. Er hat vergessen [wen er anrufen sollte].
 he has forgotten who he call should
 ‘He forgot who he was supposed to call.’
- b. Er hat entschieden [wen er anruft].
 he has decided who he calls
 ‘He decided who he will call.’

In addition to the absence of embedded *wh*-scope, infinitival constructions in German are categorically incompatible with the presence of a complementizer like *dass* ‘that’:

- (392) a. dass der Mechaniker [(**dass*) den Wagen zu reparieren] vergessen hat
 that the mechanic that car to repair forgotten has
 ‘that the mechanic forgot (**that*) to repair the car’
- b. dass der Mechaniker [(**dass*) den Wagen zu reparieren] abgelehnt hat
 that the mechanic that the car to repair refused has
 ‘that the mechanic refused (**that*) to repair the car’

There is, moreover, no equivalent of the English *for*, which is often treated as an infinitival C⁰ head. I conclude from the complete absence of any evidence of a CP layer that infinitival clauses in German are obligatorily CP-less. This conclusion applies to coherent and non-coherent infinitives equally. See Tappe (1984) for the same conclusion and a discussion of apparent counterevidence.⁷

⁷ Note, however, that this is not the traditional analysis of the non-coherent infinitives in German, which are often taken to be CPs in the literature on German. An alternative account compatible with the basic approach taken here would be to assume an extended CP structure, with non-coherent infinitives being pruned lower than finite clauses.

Wurmbrand (2001) argues that coherent infinitives are restructuring infinitives and that they contain less functional clause structure than non-coherent (non-restructuring) infinitives (also see the references listed in Wurmbrand 2001: 15 and Wurmbrand 2014). In particular, she argues that coherent infinitives are tenseless, whereas non-coherent ones are not. As evidence for this claim, she puts forth the observations that future-oriented adverbials are impossible in the complement of verbs like *versuchen* ‘try’ (see Cremers 1983 for related observations in Dutch).

- (393) a. *Hans versuchte [Maria in zwei Monaten zu besuchen].
 Hans tried Maria in two months to visit
 Intended: ‘Hans tried to visit Maria in two months.’
- b. Hans hat versucht (*morgen) zu verreisen.
 Hans has tried tomorrow to go.on.a.trip
 ‘Hans has tried to go on a trip (*tomorrow).’ (Wurmbrand 2001: 71,73)

The same restriction holds for the verb *vergessen* ‘forget’:⁸

- (394) Hans hat vergessen [(*)morgen] Maria zu besuchen]
 Hans has forgotten tomorrow Maria to visit
 ‘Hans forgot to visit Mary (*tomorrow).’

⁸ The verb *vergessen* ‘forget’ is also compatible with a finite complement clause. In this case, a future-oriented adverb is possible:

- (i) a. Hans hat vergessen, dass er morgen Maria besucht.
 Hans has forgotten that he tomorrow Maria visits
 ‘Hans forgot that he will visit Maria tomorrow.’
- b. Hans hat vergessen, dass er morgen Maria besuchen muss.
 Hans has forgotten that he tomorrow Maria visit must
 ‘Hans forgot that he must visit Maria tomorrow.’

The impossibility of (394) is hence not simply due to the semantics of *vergessen*.

A qualification that Wurmbrand (2001) points out is that a future-oriented adverb inside the infinitival clause does become available if the matrix clause appears in the future tense. The contrast is illustrated in the pair in (ii):

- (ii) a. Hans hat versucht (*morgen) zu verreisen.
 Hans has tried tomorrow to go.on.a.trip.
- b. ?Hans wird versuchen morgen zu verreisen.
 Hans will try tomorrow to go.on.a.trip (Wurmbrand 2001: 78)

Wurmbrand (2001) argues that this fact is compatible with her conclusion that restructuring infinitives are tenseless, because the tense of these infinitives is determined by the matrix tense. This is in line with her observation that the only reading available for (ii.b) is a simultaneous reading. Since nothing hinges on this complication, I will put it aside here. See Wurmbrand (2001, 2014) for discussion.

With verbs whose infinitival complement clause is non-coherent, a future-oriented adverb is possible:

- (395) Hans hat abgelehnt/angekündigt [Maria in zwei Monaten zu besuchen].
Hans has refused/announced Maria in two months to visit
'Hans refused/announced to visit Maria in two months.'

For Wurmbrand (2001: 72), infinitives that do not license a temporal modifier are quite literally tenseless. She proposes that they lack a TP projection, a claim that I will adopt here. For the sake of concreteness, I will treat coherent infinitives as *vPs*.⁹

- (396) a. *Structure of non-coherent infinitives*
[TP ... [*vP* ... [*VP* ...]]]
b. *Structure of coherent infinitives*
[*vP* ... [*VP* ...]]

A few remarks regarding (396) are in order: First, the category label assigned to the two types of infinitives is inconsequential. For the sake of concreteness, I will treat them as TPs and *vPs*, respectively, but nothing hinges on this. What is important is that the structure of coherent infinitives is a proper subpart of the structure of non-coherent infinitives. A related structural asymmetry between coherent and non-coherent infinitives in German has been proposed by Fanselow (1989), who treats coherent infinitives as IPs and non-coherent ones as CPs (also see Li 1990 for a similar proposal regarding infinitival constructions in Romance that do or do not allow clitic climbing).

Second, the structures in (396) have the consequence that *zu* cannot be a realization of T^0 because *zu* is possible in coherent infinitives. Indeed, Wurmbrand (2001: 109–115) argues against the common assumption that either *zu* or its English counterpart reside in T^0 . Instead, she proposes that *zu* is part of the projection of the lexical V. Pullum (1982) and Pollard & Sag (1994) argue for a related position for English *to*.

⁹ Wurmbrand (2001) treats them as AspPs or smaller and this view is equally compatible with everything that is to come. Also see Thráinsson (1984) for a related proposal for Icelandic infinitives.

Third, Wurmbrand (2001) in fact argues for a more fine-grained structural typology of clausal complementation in German. The bipartition of infinitival structures in (396) does not, of course, preclude the possibility that further structural distinctions exist.

4.2.2 V-final finite clauses

We will now turn to finite embedded clauses, in which the finite verb appears in the clause-final position. These sentences license nominative case, verbal agreement and tense morphology, and are normally introduced with an overt complementizer. (397) provides an illustrative examples.

- (397) Er glaubt [dass die Maria den Fritz getroffen hat].
 he believes that the Maria.NOM the Fritz.ACC met has
 ‘He believes that Maria met Fritz.’

V-final finite clauses exhibit a somewhat different extractability profile than infinitival clauses, which will be laid out in the following subsections.

4.2.2.1 Scrambling

A standard observation in German syntax going back to at least Bierwisch (1963) and Ross (1967) is that finite clauses categorically block scrambling out of them.

- (398) *Er glaubt den Fritz_i [dass die Maria _{t_i} getroffen hat].
 he believes the Fritz.ACC that the Maria.NOM met has
Intended: ‘He believes that Maria met Fritz.’

4.2.2.2 Relativization

Less widely known is the fact that relativization is likewise impossible out of finite clause (see Plank 1983: 11, Lühr 1988: 77, Bayer & Salzmann 2013: 310–311, Müller 2014b: 130–131):¹⁰

¹⁰ To express the meaning intended in (399), a resumptive pronoun would be used, a construction that Salzmann (2006) refers to as ‘resumptive prolepsis’.

- (i) der Mann [von dem er glaubt [dass die Maria ihn getroffen hat]]
 the man of whom he believes that the Maria him met has
 ‘the man who he believes that Maria met’

- (399) *der Mann [den_i er glaubt [dass die Maria t_i getroffen hat]]
 the man whom he believes that the Maria.NOM met has
Intended: ‘the man who he believes that Maria met’

4.2.2.3 *Wh*-movement into V-final clause

In contrast to relativization, *wh*-movement is able to leave a V-final finite clause. (400) illustrates this for *wh*-movement into a higher V-final clause.

- (400) Ich weiß nicht [wen er glaubt [dass die Maria t_i getroffen hat]].
 I know not whom he believes that the Maria met has
 ‘I don’t know who he believes that Maria met.’

In (400), movement proceeds from a V-final clause into a V-final clause. I should note that such movement out of a V-final clause is not possible for all speakers. I will be primarily concerned here with dialects in which such extraction is possible. The account I will propose can be adjusted to accommodate speakers for whom (400) is ungrammatical.

4.2.2.4 Movement into V2 clause

Finally, movement into a V2 clause can likewise proceed out of a V-final clause. This is demonstrated in (401) for movement of a *wh*-element, and in (402) for regular topicalization of a non-*wh* element.

- (401) Wen_i glaubt er [dass die Maria t_i getroffen hat] ?
 whom believes he that the Maria met has
 ‘Who does he believe that Maria met?’
- (402) Den Fritz_i glaubt er [dass die Maria t_i getroffen hat] .
 the Fritz.ACC believes he that the Maria.NOM met has
 ‘Fritz, he believes that Maria met.’

The extraction profile of V-final finite clauses is summarized in (403):

- (403) *Locality of V-final finite clauses*
 V-final finite clauses are opaque to scrambling and relativization, but transparent to *wh*-movement into a higher V-final clause as well as topicalization and *wh*-movement into a V2 clause.

4.2.2.5 The structure of V-final clauses

In direct contrast to infinitival clauses, V-final finite clauses exhibit all signs of the presence of a CP layer that nonfinite clauses lack. First, they allow (and in the absence of *wh*-movement force) complementizers like *dass* ‘that’ or *ob* ‘if’, as in the various examples just seen. Moreover, they provide interrogative scope:

- (404) Er hat entschieden [wen er anruft]. [= (391b)]
he has decided who he calls
‘He decided who he will call.’

In some colloquial varieties of German, it is possible for *wh*-movement and a complementizer to co-occur:

- (405) Arabella weiß [wer dass den Hans gesehen hat].
Arabella knows whom that the Hans seen has
‘Arabella know who Hans saw.’ (Bayer 1989: 8)

Moreover, finite clauses contain the full range of tense morphology and verbal agreement, as noted above. I will therefore take V-final finite clauses to be CPs, hence structurally larger than infinitival clauses:

- (406) *Structure of V-final finite clauses*
[CP ... [TP ... [_vP ... [VP ...]]]]

4.2.3 V2 clauses

The last type of clause to be considered here are embedded clause that have the shape of root clauses: the finite verb occupies the second position in the clause and is preceded by one constituent. A number of embedding verbs, like *sagen* ‘say’, *glauben* ‘believe’, *behaupten* ‘claim’, and *wünschen* ‘wish’, allow a V2 clause to be embedded under them, as illustrated in (407).¹¹

¹¹ The set of verbs that allow V2 complements is not arbitrary. As Haider (1984: 79) has observed, only bridge verbs belong to this set. Deriving this correlation remains a desideratum at this point and I will have nothing to say about it or about the distinction between bridge verbs and non-bridge verbs more generally, though see Williams (2013: 102) for an account of bridge verbs that is compatible with the horizons account proposed here.

- (407) Er hat gesagt [die Maria hat den Fritz getroffen].
 he has said the Maria.NOM has the Fritz.ACC met
 ‘He said Maria met Fritz.’

Interestingly, their extraction profile differs from that of V-final finite clauses, as detailed in the next sections.

4.2.3.1 Scrambling

Just like V-final finite clauses, V2 clauses are completely opaque to scrambling. This restriction holds irrespective of whether the scrambled element is moved out of the embedded clause in one fell swoop or successive-cyclically through the pre-verbal position of the lower clause. In (408a), the embedded subject *die Maria* ‘Maria’ fills the pre-verbal position and the object *den Fritz* ‘Fritz’ is directly scrambled into the matrix clause. In (408b), by contrast, *den Fritz* first fills the pre-verbal position of the embedded clause and subsequently scrambles into the matrix clause. Both structures are ungrammatical.

- (408) a. *Er hat den Fritz_i gesagt [die Maria hat t_i getroffen].
 he has the Fritz.ACC said the Maria.NOM has met
- b. *Er hat den Fritz_i gesagt [t_i hat die Maria t_i getroffen].
 he has the Fritz.ACC said has the Maria.NOM met
Intended: ‘He said that Maria met Fritz.’

4.2.3.2 Relativization

Embedded V2 clauses are likewise opaque to relativization out of them (e.g., Sternefeld 1989:119), as (409) attests. As in the case of scrambling, one-fell-swoop and successive-cyclic extraction are both impossible.

- (409) a. *der Mann [den_i er gesagt hat [die Maria hat t_i getroffen]]
 the man whom he said has the Maria has met
- b. *der Mann [den_i er gesagt hat [t_i hat die Maria t_i getroffen]]
 the man whom he said has has the Maria met
Intended: ‘the man who he said that Maria met’

4.2.3.3 *Wh*-movement into V-final clause

A surprising but nonetheless robust aspect of embedded V2 clauses is that they do not allow *wh*-movement out of them to land in a V-final finite clause, a constraint that has been previously observed and analyzed by Haider (1984), Reis (1985, 1996), Sternefeld (1989), Staudacher (1990), Müller & Sternefeld (1993), and Müller (1995, 2010a), among others. This restriction is illustrated in (410). Here a V2 clause is embedded inside a V-final clause. Because V-final clauses do not normally appear in root contexts, this V-final clause is itself embedded. Extraction out of the innermost V2 clause into the intermediate V-final clause is impossible, regardless of whether this extraction proceeds in one fell swoop (410a) or successive-cyclically (410b).¹²

- (410) a. *Ich weiß nicht [wen_i er gesagt hat [die Maria hat t_i getroffen]]
I know not whom he said has the Maria has met
- b. *Ich weiß nicht [wen_i er gesagt hat [t_i hat die Maria t_i getroffen]]
I know not whom he said has has the Maria met
Intended: ‘I don’t know who he said that Maria met.’

The ungrammaticality of (410) is due to the extraction. A V2 clause embedded inside a V-final clause is possible in the absence of movement, as (411) shows. I also give the example with a subjunctive verb as well to ensure that (411) does not involve direct quotation.

- (411) Ich denke [dass er gesagt hat [die Maria hat/habe den Fritz
I think that he said has the Maria.NOM has/has.SUBJ the Fritz.ACC
getroffen]].
met
‘I think that he said that Maria met Fritz.’

¹² One might wonder at this point whether extraction into the matrix clause would be possible in (410). It is not. I will put this fact aside for the time being, but return to it in chapter 6, where it will provide an argument for the need for CP phases in the account of selective opacity effects.

It is also worth noting that the badness of (410b) is not due to the linear adjacency of two *hat* auxiliaries. (i) is equally ungrammatical:

- (i) *Ich weiß nicht [wen_i er meint [t_i hat die Maria t_i getroffen]]
I know not who he thinks has the Maria met
Intended: ‘I don’t know who he said that Maria met.’

V2 clauses thus differ from V-final finite clauses in their locality profile: While movement out of a V-final finite clause can target a higher V-final clause (see (400)), movement out of a V2 clause cannot. This generalization holds for arguments and adjuncts alike. Moreover, the relativization facts in (409) in fact fall under the same generalization because relative clauses in German are always V-final.

4.2.3.4 Movement into V2 clause

The previous discussion has demonstrated that V2 clauses are opaque to scrambling, relativization, and *wh*-movement into a V-final clause. It is not simply the case, however, that V2 clauses are completely opaque to all extraction. As is well-known, *wh*-movement from a V2 clause can land inside a higher V2 clause:¹³

- (412) Wen_i hat er gesagt [t_i hat die Maria t_i getroffen]?
 whom has he said has the Maria met
 ‘Who did he say that Maria met?’

One restriction on movement from a V2 clause into a higher V2 clause is that it has to proceed successive-cyclically through the initial position of the lower clause, as in (412). One-fell-swoop extraction out a V2 clause is categorically out:

- (413) *Wen_i hat er gesagt [die Maria hat t_i getroffen]?
 whom has he said the Maria has met
Intended: ‘Who did he say that Maria met?’

This pattern not only holds of movement of a *wh*-element, but also governs topicalization of non-*wh* constituents.

- (414) Den Paul_i hat er gesagt [t_i hat die Maria t_i getroffen].
 the Paul.ACC has he said has the Maria.NOM met
 ‘Paul, he said that Maria met.’

¹³ See, however, Reis (1996) for a view to the contrary. She proposes that V2 clauses *are* completely opaque to extraction and that sentences like (412) constitute parenthetical constructions. See Tappe (1981) and Grewendorf (1988).

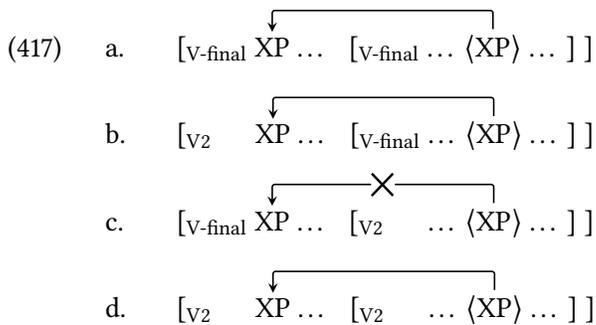
- (415) *Den Paul_i hat er gesagt [die Maria hat t_i getroffen].
 the Paul.ACC has he said the Maria.NOM has met
Intended: ‘Paul, he said that Maria met.’

The extraction profile of V2 clauses is thus unique and differs from the other three clause types just considered:

(416) *Locality of V2 clauses*

V2 clauses are selectively opaque to scrambling, to relativization, and to *wh*-movement that lands inside a V-final clause. They are transparent to topicalization and *wh*-movement that lands inside a V2 clause.

The contrast between finite V-final and V2 clauses is striking and deserves to be emphasized. Extraction out of a finite V-final clause can land in V-final as well as V2 clauses. Movement out of a V2 clause, on the other hand, can land in a V2 clause, but is barred from landing in a V-final clause.



What makes this pattern especially remarkable is that the extraction is *wh*-movement in all cases. It is hence not possible to attribute the asymmetry to differences in the moving element or the semantics of the movement type.

4.2.3.5 The structure of V2 clauses

The question of what the clause structure of V2 clauses relative to V-final finite clauses is difficult to answer. It is evident that their internal structures differ: V2 clauses comprise verb raising to the second position and topicalization over the finite-verb position. On the other hand, V2 clauses lack complementizers.

There has been a long-standing tradition in the literature on German that V2 is an indicator of illocutionary force (Wechsler 1991, Brandner 2004, Lohnstein & Bredel 2004, Meinunger 2004)

or ‘proto-force’ in Gärtner’s (2002) and Truckenbrodt’s (2006) sense. Truckenbrodt (2006) argues that V2 in unembedded as well as embedded sentences is triggered by the presence of a *context index* ⟨Epist⟩ on C⁰, which semantically results in assertional proto-force.¹⁴ I propose to combine Truckenbrodt’s (2006) proposal with Rizzi’s (1997) proposal that illocutionary force is encoded on a designated head Force⁰ high in the left periphery (an idea that harks back to Ross 1970). Specifically, I suggest that V2 clauses contain a Force projection that V-final clauses lack. This Force⁰ head contains the ⟨Epist⟩ feature and attracts the finite verb to it and another element into its specifier, thus giving rise to V2 order. This is essentially den Besten’s (1983) seminal analysis of V2 but with ForceP instead of CP as the locus of the two movements. V-final clauses simply lack a Force projection on this proposal and consequently no V2 movement takes place. This proposal is reminiscent of one made by Antomo (2012) for the contrast between V-final and V2 clauses under *weil* ‘because’. Parallel proposals to the effect that V2 clauses are structurally larger than V-final clauses are made by Sternefeld (1992), Williams (2003: 78–79) and Frey (2011).¹⁵ The proposed phrasal skeleton for V2 clauses is thus (418):

(418) *Structure of V2 clauses*
 [ForceP ... [CP ... [TP ... [_{vP} ... [VP ...]]]]]

A few remarks regarding (418): First, on Rizzi’s (1997) proposal, Force is one projection that makes up the CP-domain. The structure in (418) may beg the question in which sense a CP coexists with a ForceP. The label assigned to these projections is of no consequence to the analysis and may hence

¹⁴ As Gärtner (2002) and Truckenbrodt (2006) emphasize, V2 correlates with assertional force only in unembedded sentences. In embedded sentences, it is possible to have V2 without the speaker committing to the truth of this embedded sentence:

- (i) Er sagt, es wird morgen regnen.
 he says it will tomorrow rain

(i) does not entail or require that the person uttering it believes that it will rain tomorrow. In order to limit the effects of V2 in embedded clauses, Gärtner (2002) and Truckenbrodt (2006) propose that V2 only correlates with assertional *proto-force*. In unembedded clauses, this proto-force translates into assertional force. In embedded clauses, on the other hand, it may be “absorbed” under predicates that denote epistemic states. See Truckenbrodt (2006) for an account of various constraints on embedded V2 in terms of proto-forces and absorption.

Moreover, Truckenbrodt (2006) argues that the context index ⟨Epist⟩ triggers V2 in both embedded and unembedded contexts, but that there is another context index ⟨Deont⟩, which can appear in unembedded V2 clause but not in embedded ones. As I am solely concerned with embedded V2 clauses here, I will ignore ⟨Deont⟩.

¹⁵ Frey (2011) proposes that V2 clauses are always ForceP and that V-final clause are generally smaller structurally. Following work by Haegeman (2006) and references there on peripheral adverbial clauses, Frey (2011) also argues that certain adverbial clauses can contain a Force projection but still be V-final. For him, then, the presence of a Force layer is a necessary but not a sufficient condition for V2.

freely be replaced. I use the label ‘CP’ here simply because the head of this projection contains a complementizer. The label ‘CP’ is nothing more than mnemonic for this property. Relatedly, it is not crucial for my proposal whether the structural difference between V-final and V2 clauses is merely one projection (Force) or whether there are several projections that are present in V2 clauses but absent in V-final clauses. The key proposal is that V2 clauses are structurally larger than V-final clauses, but the details of the implementation are irrelevant for our present concerns.

Second, according to the structure in (418), complementizers reside in C^0 and the finite verb in a V2 clause is attracted to Force^0 . This raises the question of why the two are in complementary distribution:

(419) *V2 and complementizers are in complementary distribution*

- a. Ich glaube [dass Maria ein Buch liest].
I believe that Maria a book reads
- b. Ich glaube [Maria liest ein Buch].
I believe Maria reads a book
- c. *Ich glaube [dass Maria liest ein Buch].
I believe that Maria reads a book
- d. *Ich glaube [Maria dass liest ein Buch].
I believe Maria that reads a book

The standard account of this incompatibility between V2 and the presence of a complementizer, following den Besten (1983), is that the finite verb moves into the head that otherwise contains the complementizers and thereby bleeds the occurrence of the latter. Within the structure in (418), an analogous explanation is readily available on the null assumption that V2 movement obeys the Head Movement Constraint (Travis 1984). This entails that the finite verb has to pass through C^0 on its way to Force^0 and form a complex head with it. It is this complex head that undergoes subsequent movement into Force^0 . The ungrammaticality of (419c,d) then follows in exactly the same way it does on the standard account: Incorporation of a verb into C^0 bleeds the presence of a complementizer in that position.

Third, some independent evidence that V-final finite clauses and V2 clauses differ in their syntactic makeup comes from coordination. It is possible to coordinate two V2 clauses (420b) and

two V-final clauses (420a), but mixtures are impossible (420c,d) (Grewendorf 1988: 211, Truckenbrodt 2006: 280).

- (420) a. Peter glaubt [[_{V2} er habe genug gearbeitet] und [_{V2} er könne sich
Peter believes he has enough worked and he could himself
jetzt zur Ruhe setzen]]
now retire
'Peter believes that he has worked enough and that he can retire now.'
- b. Peter glaubt [[_{V-final} dass er genug gearbeitet habe] und [_{V-final} dass er
Peter believes that he enough worked has and that he
sich jetzt zur Ruhe setzen könne]]
himself now retire could
- c. *Peter glaubt [[_{V-final} dass er genug gearbeitet habe] und [_{V2} er könne
Peter believes that he enough worked has and he could
sich jetzt zur Ruhe setzen]]
himself now retire
- d. *Peter glaubt [[_{V2} er habe genug gearbeitet] und [_{V-final} dass er sich
Peter believes he has enough worked and that he himself
jetzt zur Ruhe setzen könne]]
now retire could (Grewendorf 1988: 211)

If V2 clauses encode Force, which V-final clauses lack, the contrast in (420) follows from the general constraint that coordination has to combine alike elements or at least elements with similar semantics.

A distinction between the syntactic category of verb-final and V2 clauses is also plausible in light of well-known distributional differences between the two types of clauses. Reis (1985: 285), Cinque (1989) and Webelhuth (1989, 1992), for example, draw attention to the fact that V2 clause are blocked from occupying the subject position of another clause or act as the topic of a clause.

(421) *Subject clauses*

- a. [_{V-final} Dass Hans verschlafen hat] ist schwer zu glauben.
that Hans overslept has is hard to believe
'It is hard to believe that Hans overslept.'
- b. * [_{V2} Hans hat verschlafen] ist schwer zu glauben.
Hans has overslept is hard to believe

(422) *Topicalized clauses*

- a. [_{V-final} Dass Hans verschlafen hat]_i hat Maria nicht *t_i* gesagt.
that Hans overslept has has Maria not said
'Maria didn't say that Hans overslept.'
- b. * [_{V2} Hans hat verschlafen]_i hat Maria nicht *t_i* gesagt.
Hans has overslept has Maria not said

This asymmetry between V-final and V2 clauses can then be straightforwardly modeled as category difference between CPs and ForcePs.

4.2.4 Interim summary

This section has discussed four types of embedded clauses in German with particular attention to their locality profiles and internal clause structure. We saw that the four clause types of interest (coherent and non-coherent infinitives, V-final finite clauses, and V2 clauses) differ in the amount of functional structure that they contain. These differences in clause structure coincide with differing extraction profiles. A schematic summary of the generalizations arrived at thus far is provided in (423):

(423) *Selective opacity in German (to be extended)*

<i>Operation</i>	<i>Size of embedded clause</i>			
	<i>nonfinite</i>		<i>finite</i>	
	<i>coherent</i> (vP)	<i>non-coherent</i> (TP)	<i>V-final</i> (CP)	<i>V2</i> (ForceP)
scrambling	✓	✗	✗	✗
relativization	✓	✓	✗	✗
<i>wh</i> -movement into V-final clause	✓	✓	✓	✗
topicalization/ <i>wh</i> -movement into V2 clause	✓	✓	✓	✓

(✓ : transparent; ✗ : opaque)

Before developing an account of selective opacity in German within the general framework of horizons, I will argue in the following sections that the German pattern in (423) adheres to and thereby corroborates the empirical meta-generalizations of selective opacity observed in the preceding chapters, namely Upward Entailment and the Height–Locality Connection.

4.3 Regularities of selective opacity

Just like selective opacity in Hindi, locality differences between operations are not distributed randomly in German, but adhere to the generalizations identified in the previous literature (Müller & Sternefeld 1993, Müller 1995, 2014a,b, Williams 2003, 2011, 2013, Abels 2007, 2009, 2012a). This section corroborates this claim.

4.3.1 Opacity layering

One obvious property of selective opacity in German is that embedded clauses induce layers of opacity. That is, it is not the case that there are simply two types of locality profiles, roughly A- and \bar{A} -locality. Instead, (423) instantiates four different locality profiles. Locality, then, is not simply a bipartition, but manifold. Furthermore, a differentiation between A- and \bar{A} -locality is of little help in accounting for the German facts, simply because it is not clear whether any movement type in (423) is an instance of A-movement. Movement to the clause-initial position (topicalization) as well as *wh*-movement in both V-final and V2 clauses are instances of \bar{A} -movement and so is relativization. The status of scrambling is more controversial. There is good evidence that scrambling does not behave like prototypical A-movement in several respects (Webelhuth 1992, Müller & Sternefeld 1994, Müller 1995, Grewendorf & Sabel 1999). But if scrambling is an instance of \bar{A} -movement, then all four movement types in (423) involve \bar{A} -movement and any appeal to differences between A- and \bar{A} -movement is of no utility for the German facts. What we need, then, is an account of selective opacity that allows us to make more fine-grained distinction between movement types than just an overly coarse bipartition into A- and \bar{A} -processes.

4.3.2 Upward Entailment

A cursory inspection of (423) makes it clear that the transparency/opacity of the various structures for the different movement types is not distributed randomly. A striking pattern is that structurally smaller clauses are transparent to more operations than structurally larger clauses. In other words, structurally small clauses like coherent infinitives, which we have seen evidence are bare *v*Ps, are transparent to all four movement types, whereas non-coherent infinitives, which are structurally larger TP clauses, are opaque to scrambling, but transparent to the other three movement types. Even larger clauses like V-final finite clause are opaque to scrambling and relativization, but not *wh*-movement and topicalization, and so on. It is thus plain that the internal structure of a clause is directly related to its locality profile.

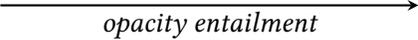
Relatedly, while the addition of functional structure leads to a decrease in transparency, the reverse never holds. It is never the case that a structurally larger clause is transparent to an operation that a structurally smaller clause is opaque for. This pattern is another example of Upward Entailment, a general and systematic property of selective opacity, which I have called Upward Entailment.

- (424) **Upward Entailment** [=(212),(264)]
If a clause of a certain structural size is opaque to an operation, then clauses that are structurally larger are also opaque to this operation.

A schematic illustration of Upward Entailment for the German facts is given in (425). If a row has a cross in a cell, then all cells to its right also contain a cross. Conversely, if a cell contains a check mark, all cells to its left likewise contain check marks. Upward Entailment is a statement of this regularity.

(425) *Upward Entailment in German selective opacity*

	vP	TP	CP	ForceP
scrambling	✓	✗	✗	✗
relativization	✓	✓	✗	✗
<i>wh</i> -movement into V-final clause	✓	✓	✓	✗
topicalization/ <i>wh</i> -movement into V2 clause	✓	✓	✓	✓



The pattern in (425) is unlikely to arise by chance alone. In addition, the fact that it is also observed for selective opacity in Hindi (see section 2.5.1 in chapter 2) and crosslinguistically (see chapter 1) makes it clear that Upward Entailment is a systemic property of selective opacity that any account has to capture. I have demonstrated in section 3.3.2 of chapter 3 how Upward Entailment can be derived in the horizons framework from category inheritance, an independent property of extended projections.

4.3.3 The Height–Locality Connection

A second systemic property of selective opacity in addition to Upward Entailment identified in the previous literature is the Height–Locality Connection, repeated in its probe-based formulation in (426). The Height–Locality Connection (HLC) states that a probe’s location in the clausal spine is related to its locality profile: The higher the location of the probe, the more types of embedded clauses are transparent to this probe.

- (426) **Height–Locality Connection** (*probe-based formulation*) [=(241),(327)]
 The higher the location of a probe is in the clausal structure, the more kinds of structures are transparent to this probe.

I have argued at length in section 2.5.2 of chapter 2.2 that the HLC holds in Hindi. In this section, I will demonstrate that it also holds for German. In order to evaluate the HLC in the German context, it is necessary to determine the landing sites of the four movement types under investigation, which will correspond to the structural positions of the probes triggering these movements.

4.3.3.1 The landing site of scrambling

It is clear that scrambling lands in a relatively low position in the clausal spine, certainly lower than C^0 . As the examples in (427) attest, scrambling may not move an element over a complementizer (427a), over a *wh*-moved element (427b) or over a relative pronoun (427c).

(427) *Scrambling lands lower than C^0*

- a. *Ich weiß [die Maria_i dass niemand t_i gesehen hat].
I know the Maria.ACC that nobody seen has
Intended: ‘I know that no one has seen Maria.’
- b. *Ich weiß [die Maria_i wer_j t_j t_i gesehen hat].
I know the Maria.ACC who seen has
Intended: ‘I know who saw Maria.’
- c. *der Mann [die Maria_i der_j t_j t_i gesehen hat]
the man the Maria.ACC who seen has
Intended: ‘the man who saw Maria’

It is thus clear that the probe underlying scrambling is located on a head lower than C^0 . The exact location of scrambling is much harder to pinpoint and the choice will be irrelevant for what is to come, so long as the probe is located lower than C^0 . Because objects can scramble over subjects of transitive verbs, I will assume here that scrambling targets at least SpecTP and hence that the relevant probe is located on T⁰. There may likely be additional scrambling sites below vP . While I will lay those aside here, they are fully compatible with what is to come.¹⁶

4.3.3.2 The landing site of relativization

Relativization unambiguously lands in a position higher than that targeted by scrambling. First, when relativization and scrambling both apply within the same clause, the relativized element has to appear to the left of the scrambled element:

¹⁶ Grewendorf & Sabel (1999) propose that scrambling targets an Agr projection and Müller (1998, 2010b, 2011) proposes that it lands in a verbal projection, VP or vP .

(428) *Relativization lands higher than scrambling*

- a. das Buch [das_i der Maria_j niemand t_j t_j empfohlen hat]
that book which.ACC the Maria.DAT nobody.NOM recommended has
'the book that no one has recommended to Maria'
- b. *das Buch [der Maria_j das_i niemand t_j t_i empfohlen hat]
the book the Maria.DAT which.ACC nobody.NOM recommended has

Second, there is indirect evidence that relativization targets SpecCP. Recall from the discussion in section 4.2.1 that there is good evidence that infinitival clauses in German lack a CP layer, unlike their English counterpart. It is instructive, then, that German does not allow infinitival relatives, unlike English, regardless of the presence or absence of an overt relative pronoun:¹⁷

(429) *No infinitival relatives*

- a. *der Mann [(der) das Waschbecken zu reparieren]
the man (who) the sink to repair
Intended: 'the man to fix the sink'
- b. *das Buch [(das) zu lesen]
the book (which) to read
Intended: 'the book to read'

If relativization targets SpecCP in German, the ungrammaticality of (429) follows without further ado from independently motivated structural properties of infinitival clauses in German. They are structurally too small to host a landing site for relativization. This contrasts with the structure of infinitival clauses in English, which can be CPs and hence allow the equivalent of (429). Unsurprisingly, the existence of infinitival *wh*-questions in English and their absence in German likewise follows if English infinitival clauses can be CPs, but German ones cannot.

Third, relative clauses are obligatorily V-final in German, as shown in (430):¹⁸

¹⁷ (429b) is grammatical without the relative pronoun on an irrelevant infinitival clause construal (i.e., *to read the book*). For discussion and analysis of infinitival relatives in English, see Bhatt (2006).

¹⁸ A qualification is in order. Gärtner (2001) discusses structures that appear to be V2 relative clauses. An example is given in (i):

(430) *No V2 relatives*

- a. Ich kenne keinen [der *Krieg und Frieden* gelesen hat].
I know nobody who War and Peace read has
'I know no one who has read War and Peace.'
- b. *Ich kenne keinen [der hat *Krieg und Frieden* gelesen].
I know nobody who has War and Peace read

This array of facts is readily accounted for if relativization lands in SpecCP in German, hence if relativization is triggered by a probe on C⁰.

4.3.3.3 The landing site of *wh*-movement in V-final clauses

There is likewise good evidence that *wh*-movement in V-final clauses lands in SpecCP as well. As shown by (431), if *wh*-movement and scrambling co-apply in a clause, *wh*-movement obligatorily lands to the left of scrambling:

(431) *Wh-movement lands higher than scrambling*

- a. Ich weiß [welches Buch_i der Maria_j niemand $t_j t_i$ empfohlen
I know which book.ACC the Maria.DAT nobody.NOM recommended
hat].
has
'I know which book no one has recommended to Maria.'
- b. *Ich weiß [der Maria_j welches Buch_i niemand $t_j t_i$ empfohlen
I know the Maria.DAT which book.ACC nobody.NOM recommended
hat].
has

Furthermore, as mentioned above, some nonstandard varieties of German are not subject to the Doubly-filled Comp Filter. In these cases, the *wh*-element appears to the left of the complementizer:

(i) Das Blatt hat eine Seite [die ist ganz schwarz].
the sheet has a side that is completely black

(Gärtner 2001: 98)

This construction is limited to colloquial German and tightly constrained syntactically (see the discussion in Gärtner 2001). I will put it aside here but would like to note that (i) is not incompatible with the analysis proposed here, because the account of V2 *wh*-questions could be carried over to relativization as well.

- (432) Arabella weiß [wer dass den Hans gesehen hat]. [= (405)]
 Arabella knows whom that the Hans seen has
 ‘Arabella know who Hans saw.’ (Bayer 1989: 8)

Finally, as mentioned in section 4.2.1 above, *wh*-movement cannot take place in infinitival clauses in German, which lack a CP layer.

- (433) *Er hat vergessen [wen anzurufen]. [= (390a)]
 he has forgotten who to.call
Intended: ‘He forgot who to call.’

This set of facts follows straightforwardly if *wh*-movement lands in SpecCP, hence if it is triggered by a probe on C^0 .

4.3.3.4 The landing site of topicalization and *wh*-movement in V2 clauses

In section 4.2.3, I have proposed that V2 clauses are structurally larger than finite V-final clauses. Concretely, I suggested that V2 clauses are ForcePs, whereas V-final clauses are CPs. On this view, verb movement in V2 structures ultimately targets Force⁰. Because the topicalized element precedes this verb, it follows that topicalization must land in SpecForceP and be triggered by a probe on Force⁰.¹⁹ It follows from this view that verb movement and topicalization go hand in hand in regular declarative clauses. Moreover, the fact that both are absent in V-final clauses follows if these clauses are merely CPs, hence lacking Force⁰'s probes for verb raising and topicalization.

4.3.3.5 Section summary

Adding the probe locations determined in this section to the summary table in (423), yields the final summary table in (434):

¹⁹ Also see Grewendorf's (2003) distinction between operator \bar{A} -movement and non-operator \bar{A} -movement.

(434) *Summary: Selective opacity in German* [=(367)]

<i>Operation</i>	<i>probe location</i>	<i>Size of embedded clause</i>			
		nonfinite		finite	
		coherent (vP)	non-coherent (TP)	V-final (CP)	V2 (ForceP)
scrambling	T ⁰	✓	✗	✗	✗
relativization	C ⁰	✓	✓	✗	✗
<i>wh</i> -movement into V-final clause	C ⁰	✓	✓	✓	✗
topicalization/ <i>wh</i> -movement into V2 clause	Force ⁰	✓	✓	✓	✓

(✓ : transparent; ✗ : opaque)

We are now able to evaluate whether the Height–Locality Connection (HLC), repeated in (435) for convenience, also holds in German.

(435) **Height–Locality Connection** (*probe-based formulation*) [=(426)]

The higher the location of a probe is in the clausal structure, the more kinds of structures are transparent to this probe.

It is clear from (434) that the HLC holds in German as well. The higher the landing site of a movement type, the more types of clauses are transparent to this movement. Scrambling, which targets a structurally low position, is possible only out of structurally small clauses. By contrast, topicalization, which lands in a structurally high position, is able to proceed out of structurally small as well as structurally large clauses.

At the same time, it is clear from (434) that the HLC is not a one-to-one relation. Just as in Hindi, two movement types that target the same position may differ slightly in their locality. In German, both relativization and *wh*-movement in V-final clauses target SpecCP, but only *wh*-movement is able to occur out of a CP clause. This mirrors the Hindi pattern discussed in chapter 2, where \bar{A} -movement as well as *wh*-licensing have their underlying probes on C⁰, but only \bar{A} -movement may proceed across a CP boundary. Height and locality are thus linked to each other, but one does not completely predict the other. We will see in the next section that horizons are able to express and derive this relationship in exactly the same way they did for the Hindi pattern.

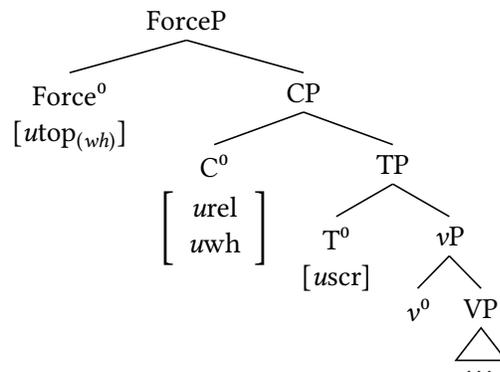
The HLC thus holds in German in the same way it does in Hindi and the various cases discussed in chapter 1. This provides additional evidence for the conclusion that the HLC is a property of selective opacity in need of explanation. In the next section, I will apply the horizons analysis developed in chapter 3 on the basis of Hindi to the German pattern. Just as in the case of Hindi, horizons are able to account for locality differences between movement types and at the same time impose a limit on such differences, thus deriving Upward Entailment and the HLC.

4.4 A horizon-based account

4.4.1 The probes

Following the framework developed in chapter 3, I will assume that the four movement operations under discussion are triggered by four distinct probes, the location of which has been discussed in section 4.3.3. Scrambling is triggered by [*uscr*], relativization by [*urel*], *wh*-movement into a V-final clause by [*uwh*], and movement to the clause-initial position in a V2 clause by [*utop_(wh)*], a probe to which I will return immediately. The claim that scrambling in German is triggered by a feature whose presence is optional is by no means new and has been argued for by Müller (1996, 2010b, 2011), Grewendorf & Sabel (1999), Sauerland (1999), Putnam (2007), and others.²⁰

(436) *Schematic clause structure and probe locations*



Before making a proposal regarding the locality properties of these probes, some remark regarding *wh*-movement are in order. We saw in the preceding section that there is evidence for distinguishing

²⁰ Müller (1996, 2010b, 2011) and Grewendorf & Sabel (1999) notate this feature as [Σ], though of course nothing hinges on the notation.

at some level between *wh*-movement in a V-final clause and *wh*-movement in a V2 clause. First, they target different landing sites (SpecCP vs. SpecForceP). Second, they differ in their locality profiles in that *wh*-movement in a V2 clause is able to leave a lower V2 clause, whereas *wh*-movement into a V-final clause is not. Both considerations suggest that the two are triggered by distinct probes. Crucially, however, *wh*-movement is obligatory in regular V2 clauses in German:²¹

- (437) a. Wen hat Maria gesehen?
 whom has Maria seen
- b. *Maria hat wen gesehen?
 Maria has whom seen

Recall that on the structure proposed in the preceding section, the finite verb in a V2 clause has moved into Force⁰ and the preverbal constituent resides in SpecForceP. We thus have to rule out a structure for (437b) in which *wen* has *wh*-moved to SpecCP, a movement masked by subsequent topicalization of *Maria* and verb movement.

- (438) *_{[ForceP Maria_i [_{Force⁰} Force⁰ hat] [_{CP} wen_j C⁰_[wh] [_{TP} ... t_i t_j ...]]]]}

What has to be ensured, then, is that a *wh*-element moves to SpecForceP of such clauses. If no such ForceP is present, as in the case of embedded V-final clauses, which are bare CPs, the *wh*-element moves to SpecCP.

There are a varieties of options of achieving this outcome and the choice will be inconsequential for the account of the selective opacity pattern in German. One option is to require that the Force⁰ head has to match the interrogative force of its complement CP. That is, in an interrogative V2 clause, both C⁰ and Force⁰ bear a *wh*-feature, a relation that may be implemented in terms of Agree. As a result, the Force⁰ head in (437) will itself carry a *wh*-feature and hence attract *wen* to its specifier. This will ensure the surface order in (437a) and rule out the order in (437b). Consequently, the Force⁰ head contains a regular (non-*wh*) probe [*utop*] in declarative clauses, but is endowed with a *wh*-probe [*utop_{wh}*] in interrogative clauses. This has the correct effect that [+*wh*] V2 clauses will

²¹ (437b) is grammatical on an echo question interpretation or if *wen* is interpreted as a weak indefinite (*Maria saw someone*).

have a *wh*-element in their preverbal position and that $[-wh]$ V2 clauses have a non-*wh* element in this position. I have indicated this featural variability as $[u_{top(wh)}]$ in (436).

Alternatively, one may assume that the $[\pm wh]$ specification of a clause is always a property of the highest projection of that clause. If the clause is a bare CP, it is C^0 that will bear the *wh*-feature. If the clause is a ForceP, the *wh*-feature will be on $Force^0$. On this view, clause typing features are not an inherent property of a given head, but only instantiated if this head is the highest head in the extended projection. Just as in the first option, the net effect of this alternative is that the clause-initial position in a V2 clause is obligatorily filled by a $[+wh]$ element and by some other element otherwise.

A third option to explore is to assume that C^0 's *wh*-feature appears on $Force^0$ as a result of head incorporation. Recall that the projection of a ForceP goes hand in hand with verb movement into $Force^0$, producing a V2 clause. Recall also that verb movement into $Force^0$ proceeds through all intervening heads, as a result of the Head Movement Constraint. As a consequence, C^0 will be incorporated into $Force^0$ alongside the verb. The *wh*-feature on C^0 will then probe from $Force^0$ and attract a *wh*-element in its domain to SpecForceP. Like the other two options just discussed, this account allows (terminal) *wh*-movement to SpecCP only in V-final clauses but not in V2 clauses. The last line of analysis would have to ensure that head movement into $Force^0$ takes place before probing of $[uwh]$. As mentioned, the choice between these options is irrelevant for what is to come. I will continue to use the notation $[u_{top(wh)}]$ as an indicator that $Force^0$ probes for a *wh*-element in $[+wh]$ clauses and for a non-*wh* element in $[-wh]$ clauses.

4.4.2 Probe locality

With these remarks in place, let us now turn to the locality properties of these probes. I propose that their different locality profiles are the result of different horizon specifications. The horizons I propose are given in (439). The subscript next to the probe indicates the syntactic head that this probe is located on.

(439) *German probes and their horizons*

- a. *Scrambling:*
[uscr]_{TP} -|| T → terminates at TP, CP or ForceP
- b. *Relativization:*
[urel]_{C°} -|| C → terminates at CP or ForceP
- c. *Wh-movement in V-final clause:*
[uwh]_{C°} -|| Force → terminates at ForceP
- d. *Topicalization/ Wh-movement into V2 clause:*
[utop_(wh)]_{Force°} -|| ∅ → no horizon

Recall from chapter 3 that if a probe [uF] has a category X as its horizon (notated as '[uF] -|| X'), [uF]'s search terminates if it encounters a node of type X. Further search into X is then impossible.

- (440) **HORIZONS** [= (260)]
If a category label X is a horizon for probe [uF] (notated as '[uF] -|| X'), then a [uF]-initiated search terminates at XP.

For example, the scrambling probe [uscr] will hence terminate its search when it encounters a TP node because of [uscr] -|| T. Because a probe's search terminates as soon as it encounters that probe's horizons, any element more deeply embedded will be invisible to this probe. As discussed at length in chapter 3, there is no special status of the edge of a domain. For example, if a probe has C as its horizon, this probe is unable to search past a CP node. All elements dominated by this CP node, including an element in SpecCP, are inaccessible to that probe. As a result, this probe cannot trigger movement out of CPs, be it successive-cyclic or in one fell swoop.

Moreover, recall from chapter 3 that category features are inherited up through an extended projection:

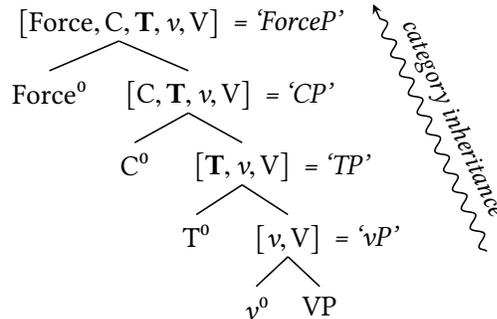
- (441) **CATEGORY INHERITANCE** [= (272)]
Given an extended projection $\Phi = \langle \Pi_n > \Pi_{n-1} > \dots > \Pi_1 \rangle$, where Π_x 's are all phrases, the categorial features of Π_m are inherited up to Π_{m+1} .

A CP node, for instance, contains a categorial specification not only for C, but also for T (along with ν and V). As a result, CPs will likewise induce search termination for probes whose horizons is T, a result that I have called the Horizon Inheritance Theorem in chapter 3:

- (442) **HORIZON INHERITANCE THEOREM** [=(276)]
 Given a probe $[uF]$ and an extended projection $\Phi = \langle \Pi_n > \Pi_{n-1} > \dots > \Pi_1 \rangle$, if $\Pi_m \in \Phi$ is a horizon for $[uF]$, then all projection Π_{m+1}, \dots, Π_n are likewise horizons for $[uF]$ (due to category inheritance (441)).

As discussed at length in section 3.3.2 of chapter 3, the Horizon Inheritance Theorem has the effect that if a clause of a particular structural size is impenetrable to a probe, then structurally larger clauses will likewise be impenetrable. This derives Upward Entailment (424). To give just one example, the scrambling probe $[uscr]$ has T as its horizon in German. As a result of (442), TP clauses, CP clauses as well as ForceP clauses will induce search termination and hence be impenetrable to this probe. This is schematized in (443). Because embedded ForceP clauses, CP clauses and TP clauses all contain a T projection, their topmost label will invariably contain a T feature, which leads to termination of $[uscr]$ search. For example, a CP node literally contains the category feature T as part of its feature content. Given $[uscr] \dashv\vdash T$, a CP node will terminate $[uscr]$ search. The same holds for ForceP nodes. Embedding a bare vP clause, on the other hand, will not lead to termination of $[uscr]$ because the label of this clause will not contain a T feature. Scrambling is hence impossible out of ForceP clauses, CP clauses, and TP clauses, but possible out of vP clauses.

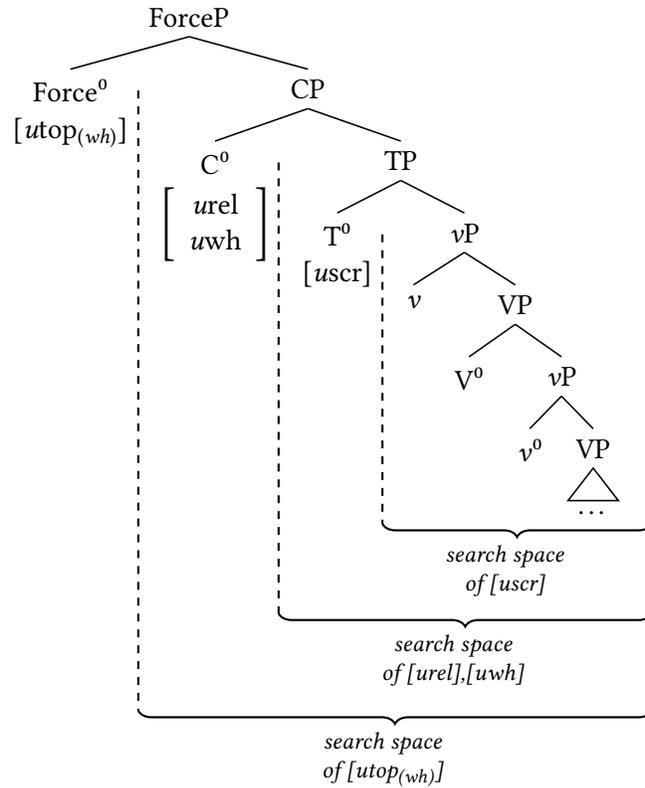
- (443) *Schematic example of category inheritance*



To see the concrete effects of (439), let us consider schematic representations of embedded clauses of varying sizes and their locality profiles that emerge from (439). Consider first coherent infinitives, which I have argued in section 4.2.1 are bare vP . The schematic structure of vP embedding is provided in (444). Specifiers are not shown for readability. Because the embedded clause contains neither a categorial specification for T nor for C nor for Force, none of the four probes in (439) will have their search prematurely terminated. All four are thus able to search into the embedded

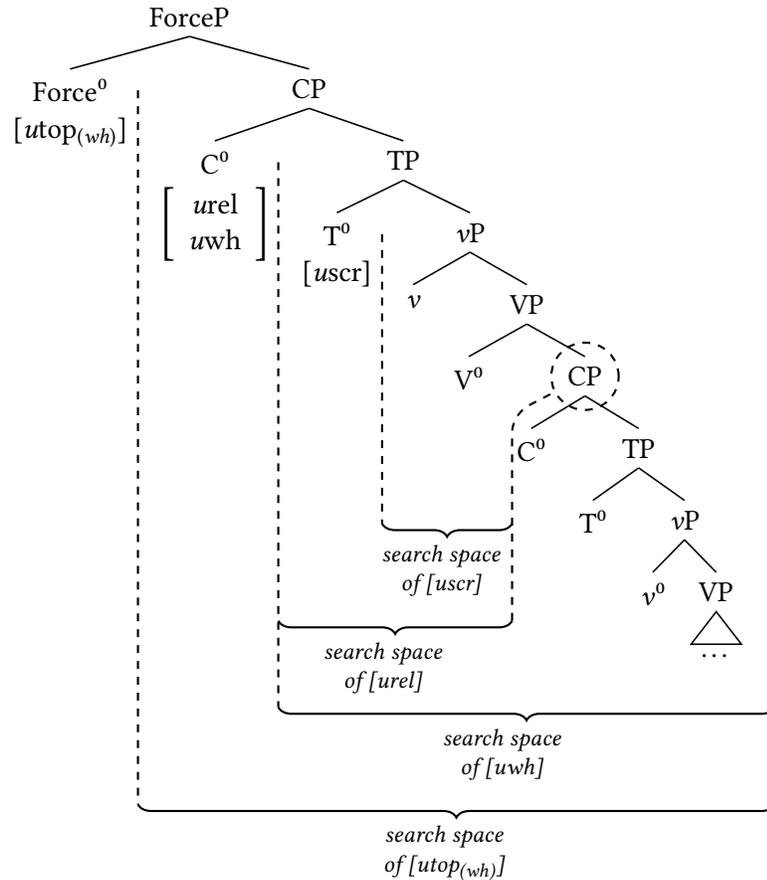
clause. As a result, coherent infinitives are transparent to scrambling, relativization, *wh*-movement, and topicalization.

(444) *Coherent infinitive (vP) embedding*



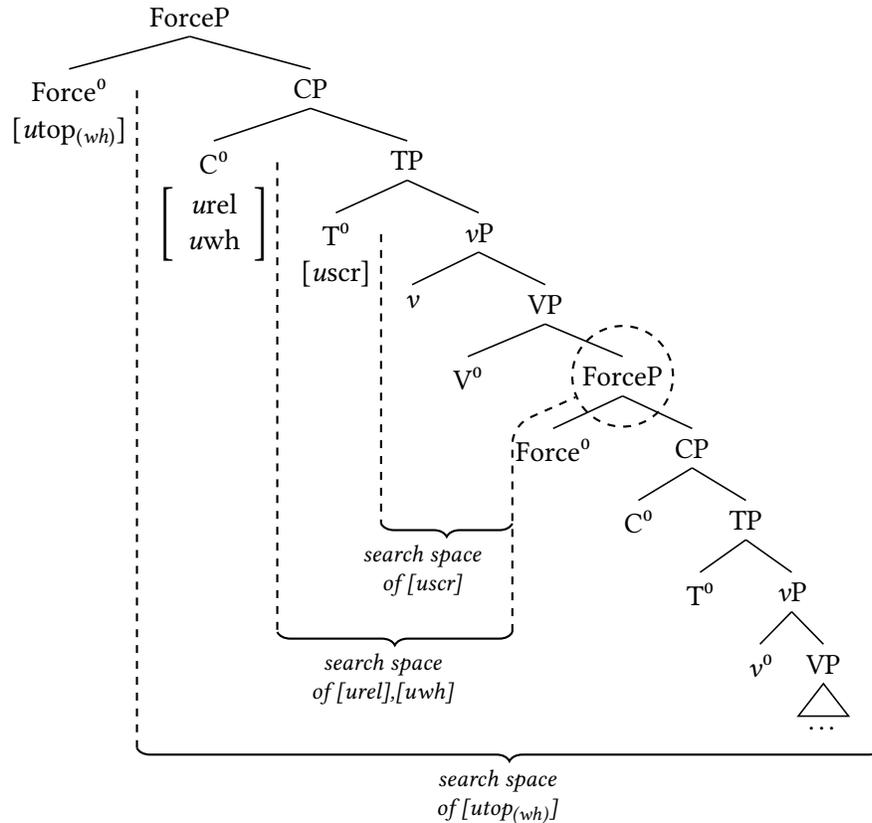
Contrast this with non-coherent infinitives, which I have argued to be TPs. The structure of TP embedding is shown in (445). Again, no specifiers are shown. Because $[uscr] \neq T$, the search initiated by the scrambling probe terminates at the embedded TP node. $[uscr]$ hence cannot search into the embedded clause and is unable to agree with any element inside that clause. As a consequence, scrambling out of a TP clause is impossible, as desired. The other three probes do not have T as their horizon and their search space is therefore not delimited by the TP node. Their search hence proceeds into the TP clause and relativization, *wh*-movement as well as topicalization out of TP clauses is therefore possible.

(446) *Finite V-final (CP) clause embedding*



The final configuration of interest is the embedding of a V2 clause, which I have argued are ForcePs, hence structurally larger than CPs. The resulting schematic structure is shown in (447). The ForceP node dominating the embedded clause induces search termination for [uscr] and [urel] by (442) and for [uwh] because of [uwh] \neq Force. These three probes are therefore unable to search into the embedded clause, including its edge, and scrambling, relativization and *wh*-movement into a V-final clause are all impossible out of V2 clauses. The only probe whose search space remains unaffected by the ForceP node is [utop_(wh)]. This probe is hence the only one that can search into the embedded clause. Consequently, only topicalization and *wh*-movement into a V2 clause can leave an embedded V2 clause.

(447) *V2 (ForceP) clause embedding*



This account straightforwardly captures one particularly peculiar property of the German selective opacity pattern, the restriction that movement from within a V2 clause must land in a higher V2 clause.

(448) *Locality of V2 clauses* [=(416)]

V2 clauses are selectively opaque to scrambling, to relativization, and to *wh*-movement that lands inside a V-final clause. They are transparent to topicalization and *wh*-movement that lands inside a V2 clause.

The account just presented derives (448) in a way parallel to other selective opacity facts. V2 clauses are ForcePs. The only probe that is able to search into a ForceP is $[utop_{(wh)}]$, which is itself located on Force⁰. As a result, because $[utop_{(wh)}]$ is the only probe able to extract an element out of an embedded ForceP, such extraction has to invariably land in SpecForceP of the higher clause. If the higher clause is not a V2 clause and hence lacks a Force projection, extraction out of a lower ForceP

will be impossible as the higher clause cannot contain a probe that is able to search into a ForceP. This derives (448).

Important to the account of (448) is the fact that *wh*-movement into a V2 clause is triggered by a probe distinct from the one that triggers *wh*-movement into a V-final clause ($[u_{\text{top}(\text{wh})}]$ vs. $[u_{\text{wh}}]$), as already motivated in section 4.4.1. This section discussed three ways of ensuring that *wh*-movement in a V2 clause ends up in SpecForceP. The first two options are to assume that in interrogative V2 clauses, Force⁰ bears the feature $[u_{\text{top}(\text{wh})}]$, which attracts a *wh*-element to its specifier, either directly or mediated by movement to SpecCP. In an interrogative V-final clause, on the other hand, *wh*-movement is driven by $[u_{\text{wh}}]$ on C⁰. If the probes underlying the two movements are distinct, the locality difference between the two with respect to embedded V2 clauses is easily implemented. I will tentatively adopt this solution here and continue to notate the topicalization feature as $[u_{\text{top}(\text{wh})}]$ to indicate that it reflects clause typing and hence searches for a $[+wh]$ or $[-wh]$ goal.²³

4.4.3 Item-based vs. domain-based approaches

As I have already discussed in section 3.4.2 of chapter 3 and as I will take up in greater detail in chapter 5, one of the core differences between the horizons account and virtually all previous approaches to selective opacity is that the horizons account is *domain-based*, whereas previous approaches are *item-based*.²⁴ The horizons account is domain-based because horizons allow us to

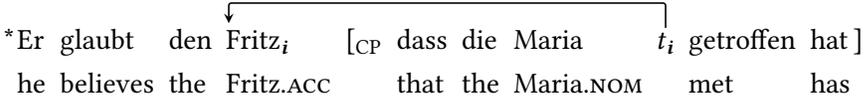
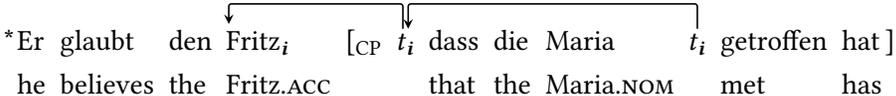
²³ The third possibility raised in section 4.4.1 presents a greater analytical challenge, though not an insurmountable one. This third account assumes that only C⁰ bears a *wh*-specific feature. The difference in the landing sites of *wh*-movement in V2 and V-final clauses is then attributed to the fact that C⁰ undergoes incorporation into Force⁰ in V2 clauses, a general by-product of the verb raising cyclically to Force⁰. While the difference in the landing site of *wh*-movement in the two clauses is thus derived, the difference in their locality remains to be accounted for. The core of this challenge is that *wh*-movement in V-final clauses is triggered by the same feature as *wh*-movement in V2 clauses. Assuming that a probe's horizon is an inherent property of that probe and not context-dependent, it is then not clear why *wh*-movement in a V2 clause should differ in its locality from *wh*-movement in a V-final clause. One might take this as an argument against this account. Alternatively, one might explore the option that the $[u_{\text{wh}}]$ -probe can form a *composite probe* with $[u_{\text{top}}]$ if they are part of the same head, as suggested in a different context by van Urk & Richards (2015) and van Urk (2015). They propose that a composite probe has the locality of the less restrictive probe. Applying their proposal to the German case at hand, C⁰-to-Force⁰ movement would put $[u_{\text{wh}}]$ and $[u_{\text{top}}]$ on the same complex head and allow the formation of a composite probe. This probe would then have the locality of $[u_{\text{top}}]$, hence be able to search in a V2 clause. This likewise derives the fact that *wh*-movement has different locality properties in the two configurations.

²⁴ Though see Williams (2003, 2011, 2013), discussed in section 5.7 of chapter 5, for another domain-based account. Abels (2007, 2009), discussed in section 5.5 of chapter 5, falls in between item-based and domain-based accounts.

directly state that a given domain is opaque to an operation. The fact that items inside that domain cannot undergo this operation then follows immediately, but these items do not directly enter into the account. The impossibility of scrambling out of a CP clause, for example, follows from the fact that CP constitutes a horizon for [*uscr*]. This entails that elements inside the CP cannot be scrambled out of the CP, but these items do not themselves enter into the account. Standard approaches, by contrast, are item-based because they crucially refer to the properties or positions of the moving element. For the sake of concreteness, let us contrast these two lines of account on the basis of the ban on scrambling out of finite clause (see sections 4.2.2 and 4.2.3). A relevant example is repeated in (449).

- (449) *No scrambling out of finite clause* [=(398)]
 *Er glaubt den Fritz_i [dass die Maria t_i getroffen hat].
 he believes the Fritz.ACC that the Maria.NOM met has
Intended: ‘He believes that Maria met Fritz.’

The standard account of why scrambling out of a finite clause is impossible, which is due to Müller & Sternefeld (1993), Müller (1995, 2014a,b), and Abels (2007, 2009), invokes a designated constraint on possible sequences of movement steps of a single element. There are at least two derivations that need to be blocked in order to successfully rule out (449). The first derivation involves one-fell-swoop movement of *den Fritz*, as in (450a), the second derivation involves successive-cyclic extraction through the edge of the lower clause, as in (450b).

- (450) a. *One-fell-swoop derivation of (449):*

 *Er glaubt den Fritz_i [CP dass die Maria t_i getroffen hat]
 he believes the Fritz.ACC that the Maria.NOM met has
 b. *Successive-cyclic derivation of (449):*

 *Er glaubt den Fritz_i [CP t_i dass die Maria t_i getroffen hat]
 he believes the Fritz.ACC that the Maria.NOM met has

On this standard account, the two derivations are ruled out by a conspiracy of two constraints, in analogy to English superraising. The one-fell-swoop derivation is ruled out by a constraint barring such movement, like subjacency or phases. Crucially, the cyclic derivation in (450b) is ruled out

by a bleeding relationship between \bar{A} -movement and scrambling: An element that has undergone \bar{A} -movement cannot subsequently undergo scrambling (Müller & Sternefeld 1993, Müller 1995, 2014a,b, Abels 2007, 2009), as this would constitute an improper sequence of movement steps:

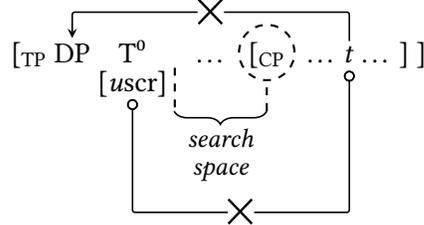
(451) *Item-based constraint on scrambling*

An \bar{A} -moved element may not subsequently undergo scrambling.

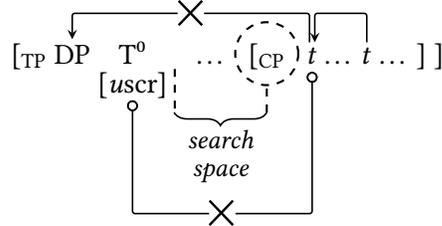
I will refer to approaches along these lines as item-based, because they impose a designated constraint on possible sequences of movement steps that an item undergoes. Because (451) rules out only one of the two possible extraction paths in (450), an item-based constraint must invoke a conspiracy of two constraints to derive the fact that scrambling out of a finite clause is impossible.

The horizons-based account, on the other hand, does not require a constraint like (451). Both the one-fell-swoop derivation and the successive-cyclic one are ruled out in a completely uniform way: Because $[uscr]$ cannot search into a CP, a DP inside a CP is inaccessible, regardless of whether this DP remains in its base position (as in (450a)) or has undergone \bar{A} -movement (as in (450b)). This result is schematized in (452).

(452) a. *One-fell-swoop movement:*



b. *Successive-cyclic movement:*



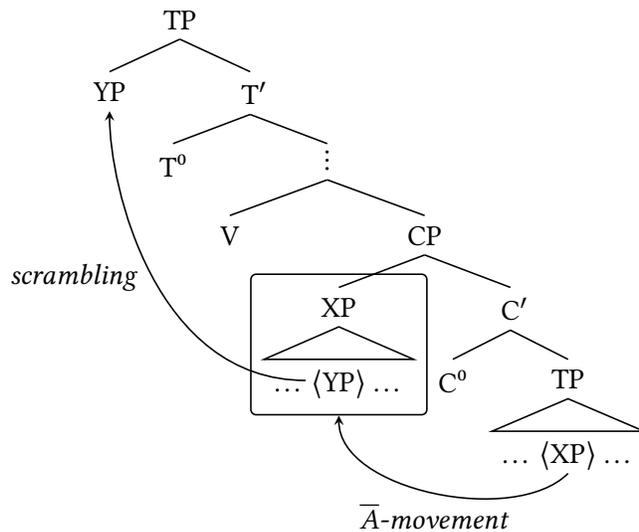
This contrast between horizons and traditional item-based accounts has two immediate consequences. First, the horizons account does not require direct constraints on interactions between movement types such as (451). Constraints like (451) are convenient descriptive generalizations, but not theoretical primitives. On the horizons account, \bar{A} -movement cannot feed scrambling

because \bar{A} -movement entails the presence of a CP structure in the embedded clause, which then constitutes a horizon to [*uscr*], hence blocking scrambling over it. The account of movement type interactions is therefore completely analogous to the account of improper movement in chapter 3. These considerations of course extend to the various other selective opacity effects observed in this chapter. In all of these cases, movement types interactions emerge as epiphenomena. This line of account is made available by the domain-based nature of horizons.

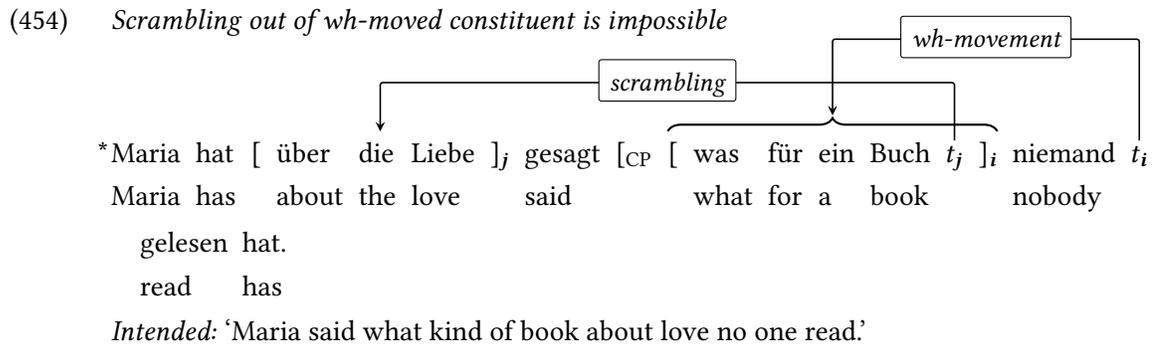
A second consequence is that horizons provide a uniform account of the locality of, e.g., scrambling, whereas traditional item-based accounts require a conspiracy of unrelated constraints. As just discussed, an item-based constraint requires two (and as we will see shortly, even three) constraints to rule out scrambling out of a finite clause, one enforcing successive-cyclic movement and one blocking scrambling from an \bar{A} -position (see (451)). Because the horizon account rules out both derivations in exactly the same way, it affords a truly unified account of the locality of scrambling, and by extension of selective opacity.

An important further argument in favor of a domain-based account comes from smuggling derivations, already discussed in the context of superraising in section 3.4.2 of chapter 3. Continuing to use scrambling as an illustrative example, a smuggling derivation involves \bar{A} -movement of an XP to the edge of a CP clause, followed by scrambling of YP out of XP into the higher clause, as schematized in (453).

(453) *Illicit smuggling derivation*



Smuggling derivations like (453) are impossible, as Abels (2007, 2009) argues in detail. The example in (454) demonstrates this:

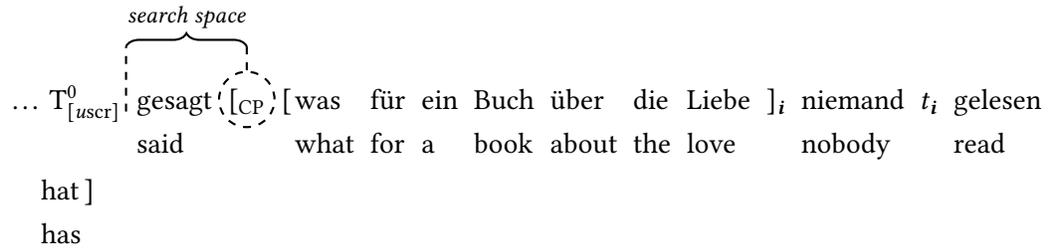


(454) involves scrambling over a finite clause boundary and its ungrammaticality should therefore be connected to that of (450). Only a domain-based account achieves this objective, as I will now show.

Consider first item-based accounts. Such accounts require a third locality constraint to block the derivation in (454) because element that undergoes \bar{A} -movement in (454) is not the same as the element that is scrambled. An item-based constraint like (451) that simply rules out scrambling of an \bar{A} -moved element therefore fails to extend to (454). A separate constraint is thus needed on an item-based account. One possibility of such a constraint is *freezing* (Wexler & Culicover 1980), which prohibits extraction out of a moved constituent. But the invocation of freezing only underscores the fact that the simple and elegant generalization that scrambling cannot leave a finite clause is nothing more than an accident on an item-based account, as it results from a conspiracy between (i) a constraint forcing successive cyclicity (to rule out (450a)), (ii) a constraint blocking scrambling of \bar{A} -moved elements (to block (450b)), and (iii) a freezing constraint (to block (454)).

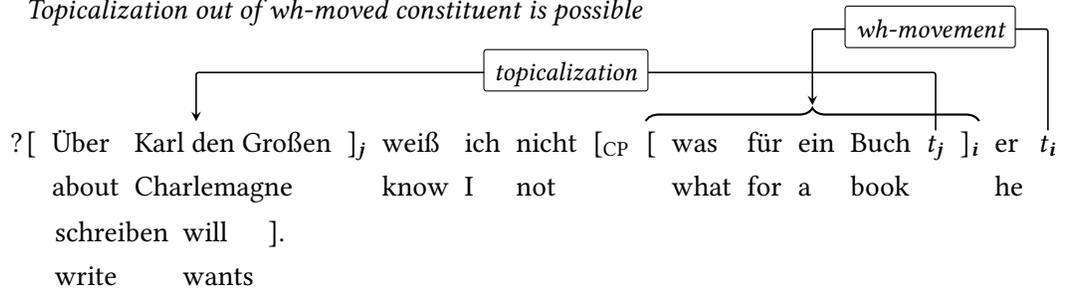
The horizons account, on the other hand, extends to smuggling derivations like (454) without further ado, simply because (454) involves scrambling over a CP node, hence a violation of [uscr]’s horizon. This is schematized in (455). The constituent *über die Liebe* ‘about love’ lies beyond [uscr]’s CP horizon and hence cannot be scrambled. No conspiracy of multiple unrelated locality principles is thus required. The domain-based nature of horizons offers a unified account of the locality of scrambling in all of its manifestations.

(455) *Impossible scrambling out of wh-moved constituent*



In addition to this conceptual advantage of the horizons account, there is empirical evidence against a blanket freezing constraint that rules out all movement out of moved constituents, hence undermining the item-based account of (454). The situation here is entirely analogous to that for superraising discussed in section 3.4.2 of chapter 3 and I take this convergence to be significant. As Abels (2007, 2009) emphasizes, topicalization out of a *wh*-moved element is much better than scrambling out of such an element. Thus, there is a contrast between (454) and (456).²⁵

(456) *Topicalization out of wh-moved constituent is possible*



‘About Charlemagne I don’t know what kind of book he wants to write.’

(based on Abels 2007: 78)

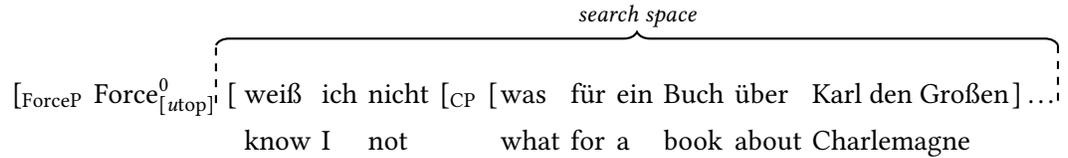
(456) is substantially better than (454). If the ungrammaticality of (454) were due to a general ban on extraction out of a moved element, both structures should have the exact same grammaticality status, contrary to fact. The problem here is that smuggling derivations themselves instantiate selective opacity. Item-based accounts are unable to accommodate this fact because descriptively smuggling derivations involve interactions between two distinct items.

The horizons account, on the other hand, straightforwardly derives the contrast between (454) and (456), as it is based on domains instead of items. Because [uscr] ⊨ T, CPs are entirely opaque to scrambling (see (455)). By contrast, because [utop] ⊨ ∅, CP clauses are not opaque to [utop], as

²⁵ Abels (2007) rates (456) as ‘??’, I find it quite good.

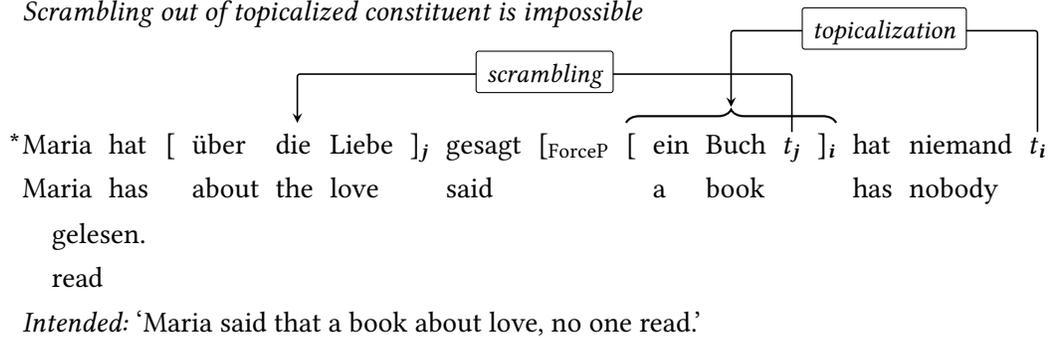
schematized in (457). Topicalization out of CPs is therefore possible and (456) is ruled in, as desired. Selective opacity in smuggling derivations is thus entirely expected on a horizons account and does not require any new analytical machinery.

(457) *Possible topicalization out of wh-moved constituent*



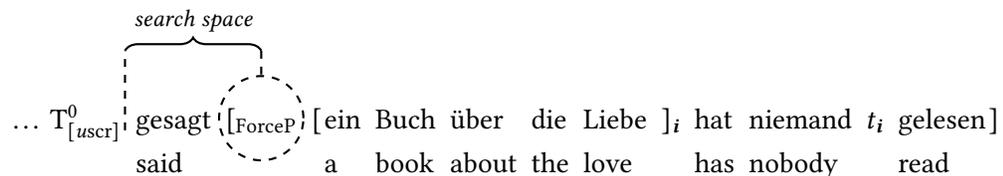
The horizons account makes immediate predictions about possible and impossible smuggling derivations. For example, because [uscr] \neq T, scrambling out of a topicalized constituent is likewise predicted to be ungrammatical. This is because topicalization targets SpecForceP and hence requires the embedded clause to be a ForceP, creating a horizon for [uscr]. This prediction is indeed borne out (Abels 2007, 2009).

(458) *Scrambling out of topicalized constituent is impossible*



Just as in the case of (454), there is no direct interaction between topicalization and scrambling. There interaction follows indirectly from horizons.

(459) *Impossible scrambling out of topicalized constituent*



Furthermore, the horizons account makes the interesting prediction that inverting the sequence of movement types in impossible smuggling derivations should render them grammatical. For

example, we have seen in (454) above that scrambling out of a *wh*-moved constituent is impossible, a horizon effect. The inverse, i.e. *wh*-movement out of a scrambled constituent, is predicted to be possible because scrambling does not implicate a horizon for the *wh*-movement probe. Asymmetries of this type have been documented in detail by Abels (2007, 2009). The fact that *wh*-movement out of a scrambled constituent is possible is demonstrated in (460a). Here, the complex DP *ein fertiges Manuskript über welchen deutschen Kaiser* ‘a done manuscript about which German emperor’ is scrambled above the subject, followed by *wh*-subextraction of *über welchen deutschen Kaiser* ‘about which German emperor’ out of it. As (460b) and (460c) demonstrate, smuggling derivations are also possible if the second movement step is driven by [*utop*_(wh)], be it topicalization of a non-*wh* element or *wh*-movement into a V2 clause. All three structures are grammatical.

(460) a. *Wh-movement out of scrambled constituent*

Maria hat gefragt [_{CP} [über welchen deutschen Kaiser]_j [ein fertiges
 Maria has asked about which German emperor a done
 Manuskript *t_j*]_i leider keiner der anwesenden Historiker *t_i*
 manuscript unfortunately none of.the present historians
 anzubieten hatte].
 to.offer had

‘Maria asked about which German emperor none of the historians present could offer a complete manuscript.’

b. *Wh-topicalization out of scrambled constituent*

[Über welchen deutschen Kaiser]_j hatte [ein fertiges Manuskript *t_j*]_i
 about which German emperor had a done manuscript
 leider keiner der anwesenden Historiker *t_i* anzubieten?
 unfortunately none of.the present historians to.offer

‘About which German emperor could none of the historians present offer a completed manuscript?’ (Abels 2007: 77)

c. *Topicalization out of scrambled constituent*

[Über Karl den Großen]_j hatte [ein fertiges Manuskript *t_j*]_i leider
 about Charlemagne had a done manuscript unfortunately
 keiner der anwesenden Historiker *t_i* anzubieten?
 none of.the present historians to.offer

‘About Charlemagne, none of the historians present could offer a completed manuscript.’

(Abels 2007: 77)

Horizons account for the grammaticality of (460) because in all three cases, the first movement step (scrambling) does not entail the existence of a horizon for the probe underlying the second movement step and subextraction is consequently allowed.

Analogously, we saw in (456), repeated here, that topicalization out of a *wh*-element is possible, because the CPs necessary for the application of *wh*-movement do not create horizons for topicalization. Horizons predict that the inverse derivation, i.e., *wh*-movement out of a topicalized constituent is impossible. This is correct, as (462) demonstrates.

- (461) *Topicalization out of wh-moved constituent* [= (456)]
 ?[Über Karl den Großen]_j weiß ich nicht [CP [was für ein Buch t_j]_i er t_i
 about Charlemagne know I not what for a book he
 schreiben will].
 write wants
 ‘About Charlemagne I don’t know what kind of book he wants to write.’

- (462) *Wh-movement out of topicalized constituent*
 *Ich weiß nicht [[über welchen deutschen Kaiser]_j er gesagt hat [ForceP [ein
 I know not about which German emperor he said has a
 Buch t_j]_i will keiner t_i schreiben]]
 book wants nobody write
Intended: ‘I don’t know about which German emperor he said that no one will write a book
 about.’

This asymmetry again follows from horizons: Topicalization implicates a ForceP structure, which constitutes a horizon to [*uwh*] in the higher clause:

- (463) *Impossible wh-movement out of topicalized constituent*
-
- ... C<sub>[uwh]}⁰ er gesagt hat [ForceP] [ein Buch über welchen deutschen Kaiser]_i
 he said has a book about which German emperor
 will keiner t_i schreiben]
 wants nobody write</sub>

These contrasts in the possibility of smuggling derivations are similar to those observed in section 3.4.2 in chapter 3 for English. There we have seen that A-movement out of an \bar{A} -moved

constituent is altogether impossible, whereas the reverse is grammatical. The horizons account of these facts is exactly the same as for the German facts just discussed.

By contrast, a simple freezing account of illicit cases of smuggling like (454) and (458) predicts that *all* smuggling derivations are ruled out. This prediction is evidently incorrect. An immediate benefit of the horizons account is that it provides an account of when smuggling is possible and when it is not: Smuggling is impossible if the first movement step entails the presence of a horizon for the probe triggering the second. The intricate asymmetries just observed then follow.²⁶

The non-identity cases considered so far all involve subextraction of an element out of a moved constituent. Analogous constraints on feeding and bleeding relationships have been observed for remnant movement (see Grewendorf 2003, 2015 and Abels 2007, 2009, who make this claim for German, and Williams 2003 for general remarks to the same effect). As documented in great detail by Müller (1996, 1998), a remnant created by scrambling can undergo topicalization and *wh*-movement, as shown by the three examples in (464), which are based on Müller (1996: 362–363).

²⁶ While horizons thus make clear predictions about impossible smuggling derivations, it is possible and in fact likely that constraints other than horizons enter the empirical picture and rule out instances of smuggling that do not violate horizons. One pertinent example of such a constraint is that, e.g., topicalization out of a topicalized element is ungrammatical:

- (i) * [Über Karl den Großen]_j hat er gesagt [_{ForceP} [ein Buch *t_j*]_i will keiner *t_i* schreiben]
 about Charlemagne has he said a book wants nobody write
Intended: ‘About Charlemagne, he said that a book nobody wants to write.’

As far as horizons are concerned, (i) should be well-formed because ForcePs do not constitute horizons for the higher [*utop_(wh)*] probe.

The ungrammaticality of (i) is arguably part of a larger empirical generalization discovered by Takano (1994) and Müller (1996) in the context of remnant movement and sometimes called the *Müller–Takano Generalization*. According to this generalization, movement out of constituent is blocked if that constituent undergoes the same movement type. That, scrambling out of constituent that itself undergoes scrambling is impossible, topicalization out of a constituent that itself undergoes topicalization is impossible, and so on. Müller (1996) calls this the *Principle of Unambiguous Domination* and proposes that it follows from the F-over-F Principle (for which see section 3.2.2 in chapter 3).

- (ii) *Principle of Unambiguous Domination* (Müller 1996: 375)
 An α -trace must not be α -dominated.

The term ‘ α ’ refers to movement type features: A trace created by, e.g., scrambling may not be dominated by a node that bears a scrambling feature, i.e., that itself undergoes scrambling. The Principle of Unambiguous Binding derives the ill-formedness of (i) above.

It is important to note that (ii) does not capture any of the movement type interactions discussed in the main text as those involve two elements undergoing distinct movement steps. Horizons and the Principle of Unambiguous Domination thus complement each other in their empirical coverage.

- (464) a. *Topicalization of remnant created by scrambling*
 [Ein Buch t_i]_j hat Petra [über die Liebe]_i sicher nicht t_j gelesen.
 a book has Antje about the love surely not read
 ‘Petra surely did not read a book about love.’
- b. *Wh-topicalization of remnant created by scrambling*
 [Was für ein Buch t_i]_j hat Petra [über die Liebe]_i sicher nicht t_j
 what for a book has Petra about the love surely not
 gelesen?
 read
 ‘What kind of book about love did Petra surely not read?’
- c. *Wh-movement of remnant created by scrambling*
 Maria hat gefragt [CP [was für ein Buch t_i]_j Petra [über die Liebe]_i
 Maria has asked what for a book Petra about the love
 sicher nicht t_j gelesen hat]
 surely not read has
 ‘Maria asked what kind of book about love Petra surely didn’t read.’

The inverse, on the other hand, is impossible. Structures in which a remnant created by either topicalization, as in (465a), or *wh*-movement, as in (465b), undergoes scrambling are completely ungrammatical.

- (465) a. *Scrambling of remnant created by topicalization*
 *Maria hat [ein Buch t_j]_i gesagt [ForceP [über die Liebe]_j hat Hans t_i
 Maria has a book said about the love has Hans
 gelesen]
 read
Intended: ‘Maria said that Hans read a book about love.’
- b. *Scrambling of remnant created by wh-movement*
 *Maria hat [ein Buch t_j]_i gesagt [CP worüber_j Hans t_i gelesen hat]
 Maria has a book said about.what Hans read has
Intended: ‘Maria said about what Hans read a book.’

The contrast between the grammatical instance of remnant movement in (464) and the ungrammatical ones in (465) straightforwardly falls out from the horizons account. In (465), the remnant-creating movements entail the presence of a ForceP or CP projection in the lower clause. These projections

then constitute horizons for the matrix [*uscr*] probe. Remnant scrambling is thus impossible. In (464), on the other hand, the remnant-creating scrambling does not entail the presence of a horizon for [*uwh*] and [*utop_(wh)*], which subsequently move the remnant. Remnant movement is thus possible in these cases.²⁷

In sum, the horizons account proposed here offers a comprehensive account of selective opacity effects in German. It successfully captures the facts (i) that selective opacity is not a binary distinction between movement types but manifold, (ii) that there exist entailments between the transparency and/or opacity of various clauses, (iii) that certain movement types descriptively bleed the application of other movement types, and (iv) that such bleeding even arises if the elements undergoing the two movement steps are not the same (i.e., in smuggling and remnant movement derivations). Crucial to these explanations is the domain-based nature of horizons accounts proposed here. The account does not in any way invoke properties of the moving element, such as its internal features, the position it occupies, etc. Rather, the account is couched exclusively in terms of probes and their horizons. I have shown that this shift in perspective away from a more traditional item-based account provides a conceptually more elegant account because it allows us to capture selective opacity effects in a truly uniform way, rather than a conspiracy of at least three unrelated constraints. Moreover, the account has a significantly improved empirical coverage because it immediately extends to selective opacity in smuggling derivations, which as I have shown are out of reach of a purely item-based account. The conceptual appeal of the account is hence accompanied by a clear gain in empirical scope.

In the remainder of this chapter, I will briefly address a data point that does not follow from the horizons account, which I take to show that phases and horizons constitute independent constraints on syntactic locality (section 4.5). I will then illustrate how the horizons account not only allows us to capture the various instances of selective opacity evident in German, but how it also imposes systematic constraints on possible selective opacity patterns in the form of the Height–Locality

²⁷ Just as in the case of smuggling derivation, something else needs to be said about configurations in which a remnant undergoes the same movement type that created the remnant, e.g., scrambling of a remnant created by scrambling. Müller (1996, 1998) argues that such derivations are impossible, a fact that he attributes to his *Principle of Unambiguous Domination*, discussed in fn. 26 on p. 302. As noted there, horizons and the Principle of Unambiguous Domination complement each other in their empirical coverage.

Connection (section 4.3.3). In an appendix, I will then briefly consider and reject an additional instance of selective opacity sometimes argued for in the literature on German (section 4.7).

4.5 The role of phases

One pertinent generalization discussed in section 4.2.3 concerns extraction out of a V2 clause. We have seen there that scrambling, relativization, and *wh*-movement into a V-final clause are all impossible out of a V2 clause. The horizons account developed in the preceding section captures these restrictions. We have also seen that topicalization or *wh*-movement into a V2 clause is possible out of a V2 clause, a fact that is accounted for because $[u\text{top}_{(wh)}] \dashv\!\! \dashv \emptyset$, rendering ForcePs transparent to $[u\text{top}_{(wh)}]$. An additional constraint, which is not accounted for on the analysis so far developed, concerns the path of licit extraction out of a V2 clause. As shown by the examples in (466), such extraction is possible if it is successive-cyclic, i.e., if the moving element creates an intermediate landing site at the edge of the lower V2 clause. This holds for *wh*-movement and topicalization alike.

- (466) *Successive-cyclic extraction out of V2 clause* [=(412),(414)]
- a. Wen_i hat er gesagt [*t_i* hat die Maria *t_i* getroffen]?
 whom has he said has the Maria met
 ‘Who did he say that Maria met?’
- b. Den Paul_i hat er gesagt [*t_i* hat die Maria *t_i* getroffen].
 the Paul.ACC has he said has the Maria.NOM met
 ‘Paul, he said that Maria met.’

Crucially, this intermediate landing site is obligatory. One-fell-swoop extraction is completely ungrammatical:

- (467) *One-fell-swoop extraction out of V2 clause* [=(413),(415)]
- a. *Wen_i hat er gesagt [die Maria hat *t_i* getroffen]?
 whom has he said the Maria has met
Intended: ‘Who did he say that Maria met?’

- b. *Den Paul_i hat er gesagt [die Maria hat *t_i* getroffen].
 the Paul.ACC has he said the Maria.NOM has met
Intended: ‘Paul, he said that Maria met.’

It is easy to demonstrate that (466) indeed involves an embedded V2 clause whose preverbal position is filled with the intermediate trace of the moving element, and not just a verb-initial clause. The reason is that in the absence of extraction, a verb-initial embedded clause is impossible:

- (468) a. Er hat gesagt [die Maria hat den Paul getroffen].
 he has said the Maria.NOM has the Paul.ACC met
 ‘He said that Maria met Paul.’
- b. *Er hat gesagt [hat die Maria den Paul getroffen].
 he has said has the Maria.NOM the Paul.ACC met
Intended: ‘He said Maria met Paul.’

Extraction out of a V2 clause, then, has to obligatory proceed through the edge of that clause, a standard view in German syntax (Thiersch 1978, *et seq.*). This requirement does not follow from the horizons account developed in the preceding sections. The reason is that the relevant probe [$u_{top(wh)}$] does not have a horizon and, all else equal, should be able to probe indefinitely deep into the lower clause, agreeing with the embedded object in its base position and extracting it in one fell swoop. Conversely, if [$u_{top(wh)}$] had Force as its Horizon, then it would not be able to search into the embedded clause at all, not even its edge, and no extraction out of a V2 clause would be possible, contrary to fact.

More generally, edge effects and successive cyclicity cannot be attributed to horizons, a point which I will discuss in greater depth in chapter 6. For now, it suffices to note that horizons determine whether a given clause is transparent or opaque for a given probe. Constraints such that only the edge of a clause is accessible to a probe are compatible with horizons, but cannot be attributed to them. This is unsurprising, of course, as it merely shows that there are locality constraints beyond horizons, an unremarkable state of affairs.

To derive the contrast between (466) and (467), I will invoke the standard account of edge effects in terms of PHASES (Chomsky 2000, 2001, *et seq.*). Assuming for now that Force⁰ constitutes a phase head (though see chapter 6 for a somewhat different view), only an element in the specifier

of ForceP will be accessible to operations outside of ForceP because the complement of Force⁰ and is spelled-out and its contents hence removed from the syntactic workspace. See section 3.2.1 in chapter 3 and chapter 6 for more background on phases.

(469) **Phase Impenetrability Condition** (Chomsky 2000: 108)

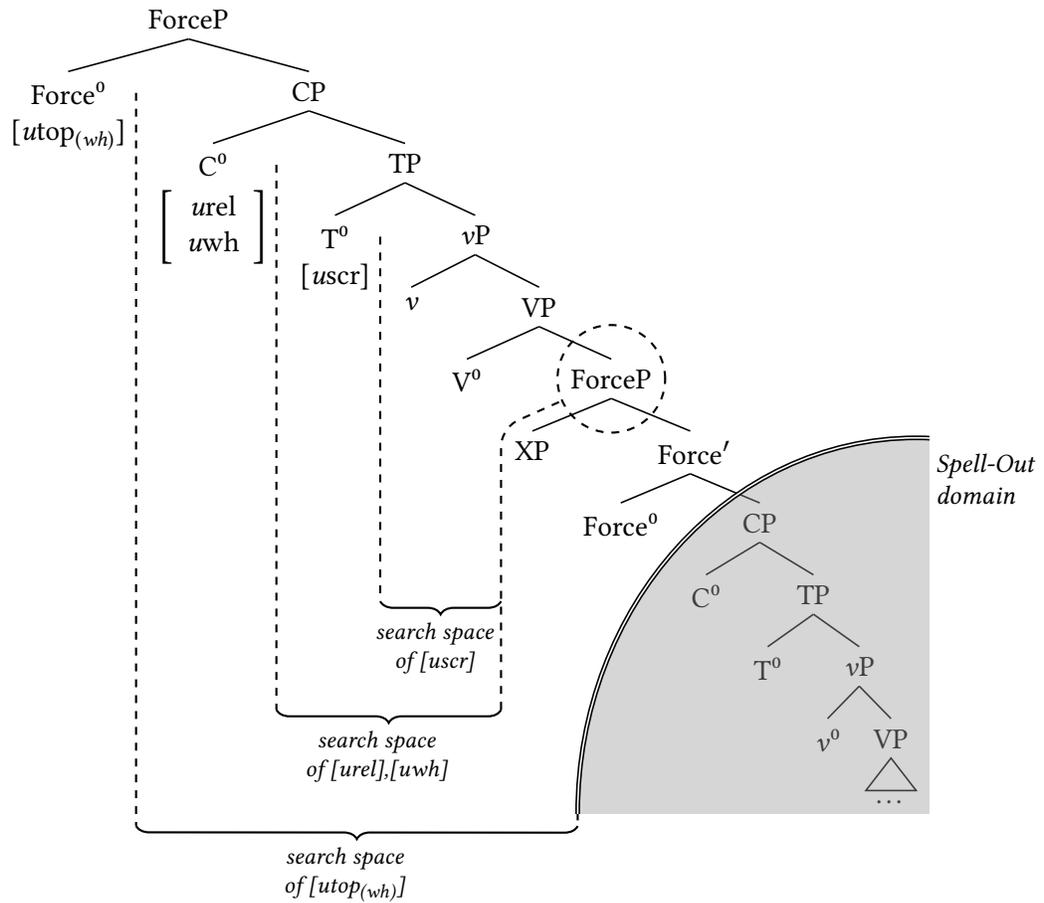
In phase α with head H, the domain of H is not accessible to operations outside of α , only H and its edge are accessible to such operations.

Phases differ from horizons in two fundamental aspects. First, phase-based locality is absolute and applies to all operations alike. If a part of a structure is removed from the workspace, it becomes inaccessible to *all* operations alike. Horizons, on the other hand, can crucially differ between different probes, giving rise to locality differences between various operations. Put somewhat differently, while horizons give rise to *selective* opacity, phases induce *unselective* opacity. Second, whereas horizons render an embedded clause entirely opaque to a given probe, phases do not, because the edge of the embedded clause remains accessible to outside probes. As such, phases are porous, but horizons are not.

The joint effect of horizons and phases is schematized in (470), which combines the horizons in (447) above with an indication of the Spellout domain of a ForceP phase. The specifier of that ForceP is filled by some element XP.²⁸

²⁸ In chapter 6, I will provide evidence that in German it is C⁰ that defines a phase head, not Force⁰. This distinction is inconsequential for the case at hand. I will stick to Force⁰ as the phase head here as it allows me to make the point more easily.

(470) *Horizons and phases in ForceP embedding*



Consider the probe $[utop_{(wh)}]$ first. Because $[utop_{(wh)}] \dashv \emptyset$, there is no horizon for this probe and its search can enter the embedded ForceP clause. However, because $Force^0$ by assumption is a phase head, the CP and everything inside it will be spelled out by the time the matrix is being built. As a result, all elements inside the CP will have been removed from the search space and hence rendered inaccessible for all probes in the matrix clause, including $[utop_{(wh)}]$. Only XP in SpecForceP will be accessible to $[utop_{(wh)}]$ and thus able to topicalize into the matrix clause. This derives the contrast between (466) and (467). Specifically, the one-fell-swoop structure in (467) is ruled out by the PIC in (469).

It is important to note that the invocation of phases does not supersede or render invalid the horizon account developed here. It is still necessary to ensure that no movement other than (*wh*-)topicalization can leave an embedded V2 clause. Thus, while XP in SpecForceP remains accessible to the matrix $[utop_{(wh)}]$ probe in (470), XP must be inaccessible for the other probes $[uscr]$, $[urel]$, and $[uwh]$. As shown in section 4.2.3, scrambling out of an embedded V2 clause is impossible

even if it successive-cyclic and therefore PIC-compliant. The same holds for relativization and *wh*-movement into a V-final clause, as has been demonstrated in section 4.2.3 above.

- (471) *Er hat den Fritz_i gesagt [_{t_i} hat die Maria _{t_i} getroffen]. [= (408b)]
he has the Fritz.ACC said has the Maria.NOM met
Intended: ‘He said that Maria met Fritz.’

This is the by now familiar effect of horizons: The search of these latter three probes terminates at the ForceP node and elements at the edge of the lower clause are out of reach for these probes. Thus, (471) is ruled out because an element in SpecForceP is inaccessible to [*uscr*], [*urel*], and [*uwh*] because of their horizons.

Summing up, phases constitute a second constraint on movement dependencies in addition to horizons. The two have distinct empirical signatures. While horizons determine whether an operation is possible across a clause boundary, phases determine the successive-cyclic extraction path of dependencies allowed by horizons. To be well-formed, a movement has to conform to both the constraints imposed by horizons and the PIC. Thus, scrambling out of a V2 clause such as (471) is excluded by horizons, whereas one-fell-swoop movement out of a V2 clause is ruled out by phases. No redundancy arises between horizons and phases. The inclusion of phases supplements, but does not otherwise affect, the account of selective opacity in terms of horizons laid out in the previous section.

The role of phases in the horizon system is discussed at greater length in chapter 6. There I not only provide further evidence for the analytical role of phases in addition to horizons, I also show that only CP-domain phases, but not *v*P-domain phases, yield a consistent account. As such, I will show that selective opacity effects in general and horizons in particular provide a new window into the distribution of phases, whose conclusions are supported by evidence from independent domains.

4.6 Constraining selective opacity: The Height–Locality Connection

One surprising general property of selective opacity effects is that the locality profile of a syntactic operation is not arbitrary, but instead partially correlated with the structural location of the probe that underlies this process, a correlation that I have called the Height–Locality Connection (see

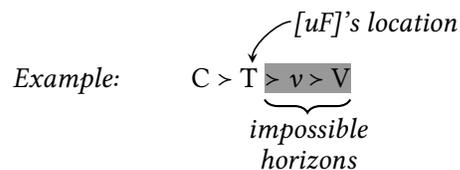
the discussion in section 1.3.1 of chapter 1, section 2.5.2 of chapter 2, section 3.5 of chapter 3, and section 4.3.3 of the present chapter).

- (472) **Height–Locality Connection** (*probe-based formulation*) [=(435)]
 The higher the location of a probe is in the clausal structure, the more kinds of structures are transparent to this probe.

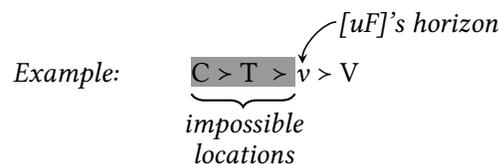
I have shown in section 3.5.1 of chapter 3 that horizons derive a systematic link between height and locality that has a version of (472) as its result. Specifically, we have seen there that the horizon account gives rise to the Height–Locality Theorem (HLT) in (473), which states that the syntactic location of a probe restricts its possible horizons (473a), and that, conversely, a given horizon setting limits the range of possible structural positions for a probe (473b). The HLT derives a version of (472). Importantly, (473) is not a stipulation superimposed on the system, but rather a consequence emerging from it (hence the label ‘theorem’). As discussed at length in section 3.5.1 of chapter 3, pairings of height and locality that do not conform to (473) are possible, but produce vacuous probes, i.e., probes whose search is so deprived that they are unable to trigger any movement or agreement operations, and which are. For all probes that are able to establish movement and agreement dependencies that are hence detectable, height and locality must stand in the relation specified by (473). Thus, (473) is an emergent property of the horizon systems and hence genuinely derived in it. For extensive discussion, see section 3.5 of chapter 3.

- (473) **HEIGHT–LOCALITY THEOREM** [=(333)]
 Given an extended projection $\Phi = \langle \Pi_n > \Pi_{n-1} > \dots > \Pi_1 \rangle$, for any non-vacuous probe $[uF]$,

- a. If $[uF]$ is located on Π_m , then a projection $\in \{\Pi_{m-1}, \dots, \Pi_1\}$ cannot be a horizon for $[uF]$.



- b. If $[uF]$ has Π_m as a horizon, $[uF]$ cannot be located on a projection $\in \{\Pi_n, \dots, \Pi_{m+1}\}$.



Horizons have the distinctive property that they impose a link between a probe's location and its locality. Determining one immediately narrows down the possible range of the other. Because horizons determine which types of embedded clauses are opaque to a probe, the fact that the location of a probe delimits its possible horizon settings automatically makes predictions about which types of clauses are transparent to this probe. For example, if a probe does not have a TP as its horizon by (473), then TP clauses are necessarily transparent to this probe.

It can be easily verified that the HLT (473) holds for the account of German developed in this section. To see this, consider again the four relevant probes:

- (474) *German probes and their horizons* [=(439)]
- a. *Scrambling:*
[uscr]_{T⁰} -|| T
 - b. *Relativization:*
[urel]_{C⁰} -|| C
 - c. *Wh-movement (in V-final clause):*
[uwh]_{C⁰} -|| Force
 - d. *(Wh)-Topicalization:*
[utop_(wh)]_{Force⁰} -|| ∅

The HLT and the horizon account that gives rise to it provide an explanation for various properties of (474):

- (475) *Empirical effects of the Height–Locality Theorem (473) in German*

a. **Height → Locality (473a):**

i. *(Wh)-Topicalization*

Section 4.3.3 has presented evidence that [utop_(wh)] is located on Force⁰. It follows from (473a) that neither V nor *v* nor T nor C can be a horizon for [utop_(wh)]. Consequently, *v*P clauses, TP clauses, and CP clauses are necessarily transparent to this probe. It follows that topicalization *must* be possible out of coherent infinitives, non-coherent infinitives, and V-final finite clauses.

ii. *Wh-movement*

Section 4.3.3 has shown that [uwh] is located on C⁰. It then follows from (473a) that neither V nor *v* nor T can be horizons for [uwh]. Consequently, *v*P and TP clauses *must* be transparent to [uwh]. In other words, *wh*-movement *must* be possible out of coherent and non-coherent infinitives.

iii. *Relativization*

Because [*urel*] is located on C^0 (see section 4.3.3), the same reasoning applies as for [*uwh*]. (473a) entails that coherent infinitives and non-coherent infinitives *must* be transparent to relativization.

iv. *Scrambling*

Given that scrambling can land in SpecTP, [*uscr*] must be located on T^0 .²⁹ (473a) then entails that [*uscr*]’s horizon cannot be v or V. Since non-coherent infinitives are v Ps, it follows that scrambling out of such infinitives *must* be possible.

b. **Locality → Height (473b):**

i. *Relativization*

Relativization is impossible out of CP clauses (section 4.2.2) and therefore must have C as its horizons. (473b) then entails that [*urel*] cannot be located on Force⁰. This in turn entails that relativization *must* be possible in a V-final finite clause. As we have seen, this is correct.

ii. *Scrambling*

Because scrambling is impossible out of non-coherent infinitives and thus TPs (section 4.2.1), it must have T as its horizons. (473b) then entails that the location of [*uscr*], and hence the landing site of scrambling, *cannot* be higher than TP, i.e., on C^0 or Force⁰. This not only implies that scrambling must target a structurally low position, it also entails that scrambling *must* be possible in infinitival clauses, unlike *wh*-movement or relativization. This prediction is borne out, as (476) shows, where *das Buch* ‘the book’ is scrambled inside an infinitival clause.

- (476) Fritz hat t_j versucht [*das Buch*_{*i*} *der Maria* t_i zu geben]_{*j*}.
Fritz has tried the book.ACC the Maria.DAT to give
‘Fritz tried to give Maria the book.’

The various entailment relationships between height and locality that arise from the HLT (473) specified in (475) are indeed all borne out. On the horizon account, they are derived and as such necessary properties of the selective opacity pattern of German. In other words, horizons provide an explanation for *why* many of the relations detailed in (475) hold in first place. Just as in the case of Hindi, horizons not only provide an account of locality differences between movement types. At

²⁹ In addition to possibly lower heads, see section 4.3.3.

the same time, they impose a limit on such locality differences, making empirical predictions about what is and what is not a possible selective opacity pattern.³⁰

4.7 Detour: Topic islands and *wh*-islands

Before closing this chapter, I would like to briefly mention an additional layer of selective opacity that has sometimes been argued for in the literature on German, but that I have not integrated into the account developed in this chapter. This additional layer is based on the interaction between *wh*-movement and topicalization. Based on work by Fanselow (1987) and Bayer (1989), Sternefeld (1992), and Müller & Sternefeld (1993) argue that *wh*-islands are marginally transparent for topicalization, but opaque to *wh*-movement, whereas topic islands (i.e., clauses in which topicalization has taken place) are categorically opaque to both, a conclusion that is also reached by Müller (1995, 2011), d’Avis (1996), and Abels (2007, 2009). The conclusion is based on contrasts like the one in (477), in which the lower clause is a *wh*-islands. As shown in (477a), topicalization out of this island is marginally possible, while (477b) is taken to show that *wh*-extraction is impossible.

(477) a. ??Radios_i weiß ich nicht [wie (dass) man *t_i* repariert].
 radios know I not how that one repairs
 ‘Radios, I don’t know how one repairs.’

b. *Welches Radio_i weißt du nicht [wie (dass) man *t_i* repariert]?
 which radio know you not how that one repairs
 Intended: ‘Which radio do you not know how one repairs?’

(Müller & Sternefeld 1993: 485)

Topic islands, on the other hand, are uncontroversially opaque to all extraction types:

(478) a. *Radios_i glaube ich [gestern hat Ede *t_i* repariert].
 radios believe I yesterday has Ede repaired
 Intended: ‘Radios, I believe Ede repaired yesterday.’

³⁰ It is worth reiterating that the connection between height and locality that derives from horizons is not a one-to-one mapping between the two. A probe’s location imposes limits on the possible values for its horizons and vice versa, but neither completely predicts the other. See section 3.5 in chapter 3 for extensive discussion.

- b. *Was_i glaubst du [gestern hat Ede t_i repariert]?
 what believe you yesterday has Ede repaired
Intended: ‘What do you believe Ede repaired yesterday?’

(Müller & Sternefeld 1993: 485)

Because all extractions in (477) and (478) land in a V2 clause, these data would appear to require a further analytical separation of topicalization and *wh*-movement into a V2 clause.

On the account developed here, topicalization and *wh*-movement into a V2 clause are triggered by the same feature [*utop_(wh)*] and no locality differences between the two are therefore expected. The account could be conservatively modified to accommodate a contrast between topicalization and *wh*-movement by simply enriching the array of movement-driving features. As such, asymmetries between the two fall within the analytical range of the account just given.

It seems to me, however, that the empirical evidence for a difference between topicalization and *wh*-movement into V2 clauses is rather weak. It is evident that the crucial contrast in (477) is not a minimal pair. While (477a) involves the extraction of a kind-denoting element, (477b) is based on extracting a specific *wh*-element. It is not clear from (477) whether the grammaticality difference between the two is a result of the difference in extraction type or of the semantic properties of the moving elements. A more suitable pair is given in (479). In both sentences, the moved element is non-specific. In this case, the *wh*-movement structure in (479b) improves considerably, to the point where both sentences have the same marginal grammaticality status.

- (479) a. ??Radios_i weiß er nicht [wie man t_i repariert].
 radios knows he not how one repairs
 ‘Radios, he doesn’t know how one repairs.’
- b. ??Was für Radios_i weiß er nicht [wie man t_i repariert]?
 what for radios knows he not how one repairs
 ‘What kind of radios does he not know how one repairs?’

Once the specificity confound inherent in (477) is controlled for, there is no indication that there is a locality asymmetry between topicalization and *wh*-movement into a V2 clause with respect to *wh*-islands.

This raises the second question of whether there exists an extraction difference between the *wh*-islands and topic islands. While the contrast between (478) and (479) suggests this, it should

be noted that the difference between *wh*-islands and topic islands is confounded with a second difference, namely whether the embedded clause is V-final or V2. This confound arises first because topicalization takes place only in V2 clause. Topic islands like (478) are thus necessarily V2 clauses. Second, it is a general fact of German that it is not possible to have *wh*-movement in embedded V2 clauses (e.g., Tappe 1981, Haider 1984, Reis 1985). Put differently, embedded interrogative clauses, and hence *wh*-islands like (477) and (479) have to be V-final. Even a verb like *sagen* ‘say’ which allows embedding of both a V2 clause (see (480a)) and an interrogative clause (see (480b)) does not allow for an interrogative V2 clause (see (480c)).

- (480) a. Maria hat gesagt [der Fritz ist angekommen].
 Maria has said the Fritz is arrived
 ‘Maria said that Fritz arrived.’
- b. Maria hat gesagt [wer angekommen ist].
 Maria has said who arrived is
 ‘Maria said who arrived.’
- c. *Maria hat gesagt [wer ist angekommen].
 Maria has said who is arrived

This restriction is peculiar and I will have nothing to say about how to derive it. For our present concerns, it suffices to note that *wh*-island are confounded with verb position, just like topic islands. While topic islands are always V2, *wh*-islands are always V-final. As a consequence, it is questionable whether *wh*-movement and topicalization differ in the strictness of the locality domain they produce or whether the independently necessary locality difference between V-final and V2 clauses is sufficient to capture the relevant contrasts. Thus, extraction out of a V2 clause that has its edge filled is completely impossible, whereas extraction out of V-final clause with a filled SpecCP is marginally possible.³¹ There is thus no evidence for a distinction between *wh*-islands and topic islands above and beyond that between V2 and V-final clauses as structures that could

³¹ This contrast between V-final and V2 clauses is interesting in its own right. But since it does not involve a movement type asymmetry it is arguably unrelated to selective opacity and an investigation would lead us too far afield. A straightforward, though non-explanatory account would be to stipulate that CPs can marginally have more than one specifier, while ForcePs can have only one (see Coon, Pedro & Preminger 2014 for a proposal based on such differences in a different domain). Another possibility is that ForceP is a phase but CP is not, so that successive-cyclic extraction is required only out of ForcePs. We will see in chapter 6, however, that there is evidence that CP is a phase in German.

motivate such a distinction (i.e., V-final topic islands and V2 *wh*-islands) are ruled out for general and independent reasons.

I conclude from these considerations that there is insufficient evidence to motivate a locality difference between topicalization and *wh*-movement into a V2 clause, either in the strictness of the island they produce or in their ability to leave embedded islands.

4.8 Chapter summary

This chapter has provided another case study of selective opacity. I have shown that in German, embedded clauses of various types differ with respect to their locality for several movement types, but in nonetheless systematic ways. The empirical evidence came from four movement types, which were crossed with four types of embedded clauses. The resulting pattern reveals a considerable amount of selective opacity. Each of the four movement types displayed a unique locality profile across the four clause types. At the same time, the distribution of opacity is not random, but clearly exhibits the signatures of Upward Entailment and the Height–Locality Connection. The fact that an investigation into selective opacity in German corroborates the key conclusions reached for Hindi in chapter 2, which is in turn consistent with other selective opacity effects observed in the previous literature (see chapter 1) provides, striking support for the empirical validity of these generalizations.

In the analysis of the German facts, I have emphasized the key distinctive features of the horizons account, which can be summarized as follows:

- (481)
- a. There is no direct interactions between movement types.
 - b. Horizons are domain-based in that they exclusively invoke probes and their horizons. The moving item itself is irrelevant.
 - c. Horizons extend beyond the binary A/\bar{A} -distinction.
 - d. Horizons extend to selective opacity in non-identity cases (smuggling and remnant movement).
 - e. Horizons derive Upward Entailment and a version of the Height–Locality Connection.

I have shown that a horizon-based account not only provides a uniform analysis of the various layers of selective opacity observed, but also constrains them and thereby succeeds in deriving Upward Entailment and the Height–Locality Connection. Just like in the case of Hindi in chapter 3, meta-generalizations are emergent properties of the framework. Their consistent presence in a variety of typologically distinct languages is thus explained as a consequence of horizons.

Finally, we have seen evidence that horizons coexist with phases, at least in the CP domain. Crucially, horizons and phases have distinct syntactic effects and do not give rise to a redundancy: Horizons determine whether extraction out of a given domain is possible or not, whereas phases dictate the successive-cyclic path that licit extractions must take. Both complete opacity and successive cyclicity play a crucial role in the data. The extraction options for embedded V2 clauses, for example, arise from the interplay of horizons and phases. Only extractions that satisfy both are admissible. The role and distribution of phases in general and in particular in the context of the horizon framework developed here is taken up again in greater detail in chapter 6.

CHAPTER 5

A COMPARISON OF HORIZONS TO PREVIOUS ACCOUNTS

5.1 Introduction

The preceding chapters proposed and developed in detail the concept of horizons as a comprehensive account for selective opacity effects. This chapter is devoted to placing the concept into the broader context of analyses of selective opacity effects. It is instructive to compare the horizons account to previous accounts of locality mismatches between operations in order to highlight similarities and differences between these accounts on a conceptual, theoretical, and empirical level. Historically, selective opacity effects have relatively rarely been the subject of systematic investigation, but particular instances of selective opacity, in particular the impossibility of superraising in English, have posed questions about how different movement types can be assigned different locality profiles, which are hence worthwhile to compare with the approach taken here.

An important distinction that I have emphasized repeatedly in the preceding chapters and that will continue to play a role in this one lies between *item-based* and *domain-based* approaches to selective opacity. In a nutshell, item-based approach impose a constraint on the movement paths of individual syntactic elements. For example, the classical ban on improper movement states that an item may not move from an \bar{A} - to an A-position. On a domain-based approach, the relevant constraint is stated at the level of the syntactic domain that blocks extraction out of it,

not at the level of the moving element per se. The horizon account is thoroughly domain-based, and so is Williams (2003, 2011, 2013). Traditional approaches like the ban on improper movement or unambiguous binding (Müller & Sternefeld 1993) are item-based. Yet other approaches, like Abels (2007, 2009), lie somewhere in the middle. This chapter will also highlight the empirical and theoretical issues surrounding the choice.

The key properties of the horizon account that I will contrast with previous approaches are the following:

- (i) Horizons are not confined to the binary A/\bar{A} -distinction
- (ii) they extend to non-movement dependencies like agreement,
- (iii) they do not involve direct interactions between movement types,
- (iv) the relevant constraint is purely domain-based and does not at all invoke the moving item itself,
- (v) they extend to selective opacity in non-identity cases (smuggling and remnant movement),
- (vi) they derive Upward Entailment and a version of the Height–Locality Connection,
- (vii) they capture variability in horizon settings both within a language and across languages.

The main purpose of this chapter is to compare the crucial analytical decisions of the horizons accounts to previous accounts and to highlight the consequences of different analytical choices. This comparison will serve to situate horizons in the broader context of approaches to selective opacity and to illustrate in which domains the horizon account represents an empirical or theoretical advancement over more traditional accounts.

5.2 Case-based approaches

One account of the impossibility of superraising in English proposed by Chomsky (2000, 2001) is based on the case of the moving element. The central claim is that only DPs with an unvalued case feature may undergo A-movement. DPs that have been assigned case subsequently may only undergo \bar{A} -movement. In more technical terms, a DP whose case feature is valued becomes **INACTIVE** for A-processes, a restriction enshrined in the **ACTIVITY CONDITION** (482):

(482) **Activity Condition**

DPs whose case feature is valued become inactive and thereby invisible to subsequent A-processes.

The Activity Condition rules out superraising as in (483a) as follows. Because of the finiteness of the lower clause, the embedded T^0 assigns nominative case to its specifier *John*. After case has been assigned to it, *Sue* becomes inactive and hence invisible to the EPP-probe on matrix T^0 , thus blocking raising into this position. If, on the other hand, the embedded clause is nonfinite, as in (483b), it does not assign nominative case to its subject, which may hence A-move into the higher clause.

- (483) a. **Sue_i seems [t_i likes oatmeal]*.
b. *Sue_i seems [t_i to like oatmeal]*.

A distinctive property of this account is that it does not refer to types of positions like A- or \bar{A} -positions, but only to independently motivated *internal* properties of the moving element. In this regard, it differs substantially from all other accounts discussed in this section, including horizons. Since its inception, the Activity Condition has been routinely employed in accounts of superraising (Nunes 2008, 2010, Carstens 2011, Diercks 2012, Halpert 2012). Interestingly, even several accounts that eschew the Activity Condition in the form of (482) still attribute the impossibility of (483a) to considerations of case (e.g., Nevins 2005, Carstens 2010, Obata 2010, Obata & Epstein 2011).

The main virtue of the Activity Condition and case-based accounts in general is the fact that they attempt to reduce the distribution of A-movement to the distribution of independently observable case features. An immediate consequence of this aspect of the account is that it is parochially confined to the A-/ \bar{A} -distinction. Consider, for instance, the fact discussed at length in section 4.2.2 of chapter 4 that finite clauses in German allow *wh*-movement out of them (484a), but not relativization (484b):

- (484) *Wh-movement vs. relativization in German* [= (399), (400)]
a. Ich weiß nicht [wen er glaubt [dass die Maria t_i getroffen hat]].
I know not whom he believes that the Maria met has
'I don't know who he believes that Maria met.'

- b. *der Mann [den_i er glaubt [dass die Maria t_i getroffen hat]]
 the man whom he believes that the Maria.NOM met has
Intended: ‘the man who he believes that Maria met’

Another example is *wh*-extraction out of a V2 clause in German, which may land in a higher V2 clause, but not in a higher V-final clause:

(485) *Extraction out of V2 clauses in German* [= (410b), (412)]

- a. Wen_i hat er gesagt [t_i hat die Maria t_i getroffen]?
 whom has he said has the Maria met
 ‘Who did he say that Maria met?’
- b. *Ich weiß nicht [wen_i er gesagt hat [t_i hat die Maria t_i getroffen]]
 I know not whom he said has has the Maria met
Intended: ‘I don’t know who he said that Maria met.’

It is clear that the Activity Condition has nothing to say about this contrast, as the moving element is case-marked in all examples and both relativization and *wh*-movement are \bar{A} -movement. Consequently, *all* instances of locality mismatches between operations other than A- and \bar{A} -movement fall outside the purview of the Activity Condition. Examples of mismatches that are beyond the Activity Condition are scrambling vs. relativization vs. *wh*-movement into a V-final clause and the contrast between *wh*-movement into a V-final clause and *wh*-movement into a V2 clause (see chapter 4), \bar{A} -movement vs. *wh*-licensing in Hindi (chapter 2), the locality mismatches in the Italian left periphery documented by Abels (2012a) and presented briefly in chapter 1, etc. The Activity Condition thus furnishes an account of only a very small subset of selective opacity effects. All instances of selective opacity outside this small set require separate accounts. But as we will see in the remainder of this chapter, all accounts of selective opacity that are general enough to extend beyond the A/ \bar{A} -distinction also capture the distribution of A- and \bar{A} -movement in English without further ado. A separate case-based account of superraising creates undesirable redundancy.

It is furthermore unclear whether a case-based account provides a satisfactory account even of the A/ \bar{A} -distinction in languages other than English. A- and \bar{A} -movement in Hindi provide an example. As originally recognized by Gurtu (1985, 1992) and Mahajan (1990) and discussed

extensively in section 2.2.1 in chapter 2, Hindi has both A- and \bar{A} -movement, which differ in their locality profile in a way similar to their English counterparts:¹

(486) A-movement in Hindi cannot leave a finite clause, but \bar{A} -movement can.

This is illustrated by (487). The example shows that extraction out of a finite clause is possible, but subject to weak crossover, hence \bar{A} -movement. Movement out of a finite clause is possible (487a), but it is invariably \bar{A} , hence subject to weak crossover (487b).

(487) *No A-extraction out of finite clauses in Hindi* [= (76)]

a. har laṛke-ko_i [us-kii_j bahin] soc-tii hai [ki raam-ne t_i
 every boy-ACC 3SG-GEN sister think-IPFV.F.SG be.PRES.3SG that Ram-ERG
 dekh-aa]
 see-PFV.M.SG
 ‘His/her_j sister thinks that Ram saw every boy_i.’

b. *har laṛke-ko_i [us-kii_i bahin] soc-tii hai [ki raam-ne t_i
 every boy-ACC 3SG-GEN sister think-IPFV.M.SG be.PRES.3SG that Ram-ERG
 dekh-aa]
 see-PFV.M.SG
Intended: ‘For every boy *x*, *x*’s sister thinks that Ram saw *x*.’

Just as in English, it is not the case that Hindi A-movement is simply clause-bounded, as it is possible out of nonfinite clauses:

(488) *A-extraction out of nonfinite clauses in Hindi* [= (79b)]

har laṛke-ko_i [us-kii_i bahin] [t_i dekha-naa] caah-tii thii
 every boy-ACC 3SG-GEN sister see-INF.M.SG want-IPFV.F.SG be.PST.F.SG
 ‘For every boy *x*, *x*’s sister wants to see *x*.’

Unlike in English, however, there is good evidence that extraction out of finite and nonfinite clauses does *not* differ with respect to case in Hindi. Neither can feed case assignment and case is invariably assigned in the lower clause. I will provide three arguments for this claim. First, crossclausal movement in Hindi can affect elements with a variety of different structural or lexical cases, but

¹ (486) is somewhat of an oversimplification. I have shown in section 2.3 of chapter 2 that nonfinite clauses in Hindi come in two varieties and that only one of them (*v*P’s on the analysis proposed there) are transparent to A-movement. This complication is irrelevant for the point made in the main text here and hence put aside for the sake of streamlining the discussion.

the case of the element in its landing site has to match the case it receives in its base position. Such case connectivity effects indicate that the case of the moving element is assigned before movement takes place, i.e., in the lower clause. This holds for movement out of finite and nonfinite clauses alike. As (489) shows, a proper name in direct object position has to be marked with the accusative case marker *-ko* and this case marking has to be preserved under movement.

(489) *Case connectivity: Direct objects* [Hindi]

a. *Baseline*

siitaa-ne Raam- $\{ko/*se/*kaa\}$ dekh-aa
 Sita-ERG Ram- $\{ACC/*INSTR/*GEN\}$ see-PFV.M.SG
 ‘Sita saw Ram.’

b. *Movement out of nonfinite clause*

raam- $\{ko/*se/*kaa\}_i$ siitaa-ne [t_i dekh-naa] caah-aa
 Ram- $\{ACC/*INSTR/*GEN\}$ Sita-ERG see-INF.M.SG want-PFV.M.SG
 ‘Sita wanted to see Ram.’

c. *Movement out of finite clause*

raam- $\{ko/*se/*kaa\}_i$ siitaa-ne soc-aa hai [ki prataap-ne t_i
 Ram- $\{ACC/*INSTR/*GEN\}$ Sita-ERG think-PFV.M.SG be that Pratap-ERG
 dekh-aa]
 see-PFV.M.SG
 ‘Sita thought that Pratap saw Ram.’

The same holds for lexical cases like the instrumental in (490):

(490) *Case connectivity: Instrumentals*² [Hindi]

a. *Baseline*

prataap siitaa- $\{se/*ko/*kaa\}$ mil-aa hai
 Pratap Sita- $\{INSTR/*ACC/*GEN\}$ meet-PFV.M.SG be.PRES.3SG
 ‘Pratap met Sita.’

b. *Movement out of nonfinite clause*

siitaa- $\{se/*ko/*kaa\}_i$ prataap-ne [t_i milaa-naa] caah-aa
 Sita- $\{INSTR/*ACC/*GEN\}$ Pratap-ERG meet-INF.M.SG want-PFV.M.SG
 ‘Pratap wanted to meet Sita.’

² The accusative marker *-ko* is marginally possible in (490) under the reading ‘Sita found Pratap’. The possibility of *-ko* and the distribution of its reading is not affected by movement.

- c. *Movement out of finite clause*
- siitaa- $\{se/*ko/*kaa\}_i$ raam-ne soc-aa [ki prataap t_i
 Sita- $\{INSTR/*ACC/*GEN\}$ Ram-ERG think-PFV.M.SG that Pratap-ERG
 mil-aa hai]
 meet-PFV.M.SG be.PRES.3SG
 ‘Ram thought that Pratap met Sita.’

The second argument comes from genitive case assignment. As Bhatt (2005) emphasizes, genitive case in Hindi is only ever assigned within the nominal domain, never to internal arguments of verbs. Moreover, Hindi is a language that allows left branch extraction and hence the movement of possessors out of their host DP. Crucially, if a possessor is moved, it has to retain its genitive case. As before, this holds for movement out of finite and nonfinite clauses alike.

(491) *Case preservation: Possessors* [Hindi]

- a. *Movement out of nonfinite clause*
- siitaa- $\{kaa/*ko/*se\}_i$ raam-ne [[_{DP} t_i lekh] parh-naa] caah-aa
 Sita- $\{GEN/*ACC/*INSTR\}$ Ram-ERG article read-INF.M.SG want-PFV.M.SG
 ‘Ram wanted to read Sita’s article.’
- b. *Movement out of finite clause*
- siitaa- $\{kaa/*ko/*se\}_i$ raam-ne soc-aa [ki prataap [_{DP} t_i lekh]
 Sita- $\{GEN/*ACC/*INSTR\}$ Ram-ERG think-PFV.M.SG that Pratap article
 parh-taa hai]
 read-IPFV.M.SG be.PRES.3SG
 ‘Ram thought that Pratap reads Sita’s article.’

Because genitive case is restricted to the nominal domain, it must be assigned to the possessor inside the contained DP in the lower clause, regardless of whether the possessor subsequently moves out of a finite or nonfinite clause.

A third argument for the same conclusion comes from genitive agreement. In Hindi, the morphological form of the genitive marker in fact agrees in ϕ -features with the head noun of the contained DP. This is illustrated in (492). In (492a), the head noun is *lekh* ‘article’, which is masculine singular and the genitive marker on the possessor takes the form *-kaa*. In (492b), by contrast, the head noun is *kitaab* ‘book’ and the form of the genitive marker is *-kii*.

(492) *Genitive agreement* [Hindi]

- a. raam [DP siitaa-**{kaa/*kii}** lekh] paṛh-taa thaa
 Ram Sita-**{GEN.M.SG/*GEN.F.SG}** article.M read-IPFV.M.SG be.PST.3SG
 ‘Ram read Sita’s article.’
- b. raam [DP siitaa-**{*kaa/kii}** kitaab] paṛh-taa thaa
 Ram Sita-**{*GEN.M.SG/GEN.F.SG}** book.F read-IPFV.M.SG be.PST.3SG
 ‘Ram read Sita’s book.’

This agreement has to persist under crossclausal movement:

(493) *Genitive agreement: Movement out of nonfinite clauses* [Hindi]

- a. siitaa-**{kaa/*kii}**_i raam-ne [[DP t_i lekh] paṛh-naa]
 Sita-**{GEN.M.SG/*GEN.F.SG}** Ram-ERG article.M read-INF.M.SG
 caah-aa
 want-PFV.M.SG
 ‘Ram wanted to read Sita’s article.’
- b. siitaa-**{*kaa/kii}**_i raam-ne [[DP t_i kitaab] paṛh-naa]
 Sita-**{*GEN.M.SG/GEN.F.SG}** Ram-ERG book.F read-INF.M.SG
 caah-aa
 want-PFV.M.SG
 ‘Ram wanted to read Sita’s book.’

(494) *Genitive agreement: Movement out of finite clauses* [Hindi]

- a. siitaa-**{kaa/*kii}**_i raam soc-taa hai [ki prataap [DP
 Sita-**{GEN.M.SG/*GEN.F.SG}** Ram think-IPFV.M.SG be.PRES.3SG that Pratap
 t_i lekh] paṛh-taa hai]
 article.M read-IPFV.M.SG be.PRES.3SG
 ‘Ram thinks that Pratap reads Sita’s article.’
- b. siitaa-**{*kaa/kii}**_i raam soc-taa hai [ki prataap [DP
 Sita-**{*GEN.M.SG/GEN.F.SG}** Ram think-IPFV.M.SG be.PRES.3SG that Pratap
 t_i kitaab] paṛh-taa hai]
 book.F read-IPFV.M.SG be.PRES.3SG
 ‘Ram thinks that Pratap reads Sita’s book.’

For agreement between the genitive marker and the head of the container DP to be established, genitive case must be assigned inside the container DP, hence before crossclausal movement takes

place. Because agreement is obligatory, we can conclude that genitive case assignment in fact *must* take place in the container DP, before movement.

Nonetheless, the locality condition on A- and \bar{A} -movement in (486) also holds for possessor extraction. Thus, it is possible to A-move a possessor out of a nonfinite clause, as demonstrated for weak crossover in (495a) and for reciprocal binding in (495b), two diagnostics of A-movement (see section 2.2.1 in chapter 2).

- (495) *A-movement of possessor out of nonfinite clause* [Hindi]
- a. har larke-kaa_i us-kii_i bahin-ne [[DP t_i lekh] paṛh-naa] caah-aa
 every boy-GEN 3SG-GEN sister article read-INF.M.SG want-PFV.M.SG
 ‘For every *x*, *x*’s sister wanted to read *x*’s article.’
- b. [raam aur prataap]-ke_i ek-duusre-kii_i bahinõ-ne [[DP t_i lekh] paṛh-ne]
 [Ram and Pratap]-GEN each other’s sisters-ERG articles read-INF.M.PL
 caah-e
 want-PFV.M.PL
 ‘Ram and Pratap, each’s sisters wanted to read the other’s articles.’

Possessor extraction out of a finite clause is possible, but has to be \bar{A} -movement:

- (496) *No A-movement of possessor out of finite clause* [Hindi]
- a. har larke-kaa_i us-kii_j/^{*}_i bahin-ne kah-aa [ki raam-ne [DP t_i lekh]
 every boy-GEN 3SG-GEN sister-ERG say-PFV.M.SG that Ram-ERG article
 paṛh-aa]
 read-PFV.M.SG
 ‘His/her sister said that Ram read every boy’s article.’ (*bound reading impossible*)
- b. *[raam aur prataap]-ke_i ek-duusre-kii_i bahinõ-ne soc-aa [ki
 [Ram and Pratap]-GEN each other’s sisters-ERG think-PFV.M.SG that
 monaa-ne [DP t_i lekh] paṛh-e]
 Mona-ERG articles read-PFV.MASC.PL
Intended: ‘Ram and Pratap, each’s sister thought that Mona read the other’s articles.’

We have seen evidence that the case of the moving possessor is invariably assigned in its base position, regardless of whether the embedded clause is nonfinite, as in (495), or finite, as in (496). Yet in the former case A-movement of the possessor is possible, while in the latter it is not.

Because the case of the possessor is assigned before movement in all of these cases, the Activity Condition would entail that the moving elements are necessarily inactive for A-movement and that A-movement is thus impossible out of both finite and nonfinite clauses in Hindi. As I have shown, this is incorrect. The Activity Condition is thus overly restrictive.

One might explore the view that the Activity Condition for some reason does not apply to Hindi, perhaps along the lines of recent work on Bantu where it is argued that in some languages, case never deactivates a DP (see Carstens 2010, 2011, Obata 2010, Obata & Epstein 2011, Diercks 2012, Halpert 2012, Carstens & Diercks 2013). Yet such an account would be too unrestrictive for Hindi, as it leaves a DP active for A-processes indefinitely. The ban on A-movement out of finite clauses remains unaccounted for.

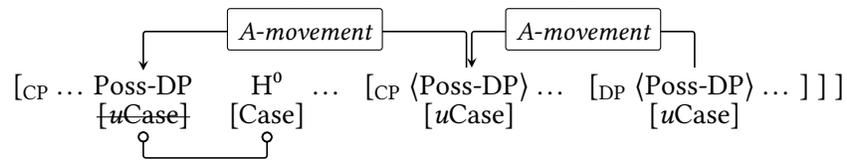
This conclusion that emerges is perfectly general: Because extraction out of finite and nonfinite clauses does not differ with respect to case in Hindi, a case-based approach has no means of distinguishing between them. As a result, it is necessary to invoke some other, non-case-based restriction on A-movement to rule out A-movement out of finite clauses in Hindi, be it horizons, the ban on improper movement, or some other constraint. Importantly, non-case-based constraints evidently also succeed to prohibit superraising in English, making superfluous an additional case-based account of English. The Activity Condition can hence be retired with no loss in empirical coverage.

It is possible to explore a more abstract version of the Activity Condition, according to which activity of a DP is determined on the basis of abstract case. For instance, it is well-known that dative subjects in Icelandic undergo A-movement to SpecTP despite their lexical-dative case marking (see, e.g., Sigurðsson 2004). While Nevins (2005) takes this as evidence against the Activity Condition, Bošković (2002) contends that dative subjects are assigned an abstract nominative case in SpecTP, hence obeying the Activity Condition. One immediate concern about such an account is that it renders the Activity Condition very weak, because the feature that distinguishes DPs that can undergo A-movement and those that cannot runs the risk of becoming an analytical diacritic whose sole purpose is to constrain the effects of the Activity Condition.

The Hindi data just considered are instructive as even an appeal to abstract case is unlikely to provide an empirically adequate restriction. To see this, recall that possessors may be A-moved out of their host DPs, as in (495). On the abstract-case account, this fact would entail that a possessor's

abstract case does not have to be assigned within the host DP, but that it can be assigned in a higher clause, after possessor extraction has taken place. But possessors are, of course, optional in Hindi. This in turn means that the purported abstract case assigner in the higher clause in (495) must be optional as well. This follows because Bošković (2002) assumes the Inverse Case Filter, the requirement that a case feature on a functional head must be assigned. If the possessor that receives abstract case in the higher clause in (495) is optional, this abstract case must consequently be optionally present on the head assigning it. But this entails that abstract case assigners must be optional in Hindi, the only limit being that at the end of the derivation, all DPs have received case. This opens up the structural possibility in (497). Here the possessor does not receive abstract case in its containing DP and the embedded finite clause does not contain an abstract case assigner, but the topmost finite clause does (indicated as ‘H⁰’ in (497)). In this case, the possessor’s abstract case feature remains unassigned until it enters the matrix clause, a movement step that should qualify as A-movement. As we have seen on the basis of (496), this is impossible. Only \bar{A} -movement out of a finite clause is possible.

(497) *Illicit A-movement out of finite clause conforming to Activity Condition*



In sum, the very mechanism that allows the delay of valuing a DP’s abstract case, hence allowing A-movement in (495), invariably also allows such a delay in (496), generating A-movement out of a finite clause. I thus conclude that a version of the Activity Condition that appeals to abstract case is no more successful for Hindi than one that is based on visible case.

I have thus argued that the Activity Condition does not extend to the distribution of A- and \bar{A} -movement in languages like Hindi. Because case assignment is subject to different conditions in Hindi than it is in English, but the distribution of A-movement is similar, the Activity Condition has nothing to say about Hindi. Finally, the Activity Condition arguably fails to adequately restrict the distribution of A-movement even in English. Consider the examples in (498)–(500). In both cases, the A-moved element *John* does not receive case in its base position or anywhere prior to

its final A-landing site. The Activity Condition is hence satisfied in both examples. Yet both are ungrammatical.

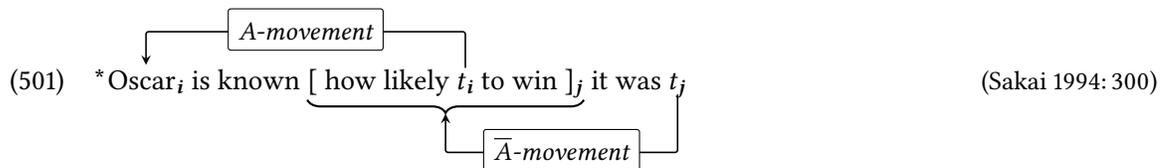
(498) *John_i seems [that it is certain [t_i to like ice cream]]. (Chomsky 1981: 58)

(499) *John_i seems [that it was told t_i [that Mary was a genius]]. (Lasnik & Saito 1992: 192)

(500) *Who_i is believed [that Mary knows/remembered [t_i to invite Bill]]?

The ungrammaticality of (498)–(500) follows without further ado from horizons because both involve A-movement out of a CP clause. (498)–(500) thus demonstrate that A-movement out of finite clause is impossible regardless of the case properties of the moving element. This provides evidence against case-based accounts of superraising even for English.³

A final point worth considering is that the Activity Condition does not cover non-identity cases of superraising such as the smuggling derivation in (501). Here an embedded nonfinite clause is \bar{A} -moved to the edge of a finite clause, followed by A-subextraction of the subject. Because the Activity Condition is satisfied in (501), the fact that the sentence is nonetheless ungrammatical remains unaccounted for.



A horizon-based account captures (501) straightforwardly as discussed in section 3.4.2 of chapter 3 for English and section 4.4.3 of chapter 4 for analogous facts in German.

³ It is possible to appeal to a Relativized Minimality effect to rule out (498)–(500), according to which A-movement over a subject (*it* in (498) and (499) and *Mary* in (500)) is impossible. Note that such intervention would have to be defective (Chomsky 2001) as these subjects, being case-marked, should be invisible to A-processes from the point of view of the Activity Condition. What emerges as a puzzle on this account is what Lasnik & Boeckx (2006) have called the *experiencer paradox*, because experiencers do not intervene for A-movement:

- (i) Mary_i seems to John [t_i to be a genius].

As is well-known, *John* in (i) behaves as if it c-commands into the embedded clause. Moreover, even bare experiencer DPs fail to induce an intervention effect:

- (ii) John_i strikes Mary [as t_i being the best]. (Lasnik & Boeckx 2006: 123)

Without an explicit theory of defective intervention, the cases in (498)–(500) pose a problem for the Activity Condition. On a horizon account, the contrast between (498)–(500) on the one hand and (i) and (ii) on the other is unremarkable of course. A-movement out of a CP is impossible, barring (498)–(500), but allowing (i) and (ii).

In sum, an account that ties the availability of A-movement out of a domain to the case properties of the moving element has very limited utility. First, it does not extend beyond the A/ \bar{A} -distinction. Second, it does not extend to the A/ \bar{A} -distinction in languages like Hindi. Third, it overpredicts the availability of A-movement in English. Non-case-based accounts of superraising are thus necessary to accommodate the full range of facts. But since such accounts automatically capture the range of facts that fall under the Activity Condition, a case-based principle like the Activity Condition creates substantial redundancy with no gain in empirical coverage.

5.3 Ban on improper movement

The classical account of superraising in English takes as its point of departure the distinction between A- and \bar{A} -positions in English and imposes a well-formedness constraint on how these positions may be distributed over the movement chain of an element (Chomsky 1973, 1981, May 1979, Lasnik & Saito 1992, Fukui 1993). This constraint is standardly referred to as the **BAN ON IMPROPER MOVEMENT**.

(502) **Ban on Improper Movement**

An element may not be moved from an \bar{A} - to an A-position.

The constraint in (502) has to be combined with a second locality constraint (subjacency or phases) that requires extraction out of a finite clause to proceed through the edge of that clause, an \bar{A} -position. Superraising is then ruled out as a conspiracy of these two constraints. Subjacency rules out one-fell-swoop extraction, and (502) excludes the successive-cyclic derivation in (503).

(503) $\begin{array}{c} \text{A} \qquad \qquad \bar{A} \\ \swarrow \qquad \quad \searrow \\ \text{Sue}_i \text{ seems } [_{\text{CP}} \text{ } t_i \quad t_i \text{ likes oatmeal }] . \end{array}$

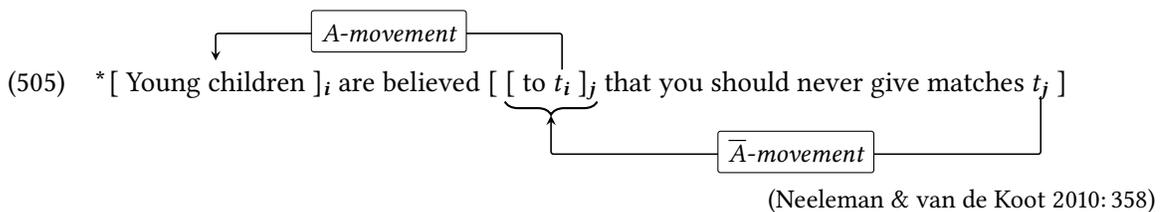
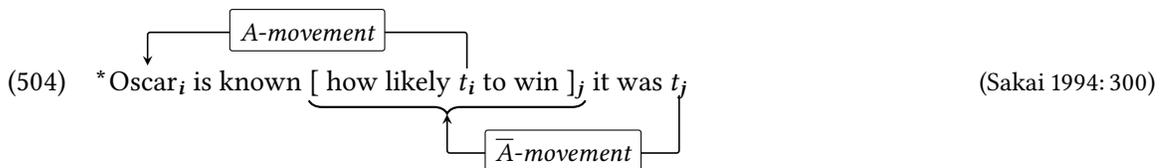
There are a number of ways to technically impose the constraint in (502). May (1979) proposes that the traces left by \bar{A} -movement are variables, subject to Principle C of the binding theory and hence required to be globally A-free, an analysis adopted by Chomsky (1981). Alternatively, (502) can be stated as a constraint on movement chains (e.g., Lasnik & Saito 1992).

The Ban on Improper Movement differs substantially from the case-based accounts reviewed in the previous section in that the constraint does not refer to properties internal to the moving

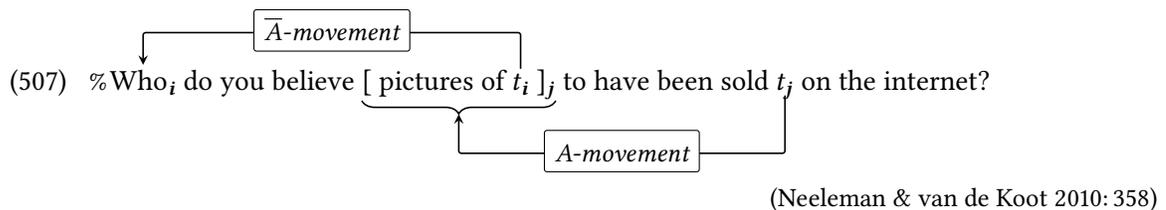
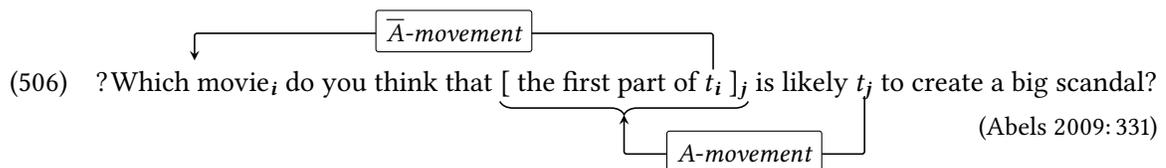
element, but rather to the syntactic positions this element occupies. This enables it to circumvent the various problems that a case-based account faces and which were discussed in the previous section. Thus, it extends to Hindi, where the locality of A-movement cannot be reduced to properties of case assignment.

As Müller & Sternefeld (1993) discuss in detail, the Ban on Improper Movement is of confined utility because it is narrowly restricted to the A/\bar{A} -distinction. Yet as we have seen throughout this dissertation, locality mismatches extend far beyond to this binary distinction. A central limitation of (502) is thus that it does not extend to other locality mismatches.

A second limitation, noted by Sakai (1994), Grewendorf (2003, 2015), Williams (2003), Abels (2007, 2009), Neeleman & van de Koot (2010) and others and already discussed in section 3.4.2 of chapter 3 and section 4.4.3 of chapter 4, revolves around the fact that the Ban on Improper Movement only applies to movement paths of individual items. This is demonstrably insufficient, as there are non-identity cases of improper movement in which movement of one constituent affects the movement options of another constituent. We have already one such example in (501), repeated here as (504), where a clausal constituent is moved to the edge of lower clause, followed by A-subextraction out of that constituent. Because it is the clausal constituent that resides in the \bar{A} -position, not the A-moved element itself, the structure in (501) does not constitute improper movement and should hence be well-formed, contrary to fact. Another example that makes the same point is given in (505), where a PP is \bar{A} -moved, followed by A-subextraction of the DP inside it in a process akin to pseudo passivization. Again, the result is ill-formed, but the Ban on Improper Movement fails to exclude it.



This limitation of improper movement stems from its *item-based* nature in the sense that the principle only applies to individual items and their movement derivations. Because in smuggling derivations like (501) and (505) each movement step is taken by a distinct element, a constraint that only focuses on individual elements does not extend to smuggling configurations. A second, unrelated constraint is necessary to exclude such derivations. Freezing provides a plausible candidate (Wexler & Culicover 1980). Yet such an account faces two serious drawbacks. First, superraising is then attributed to a conspiracy of *three* different principles: (i) one forcing successive-cyclic movement out of finite clauses, (ii) another excluding A-movement from an \bar{A} -position, and (iii) one that excludes smuggling derivations. On such an account, the simple and pervasive empirical generalization that A-movement out of finite clause is impossible would then be merely the result of a conspiracy of three unrelated constraints on this account. Second, Abels (2007, 2009) and Neeleman & van de Koot (2010) emphasize that \bar{A} -subextraction out of an A-moved element, as in (506) and (507), is considerably better than the inverse ordering in (504) and (505).



Whether or not subextraction out of a moved constituent is possible or not thus depends on the movement types of the two. A-movement out of an \bar{A} -moved constituent is impossible (see (504) and (505)), whereas the inverse is at least marginally possible (see (506) and (507)). This pattern bears an obvious parallelism to standard cases of superraising, but the Ban on Improper Movement has no way of covering them.

The horizon account differs significantly from the improper-movement account because it is fundamentally *domain-based*. It does not state that a specific item in a specific position cannot undergo a specific movement type, but rather that an entire domain is inaccessible to a probe.

The constraint thus only targets the relation between a probe and a syntactic domain. As a result, individual elements within this domain will be inaccessible to the probe, but these elements do not themselves play any role in the account. As discussed at length in section 3.4.2 of chapter 3 and section 4.4.3 of chapter 4, this property enables it to extend without further ado to the non-identity cases just discussed.

A third important aspect of the improper-movement account is that it is strictly movement-based. It hence does not extend to syntactic operations that are established long-distance, lacking movement. As I have argued in chapter 2 based on ϕ -agreement and *wh*-licensing in Hindi, this is empirically insufficient.

To give just one illustration of this point, consider May's (1979) classic account of improper movement, adopted in Chomsky (1981), according to which traces left by \bar{A} -movement are variables, subject to Principle C, hence globally A-free. This account quite elegantly explains why A-movement from an \bar{A} -position is impossible, a restriction that holds in English and Hindi alike. The limitation of this account is that it has nothing to say about ϕ -agreement. We saw in chapter 2 that elements in SpecCP are not only unable to undergo A-movement, but are also unable to control ϕ -agreement or be *wh*-licensed by a C^0 head in a higher clause. Treating \bar{A} -traces as variables has nothing to say about these restriction. A uniform theory of selective opacity is thus out of reach for an improper-movement account, precisely because it is solely based on movement chains.

5.4 Unambiguous binding

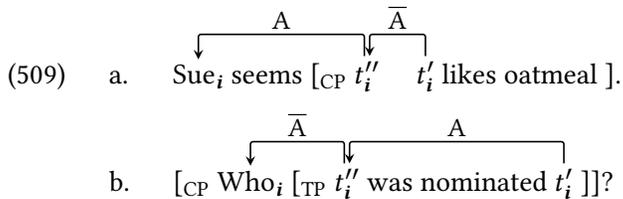
The approach to selective opacity effects developed by Müller & Sternefeld (1993, 1996) and Müller (1995) to my knowledge constitutes the first systematic attempt to generalize the locality distinction between A- and \bar{A} -movement in English to locality asymmetries between other movement types. The core contribution of these accounts is the view that A/ \bar{A} -locality differences in English are not the result of an essentially construction-specific locality constraint, but a particular instance of a much more pervasive and general pattern that holds across languages and constructions.

Müller & Sternefeld (1993, 1996) and Müller (1995) propose the **PRINCIPLE OF UNAMBIGUOUS BINDING** in (508), which imposes a very general constraint on the shape of movement paths:

(508) **Principle of Unambiguous Binding**

A variable that is α -bound must be β -free in the domain of the head of its chain (where α and β refer to different types of positions).

The term ‘variable’ in (508) refers to traces left by \bar{A} -movement. What (508) states is that a variable that is bound from one type of position cannot also be bound from another type of position, where binding refers to being c-commanded by a coindexed element. In combination with the requirement that long extraction proceed successive-cyclically, (508) gives rise to selective opacity. Using superraising as an example, in the derivation in (509a), the trace t'_i of *Sue* is bound by two elements: (i) by the intermediate trace t''_i in SpecCP, and (ii) by the ultimate landing site of *Sue* in SpecTP. This leads to a violation of (508), because t'_i is both A- and \bar{A} -bound. The inverse ordering in (509b) is allowed because t'_i is not a variable, as it is the result of A-movement. The only variable (t''_i) is unambiguously bound by *who*.



Finally successive-cyclic \bar{A} -movement is compatible with (508) because all relevant members of the chain will occupy the same type of position. That is, all traces left by movement are \bar{A} -bound. The gist of the account is that the escape hatches employed during the course of a derivation have to match the type of the final landing site. Selective opacity is the result of the fact that certain domains only provide certain types of escape hatches.

Crucially, (508) makes no reference to A- and \bar{A} -positions per se, but rather to syntactic positions more generally, making the principle extremely versatile, flexible, and broad in its scope. Müller & Sternefeld (1993) discuss a wealth of cases covered by (508), but I will use just one for illustration. As discussed in chapter 4, scrambling in German is unable to leave a finite clause, an observation that goes back to Bierwisch (1963) and Ross (1967):

- (510) *Er glaubt den Fritz_i [dass die Maria _{t_i} getroffen hat]. [German]
 he believes the Fritz.ACC that the Maria.NOM met has
Intended: ‘He believes that Maria met Fritz.’

Two relevant derivations of (510) have to be ruled out. First, one-fell-swoop movement of *den Fritz* is ruled out by subjacency. The second derivation involves movement to the edge of the lower clause, followed by scrambling into the matrix clause. Due to the fact that these two movement steps target different positions, this derivation is successfully ruled out by (508). As a result, scrambling from an \bar{A} -position is ruled out for exactly the same reason that A-movement from an \bar{A} -position is precluded: it involves an illegitimate mixing of positions in a movement chain.

The account developed in this dissertation clearly shares with unambiguous binding the contention that locality mismatches can be manifold and that the principle underlying them cannot be strictly binary. In several other respects, the two lines of account diverge substantially. First, the Principle of Unambiguous Binding is *symmetric*. If a particular combination of positions in a movement path lead to a violation of (508), the order in which these positions are targeted is irrelevant. Thus, as just discussed, (508) prevents an element from undergoing \bar{A} -movement followed by scrambling. But the inverse, \bar{A} -movement of a scrambled constituent, is also ruled out. This stands in stark contrast to horizons, in which only the former feeding relationship would be ruled out. It is clear that an asymmetric constraint is required for English, as \bar{A} -movement followed by A-movement is impossible whereas the inverse is unquestionably allowed. This asymmetry is captured in (508) because the principle only applies to variables, therefore exempting traces left by A-movement. The question that arises is whether other instances of selective opacity or also asymmetrical or not.

One argument that scrambling may feed *wh*-movement in German comes from superiority effects. It is well-known that German does not, at first glance, exhibit superiority effects (Haider 1983: 99, 1986: 113–114, Grewendorf 1988: 320, Fanselow 1991: 330).⁴ In (511), either of the two *wh*-elements may be fronted.

- (511) a. Wer hat wen gesehen? [German]
 who.NOM has who.ACC seen
- b. Wen hat wer gesehen?
 who.ACC has who.NOM seen
 ‘Who saw whom?’

⁴ Superiority, first proposed in Chomsky (1973), arguably follows from the more general *Relativized Minimality* (Rizzi 1990), the *Shortest Paths Condition* (Chomsky 1993) or the *Minimal Link Condition* (Chomsky 1995b), a view that I will adopt here without discussion.

At first glance, data like (511) may be taken to show that German, unlike English, altogether lacks superiority as a grammatical principle. This conclusion is premature, however, as Wiltschko (1997, 1998), Grohmann (1997), and Fanselow (2004), among others, argue (also see Takahashi 1993 for Japanese). These authors propose that the base-generation order may be reversed by scrambling of the object over the subject, in which case the object will be the closer goal and hence attracted to SpecCP. In the absence of scrambling, it is the subject (and only the subject) that is attracted. The observation that *wh*-movement in German seemingly violates superiority is hence attributed to the fact that German is a scrambling language whereas English is not and, moreover, that scrambling may feed *wh*-movement.⁵

Evidence for this analysis comes from complex clauses. As we have seen in (510), German does not allow scrambling out of finite clauses. As it turns out, if two *wh*-elements are separated by a finite clause boundary, superiority effects reemerge (Fanselow 1991: 331n7, Buring & Hartmann 1994: 62–64, Müller 1995: 324, Heck & Müller 2000: 225, and Fanselow 2004: 77).

- (512) a. Wen_i glaubt der Fritz [dass Maria t_i gesehen hat]? [German]
 who.ACC believes the Fritz that Maria seen has
 ‘Who does Fritz believe that Maria has seen?’
- b. Wer glaubt [dass Maria wen gesehen hat]?
 who.NOM believes that Maria who.ACC seen has
- c. *Wen_i glaubt wer [dass Maria t_i gesehen hat]?
 who.ACC believes who.NOM that Maria seen has
 ‘Who believes that Maria saw whom?’

This restriction is explained immediately if German, like English, adheres to superiority, but if this fact is systematically obscured by the availability of scrambling in simple clauses like (511). Crucial to this analysis is the fact that scrambling can feed *wh*-movement.

Converging evidence comes from weak crossover. As Grewendorf (1988: 320) observes, local *wh*-movement is not subject to weak crossover in German.

⁵ Fanselow (2001) likewise treats the superficial lack of superiority in German to be the result of scrambling. Since he treats scrambling as variability in base generation order, it is not the case under his analysis that *wh*-movement is fed by another movement step.

- (513) Wen_i mag seine_i Mutter nicht t_i? [German]
 whom likes his mother not
 ‘For which x, x’s mother doesn’t like x.’ (Grewendorf 1988: 320)

Wiltschko (1997) argues that this fact is accounted for if *wen* ‘who’ can undergo scrambling over the subject in (513) before being *wh*-moved. On this view, scrambling is able to obviate a crossover violation, but *wh*-movement is not. This account makes the straightforward prediction that in configurations in which scrambling is not an option no crossover amnesty should arise. We have already seen on the basis of (510) that finite clauses are opaque to scrambling. As expected, these cases give rise to weak crossover violations (see also Abels 2009: 334–335 and Müller 2014b: 52–53):

- (514) Wen_i hat seine_{j/*i} Mutter gesagt [dass die Maria nicht t_i mag]? [German]
 whom has his mother said that the Maria not likes
 ‘Who_i did his_j mother say that Maria does not like?’ (bound reading impossible)

Like the superiority effects discussed above, the crossover facts provide good evidence for the view that scrambling may feed *wh*-movement in German.⁶

⁶ Also see Wiltschko (1997, 1998) for an investigation of the interaction between *wh*-movement, scrambling, and D-linking in German.

Müller & Sternefeld (1993, 1996), following much of the literature (see von Stechow & Sternefeld 1988: 466, Fanselow 1990: 117–118, and also Müller 1998: 39–40, Pesetsky 2000: 78–80, Grewendorf 2001: 110n37), argue that *wh*-phrases cannot scramble in German, citing examples like (i). This restriction would follow from the Principle of Unambiguous Binding if covert *wh*-movement takes place.

- (i) a. Warum hat er den Studentinnen was gegeben?
 why has he the students.DAT what.ACC given
 b. *Warum hat er was_i den Studentinnen t_i gegeben?
 why has he what.ACC the students.DAT given
 ‘Why did he give what to the students?’ (Müller & Sternefeld 1996: 483)

The unacceptability of (i.b) is, however, plausibly due to an independent information-structural constraint on scrambling (see, e.g., Müller 1999b for an overview). In German, there is a strong preference for definite DPs to precede indefinites in the middle field, a constraint that is violated by scrambling in (i). Once this confound is controlled for, scrambling of *wh*-phrases improves considerably, as the examples in (ii) (also see Beck 1996, Wiltschko 1997: 127–129, Sauerland 1999: 171–172 and Fanselow 2004: 100–101 for discussion):

- (ii) a. Wann würde wem_i nur ein Held t_i helfen?
 when would who.DAT only a hero help
 ‘When would only a hero help whom?’ (Fanselow 2001: 414)
 b. Wer hat was_i denn schon oft t_i gesehen?
 who has what PRCL already often seen
 ‘Who has often seen what?’ (Wiltschko 1997: 128)

In sum, the Principle of Unambiguous Binding (508) is too strong because it completely rules out any mixing of scrambling and *wh*-movement within the same movement path. The empirical situation appears to be more nuanced: scrambling can feed *wh*-movement but not vice versa. Horizons provide a way of implementing this asymmetry.

Despite the substantial increase in empirical coverage that the Principle of Unambiguous Binding represents over the classical Ban on Improper Movement, it coincides with the latter in being an item-based constraint, i.e., the principle applies to the movement path of a single item. Consequently, cases in which movement of one element affect the movement options of another element, such as (501) and (505), lie outside the purview of (508). Interestingly, as emphasized by Abels (2007) and Abels (2009) and discussed in section 4.4.3 of chapter 4, such non-identity cases of selective opacity are attested outside the A/\bar{A} -distinction. (515) provides an example from German:

- (515) *Scrambling out of wh-moved constituent* [German]
 *Maria hat [über die Liebe]_j gefragt [_{CP} [was für ein Buch *t_j*]_i niemand *t_i*
 Maria has about the love asked what for a book nobody
 gelesen hat.
 read has
Intended: ‘Maria asked what kind of book about love no one read.’

Horizons, being domain-based, extend to such non-identity cases, while strictly item-based principles like (508) do not.

Furthermore, because (508) is a constraint on traces, the unambiguous-binding account does not obviously extend to syntactic relations that do not involve movement, but participate in selective opacity. We have seen, though, that they should fall under this generalization.

A final point of comparison between unambiguous binding and horizons is that the connection between a movement type’s locality and the height of the landing site it targets remains unaccounted for. There is, for instance, no expectation that movement types targeting structurally low positions

If in-situ *wh*-phrases undergo covert *wh*-movement in German, data like (ii) provide an argument that scrambling may feed *wh*-movement, in violation of unambiguous binding. It should be noted, however, that the basic assumption of covert *wh*-movement in German is dubious in light of Beck’s (2006) and Kotek’s (2014) accounts of focus intervention effects (see section 2.4.1 in chapter 2). On these accounts, focus intervention results if a focus element intervenes between a *wh*-phrase and its scope at LF. Focus intervention hence serves as a diagnostic for the absence of covert *wh*-movement. As is well-known (Beck 1996, et seq.) in-situ *wh*-phrases in German are subject to focus intervention. Thus, if there is no covert *wh*-movement in German, then data like (ii) do not provide evidence for scrambling feeding *wh*-movement, but the data discussed in the main text still do.

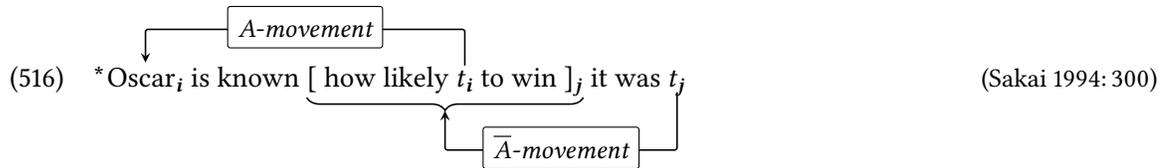
are subject to stricter locality constraints than ones that target high positions, if all that is required is uniformity of movement paths. This aspect of their proposal is similar to the horizons account, as discussed in section 3.5 of chapter 3. Where they diverge is with respect to structurally high probes. On the horizons account, structurally high probes exhibit a high degree of connectedness between height and locality. On the unambiguous-binding account, by contrast, there is no connection between the two for either structurally high or low probes.

5.5 Operational ordering

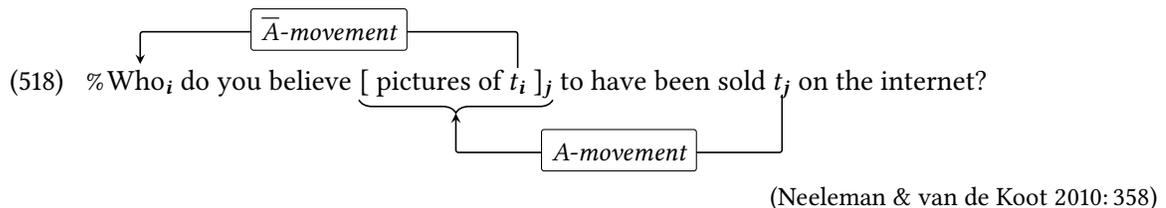
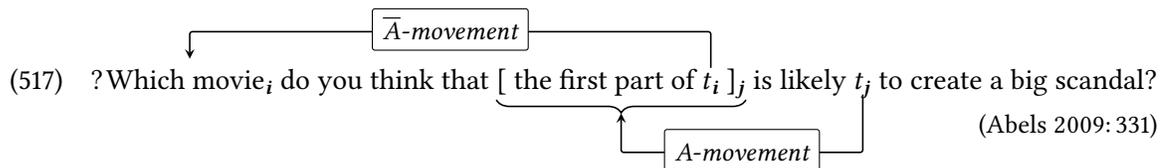
Müller & Sternefeld (1993) considerably advanced the study of improper movement by taking it as being just one instantiation of a more fundamental constraint on syntax whose effects are observable in a range of constructions and languages. Sakai (1994), Grewendorf (2003, 2015), Williams (2003), Abels (2007, 2009), and Neeleman & van de Koot (2010) extend the scope of improper movement even further and argue that it extends to non-identity cases, i.e., instances of movement type interactions in which *distinct* elements undergo the two movements. Examples of such configurations were already discussed in the preceding sections. On the standard view of improper movement, as well as Müller & Sternefeld's (1993), interactions between different movement types arise only if one and the same element undergoes two movements. Thus, on the standard view, however derived technically, \bar{A} -movement of an element prevents subsequent A-movement of that same element. The authors just cited observe that bleeding relations between movement types can also arise if the elements undergoing the two movements are not one and the same. Grewendorf (2003) and Williams (2003) investigate remnant movement cases, and Abels (2007, 2009) and Neeleman & van de Koot (2010) focus on remnant movement as well as subextraction out of a moved constituent. As already discussed above in the context of the ban on improper movement, such interactions are beyond the reach of traditional accounts of superraising. Abels' (2007, 2009) aim is thus to develop a theory that is more comprehensive in its empirical scope.

I will illustrate this point once more using A- and \bar{A} -movement in English, but the point holds more generally (see Abels 2007, 2009 as well as section 3.4.2 in chapter 3 and section 4.4.3 in chapter 4 for much additional discussion in the context of horizons). As we have already seen in

section 5.2 above, it is not possible to \bar{A} -move a constituent, followed by A-movement out of it. This was shown on the basis of (501), repeated here as (516).



Recall also from section 5.4 that Abels argues that it is not simply the case that moved constituents disallow subextraction out of them because the inverse feeding relationship – \bar{A} -movement out of an A-moved constituent – yields a much better result, as (517) and (518), repeated from (506) and (507) above, show:



The contrast between (516) on the one hand and (517) and (518) on the other is strikingly similar to improper movement. In all of these case, an application of \bar{A} -movement bleeds a subsequent application of A-movement. Yet neither the traditional account of improper movement nor unambiguous binding extend to this contrast because they only constrain movement of a single element. Because the element that undergoes A-movement is not the one that undergoes \bar{A} -movement in (517) and (518), the contrast remains unaccounted for.

A second instance of a configuration argued for in the literature just cited to instantiate improper movement without identity comes from remnant movement. While it is uncontroversial that a remnant created by A-movement can itself undergo \bar{A} -movement, as in (519), the inverse is impossible, as (520) from Abels (2009: 331) shows. In (520a), *which king* is \bar{A} -moved to the edge of a finite clause. In (520b), the remnant *a picture of t* is then A-moved into a higher clause.

To ensure that the sentence is not merely ruled out for reasons of case, the embedded clause is passivized and nonfinite in (520b) so that the remnant receives case only in the higher clause. The resulting configuration is ungrammatical. A-movement of a remnant created by \bar{A} -movement is thus impossible.

(519) [How likely t_i to win]_j is Sue_i t_j ?

(520) a. It is known [which king_i they sold [a picture of t_i]]

b. * [A picture of t_i]_j is known [which king_i to have been sold t_j]

Again, an instance of \bar{A} -movement renders impossible a later application of A-movement, whereas the inverse ordering is possible. The similarity to improper movement is uncanny, but the pattern is outside the purview of standard item-based accounts of improper movement, since the moving elements are distinct. Grewendorf (2003), Williams (2003), Abels (2007, 2009), and Neeleman & van de Koot (2010) conclude from these considerations that the classical view on improper movement misses a generalization and that a more generalized version that also encompasses non-identity cases is called for. The challenge that arises is to block movement type interactions in some cases but allow it in others. For instance, it is clear that \bar{A} -movement of an element in a lower clause does not bleed A-movement in a higher clause:

(521) Sue_i was asked t_i [who_j Mary met t_j]

Why does \bar{A} -movement of one element sometimes preclude A-movement of another element and sometimes not? Abels (2007, 2009) proposes an account of these fact in terms of an extrinsic ordering of operations. Specifically, he proposes the *Generalized Prohibition against Improper Movement* in (522). This prohibition is encouched in a theory of *affectedness* as defined in (523). The core idea underlying Abels' (2007, 2009) proposal is that movement of a constituent *affects* other constituents in the structure. The nodes that are affected are as defined in (523). The ordering of operations crucial to the working of (522) is as in (524).

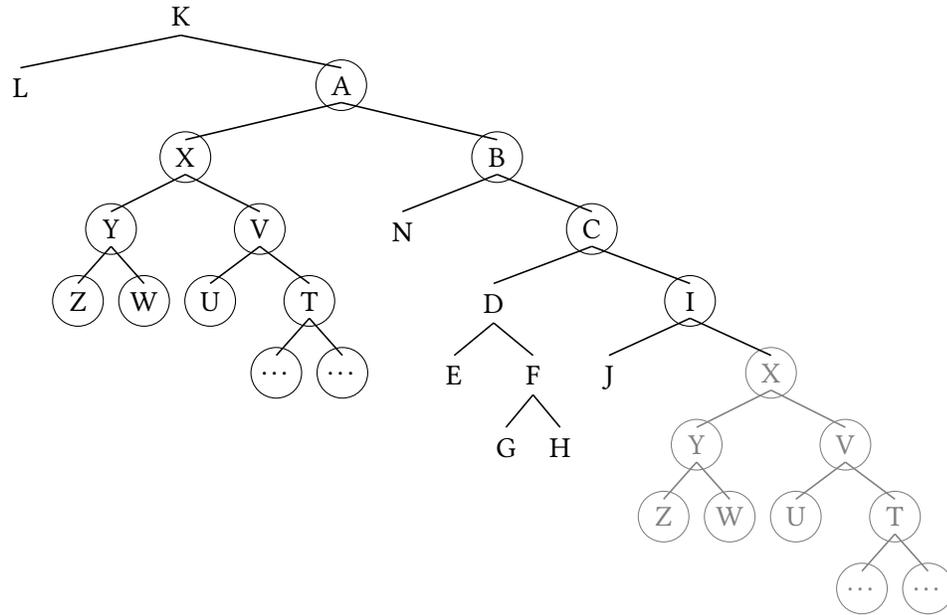
(522) **Generalized Prohibition against Improper Movement** (Abels 2007: 66)
 No constituent may undergo movement of type τ if it has been affected by movement of type σ , where $\tau \ll \sigma$ under UCOOL.

- (523) **Affectedness** (Abels 2007: 66)
 A constituent α is *affected* by a movement operation iff
- i. α is reflexively contained in the constituent created by movement, and
 - ii. α is in a (reflexive) domination relation with the moved constituent.
- (524) **Universal Constraint on Operational Ordering in Language (UCOOL)**
 $\theta \ll$ scrambling \ll A-movement \ll *wh* \ll topicalization (Abels 2007: 80)

Some clarifications regarding (523): The ‘constituent created by movement’ in (523i) designates the phrase-structure tree that is the output of the movement operation, hence the entire syntactic structure built up to this point. The term ‘domination’ in (523ii) is to be taken as meaning that α either dominates or is dominated by the constituent created by movement.

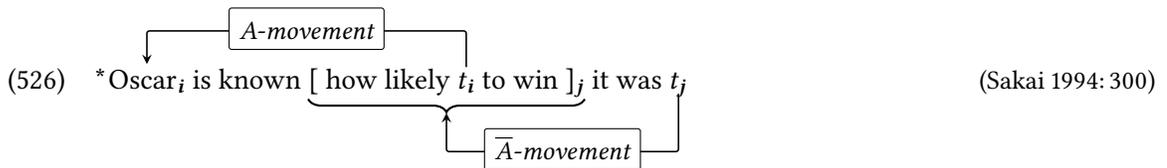
A graphic representation is helpful in understanding the distribution of affectedness. In (525), the constituent X is moved and the shading indicates the lower copy of X in its launching site. All nodes that are circled are affected by this movement of X: (523i) states that only nodes inside the tree that results from movement have to be considered. In (525), this excludes L and K, as these nodes are added only after movement of X takes place. All other nodes in (525) satisfy (523i). The condition in (523ii) limits the set of nodes to those that either dominate the moving element X (more specifically X’s copy in its base position), or are dominated by X. In the case of (525), this includes all nodes that dominate X or are dominated by X. The circled nodes in (525) correspond to the union of the two conditions in (523).

(525) Example of affected nodes (Abels 2007: 67)



(522) states that all nodes affected by some movement type cannot be targeted by a movement type lower in UCOOL (524). For example, if X has undergone *wh*-movement in (525), no circled constituent can subsequently be targeted by A-movement or scrambling. They can, however, undergo topicalization, as topicalization is ranked higher on UCOOL than *wh*-movement.

In this way, Abels (2007, 2009) and Neeleman & van de Koot (2010), who adopt this system, are able to capture the non-identity cases above. In (516), repeated here as (526), \bar{A} -movement of the infinitival clause affects *Oscar* inside of it, blocking the A-movement step. In (520), \bar{A} -movement of *which king* affects the container DP *a picture of t*, preventing A-movement of it.



The classical identity cases are likewise accounted for without further ado because movement of a constituent affects this constituent. Furthermore, the structure in (521), repeated here as (527), is allowed because \bar{A} -movement of *who* does not affect *Sue*.

(527) Sue_i was asked t_i [who_j Mary met t_j]

Finally, the non-binary nature of the hierarchy in (524) makes it clear that the account incorporates Müller & Sternefeld's (1993) discovery that movement type asymmetries are not limited to the A/ \bar{A} -distinction.

Abels' account represents a very attractive generalization to non-identity cases and is thus similar in scope to the account developed here. It is therefore instructive to compare the two, as the lines of analyses pursued by the two accounts differs significantly. The gist of Abels' account is that movement of an element affects a potentially very large set of nodes. Non-identity cases of improper movement follow from the fact that movement of one element quite literally affects other elements in the structure. The set of nodes that is affected by a given movement step is simply stipulated rather than derived. One might, for instance, easily envision a definition of affectedness in which only nodes that dominate the launching site of movement are affected by movement. On this hypothetical definition, the remnant movement case in (520) would still be ruled out, but the subextraction configuration in (516) would be well-formed. As a result, while the account offers a unification of the range of configurations in which improper movement is attested, this set of configurations does not have any privileged status amongst the logically possible configurations that could be described in the system under various possible definitions of affectedness.⁷

The horizon-based account of non-identity cases discussed in section 3.4.2 of chapter 3 ties the various configurations in which improper movement arises together in a tighter way. Recall

⁷ Müller (2014b: 119), citing Philipp Weisser (p.c.) and Grewendorf (2013), points out an interesting problem for the UCOOL account. Namely, it predicts that a CP in which *wh*-movement takes place should be unable to scramble. This is because *wh*-movement to SpecCP affects the CP node (in the visual representation in (525) movement of X within A affects A) and scrambling of that CP would then violate UCOOL. This, however, is incorrect, as (i) illustrates:

- (i) dass [_{CP} wen_i sie t_i dort getroffen hat]_j keiner der Anwesenden t_j sagen konnte
that whom she there met has nobody of the attendees say could
'that none of the attendees could say who she met there' (Müller 2014b: 119)

It seems to me that this problem for the UCOOL account is relatively superficial as it can be avoided by a slight reformulation of the definition of affectedness in (523): Removing 'reflexively' from (523-i) will ensure that the CP in (i) is not affected by *wh*-movement and scrambling is hence possible. On the other hand, the fact that such a resolution is readily available only underscores the claim in the main text that the definition of affectedness in (523) and the set of nodes that are affected by a movement step are *ad hoc*. Arbitrary reformulations of the definition with quite distinct empirical consequences are easily conceivable. Affectedness thus affords a way to encode movement interactions between different items but it leave unanswered the question of *why* we observe the interactions we do.

that on the horizon account, the probe underlying A-movement ($[uA]$) has C as its horizon and hence cannot search into a CP clause, barring A-movement out of CP. The traditional identity case of improper movement as well as the two non-identity cases in (516) (smuggling) and (520) (remnant movement) all fall under this account, simply because all involve A-movement over a CP. On the logic of the horizon account, it is impossible for horizons to give rise to a system in which improper movement arises in remnant movement configurations like (520) but not in smuggling configurations like (516) or vice versa. If probing of $[uA]$ is impossible, A-movement out of CP will be barred in identity as well as non-identity cases. If such probing is possible, it will be allowed in all three configurations. In this sense, then, horizons provide a truly unified account of the range of configurations that exhibit improper movement-style restrictions. They explain why movement of different elements interact rather than just stating such interaction as an axiom of the account.

A second central point of divergence between operational ordering and UCOOL concerns the means by which selective opacity is derived. On the operational ordering account, different movement types interact with each other directly. α -Movement may directly affect the possibility of β -movement. On the horizons account, by contrast, movement types never directly interact. Rather, \bar{A} -movement may bleed subsequent A-movement because \bar{A} -movement implicates the presence of a CP structure, which constitutes a horizon for the A-probe. As such, movement type interactions are *structurally mediated*. Interestingly, on an operational ordering account, selective opacity effects are *operationally mediated*. A-movement out of a CP is ruled out on this account because it requires an intermediate movement step to SpecCP, an \bar{A} -position, which prevents the element from undergoing further A-movement.

While the general approach to selective opacity effects and movement interactions in these two accounts is thus quite different, it is not easy to find empirical evidence that would adjudicate between the two views. One potential domain of interest involves ranking paradoxes for the UCOOL. To illustrate, consider the interaction between \bar{A} -movement and in-situ *wh*-licensing in Hindi, discussed in great detail in chapter 2. As I have shown there, *wh*-licensing may not cross a CP boundary. Thus, elements inside a CP clause cannot receive a matrix *wh*-contrual even if they are \bar{A} -moved to the edge of the CP:

- (528) *siitaa-ne soc-aa [kis-ko_i ravii-ne t_i dekh-aa] [Hindi, =(175)]
 Sita-ERG think-PFV.M.SG who-ACC Ram-ERG see-PFV.M.SG
Intended: ‘Who did Sita think that Ravi saw?’

On an operational ordering account, this would entail that \bar{A} -movement bleeds *wh*-licensing, hence that ‘*wh*-licensing \ll \bar{A} -movement’ in UCOOL. However, as discussed in section 2.4.3 of chapter 2, one way of achieving a matrix *wh*-construal of an element embedded inside a CP clause is to \bar{A} -move this element into the matrix clause:

- (529) kis-ko_i siitaa-ne soc-aa [ki ravii-ne t_i dekh-aa] [Hindi, =(184)]
 who-ACC Sita-ERG think-PFV.M.SG that Ravi-ERG see-PFV.M.SG
 ‘Who did Sita think that Ravi saw?’

The fact that \bar{A} -movement makes *wh*-licensing possible entails that the former can feed the latter, hence that ‘ \bar{A} -movement \ll *wh*-licensing’ in UCOOL.

We have thus arrived at a contradiction: (528) requires that ‘*wh*-licensing \ll \bar{A} -movement’, whereas (529) mandates that ‘ \bar{A} -movement \ll *wh*-licensing’. I do not see a way of avoiding the ranking paradox. By contrast, because on the horizon account operations do not directly interact with other operations but only with node labels in the phase structure tree, it suffices to say that *wh*-licensing has C as its horizon and that the \bar{A} -probe does not. Consequently, *wh*-licensing is blocked in (528) due to the intervening CP node, but not in (529), where \bar{A} -movement brings the *wh*-element into the matrix, effectively removing the CP barrier and enabling *wh*-licensing. Configurations of this sort thus appear to favor a structurally mediated account of selective opacity over an operationally mediated one. They deserve more study, which I will not undertake here. The case study just reported does, however, provide support for the horizons account.

A third point of comparison between horizons and operational ordering concerns the Height–Locality Connection, i.e., the observation that the locality of a movement type is related to the structural height of its landing site. We have seen empirical evidence for such a connection in chapter 1, in section 2.5.2 of chapter 2 for Hindi, and in section 4.3.3 of chapter 4 in German. It can be derived from horizons as described in section 3.5 of chapter 3. Abels (2007) and especially Abels (2009) notes that the UCOOL in (524) strongly resembles the relative height of the landing sites of the various movement types, but that this resemblance remains a coincidence on this account.

There is thus a pervasive redundancy in the system in that the hierarchy of the projections that make up the clausal spine and the UCOOL are isomorphic to each other, with no analytical link between the two. Consequently, on the UCOOL account it would be perfectly conceivable for UCOOL to be the inverse of hierarchy of landing sites. This would, however, have the effect that movement that lands structurally high is subject to stricter locality constraints than movement that lands low, exactly the inverse of the Height–Locality Connection. Abels (2007, 2009) leaves this question as an open problem, though Abels (2012a) can be seen as an attempt to solve it.

5.6 Buffers

A recent approach to selective opacity proposed by Müller (2014a,b) introduces the concept of a *buffer* that stores the (recent) movement history of a moving element and is required to conform to a given hierarchy of projections when that item reaches its final landing site. Selective opacity then results from certain movement paths being ruled out.

The core motivation for an approach that incorporates buffers comes from a locality problem that Müller (2014a,b) identifies for accounts of improper movement that adopt phases. I will illustrate the problem with scrambling and *wh*-movement in German, but it is considerably more general. Consider the standard view that CP and *v*Ps are phases and that as a consequence movement must pass through their specifiers. Assuming furthermore for the sake of concreteness that scrambling lands in Spec*v*P.⁸ The German example in (530a) illustrates *wh*-movement out of an embedded finite V-final clause. As we have seen in section 4.2.2 of chapter 4, such movement is possible. Compare this to (530b), where scrambling out of a finite V-final clause takes place and the result is ungrammatical (also see section 4.2.2 of chapter 4 for more information).

- (530) a. *Wh-movement out of CP clause* [German]
 Welches Buch_{*i*} hat [_{*v*P} *t*'_{*i*}' Karl gemeint [_{CP} *t*'_{*i*}' dass [_{*v*P} *t*'_{*i*}' jeder *t*_{*i*} lesen
 which book has Karl meant that everyone read
 möge]]]?
 should
 'Which book did Karl say that everyone should read?'

⁸ For Müller (2014a,b), every phrase is a phase, so the problem sketched in the text will arise regardless of the landing site of scrambling. The assumption in the text are made for expository purposes only.

b. *Scrambling out of CP clause*

*dass Karl [_{vP} das Buch_i glaubt [_{CP} t'_i dass [_{vP} t'_i keiner t_i liest]]]
 that Karl the book thinks that no one reads
Intended: 'that Karl believes that no one reads this book'

If vP is a phase, *wh*-movement in (530a) has to proceed from the embedded SpecCP to the matrix SpecvP. Assuming that scrambling targets SpecvP (as Müller 2014a,b does), (530b) would appear to show that precisely such a movement step is impossible.⁹ The puzzle Müller (2014a,b) identifies is to account for the fact that SpecCP-to-SpecvP movement is well-formed in (530a), where the element then moves on to the matrix SpecCP, but ruled out in (530b), where it does not. The core of the problem is that vP phases make it impossible to simply distinguish illicit CP-to-vP movement from licit CP-to-CP movement because the latter properly contains the former if vP is a phase. Müller (2014a,b) refers to this predicament as the *promiscuity problem*. It arises because an intermediate movement step from SpecCP to SpecvP that continues on to a higher SpecCP must be allowed, whereas a terminal movement step from SpecCP to SpecvP must be blocked. Standard accounts have no means of distinguishing the two. If SpecCP-to-SpecvP movement is allowed, both (530a,b) are ruled in. If SpecCP-to-SpecvP movement is blocked, both are ruled out. Either way fails to capture their grammaticality contrast.

Despite the fact that this problem arises for any account that assumes vP phases, Müller (2014a,b) is to my knowledge the only existent attempt of addressing it while maintaining vP phases. His solution is couched within the assumption that not only CP and vP are phases, but that every phrase is a phase so that extraction must pass through the specifier of every projection it leaves. This assumption is not crucial to his account but I will adopt it in laying out his proposal. In a nutshell, he proposes that moving elements contain a buffer that keeps a record of the projections that this item has moved through. Concretely, he suggests that movement of an element α to Spec π P adds ' π ' to α 's buffer. Moreover, if an element moves into a projection that already exists in its buffer, the lower occurrence of the projection on that buffer is deleted and only the higher one is kept. When an element reaches its final landing site (or criterial position), its buffer has to conform to functional sequence of projections:

⁹ The assumption that scrambling lands in SpecvP is inconsequential for the argument. Assuming some other landing site does not obviate the problem and Müller in fact assumes that scrambling may target a variety of positions.

(531) *Williams Cycle* (simplified)¹⁰

When an element reaches a criterial position, its buffer has to correspond to the functional sequence of projections f-seq.

To illustrate, in (530a), the DP *welches Buch* ‘which book’ will move through the specifier of every projection between its base-generation site and the matrix CP. This derivation is indicated in (531a), with the buffer of the moving element at various positions. Crucially, only the buffer state at the final landing site has to conform to (531). This is the case, as shown in (531a-iv), where the buffer $\boxed{\text{CTvV}}$ matches f-seq (a strikethrough indicates deleted symbols for the sake of transparency) and the entire movement path is well-formed. This contrasts with the case of long-distance scrambling in (530b), the structure of which is indicated in (531a). Here the buffer at the criterial position is $\boxed{\text{vVCT}}$ (see (531b-iii)), which does not conform to f-seq. The derivation is hence prohibited by (531).

(531) a. *Schematic structure of (530a):*

$[\text{CP}_1 \text{DP}_i [\text{TP}_1 t_i'''''' [\text{vP}_1 t_i'''' [\text{VP}_1 t_i'''' [\text{CP}_2 t_i''' [\text{TP}_2 t_i'' [\text{vP}_2 t_i' [\text{VP}_2 t_i]]]]]]]]]]$

i. Buffer at $\text{vP}_2 (t_i')$: $\boxed{\text{vV}}$

ii. Buffer at $\text{CP}_2 (t_i''')$: $\boxed{\text{CTvV}}$

iii. Buffer at $\text{vP}_1 (t_i''''')$: $\boxed{\text{vVCT}\cancel{\text{vV}}}$

iv. Buffer at $\text{CP}_1 (\text{DP}_i)$: $\boxed{\text{CTvVCT}\cancel{\text{vV}}}$ \leadsto criterial position \rightarrow conforms to f-seq

¹⁰ The specific definitions are given in (i) (Müller 2014b: 42–43):

(i) a. *Valuation*

Merge($Y: [\bullet X_\gamma \bullet]$, $Z: [F: \boxed{\delta_1 \dots \delta_n}]$) \rightarrow $Y Z: [F: \boxed{\gamma \delta_1 \dots \delta_n}]$

where F is a movement-related feature, $\delta_1, \dots, \delta_n$ is a (possibly empty) list of (category, possibly other) symbols, and γ is the category label of Y.

b. *Deletion*

$Y Z: [F: \boxed{\gamma \delta_1 \dots \delta_i \gamma \delta_j \dots \delta_n}] \rightarrow Y Z: [F: \boxed{\gamma \delta_1 \dots \delta_i \delta_j \dots \delta_n}]$

where F is a movement-related feature, $\delta_1, \dots, \delta_n$ is a (possibly empty) list of (category, possibly other) symbols, and γ is the category label of Y.

The well-formedness constraint on buffers is stated as in (ii):

(ii) *Williams Cycle*

(Müller 2014b: 44)

Information on a list of a movement-related feature β must conform to f-seq when β is checked by an inherent structure-building feature $[\bullet \beta_\pi \bullet]$ of a phase head π (i.e., in criterial positions).

See Müller (2014a,b) for the technical detail.

b. *Schematic structure of (530b):*

*[CP₁ [TP₁ [vP₁ DP_i [VP₁ t_i^{'''} [CP₂ t_i^{'''} [TP₂ t_i^{''} [vP₂ t_i['] [VP₂ t_i]]]]]]]]]]

i. Buffer at vP₂ (t_i[']): vV

ii. Buffer at CP₂ (t_i^{'''}): CTvV

iii. Buffer at vP₁ (DP_i): vVCTvV

~> *riterial position* → violates f-seq

In this way, selective opacity effects are accounted for by constraints on the content of buffers in criterial positions. It should be clear that this account is general enough to handle selective opacity cases that fall beyond the A/ \bar{A} -distinction. Rather, there are as many conceivable locality distinctions as there are projections in the clausal spine. This makes the interesting prediction that movement types that target the same structural position must have identical locality profiles. This is not the case on the analyses of Hindi in chapter 3 (where the \bar{A} -probe and the *wh*-probe both reside on C⁰ but differ in their locality) and of German in chapter 4 (where relativization and *wh*-movement target SpecCP but differ in locality).

What is less clear is whether the account also applies to non-identity cases of improper movement such as the ones discussed in the preceding section. The reason is that each buffer only registers the movement path of a single element. All else equal, then, movement of a constituent α should have no impact on the movement options of another constituent β . Yet as we have seen in the previous sections, this is not the case. The underlying problem is that the account is item-based. Just like for other item-based accounts, the non-identity cases pose a problem. Müller (2014b) is aware of the issue and proposes an extension of the buffer-based account that applies to remnant movement. I will merely sketch the gist of the account here, the reader is referred to Müller (2014b: chapter 3) for a full discussion.

The core of Müller's (2014b) proposal for remnant movement is that extraction of an element α out of a constituent β that will itself undergo movement *contaminates* β 's buffer with α 's index feature. Contaminated buffers invariably violate f-seq and lead to ungrammaticality. The contaminated buffer of β can also be *decontaminated* in a configuration in which β is located in the specifier position of some projection and α 's criterial landing site is an inner specifier of the same projection. The key consequence of this set of assumptions is that in remnant movement configurations, the final landing site of the remnant has to be higher than the final landing site of the subextracted

constituent. In this way, movement type interactions between the two elements α and β can be captured because movement of one has effects on the buffer of the other.

Incidentally, the constraint on remnant movement that emerges under a buffer-based account is different than the one argued for by Abels (2007, 2009). For example, Abels argues that *wh*-movement out of scrambled constituent is possible, as in (532). This configuration is excluded on a buffer-based account because the buffer of the remnant *einen Südkurier-Artikel* is still contaminated when it reaches its criterial position. Because (532) is grammatical, the buffer account seems too restrictive.

- (532) *Wh-movement out of scrambled constituent* [German]
 Worüber_i kann jeder Schwachkopf [einen Südkurier-Artikel *t_i*]_j am Strand *t_j*
 what about can every moron a Südkurier-article at.the beach
 verfassen?
 write
 ‘For which topic is it the case that every moron can write an article about it for the
 Südkurier when he is at the beach?’ (Faselow 1991: 188)

Müller (2014b) gives the example in (533) to show that *wh*-movement out a scrambled element (a VP in this case) is ungrammatical:

- (533) *Was_i hat [*t_i* gelesen]_j keiner *t_j*? [German]
 what has read nobody
Intended: ‘What did no one read?’ (Müller 2014b: 69)

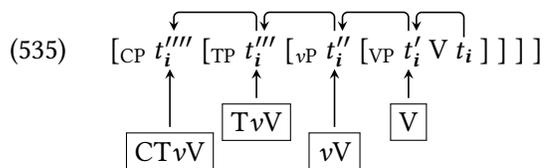
However, (533) is quite plausibly ruled out for unrelated reasons. As (534) shows, the structure is ungrammatical even if no *wh*-movement out of the scrambled VP takes place:

- (534) *Gestern hat [das Buch gelesen]_j keiner *t_j*. [German]
 yesterday has the book read nobody
Intended: ‘Yesterday no one read the book/a book.’

I conclude from these considerations that (534) is ungrammatical due to an independent restriction on VP scrambling and that Abels’ (2007, 2009) generalization, which is adopted in the horizon account, is corroborated. The undergeneration problem a buffers-based account faces thus persists.

Moreover, while a buffer-based can thus be extended to the non-identity cases of improper movement, it should be pointed out that the two are not actually unified, they are merely implemented with reference to the same concept (i.e., buffers). The concept of contamination that lies at the heart of the account of the non-identity cases plays no role in the account of the identity cases. Consequently, it would be perfectly conceivable a priori to have a system that lacks the notion of contamination and would exhibit improper movement *only* in the identity cases. Conversely, it would also be possible to describe a system using buffers that merely requires buffers to be free of indices. In such a system, improper movement would arise *only* in the non-identity cases. As such, the fact that improper movement arises in both domains is coincidental on the buffers account. Horizons provide a more unified account of identity and non-identity cases (see section 3.4.2 in chapter 3, section 4.4.2 of chapter 4, and sections 5.3 and 5.5 above).

In fact, even the fundamental claim that buffers have to be compatible with the functional sequence in criterial positions is analytically arbitrary. After all, one could have just as well imposed a number of logically conceivable constraints on buffers, such as that the buffer has to correspond to a complete functional sequence. As illustrated in (535), only traces in SpecCP represent an entire functional sequence. If the buffer were to be required to correspond to an entire sequence, movement would be well-formed only if it ends up in the topmost projection of a functional sequence, i.e., CP in the anatomy of clause structure Müller (2014a,b) assumes. Such a situation is unattested.



Conversely, one could also require buffers to be proper subsets of the functional sequence, in which case movement in (535) would be able to terminate in SpecVP, SpecvP, and in SpecTP, but not in SpecCP. Again, such a situation is unattested, but could be straightforwardly imposed using the machinery of buffers.

Most strikingly, it would be possible to require that the final form of a buffer must *not* conform to the functional sequence. This hypothetical requirement would constitute the inverse of the Williams Cycle in (531) and it would result in a situation in which only sentences that violate the Williams Cycle are grammatical. For example, movement out of a CP clause would be allowed

only if this movement lands *lower* than SpecCP in the higher clause, and so on. Such a situation is non-existent, but from the point of view of the buffers account, it is no more or less plausible than the situation that is empirically attested. As a consequence, buffers provide the means to technically implement selective opacity effects, but they do not offer a rationale for why these effects should exist or why they should pattern the way they do.

Despite its versatility, it is not obvious how an account in terms of buffers could be extended to selective opacity effects that arise in the domain of non-movement structures. As I have argued at length in chapter 2, selective opacity is not limited to movement, but also arises for in-situ relations. Because buffers and their feature structures are fundamentally linked to Merge and Move, it is not evident that the account has anything to say about relations that do not involve Merge or Move.¹¹

A final point worth mentioning is conceptual in nature. The problem that gives rise to Müller's (2014a,b) account is a very general tension between selective opacity as a phenomenon and the locality constraints imposed by clause-internal phases. What selective opacity effects have in common is that it is particular *combinations* of launching and landing sites that are ruled out (e.g., CP-to-*v*P scrambling vs. CP-to-CP movement, etc.) Clause-internal phases impose locality that makes it impossible to state such constraints directly because the launching and the landing site of movement may be in different phases and hence never simultaneously accessible. I believe that the problem is real, of great significance for the understanding of selective opacity and in need of solution. Müller (2014a,b) is, to my knowledge, the only attempt to address it. The solution he advocates strikes me as conceptually odd, however. After all, the account on the one hand assumes extremely local phase domains, but on the other hand invokes a mechanism whose sole purpose is to circumvent the effects of these local domains by keeping a record of an element's movement history. Put differently, the account assumes that syntactic information is lost periodically but at the same time makes use of a mechanism that keeps a record of this information, preventing it from getting lost. One might wonder why syntax should employ cyclic Spell-Out and a mechanism that partially negates the effects of cyclic Spell-Out. In chapter 6, I will propose a solution that is very different from Müller's (2014a,b). I will argue that the locality problem ceases to exist if there

¹¹ Incidentally, on Müller's (2014a,b) view that every phrase is a phase, there should not be long-distance relations that do not involve movement. See section 6.3.2 of chapter 6 for related discussion.

are no clause-internal phases, hence if νP is not a phase (I will, however, argue that CP is a phase). The problem Müller (2014a,b) addresses will then simply not arise in the first place.

5.7 Derivational clock

The final account I would like to briefly discuss here has been developed most clearly in Williams (2003, 2011, 2013), but goes back to Williams (1974) and van Riemsdijk & Williams (1981). The account derives a strong and extremely elegant account of improper movement in a variety of constructions from the very way that syntactic structures are built. The framework that lies at the heart of the account, *Representation Theory*, represents a major departure from the standard bottom-up mode of structure building standardly assumed in Minimalist syntax. In the interest of space, I will confine myself to the basics of Williams' account here. A brief overview is provided by Hornstein & Nevins (2005), also see Nevins (2005), who adopts a variant of the account.

The core proposal underlying Williams' (2003, 2011, 2013) account of selective opacity is that clauses are built in parallel, even if one is embedded inside the other. At first, no clause is embedded inside another and all clauses are built simultaneously. The derivational clock (or 'F-clock') refers to different points of structure building. Thus, first all clauses are built up to the VP level (this derivational point is called F_V). Next, all clauses are built up to the νP level (F_{ν}), then to the TP level (F_T), and so on. Embedding of one clause into another can take place at any point but once a clause has been embedded, it no longer increases in structural size. Clausal embedding is hence fundamentally countercyclic. As an example, consider a structure in which a CP clause is embedded inside another CP clause.¹² Initially, the two clauses are assembled in separate workspaces and grow in parallel. Thus, first both clauses are assembled to the VP level, then they are assembled to the νP level, then to the TP level, and finally to the CP level. At that point, one CP clause is embedded inside the other. Compare this to a structure in which a TP clause embedded. As before, both clauses are simultaneously built up to the TP level. At this point, one TP clause is embedded inside the other. After this embedding has taken place, the matrix TP clause continues to grow to become a CP, but the embedded clause does not, it remains a TP.

¹² I will base this exposition on the derivational version of the system developed in Williams (2013), which is based on the 'F-clock'. The version of the account developed by Williams (2003) is thoroughly representational in nature.

Selective opacity arises on this account from the familiar Extension Condition, according to which movement must target the root of the phrase marker. Consider the standard case of improper movement from SpecCP to SpecTP. Because the embedded clause is a CP, embedding must take place after *both* clauses are CPs. The Extension Condition then allows movement out of the lower clause to target the matrix SpecCP, but it rules out movement to SpecTP.¹³ TP-to-TP movement, on the other hand, is allowed in this system because embedding will take place when both clauses are TPs. At this point, the matrix TP is the root node and can hence be targeted by extraction out of the lower clause.

These considerations generalize. As Williams (2003, 2011, 2013) discusses in detail, this account has the general consequences in (536):

- (536) Given a functional clause structure $\langle \Pi_n > \dots > \Pi_1 \rangle$ (e.g., $\langle C > T > v > V \rangle$),
- a. no movement can proceed from Π_i in a lower clause to Π_j in the higher clause if $\Pi_i > \Pi_j$ (e.g., no movement from CP to TP);
 - b. no movement from a lower clause landing in Π_j of the higher clause can cross Π_i if $\Pi_i > \Pi_j$ (e.g., movement to TP cannot cross CP);
 - c. movement that lands in Π_i may be longer than movement that lands in Π_j for all $\Pi_i > \Pi_j$ (movement to CP may be less local than movement to TP).

What is remarkable about (536) is that they are all derived from the very architecture of the grammar. (536a) instantiates a generalized ban on improper movement that extends beyond the A/\bar{A} -distinction. (536b) extends this ban to the non-identity cases discussed above. (536c) instantiates a version of the Height–Locality Connection, as it imposes a link between the structural height of the landing site of a movement and its locality profile: movement that lands in a structurally high position is subject to less strict locality conditions than movement that lands low.

Williams’ account is ingenious in its elegance and generality. It manages to derive selective opacity effects as a consequence of the way syntactic structures are assembled.¹⁴ It captures the full range of selective opacity effects identified in chapter 1: In particular, it extends to a variety of movement type interactions that lie outside the A/\bar{A} -divide and it captures the non-identity cases

¹³ Incidentally, Williams notes that this account rules out successive-cyclic movement.

¹⁴ Williams (2013) in fact pursues the view that regular islands are the result of improper movement, as conceived of on the derivational clock account, further broadening the scope of the account.

in exactly the same way as the identity cases, thus providing a truly unified analysis. Furthermore, while Williams does not address in-situ relations like agreement, the generality of the account makes it possible to capture those as well. Furthermore, the account derives without further ado Upward Entailment and the Height–Locality Connection as meta-generalizations of selective opacity. The key to the generality and versatility of the account is that it is domain-based in that the account does not refer to properties of individual items, but to clauses and their structural size. Horizons share the domain-based nature of the account and therefore afford a similarly general account, albeit in technically a very different way.

As is so often the case, there is a price to pay for the virtues of Williams’ (2003, 2011, 2013) account. Deriving (536) requires a view of syntactic derivations that is inconsistent with standard principles of structure building in a bottom-up fashion. Furthermore, as Abels (2007, 2009) has argued, the account is empirically too restrictive in that it rules out a number of attested movement dependencies. This point is addressed in greater detail in the next section.

5.8 Previous approaches to the Height–Locality Connection

One important generalization of selective opacity effects across constructions and languages is that the locality profile of an operation is related to the structural height of the projection that this operation is associated with. In the domain of movement, for instance, processes that target positions high in the clausal spine are subject to less strict locality constraints than processes that land low. Evidence for this generalization, which is stated in (537), has been presented in section 1.3.1 of chapter 1, section 2.5.2 of chapter 2, and section 4.3.3 of chapter 4.

(537) **Height–Locality Connection**

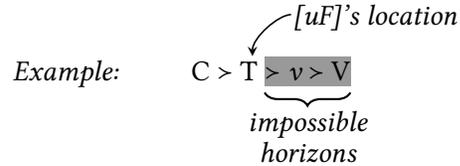
The higher the landing site of a movement type is in the clausal structure, the more kinds of structures are transparent to this movement type.

As discussed at length in section 3.5 of chapter 3 (also see section 4.6 of chapter 4), horizons derive a version of (537) because certain pairings of location and horizon on a given probe render this probe vacuous in the sense that such probe are unable to give rise to long-distance dependencies like movement or agreement for principled reasons. The theorem that emerges on a horizon-based account is repeated in (538) from (333) in chapter 3:

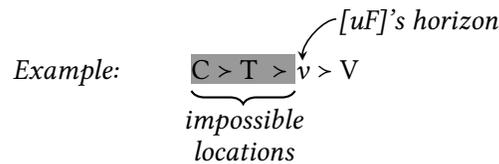
(538) **HEIGHT-LOCALITY THEOREM**

Given an extended projection $\Phi = \langle \Pi_n > \Pi_{n-1} > \dots > \Pi_1 \rangle$, for any non-vacuous probe $[uF]$,

- a. If $[uF]$ is located on Π_m , then a projection $\in \{\Pi_{m-1}, \dots, \Pi_1\}$ cannot be a horizon for $[uF]$.



- b. If $[uF]$ has Π_m as a horizon, $[uF]$ cannot be located on a projection $\in \{\Pi_n, \dots, \Pi_{m+1}\}$.



According to (538a), if a probe is located on, e.g., C^0 , this probe cannot have V , v or T as its horizon. Conversely, (538b) states that if a probe has, e.g., T as its horizon, this probe cannot be located on any head higher than T^0 , i.e., it cannot reside in the C domain. As also discussed in chapter 3, (538) is not a stipulation, but a theorem that emerges from horizons. In a nutshell, only probes for which location and horizon conform to (538) have a search space that allows them to enter into long-distance relations. It follows that all long-distance relations, be it movement or agreement, must conform to (538).

In comparing horizons to previous accounts of selective opacity, it is instructive to consider previous attempts to capture the relationship between height and locality in (537). I will focus on the approaches in Williams (2003, 2011, 2013) and Müller (2014a,b) on the one hand and Abels (2012a) on the other. Despite significant differences in execution, previous approaches pursue a similar general strategy: They derive one half of the connection from the other. Williams (2003, 2011, 2013) and Müller (2014a,b) suggest that the locality restrictions of a movement type should follow from the height of its landing site given a general theory of locality. Abels (2012a) pursues the opposite route: He argues that the height of a movement type's landing site should be a consequence of its locality properties.

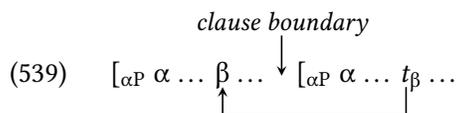
These approaches provide elegant accounts of the height-locality connection in (537) by essentially reducing one to the other. This in effect renders the connection a correlation: One side

of the connection completely determines the other: If we know a movement type’s landing site, we immediately know its locality profile and vice versa. This contrasts with the account advocated here. The theorem in (538) imposes restrictions on possible pairings of height and locality but differs from previous proposals in that it does not reduce one to the other. That is, if we know the location of a probe, certain locality profiles become impossible for principled reasons, but location does not uniquely determine a locality profile. This section will argue that previous accounts are too restrictive empirically. Neither side of the height–locality connection fully determines the other. This provides empirical evidence in favor of the more flexible account proposed here.

5.8.1 Reducing locality to height (Williams 2003, 2011, 2013)

Focusing on movement dependencies, Williams (2003, 2011, 2013) and to some extent Müller (2014a,b)¹⁵ derive a movement type’s locality from the height of its landing site: Given an extended projection, there is a general principle of grammar that establishes whether extraction from one projection into another is licit or not (see sections 5.7 and 5.6 above for a brief description of these accounts). On these accounts, a movement type’s locality restrictions fall out from the structural height of its landing site.

Because on this account locality is a function of height, the connection between the two is captured in a very elegant way. Nonetheless, the approach faces a number of substantial problems.¹⁶ Recall that on Williams’ account movement into a projection XP cannot cross a projection YP if YP is structurally higher than XP in the functional hierarchy. Thus, movement to TP cannot cross a CP node (see (536)). Abels (2007, 2009) provides strong evidence that this yields too restrictive a system, based on configurations of the general shape in (539).



¹⁵ The buffer-based system as described in section 5.6 makes the same empirical predictions as Williams’ system. Müller (2014a,b) is aware of the empirical problems that arise and proposes a weakened version, according to which symbols on a buffer can be deleted if a projection is encountered that is sufficiently similar, though not necessarily identical, to the symbol being deleted. This more nuanced account does not face the undergeneration issue discussed in this section and I will therefore focus on Williams’ stricter system instead.

¹⁶ I am very grateful to Klaus Abels, Gereon Müller, and Edwin Williams for very helpful discussion of these issues.

In this configuration, β has moved over an element α within an embedded clause into a position that is lower than α in a higher clause. Such a configuration is ruled out on Williams' (2003, 2011, 2013) account for the following reason: Because β lands to the right of α in the matrix clause, β 's landing site must be lower than α in its extended projection. On Williams' account, this entails that this movement should not be able to pass α . The account therefore make the direct prediction that (539) should be unattested. Yet this prediction is not borne out. Structures of the type in (539) are well-attested, as Williams (2003: 80) himself notes for the French *L-tous* construction. Abels (2007: 82–84, 2009: 343–345) brings to bear a host of other constructions that exemplify (539). Among them is subject-to-object raising in English, which can move an element over a negation in the lower clause but nonetheless land beneath a negation:

(540) I expect John [*not* to want Mary_i [*not* to be told t_i the truth]]. (Abels 2009: 343)

Another example from Abels (2007, 2009) comes from the Bantu language Kîtharaka. Here the *wh*-word *ûû* 'who' is moved over the complementizer *atî* 'that' of the lower clause, but nevertheless lands to the right of that complementizer in the higher clause:

(541) U- ri- thugania [atî] n- ûû_i John a- ug- ir- e [atî] Pat n- a-
 2SG- PRES- think that FOC- who John SA- say- PFV- FV that Pat FOC- SA-
 ug- ir- e [Lucy n- a- ring- ir- e t_i]]
 say- PRES- FV Lucy FOC- SA- beat- PFV- FV
 'Who do you think that John said that Pat said that Lucy beat?' (Abels 2007: 84)

(541) instantiates (539) and is hence predicted to be ungrammatical on Williams' (2003, 2011, 2013) account, contrary to fact. Another example Abels (2007, 2009) gives is scrambling out of finite clauses in Russian (see, e.g., Müller & Sternefeld 1993 for examples and discussion as well as section 3.5.2 of chapter 3). Superraising in Bantu languages, where an element A-moves out of a CP clause, provides yet another example. (483) provides an example from Zulu:

(542) *uZinhle u-bonakala [ukuthi u-xova ujeqe manje].*
 AUG.1Zinhle 1SUBJ-seem that 1SUBJ-make AUG.1steamed.bread now
 'Zinhle seems to be making steamed bread now.' (Halpert 2012: 247)

To this catalog Müller (2014a) adds superraising in Japanese and Greek. One may also add superraising in Brazilian Portuguese to the list (Nunes 2008), though see Williams (2011) for a critical assessment. Other examples attesting to the same problem are A-movement out of a CP (see the list in McCloskey 2000: 71n20). Williams' (2003, 2011, 2013) version of the Height–Locality Connection is too strict to accommodate structures with the general shape in (539), which are widely attested. The account thus faces a severe undergeneration problem.

The horizon-based approach, on the other hand, accommodates (539) without problem because it is not ruled out by the Height–Locality Theorem (538). Superraising has been discussed at length in section 3.5.2 of chapter 3 and the account there extends to the other examples just mentioned. In a nutshell, the relevant ECM-probe in English is located lower than T^0 (perhaps even on V itself) and negation, but its horizon is C. ECM is therefore impossible out of finite clauses, but allowed out of nonfinite ones and across negation. Horizons thus allow us to recast the difference between languages that do not allow superraising and ones that do as a simple parametric choice. This variability is possible in the horizons-based approach precisely because a probe's syntactic height constrains its locality but does not completely determine it. 'Leakage' between height and locality such as in (539) is thus expected.

5.8.2 Reducing height to locality (Abels 2012a)

The opposite direction of elimination is explored by Abels (2012a), who likewise focuses on movement dependencies. In an intriguing investigation of the Italian left periphery, he observes that constraints on locality that are observable *across clauses* make superfluous constraints on landing sites *within clauses*. To give just one example for Abels' (2012a) reasoning, consider the complementizer *se* 'if'. As (543) demonstrates, if focus movement takes place in a clause that contains *se*, the focused constituent has to follow *se*.

- (543) a. Mi domando se QUESTO gli volessero dire (non qualcos' altro).
 I wonder if THIS they wanted to say to him not something else

- b. *Mi domando QUESTO se gli volessero dire (non qualcos'
 I wonder THIS if they wanted to say to him not something
 altro).
 else (Rizzi 2001: 289)

Following Rizzi's (1997) seminal work, facts like these are usually accounted for by a cartographic statement, according to which the projection hosting focus movement (FocusP) is generated lower than the projection hosting *se*. Abels (2012a: 244–245) observes that such a statement is redundant because focus movement cannot leave a clause that contains *se*, as the example in (544) illustrates:

- (544) *QUESTO mi domando se gli volessero dire (non qualcos'
 THIS I wonder if they wanted to say to him not something
 altro).
 else (Ilaria Frana, p.c.)

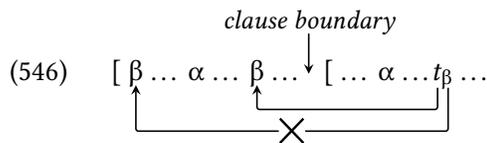
(544) cannot be accounted for in terms of a cartographic statement because the focused element and the complementizer are in distinct clauses. Abels (2012a) reasons that what (544) shows is that there is a locality constraint that prohibits focus movement from crossing the complementizer *se*. However, once such a locality constraint is acknowledged, the contrast in (543) follows immediately. To derive (543b), a focused element would have to be moved over *se*, but this is impossible, as we have seen. Abels (2012a) thus concludes that no cartographic statement about the landing position of focus movement relative to *se* is necessary. Their relative height already follows from locality. Ideally, Abels (2012a) reasons, restrictions on the height of a movement type are a direct consequence of its locality, i.e. what structures it cannot apply over. In other words, on a strong version of the approach a movement type's locality restrictions completely determine where this movement's landing site can be.

I have already shown in section 3.5.1 of chapter 3 how these facts are accounted for by horizons and the Height–Locality Theorem they give rise to. In short, based on (544), we can conclude that the head that *se* is the realization of (C_{se}^0) is a horizon for the focus probe [*uFoc*]. Because of the

Height–Locality Theorem (538), this fact entails that the Focus head must be located lower than C_{se}^0 in the same clause. Otherwise, [*uFoc*] would be vacuous. (543) then follows.¹⁷

- (545) a. [*uFoc*]_{Focus⁰} \neg | C_{se} entailment
 b. ... > C_{se} > ... > Focus > ... ←

It can be shown, however, that reducing height to locality is not tenable in all cases. Structures that are problematic have the schematic biclausal structure in (546):



In (546), movement of β passes α in the embedded clause. This entails that α cannot be a barrier for this movement step. Yet in the higher clause, β may land only underneath α , not above it. Because α does not block β -movement over it, locality constraints alone do not determine the ordering of α and β relative to each other within the same clause. In other words, movement of β over α should be possible within one clause just as it is possible across clauses. That movement of β in (546) obligatorily lands lower than α thus cannot be the result of locality restrictions on the movement type.

¹⁷ Abels (2012a) also shows that the same argument applies to the interaction of different movement types. To illustrate using relativization and topicalization, (i) shows that relativization has to land in a position higher than topicalization:

- (i) a. Un uomo a cui, il premio Nobel, lo daranno senz'altro
 a man to whom the Nobel Prize they will give it undoubtedly
 b. *Un uomo, il premio Nobel, a cui lo daranno senz'altro
 a man the Nobel Prize to whom they will give it undoubtedly (Rizzi 1997: 289)

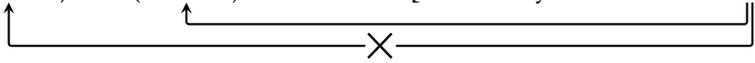
As (ii) demonstrates, relativization is possible out of clause in which topicalization has taken place, but the inverse is impossible.

- (ii) a. Questo é l'uomo, a cui tu pensi che, il premio Nobel, lo daranno senz'altro.
 this is the man to who you think that the Nobel Prize they will give it undoubtedly
 b. *A Gianna, ti parleró solo delle persone che senz'altro gli daranno
 to Gianni I will talk to you only about people who undoubtedly will give him
 il premio Nobel.
 the Nobel Prize (Abels 2012a: 235)

The contrast in (ii) demonstrates a locality constraint on topicalization: it is impossible over a relativized element. From this fact alone it follows that a topic cannot land to the left of a relative pronoun in (i).

Configurations like (546) are, however, attested. I will first give a few examples that instantiate it. After that, I will consider these examples in the context of the horizons account and demonstrate that they are not problematic. The first example of (546) is topicalization in English. The complementizer *that* is not a barrier for topicalization, yet a topicalized element cannot appear to the left of *that* in the same clause:¹⁸

(547) Sue said [*(*Barriers)* **that** (*Barriers*) Tom believes [**that** everyone should read *t*]].



The complementizer *that* hence does not block topicalization over it. The fact that topicalization cannot land above *that* in the same clause therefore cannot follow from the locality of topicalization with respect to *that*.

The second example of (546) is noted by Abels (2012a: 251) himself: In Italian the complementizer *che* may be crossed by various movement types, like topicalization in (548). However, if topicalization lands in a clause that contains *che*, it has to obligatorily occur to the right of it, as (549) shows. The facts are thus equivalent to those of topicalization in English.

(548) *Il premio Nobel_i Gianni pensa [che lo daranno a Giulia t_i senz'altro]*
 the Nobel Prize Gianni thinks that it they will give to Giulia undoubtedly
 'The Nobel Prize, Gianni thinks that they will undoubtedly give to Giulia.'
 (Ilaria Frana, p.c.)

(549) a. Gianni pensa [che *il premio Nobel* lo daranno a Giulia t_i senz'altro]
 Gianni thinks that the Nobel Prize it they will give to Giulia undoubtedly
 'Gianni thinks that the Nobel Prize, they will undoubtedly give to Giulia.'

b. *Gianni pensa [*il premio Nobel* che lo daranno a Giulia t_i senz'altro]
 Gianni thinks the Nobel Prize that it they will give to Giulia undoubtedly
 (Ilaria Frana, p.c.)

As before, because *che* does not induce a locality boundary for topicalization (548), the ordering of the two relative to each other in the same clause must follow from a cartographic statement that does not reduce to locality. I will return to a horizon analysis of these facts momentarily.

¹⁸ Note that (547) likewise constitutes a problem for Williams' (2003, 2011, 2013) system: Because topicalization has to land in a position lower than *that*, it is predicted not to be able to cross a complementizer in the case of crossclausal extraction, contrary to fact.

The third example of (546) comes from complementizer *ki* in Hindi, which exhibits analogous restrictions with respect to a displaced element. A fourth example is raising-to-object in English, which may cross a negation in the lower clause but has to land below negation in the matrix clause (see (540)). A final example is \bar{A} -movement in Hindi. As we have seen on the basis of (223) in chapter 2, \bar{A} -movement may leave nonfinite clauses, but it cannot land inside them. This yet again shows that restrictions on the landing site of \bar{A} -movement are not a mere reflex of its locality constraint and consequently cannot be reduced to it.

In sum, configurations like (546) are well attested. While Abels (2012a) makes a compelling argument that locality restrictions impose restrictions on the height of a landing site in line with the Height–Locality Connection, it is also clear that the two can mismatch in principle. It is thus not possible in all cases to deduce a movement’s landing site from its locality.

Configurations like (546) are not problematic for horizons, because horizons allow height and locality to be specified independently (thus creating the possibility of mismatches between them), but at the same time impose restrictions on the range of possible mismatches in the form of the Height–Locality Theorem (538). To illustrate, in the case of English topicalization, the probe is located on T^0 (see section 3.5.3 of chapter 3 for some remarks), but does not have a horizon and can hence search into a CP clause (550). It follows from the empty horizons setting of [*uTop*] that topicalization can cross *that* if it lands in a higher clause. At the same time, the location of [*uTop*] on T^0 ensures that topicalization cannot land to the right of *that* in the same clause. (547) then follows.

(550) *English topicalization:*
 $[uTop]_{T^0} \dashv \! \! \dashv \emptyset$

Analogous remarks apply to the relation between topicalization and *che* in Italian in (548) and (549). Assuming a decomposed left periphery (Rizzi 1997) and remaining agnostic with respect to the identity of the heads involved, assume that *che* is the realization of a ‘ C_{che}^0 ’ head, which appears above the Topic projection in the extended projection (551b). The topic probe has no horizon (551a) and is hence not blocked by an intervening *che*. The extraction option in (548) then follows from (551a) and the local ordering in (549) is the result of (551b).

- (551) *Italian topicalization:*
- a. $[u\text{Top}]_{\text{Topic}^0} \dashv\!\! \dashv \emptyset$
 - b. $\dots > C_{che} > \dots > \text{Topic} > \dots$

Crucial to this account is the partial decoupling of location and horizons.

The horizon account allows us express the difference between the facts surrounding *se* and *che* in Italian. The critical entailment that the horizon account gives rise to is that *if* an operation is blocked across a head in a lower clause, *then* this operation cannot land to the left of this head in the same clause. In the case of *se* in (545), focus movement across *se* is blocked ($[u\text{Foc}]_{\text{Focus}^0} \dashv\!\! \dashv C_{se}$). This entails that focus movement must land below *se* in the same clause (see (545)). In the case of *che*, on the other hand, topicalization over *che* is *not* blocked ($[u\text{Top}]_{\text{Topic}^0} \dashv\!\! \dashv \emptyset$), and therefore the position of the two heads relative to each other is not forced by horizons but up to a cartographic statement (see (551)). In both cases, then, the overall entailment holds: The presence of a horizon has implications for positions, whereas the absence of a horizon does not. The key feature of the horizons account is thus that they establish a link between height and locality, thus deriving the *se* facts, but this link is not one-to-one, thus allowing us to handle the *che* facts. Any account that assumes a strict correlation between height and locality sacrifices this flexibility.

5.9 Chapter summary

This chapter has placed the horizons account developed in the preceding chapters into the broader context of other approaches to selective opacity. The main goal of this chapter was to highlight in which respects horizons are similar or distinct from the various accounts proposed in the literature and to show how various empirical patterns unearthed in this literature receive an explanation using horizons. I have placed particular emphasis on distinctive features of the horizons account, such as its domain-based nature.

The overview of previous approaches has made it clear that the horizons account represents a continuation of long-standing research efforts. The original ban on improper movement was a very specific constraint that made particular reference to the movement path of an element with respect to the A/\bar{A} -distinction. Subsequent work has made it clear that it is too narrow and that selective

opacity effects extend beyond the A/ \bar{A} -distinction and are not, in fact, limited to movement steps of a single element. The horizons account encompasses both observation and adds a third layer of generalization on top of them: Selective opacity is not limited to movement, but due to a more abstract one on Agree. The original ban on movement from an \bar{A} - to an A-position emerges as the tip of an iceberg.

In other respects, horizons represent a substantial break with previous accounts. Unlike virtually all accounts (except for Williams'), the horizons account makes no reference to the moving element and its properties at all. That is, rather than stating that an element located in an \bar{A} -position cannot undergo A-movement, the approach is solely probe-based: An A-probe cannot search into a CP. The to-be-moved element is entirely irrelevant. A second important contrast between horizons and many previous approaches is that there is no reference to movement type interactions. That is, rather than stating that elements in \bar{A} -positions cannot move to an A-position, the horizon constraint is based on the CP layer that an \bar{A} -position entails. Movement type interactions are purely epiphenomenal.

Finally, I have discussed previous accounts of the Height–Locality Connection, focusing on Williams (2003, 2011, 2013) and Abels (2012a). The most appealing feature of these two alternative approaches is that they attempt to explain the Height–Locality Connection by reducing one aspect of it to the other. As we have seen, this entails that there should never be any mismatch between the two. However, this expectation is not borne out. What is required, then, is an approach that grants some potential for mismatch between height and locality and yet at the same time imposes a limit on the mismatch possibilities to capture the connection between the two. The account developed here does precisely that. The greater variability that it thereby affords allows it to be extended to cases where height and locality part ways.

CHAPTER 6

THE ROLE AND DISTRIBUTION OF PHASES

6.1 Introduction

The previous chapters have motivated, proposed and developed an account of selective opacity that is based on the concept of horizons, probe-specific limits on the search space visible to a probe. In this chapter, I will explore the relationship between horizons and the more traditional concept of phases. I will argue that horizons and CP phases coexist as independent constraints on syntactic locality. Syntactic dependencies must obey both in order to be well-formed. Crucially, horizons and phases have clearly distinct effects: First, while horizons are probe-specific, phases involve cyclic Spell-Out of syntactic structure and are hence identical for all operations and probes. Second, horizons render a given domain entirely opaque to a probe, including its edge, whereas the edge of a phase remains visible to higher operations. In a nutshell, horizons and phases have distinct empirical signatures and thus do not give rise to analytical redundancy.

While horizons and CP phases are consistent with each other and both empirically necessary, clause-internal phases can be shown to be incompatible with the horizon account proposed here. It is standardly assumed in much work following Chomsky (2000, 2001) that not only CP is a phase, but that vP is as well. While the horizon accounts developed in the preceding chapters are compatible with CP phases, they become untenable if vP is a phase as well. Interestingly, this incompatibility is not particular to horizons, but a systemic property of virtually all accounts of selective opacity. Arguably, then, selective opacity provides theory-internal evidence against vP phases. Against this

backdrop, this chapter will then provide independent evidence from other domains that dovetail with the conclusion that vP is not a phase. The evidence includes the locality of ϕ -agreement and *wh*-licensing as well as the real-time processing of movement dependencies (the latter is the subject of chapter 7), all converging on the view that vP is fundamentally different from CP in its locality in a way that is straightforwardly explained if CP is a phase, but vP is not. This conclusion has important consequences for our understanding of phase locality. Lastly, I will re-assess some previous arguments in favor of vP phases and argue that they are ultimately unconvincing, as they are too weak for empirical or theory-internal reasons to constitute compelling evidence of vP phases.

This chapter is structured as follows: In section 6.2, I show that there is evidence for CP phases as constraints on syntactic locality in addition to horizons. I will also demonstrate that a system that incorporates horizons and phases does not give rise to redundancy because horizons and phases have distinct empirical effects, making it possible to reliably distinguish the effects of one from the effects of the other. In section 6.3, I will then examine the status of vP phases in the horizons system developed here. The conclusion of this section will be that, in striking contrast to CPs, the system is incompatible with vP phases. I will then show that this incompatibility is not specific to horizons, but arises under virtually any account of selective opacity (section 6.3.1). I will then provide independent evidence from the locality of ϕ -agreement and *wh*-licensing that vP s are not phases (section 6.3.2). The conclusion of section 6.3 thus stands in clear contrast to the standard view in the literature, according to which vP does constitute a phase. It also raises the question of how to interpret previous evidence in support of vP phases. To address this question, section 6.4 reassesses some influential arguments to this effect (reconstruction, successive-cyclic movement in Dinka and *meN*-deletion in Indonesian) and demonstrates that, despite initial appearance, they are in fact inconclusive with respect to the phase status of vP , for very general reasons. As a result, they are fully compatible with the view that vP is not a phase.

6.2 The necessity of CP phases

6.2.1 The successive cyclicity of movement

One of the central discoveries in the syntax of movement has been that unbounded movement dependencies are created successive-cyclically through the edges of a clause (Chomsky 1973, 1977, 1981). Thus, there is good evidence that the crossclausal movement in (552) is not created in one fell swoop, but rather mediated by an intermediate trace in SpecCP of the lower clause.

- (552) a. Who_i does John think that Mary met *t_i*?
b. [CP Who_i does [TP John think [CP *t_i* that [TP Mary met *t_i*]]]]

There is a wealth of evidence from a variety of languages for such intermediate landing sites at clause edges. One illustrative type of evidence has already been presented in section 4.5 of chapter 4. We saw there that in German it is possible to embed a V2 clause under certain verbs. In the absence of extraction, the clause-initial position has to be overtly filled, as in (553a). A verb-initial clause as in (553b) is impossible:

- (553) a. Er hat gesagt [die Maria hat den Paul getroffen]. [German]
he has said the Maria.NOM has the Paul.ACC met
'He said that Maria met Paul.'
b. *Er hat gesagt [hat die Maria den Paul getroffen].
he has said has the Maria.NOM the Paul.ACC met
Intended: 'He said Maria met Paul.'

If an element is moved out of the embedded clause into the matrix clause, the embedded clause descriptively has to have the finite verb as its first element. Overt V2, i.e., having an overt constituent preceding the finite verb, is impossible. This fact follows straightforwardly if one fell swoop extraction out the lower clause is impossible and if the moving element has to instead pass through the edge of V2 clause (SpecForceP on the account in chapter 4), which is indeed the standard account since at least Thiersch (1978). This intermediate touchdown will fill the preverbal position with a trace, which goes unpronounced. The verb-initial order of (554a) then follows.

- (554) a. Wen_i hat er gesagt [t_i hat die Maria t_i getroffen]? [German]
 whom has he said has the Maria met
 ‘Who did he say that Maria met?’
- b. *Wen_i hat er gesagt [die Maria hat t_i getroffen]?
 whom has he said the Maria has met
Intended: ‘Who did he say that Maria met?’

The fact that (554b) is ungrammatical demonstrates that an intermediate touchdown at the clause edge is not merely an option, but in fact obligatory, a standard conclusion in studies on the syntax of German (Thiersch 1978, *et seq.*).

Striking additional support for successive cyclicity comes from languages that exhibit morphological reflexes of movement. A well-known example is Irish, discussed in great detail by McCloskey (1979, 2002). In Irish the form of a complementizer depends on whether or not movement over it has taken place. In the absence of movement the complementizer *go* is used, as in (555). If crossed by extraction, on the other hand, the complementizer surfaces as *aL*. Illustrative examples are given in (556), where (556a) shows *wh*-movement and (556b) illustrates relative clause movement.¹

- (555) Creidim gu-r inis sé bréag. [Irish]
 I-believe go-PST tell he lie
 ‘I believe that he told a lie.’ (McCloskey 2002: 185)

- (556) a. Céacu ceann a dhíol tú? [Irish]
 which one aL sold you
 ‘Which one did you sell?’
- b. an fhilíocht a chum sí ____
 the poetry aL composed she
 ‘the poetry that she composed’ (McCloskey 2002: 186, 189)

McCloskey (2002) argues that the form of the complementizer is sensitive to whether its specifier is filled by movement. If it is, as in (556), the complementizer *aL* has to be chosen. Crucially, *all* complementizers along the movement path undergo this alternation:

¹ Following the previous literature, McCloskey (2002) assumes that it is a silent relative pronoun that moves in (556b). In addition to the two complementizers discussed here, there is a third one that tracks whether resumption over it has taken place. See McCloskey (2002) and references cited there for elaboration.

- (557) an t-ainm a hinnseadh dúinn a bhí ___ ar an áit [Irish]
 the name aL was-told to-us aL was on the place
 ‘the name that we were told was on the place’ (McCloskey 2002:185)

If, as McCloskey (2002) argues, the form of the complementizer is established based on whether its specifier is filled, the fact that both complementizers in (557) are *aL* entails that movement must target the lower CP in a first step and then move into the higher CP in a second one.

A third illustrative type of evidence in favor of successive-cyclic movement through the edge of finite clauses are so-called *wh*-copying constructions. In these constructions a single *wh*-element can be multiply realized – in the SpecCP it takes scopes from as well as in lower Spec,CP along the movement path. (558) provides an example from the Algonquian language Passamaquoddy (Bruening 2006).

- (558) *Wh-copying in Passamaquoddy*
- a. [CP *Wen-il* Mali wewitaham-a-c-il [CP *wen-il*
 who-OBV Mary (3)-remember.ANO-DIR-3CONJ-PRTAGR.OBV who-OBV
 kisi-niskam-uk]]?
 PFV-dance.with.ANO-1CONJ
 ‘Who does Mary remember I danced with?’
- b. [CP *Tayuwe* kt-itom-ups [CP *tayuwe* apc k-tol-i malsanikuwam-ok]]?
 when 2-say-DUB when again 2-there-go store-LOC
 ‘When did you say you’re going to go to the store?’ (Bruening 2006:26,36)

Semantically and syntactically, there is only a single *wh*-element in (558a,b), which is, however, morphologically realized twice. In line with previous literature on related phenomena in other languages, Bruening (2006) argues that the two occurrences of the *wh*-element are distinct realizations of the same syntactic element. In other words, the lower occurrence of the *wh*-element in (558) is the realization of a trace (or copy) inside the lower SpecCP. This in turn constitutes evidence that the extraction applies successive-cyclically, passing through the lower SpecCP.

The fourth type of motivation comes from *wh*-islands. The fact that *wh*-movement out of a *wh*-island is impossible has traditionally been accounted for in terms of successive-cyclic movement. *Who* in (559) has to pass through the lower SpecCP, a position that is occupied by *why* and hence unavailable.

- (559) a. *Who_i did Susan ask why Sam was waiting for *t_i*?
 b. [CP Who_i did [TP Susan ask [CP why C⁰ [TP Sam was waiting for *t_i*]]]]

An impressive body of additional evidence in favor of successive cyclicity, of which the above examples are representative, has been accumulated over the years. *Wh*-copying similar to Passamaquoddy is found in German (Müller 1999a), Romani (McDaniel 1989), Frisian and Afrikaans (Felser 2004), Hungarian (Horvath 1997), and Ancash Quechua (Cole 1982). Other types of phenomena include reconstruction to intermediate trace positions (Barss 1986, Lebeaux 1988, Fox 1999), *wh*-agreement in Chamorro (Dukes 1992, Chung 1994, 1998, Kaplan 2005), quantifier floating in West Ulster English (McCloskey 2000), verb movement in Belfast English (Henry 1995), French (Kayne & Pollock 1978) and Iberian Spanish (Torrego 1984) and tonal downstep in Kikuyu (Clements & Ford 1979, Clements 1994).²

In all of these cases, long extraction out of a finite clause is mediated by an intermediate landing site at the edge of this clause. Traditionally, successive-cyclic extraction was derived from **SUBJACENCY** (Chomsky 1973, 1977, 1981, 1986). More recently, successive cyclicity has been accounted for as an effect of **PHASES** Chomsky (2000, 2001, 2004, 2008), according to which built syntactic structures are periodically shunted and removed from the workspace. Only the edge of a phase head, i.e., its specifier, remains accessible to outside operations. This is formulated in the **PHASE IMPENETRABILITY CONDITION** (PIC) in (560), where the ‘domain’ of a phase head corresponds to its complement, and its ‘edge’ to its specifier(s).

(560) **Phase Impenetrability Condition** (Chomsky 2000: 108)

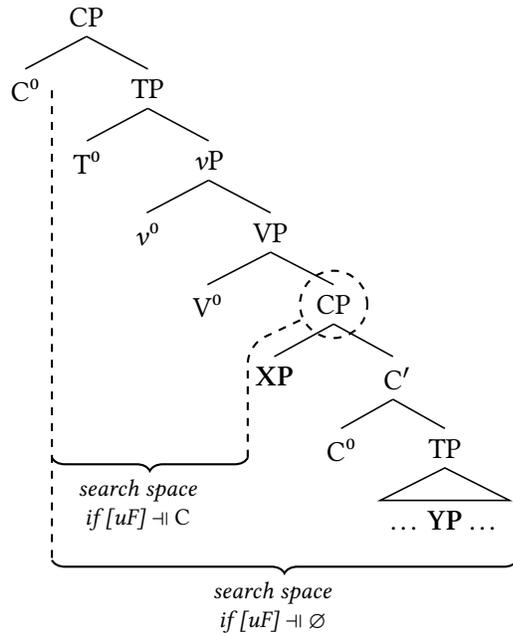
In phase α with head H, the domain of H is not accessible to operations outside of α , only H and its edge are accessible to such operations.

² The general notion that displacement dependencies are strictly local is also remarkable in that it has been widely adopted in a range of otherwise quite different syntactic frameworks, sometimes in much more radical forms such as in Generalized Phrase Structure Grammar (Gazdar, Klein, Pullum & Sag 1985) and Head-Driven Phrase Structure Grammar (Pollard & Sag 1994). The treatment of displacement in Generalized Phrase Structure Grammar and Head-Driven Phrase Structure Grammar makes use of a radical version of successive cyclicity, according to which information about the existence of a gap within a constituents is transmitted through every node between the gap and its filler. The process accomplishing this is so-called *slash feature percolation*. Roughly put, the information that a constituent contains a gap is indicated on that constituent by a slash feature. This slash feature is percolated up from a daughter node to its respective mother until it reaches the sister node of the filler, where it is deleted.

Successive cyclicity through the edge of finite clauses then follows from the PIC if C^0 is a phase head: In order to remain accessible to material in a higher clause, an element embedded inside a CP has to move to SpecCP of that embedded clause. There is a considerable debate in the literature about what the phase heads are. In addition to C^0 , v^0 is also commonly treated as one, following Chomsky (2000, 2001). As my main interest in this section is to establish the need for CP phases, I will defer a discussion of vP phases until section 6.3.

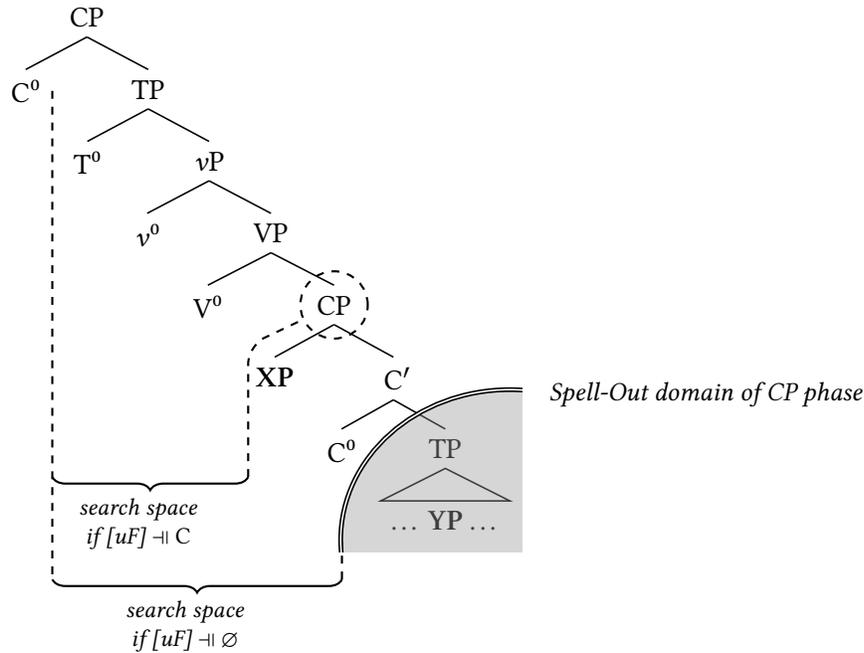
Successive-cyclic movement does not follow from horizons and thus motivates the existence of phases even in a system that adopts horizons. As already discussed in section 4.5 of chapter 4, horizons render a given domain entirely opaque to a probe. This crucially includes the edge of this domain. Consider, for example, a probe $[uF]$ for which $[uF] \dashv\vdash C$. Search by this probe will terminate as soon as it encounters a CP node. Consequently, everything dominated by this CP node will be inaccessible to $[uF]$, including an element in SpecCP. On the other hand, under $[uF] \dashv\vdash \emptyset$ the entire embedded CP clause will be accessible as far as horizons are concerned because no element along $[uF]$'s search path will compel it to terminate. Other settings, such as $[uF] \dashv\vdash T$, will either be equivalent to $[uF] \dashv\vdash C$ or yield a vacuous probe, depending on where $[uF]$ is located (see the discussion in section 3.5 of chapter 3). This state of affairs is depicted in (561), where an element XP is located in SpecCP of the lower clause and another element YP is more deeply embedded. If $[uF] \dashv\vdash C$, then $[uF]$'s search will terminate before it encounters either XP or YP and no Agree between them and $[uF]$ can be established. If $[uF] \dashv\vdash \emptyset$, on the other hand, then $[uF]$'s search space comprises both XP and YP. There is no possible horizon setting that would render XP accessible to $[uF]$ without also rendering YP accessible.

(561) Search spaces for $[uF]$ on two horizon settings



To incorporate successive cyclicity into the system, XP in SpecCP must be accessible to $[uF]$, but YP must not be. Phases accomplish this result in the traditional fashion. If C⁰ is a phase, then the TP inside the lower clause will undergo Spell-Out upon completion of the embedded CP. YP will therefore be rendered inaccessible to all material in the higher clause, including $[uF]$. XP remains accessible, being at the edge of CP.

(562) *Combined effects of horizons and CP phases*



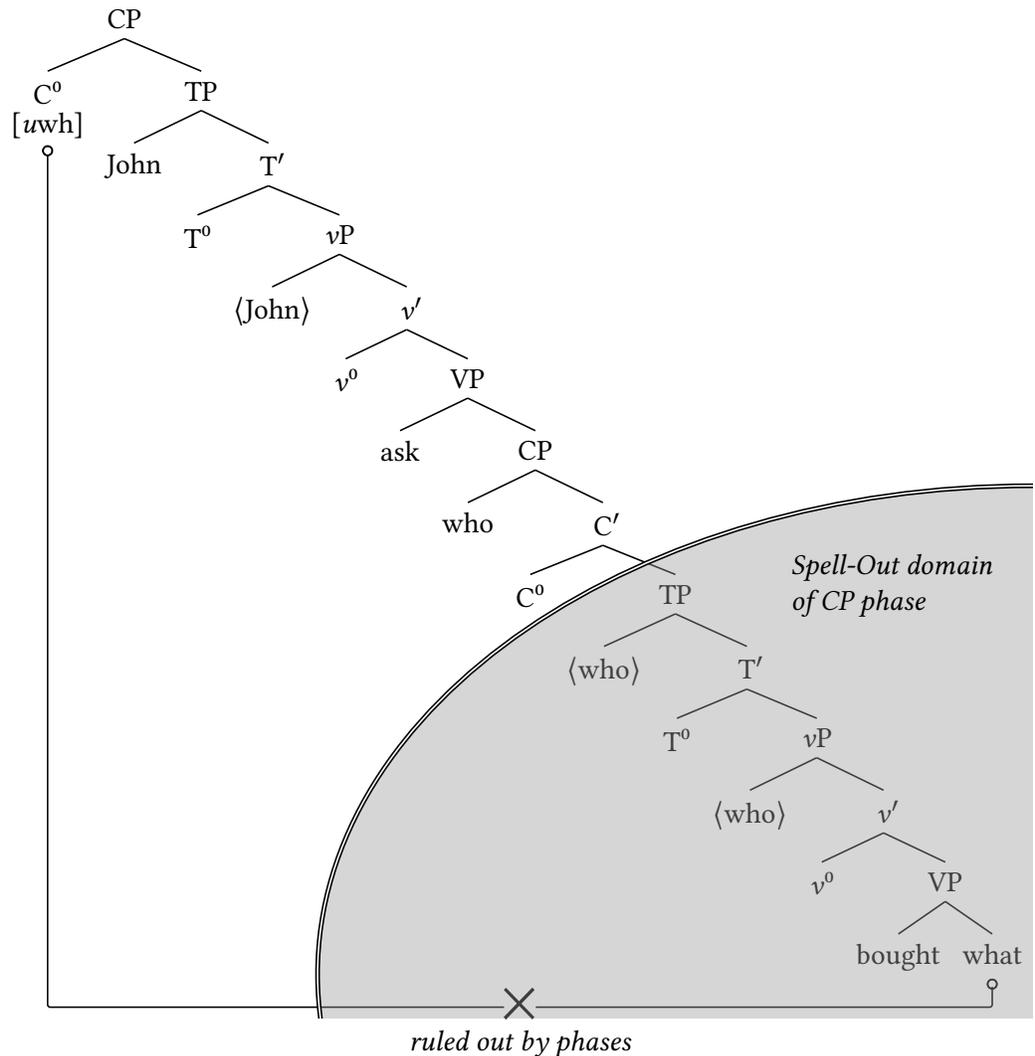
The combined effect of horizons and phases in (562) is that the edge of embedded clause is visible to $[uF]$ if $[uF] \neq \emptyset$. If $[uF] \neq C$, on the other hand, $[uF]$'s search terminates at the CP node, because this node bears a C feature. Therefore, not even the edge of the lower clause is accessible in this case, as the search terminates before it reaches the embedded SpecCP.

Horizons and phases have distinct empirical effects and signatures and combining the two does not give rise to analytical redundancy. As just discussed, horizons do not induce successive-cyclic movement, but phases do. Conversely, phases enforce movement through the phase edge, but because this edge remains accessible, nothing in the concept of phases as such would ever render a domain opaque to extraction, as noted by Boeckx & Grohmann (2007) and Abels (2012b), among many others, including Nevins (2005: 294), who attributes to Bob Franks the observation that the Phase Impenetrability Condition contains an implicit Phase *Penetrability* Condition. In other words, phases in and of themselves do not offer an account of why extraction is possible out of some domains, but impossible out of others, precisely because they are porous by design. With only phases, all extraction out of all domains would be well-formed as long as it proceeds successive-cyclically. Clearly, the concept of a phase needs to be supplemented with a theory of opacity. Horizons provide one such option.

The illustrate the division of labor between horizons and phases, consider extraction out of finite clauses in English. Descriptively, there are two constraints. First, \bar{A} -extraction is possible, but must proceed successive-cyclically. This is ensured by phases. Second, A-extraction is categorically ruled out, regardless of whether it applies in one fell swoop or successive-cyclically. This in turn is the effect of horizons. Thus, a violation of successive cyclicity as in (563) is ruled out by phases, as only SpecCP of the lower clause is visible to the *wh*-movement probe [*uwh*].

(563) *One-fell-swoop movement out of CP* ([*uwh*] $\neq \emptyset$)

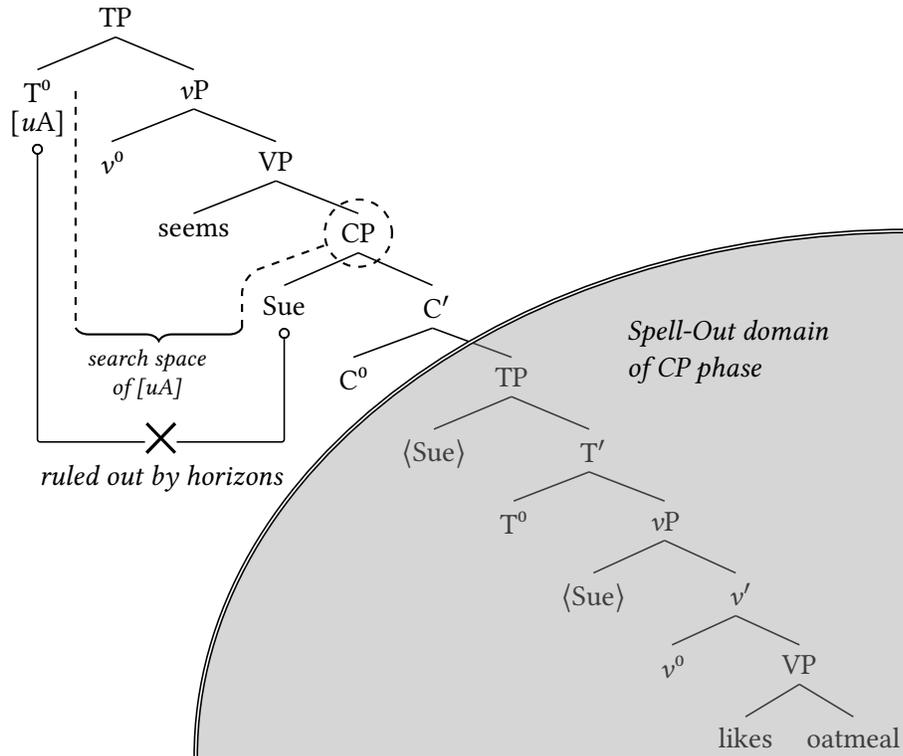
*What_{*i*} did John ask who bought *t_i*?



A-movement out of a CP clause is ruled out by horizons, because the A-movement/EPP probe [*uA*] has C as its horizon, it is unable to access an element in SpecCP of the lower clause. (564) hence

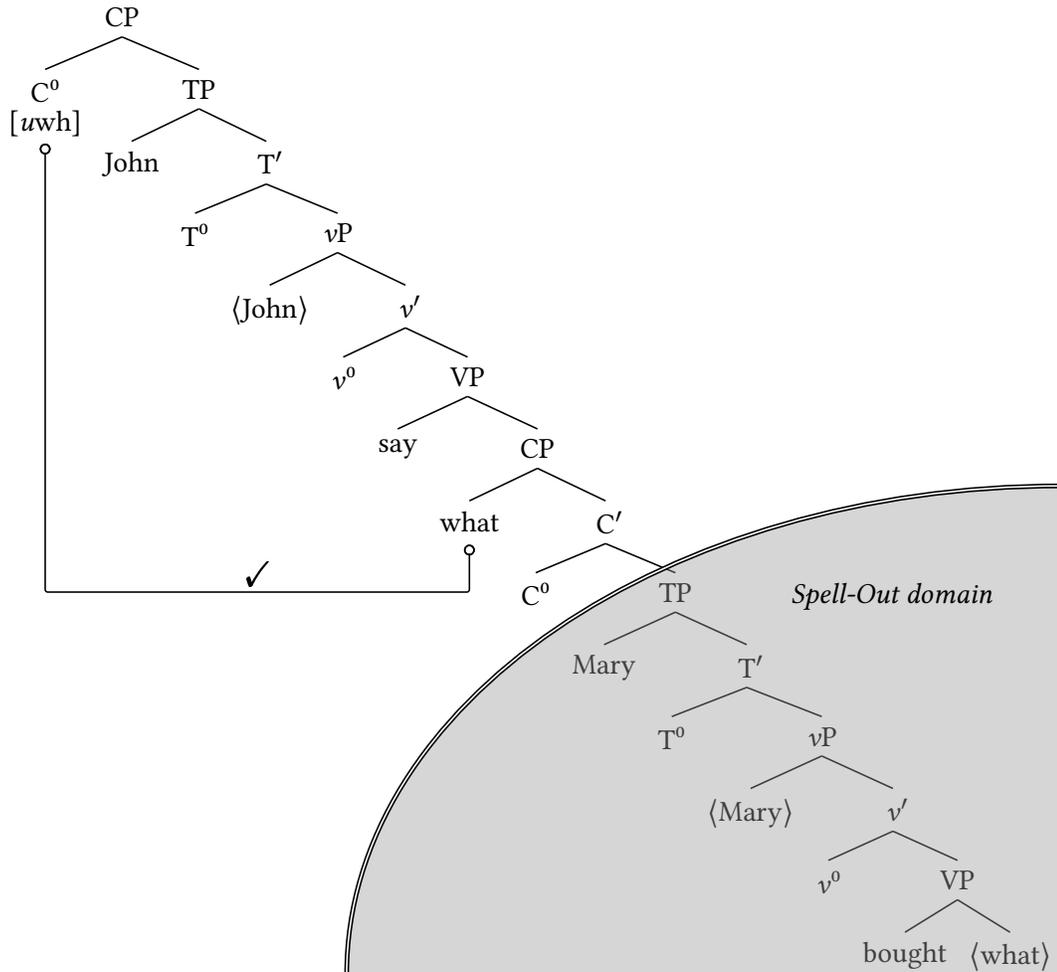
constitutes an example that would be allowed by phases but is blocked by horizons. Because Agree between $[uA]$ and *Sue* is thus impossible, superraising is ruled out.

- (564) *A-movement out of CPs* ($[uA] \rightarrow C$)
 *Sue seems (that) likes oatmeal.



Only dependencies that conform to both horizons and phases are well-formed. For the example at hand, successive-cyclic \bar{A} -movement out of a finite clause is thus allowed, as (565) shows:

- (565) *Successive-cyclic \bar{A} -movement* ([*uwh*] $\neq \emptyset$)
 What did John say that Mary bought?



The empirical effects of horizons and phases are thus clearly distinct. Horizons determine whether a given probe is able to search into the embedded clause at all. Phases determine that only their edge is accessible even to probes which can search into a clause. For movement, then, horizons determine whether a given movement type out of a clause is at all possible or not, while phases condition the successive-cyclic path of extractions allowed by horizons. Both principles are empirically necessary and do not give rise to analytical redundancy. Complex conditions on extractions follow from the conjunction of the two principles.

A second important difference between horizons and phases already mentioned in section 3.2.1 of chapter 3 concerns the possible differentiation between different operations. Because phases involve cyclic Spell-Out of syntactic structure, they do not discriminate between operations. Parts of a structure that have undergone Spell-Out are removed from the workspace and are hence

invisible to *all* operations. Phase locality is thus *absolute* in that it is identical for all operations, for very general reasons. Horizons, on the other hand, are probe-specific and hence *selective*. Their locality effect thus differs between different operations and movement types.

The two differences just discussed give rise to an interesting prediction. If horizons are responsible for complete opacity of a domain and operation-specific, whereas phases give rise to an edge effect and are not operation-specific, we predict that opacity can be selective, but the edge effect cannot be. In a situation with a selective edge effect, one movement type would have to leave a domain in a successive-cyclic fashion, while another movement type could leave the same domain in one fell swoop. This appears to be unattested. If a domain requires extraction to be successive-cyclic through its edge, it does so for all movement types. There is thus a fundamental contrast between selective opacity effects and successive cyclicity: The former is operation-specific, as documented at length in this dissertation, but successive cyclicity never is. On the view taken here, this asymmetry follows as a direct consequence of the coexistence of two very different locality principle – horizons and phases.

6.2.2 Evidence from selective opacity for CP phases:

Revisiting German V2 clauses

The preceding section has demonstrated that successive cyclicity provides evidence for the existence of phases in addition to horizons. Additional evidence for this view comes from selective opacity effects. Recall from the discussion in sections 4.2.2 and 4.2.3 of chapter 4 that in German V-final clauses are transparent for extraction into a higher V2 or V-final clause, whereas V2 clauses allow extraction into a higher V2 clause but not extraction into a higher V-final clause (Haider 1984, Reis 1985, 1996, Sternefeld 1989, Staudacher 1990, Müller & Sternefeld 1993, Müller 1995, 2010a).

(566) In German, V-final clauses are transparent to *wh*-movement into V-final and V2 clauses. V2 clauses only allow *wh*-movement into a V2 clause, not into a V-final clause.

(567) provides an example of movement out of a V-final clause, (568) gives the analogous counterparts for movement out of a V2 clause.

(567) *Movement out of V-final clause*

a. *Into V2 clause*

Wen_i glaubt er [dass die Maria t_i getroffen hat]?
 whom believes he that the Maria met has

b. *Into V-final clause*

Ich weiß nicht [wen_i er glaubt [dass die Maria t_i getroffen hat]].
 I know not whom he believes that the Maria met has

(568) *Movement out of V2 clause*

a. *Into V2 clause*

Wen_i glaubt er [t_i hat die Maria t_i getroffen]?
 whom believes he has the Maria met

b. *Into V-final clause*

*Ich weiß nicht [wen_i er glaubt [t_i hat die Maria t_i getroffen]]
 I know not whom he believes has the Maria met

The account of this curious fact proposed in chapter 4 is crucially based on horizons. *Wh*-movement into a V-final clause is triggered by a probe on C⁰ and *wh*-movement into a V2 clause by a probe on Force⁰ (as an instance of topicalization). Moreover, V2 clauses, being embedded root clauses, are structurally larger than V-final clauses. Concretely, I have taken V2 clauses to be ForceP and V-final clauses to be CPs. Against this structural background, I have proposed the horizons in (569):

(569) a. *Wh-movement (in V-final clause):* [=(439)]

[*uwh*]_{C⁰} \neg Force

b. *(Wh)-Topicalization:*

[*utop_(wh)*]_{Force⁰} \neg ∅

The ungrammaticality of (568b) then follows as a violation of horizons. To produce (568b), [*uwh*] on head C⁰ would have to probe into an embedded ForceP clause. This is impossible given that Force is [*uwh*]'s horizon. Movement out of a V2 clause into a V-final clause is hence ruled out. See chapter 4 for extensive discussion.

Against this background, an interesting problem emerges. Consider a structure in which a V2 clause is embedded inside a V-final clause, which is itself embedded inside a V2 clause. In this

structure, movement out of the innermost V2 clause into the highest V2 clause is ungrammatical, as Sternefeld (1989, 1992) and Müller (2010a) have observed. (570) illustrates:

- (570) *V2-V-final-V2*
 * $[_{V2}$ Wen_i hat Fritz gesagt $[_{V-final}$ dass er glaubt $[_{V2}$ hat die Maria t_i
 whom has Fritz said that he believes has the Maria
 getroffen]]]?
 met
Intended: ‘Who did Fritz say that he believes that Maria met?’

Reversing the order of the lower V2 and V-final clauses produces an acceptable sentence:

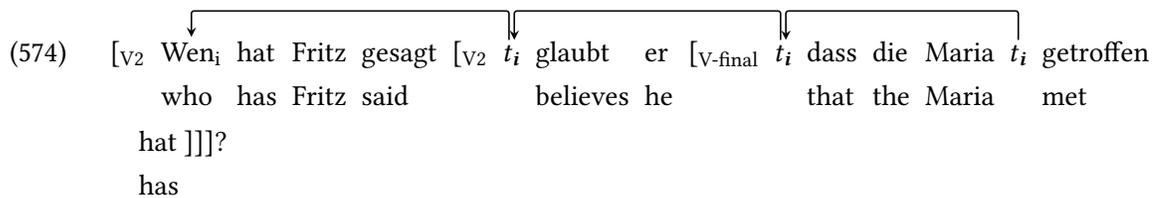
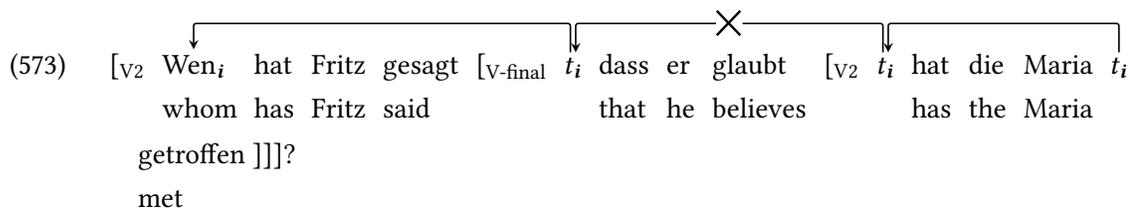
- (571) *V2-V2-V-final*
 $[_{V2}$ Wen_i hat Fritz gesagt $[_{V2}$ t_i glaubt er $[_{V-final}$ dass die Maria getroffen
 whom has Fritz said believes he that the Maria met
 hat]]]?
 has
 ‘Who did Fritz say that he believes that Maria met?’

This asymmetry is due to movement, as the control structures in (572a,b) illustrate:

- (572) a. *V2-V-final-V2*
 $[_{V2}$ Fritz hat gesagt $[_{V-final}$ dass er glaubt $[_{V2}$ die Maria hat den Hans
 Fritz has said that he believes the Maria has the Hans
 getroffen]]].
 met
- b. *V2-V2-V-final*
 $[_{V2}$ Fritz hat gesagt $[_{V2}$ er glaubt $[_{V-final}$ dass die Maria den Hans getroffen
 Fritz has said he believes that the Maria the Hans met
 hat]]].
 has

It is easy to see that the ungrammaticality of (570) does not follow from the horizons in (569) alone. We saw that *wh*-movement into a V2 clause is possible out of both V-final and V2 clauses. This follows from (569) because $[_{utop}_{(wh)}]$ has no horizon and is hence able to search into both types of clauses (modulo the effect of phases). As far as horizons are concerned, then, there is nothing wrong with the movement in (570).

What, then, causes the ungrammaticality of (570)? The contrast to (571) makes it clear that the underlying factor is that the extraction in (570) has to cross a V-final clause before proceeding into the matrix clause. Because we know independently that movement out of a V2 clause into a V-final clause is impossible (see (568b)), the ungrammaticality of (570) follows if extraction from the lowermost V2 clause into the matrix V2 clause *has* to proceed successive-cyclically through the intermediate V-final clause. This intermediate movement step would require movement from a V2 clause into a V-final clause, and the ungrammaticality of (570) follows. In (571), by contrast, movement through the edge of every clause does not violate a horizon at any point of the derivation.



The horizons account thus derives the ungrammaticality of (570) *if* movement through a finite clause is successive-cyclic. CP phases deliver this result. As a result, there is no derivation of (570) that satisfies both phases and horizons.

Selective opacity effects themselves thus provide additional evidence for CP phases as independent constraints on syntactic locality in addition to horizons. An interesting consequence of this account is that intermediate movement must be subject to horizons, in order to rule out (573). Because horizons are properties of probes, intermediate movement must itself be triggered by an agreeing probe, just like terminal movement. This conclusion is in line with works like Abels (2012b), Georgi (2014), van Urk (2015) and references cited there, to name just a few recent examples that arrive at the same conclusion.³

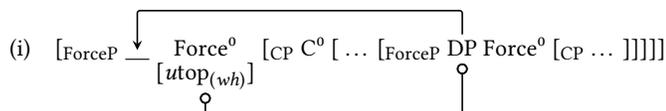
³ An interesting emerging questions for extraction from a V2 clause into a V2 clause. Is C⁰ is always an invariably a phase head, direct movement from SpecForceP to SpecForceP should be impossible. Instead, movement would have to proceed through SpecCP of the higher clause. This presents an interesting puzzle in light of the fact that movement from SpecForceP to SpecCP is impossible, as the discussion in the main text has made clear.

In sum, in addition to the traditional evidence for successive-cyclic movement, selective opacity effects also provide converging evidence for successive cyclicity. In both cases, the data follow if horizons are supplemented with phases. Because the overarching patterns emerge from the interplay of the two principles, it is clear that neither of them is dispensable.

6.2.3 Evidence from ϕ -agreement for CP phases

A third type of evidence for CP phases in addition to horizons comes from long-distance agreement (LDA). Some languages allow ϕ -agreement between a matrix verb and a DP at the edge of an embedded finite clause. Examples are Tsez (Polinsky & Potsdam 2001), Innu-aimûn (Branigan & MacKenzie 2002), and Passamaquoddy (Bruening 2001). I will illustrate this point using evidence from Tsez here. Polinsky & Potsdam (2001) and Polinsky (2003) argue that a DP inside a finite

There are at least two solutions to this problem. One is to conclude with Bošković (2014) and the references cited there that being phasal is not an inherent property of a head. Rather, Bošković (2014) proposes that it is the highest projection within an extended projection that constitutes a phase head. On this assumption, CP would be a phase head in V-final clauses, but it would not be a phase in V2 clauses, as these are ForcePs. Direct movement from SpecForceP to SpecForceP is then allowed, as in (i), while maintaining the account of (573).



The alternative solution involves the timing of phasal Spell-Out, a contended issue quite generally (compare Chomsky 2000 and Chomsky 2001). Suppose that a phasal complement is not immediately spelled-out upon merger of the next higher head, but potentially later. The domain of the phase will then remain accessible to outside material until Spell-Out applies. One option is to delay phasal Spell-Out until the next higher *phase* head is merged (see Chomsky 2001). An alternative is to assume that the CP phase domain is spelled-out upon completion of the entire left periphery, i.e., upon completion of the local clause or, equivalently, the extended projection. This latter option arguably arises as the default view if C is viewed as an abbreviation for an extended CP domain à la Rizzi (1997). Consider the following representative quotes:

“Ignored as well are the ‘peripheral’ systems outside TP; I will use C and T as surrogates for richer systems.” (Chomsky 2000: 143n31)

“The next question is: What are the phases? I will pursue the suggestion in Chomsky 2004 that they are CP and v*P, where C is shorthand for the region that Rizzi (1997) calls the ‘left periphery,’ possibly involving feature spread from fewer functional heads (maybe only one) ...” (Chomsky 2008: 143)

If C is merely an abbreviation for a whole sequence of projections, it is quite natural to assume that Spell-Out of that phasal domain applies once this sequence of projections is completed for that clause. On this view, then, the entire left periphery of a clause defines a phasal domain, whose complement TP undergoes Spell-Out upon completion of this domain, i.e., upon completion of the extended projection. In V-final clauses, by contrast, the extended projection lacks a ForceP. Phasal Spell-Out of the intermediate CP in (570) will hence apply as soon as this CP is completed, thereby successfully ruling out this derivation.

For the purposes of the account developed here, the choice between these two accounts is inconsequential and I will put the matter aside in what follows.

clause can optionally control agreement on the matrix verb out of a finite clause, as illustrated in (575), where *iyxo* ‘know’ either agrees with *magalu* ‘bread’ in gender class III or shows class IV default agreement.

- (575) enir [užā **magalu** bāc’ruli] r-/b-iyxo [Tsez]
 mother boy bread.III.ABS ate IV-/III-know
 ‘The mother knows the boy ate the bread.’ (Polinsky & Potsdam 2001: 584)

Polinsky & Potsdam (2001) provide a number of arguments that such LDA is possible only if the embedded DP is the topic of the embedded clause. To give just two of their arguments, if the embedded object is overtly marked as a topic, LDA is obligatory:

- (576) enir [už-ā **magalu-n** b-āc’-ru-li] *r-/b-iy-xo [Tsez]
 mother boy-ERG bread.III.ABS-TOP III-eat-PASTPART-NMLZ *IV-/III-know-PRES
 ‘The mother knows that the bread, the boy ate.’ (Polinsky & Potsdam 2001: 610)

Conversely, if the embedded object is overtly marked as focused, LDA is impossible:

- (577) eni-r [t’ek-kin y-igu yāl-ru-li] r-/*y-iy-xo [Tsez]
 mother-DAT book.II.ABS-FOC II-good be-PASTPART-NMLZ IV-/*II-know-PRES
 ‘The mother knows that the BOOK is good.’ (Polinsky & Potsdam 2001: 611)

Polinsky & Potsdam (2001) argue for an analysis of these and further facts according to which an embedded topic covertly moves into the specifier of an embedded Topic projection, thus placing it at the edge of the lower clause. This edge is accessible to the ϕ -probe in the matrix clause, establishing LDA. Importantly, it is only the edge of the lower clause that is visible to the ϕ -probe. If topic movement does not take place, no LDA can be established (see (577)).

Polinsky & Potsdam’s (2001) insights are important because they make it clear that edge accessibility effects are not confined to movement (in the form of successive cyclicity), but also arise in the domain of ϕ -agreement. As discussed at length above, edge effects do not follow from horizons but instead are the hallmark property of phases. The Tsez pattern thus provides converging evidence for CP phases.

Recall from the discussion above that phases differ from horizons in that their effects are identical for all operations. In other words, they do not give rise to selective opacity. If the edge

effect is indeed due to phases and not to horizons, we expect ϕ -agreement to *never* be able to proceed past the edge of an embedded finite clause. This appears to be the case, as Polinsky (2003) and Bobaljik (2008) have argued in crosslinguistic investigations into LDA:

(578) *Polinsky–Bobaljik Generalization*

There are no clear cases in the literature of agreement reaching deeper into a finite clause than to the primary topic of that clause, regardless of the overt position of that topic.

(Bobaljik 2008: 317)

The generalization in (578) again follows from CP phases. It is worth noting at this point that no such obligatory edge effect exists for nonfinite clauses. In chapter 2, I have provided evidence that LDA into nonfinite clauses in Hindi does *not* require movement of the agreement controller. Put differently, it is possible in Hindi for a ϕ -probe in the matrix clause to search all the way to the bottom of an embedded ν P clause. This provides at least suggestive evidence that there is a fundamental asymmetry between CPs and ν Ps. CPs induce an edge effect in all languages, ν Ps do not. This asymmetry is taken up again in greater depth in the subsequent sections of this chapter, where I will argue that CPs are phases, whereas ν Ps are not.

Finally, just as in the cases discussed above, an appeal to phases does not render superfluous or redundant the notion of horizons. As already mentioned in section 3.4.3 of chapter 3, not all operations are able to access the edge of an embedded clause in Tsez. Polinsky & Potsdam (2001) demonstrate that Tsez does not any allow crossclausal movement, as illustrated in (579).

(579) a. kid-bā [už-ā hibore-d bikori žāk'-ru-li] esis [Tsez]
 girl-ERG boy-ERG stick-INSTR snake hit-PASTPART-NMLZ said
 'The girl said that the boy had hit the snake with a stick.'

b. *bikori kidbā [užā hibored žāk'ruhi] esis
 snake girl boy stick hit said (Polinsky & Potsdam 2001: 590)

An element at the edge of a finite clause must thus be visible to a higher ϕ -probe, but not to a higher movement-inducing probe. Because phases simply state that all and only elements at the edge of a phase remain accessible, this difference can simply not be attributed to phases. Consequently, selective opacity cannot be the result of phases, for principled reasons. An account that invokes

both horizons and phases is able to handle these facts. The horizons in (580) deliver the desired contrast between the ϕ -agreement probe $[u\phi]$ and the movement probe $[u\mu]$.⁴

(580) *Tsez horizons*

- a. $[u\phi] \dashv\!\! \dashv \emptyset$
- b. $[u\mu] \dashv\!\! \dashv C$

Thus, ϕ -agreement is possible with an element in SpecCP, but that same element is accessible to a matrix movement probe, as it lies beyond $[u\mu]$'s horizon. As before, horizons and CP phases are both empirically supported and necessary.

6.2.4 Section summary

In this section, I have presented evidence in favor for adopting CP phases in addition to horizons. Evidence from a range of domains has made it clear that horizons do not capture edge effects, which follow from CP phases. Conversely, it is equally clear that phases do not offer an account of selective opacity effects, which follow from horizons. Both principles are thus empirically motivated. Because they have distinct empirical effects, no redundancy between the two arises: Horizons determine whether a domain is transparent or opaque for an operation, whereas phases give rise to edge effects, and hence mandate that extraction out of a domain that is allowed by horizons be successive-cyclic. Moreover, horizons by their very nature are operation-specific, whereas phase effects are stable across operations.

⁴ No locations for the two probes are given in (580), as this would require a broader range of evidence not considered here. Also, for the sake of concreteness, the ϕ -probe is not assigned a horizon in (580). On this view, its search space is only delimited by phases. There is evidence that this is not quite correct, however. Polinsky & Potsdam (2001) show that LDA into an embedded clause is impossible if that clause contains the complementizer λin :

(i) *eni-r [už-ā magalu b-ac'-si-λin] b-iyxo
 mother-DAT boy-ERG bread.III.ABS III-eat-PST.EVID-COMP III-knows
Intended: 'The mother knows that the boy ate bread.' (Polinsky & Potsdam 2001: 635)

(i) may be taken to suggest that $[u\phi]$ does have a horizon, albeit one relatively high in the left periphery, at least as high as the projection whose head λin is the overt realization of.

6.3 Against ν P phases

Having motivated the role of CP phases in addition to horizons, we now turn to the question of how the system relates to ν P phases. Traditionally, C^0 and ν^0 are both taken to be phase heads (Chomsky 2000, 2001, 2004, 2008), but it is clear that the notion of a phase in and of itself does not dictate the choice. The mechanics of phases and the PIC have implications for heads that are phasal, but it is an open question which heads constitute phase heads (see, e.g., Bošković 2002, Bošković 2014, Epstein & Seely 2002, Abels 2003, 2012b, Legate 2003, 2012, Lahne 2009, and Müller 2010b, 2011 for discussion). As discussed in this section, selective opacity as a phenomenon presents severe problems for any system that incorporates ν P phases. This conclusion is quite independent on the account of selective opacity invoked. It holds for horizons, but it also applies to virtually all previous accounts of selective opacity (with the notable exception of Müller 2014a,b). The conclusion that ν P phases are problematic will then be corroborated by evidence independent of selective opacity.

6.3.1 Selective opacity as an argument against ν P phases

The problem of ν P phases in the context of selective opacity in a nutshell is that an account that incorporates ν P phases is unable to rule out improper movement in a local way. Instead, it has to appeal to a global constraint that applies to the entire movement history of a moving element and that makes reference to information contained inside an already spelled-out ν P phase. Versions of this problem have been observed by Neeleman & van de Koot (2010) and Müller (2014a,b).

To illustrate the problem, consider first the view that all ν Ps are phases (Legate 2003) and its repercussions for superraising in English. Descriptively, movement from SpecCP to SpecTP is ungrammatical in English and any account of superraising has to derive this fact in one way or another.

(581) * $[_{TP} \text{ Sue}_i [_{\nu P} \text{ seems } [_{CP} t_i [_{TP} t_i \text{ likes oatmeal }]]]]$

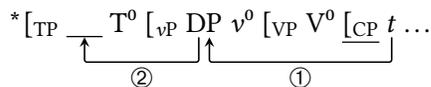
- (582) *Descriptive ban on superraising*
 Movement from SpecCP to SpecTP is impossible (in English).

The crucial property of (582) is that it is the *combination* of launching and landing site that is ruled out. Movement out of a SpecCP position is of course possible (successive-cyclic *wh*-movement), and so is movement to SpecTP (regular movement into subject position). It is only movement steps from SpecCP to SpecTP that must be ruled. In other words, information about the launching as well as the landing site have to be simultaneously available to correctly apply (582).

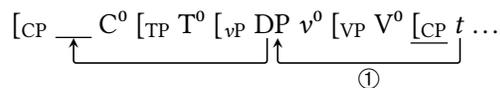
Without *vP* phases, it is straightforward to state a constraint that has (582) as its consequence in a variety of ways (see chapter 5). Consider now the status of (582) in a system with *vP* phases. In such a system, no movement proceeds directly from SpecCP to SpecTP because the intervening matrix *vP* enforces an intermediate touchdown in its specifier. Illegitimate superraising must therefore apply in two steps, first from SpecCP to Spec*vP* and then from Spec*vP* to SpecTP, as schematized in (583a). Importantly, each link of the movement chain is locally well-formed. Movement from SpecCP to Spec*vP* (①) must be allowed for \bar{A} -extraction out of a CP clause, as in (583b). And movement from Spec*vP* to SpecTP (②) must be allowed to allow A-movement out of a TP clause, as in (583c).

- (583) *Extraction paths under vP phases*

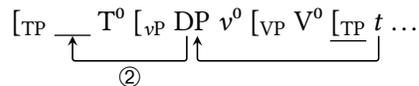
- a. *A-extraction out of finite clause*



- b. \bar{A} -extraction out of finite clause



- c. *A-extraction out of nonfinite clause*

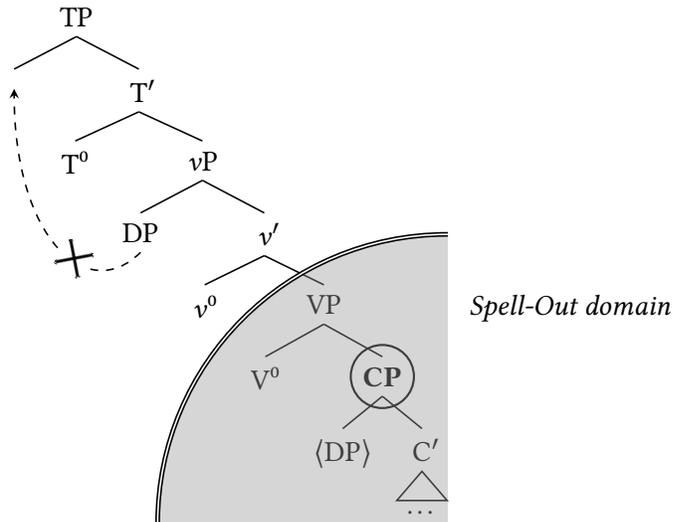


If the two movement steps that make up the derivation in (583a) are independently well-formed, it must be their combination that rules out (583a). Plainly, doing so requires a global constraint that does not just inspect individual movement steps, but an element's entire movement history. This requirement is of course accompanied by the significant increase in computational complexity that global constraints entail. This consequence is somewhat ironic, as the prime conceptual motivation

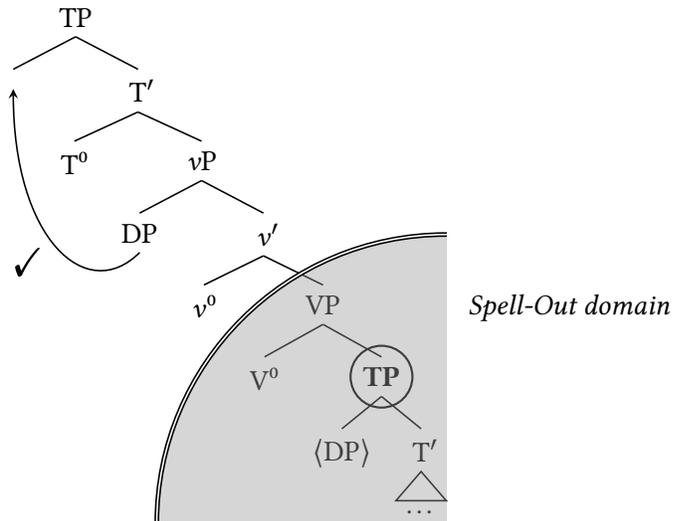
for phases is that they decrease computational complexity (Chomsky 2005). Yet at the same time *v*P phases necessitate constraints on movement that are substantially more complex than systems without *v*P phases.

Another way of conceptualizing the problem is the following. Because of (583b), we know that the initial movement step in (583a) (labeled ①) is well-formed. Subsequent movement to SpecTP (②) in must then be ruled out in (583a) but not in (583c). Evidently, the relevant distinction between the two is that the DP has been previously extracted out of a CP in (583a) but not in (583c). Crucially, this information is contained inside an already spelled-out *v*P phase in both cases and hence no longer available, as displayed in schematic form in (584) and (585).

(584) *Extraction out of finite CP clause → vP-to-TP movement impossible*



(585) *Extraction out of nonfinite TP clause → vP-to-TP movement possible*



One and the same movement step from SpecvP to SpecCP must be ruled out in (584), but ruled in in (585). The crucial difference about where the DP came from is contained in a part of the structure that has already undergone Spell-Out and is hence removed from the workspace. Thus, in addition to the added computational complexity of a global constraint on movement, it is not even clear how such a constraint could be formulated or evaluated on the standard view that phases induce cyclic removal of structure from the workspace.

I know of three strategies that have been proposed to avoid this problem. The first is to exempt vP from its phase status in precisely those environments in which the problem would arise. The second is to appeal to the case of the moving element to distinguish between (584) and (585). The third is to record the derivational history of the moving element and hence information about spelled-out phases. The first two solutions go back to Chomsky (2000, 2001), the third is due to Müller (2014a,b). I will review all three and show that they are insufficient. My own solution to the problem will be to dispense with vP phases altogether.

The first attempt to solve the locality problem is centered around the idea that not all v^0 heads are phasal. Chomsky (2000, 2001) proposes that only transitive v , or v^* , is a phase. v^* is a flavor of v that introduces an external argument. It contrasts with defective v , or v_{def} , which does not introduce an external argument and appears in passives and unaccusatives, in that v_{def} by hypothesis is not a phase head.⁵ Crucially, superraising is possible in English only if the matrix clause does not project an external argument itself (i.e., if it contains a raising predicate or a passivized ECM verb). Consequently, it is v_{def} that appears in the higher clause. Not being a phase head, no Spell-Out of its complement takes place. Moreover, no intermediate movement of the DP through its specifier is necessary. On this view that structure of (581) is not (583a), but rather (586).

(586) * [TP $\overbrace{\text{ } T^0 \text{ } [vP v_{\text{def}}^0 \text{ } [VP V^0 \text{ } [CP t \dots]]]}$]

This solution to the problem essentially strips v^0 of its phase status in precisely those environments in which the problem would emerge. There are a number of reasons to question this approach. First,

⁵ This characterization is somewhat of an oversimplification. Chomsky (2001) in fact treats all v 's as phase heads, with v^* being a 'strong' phase and v_{def} being a 'weak' phase. Only strong phases trigger Spell-Out of their complement. v_{def} crucially does not. In this sense, it mirrors the properties of non-phase heads. At least in the domain of interest here, the distinction between a weak phase and a non-phase head is terminological only.

it invokes what amounts to a conspiracy because it is simply a lexical stipulation that exempts *v* from its phase status in precisely the set of environments in which this phasal status would result in locality problems. This makes the solution conceptually unappealing.

Second, there is to my knowledge no independent evidence to support the split between a phasal *v** and a non-phasal *v*_{def}. As Legate (2003) argues, there does not appear to be a test for phasehood that discriminates between the two flavors of *v*. Legate (2003) concludes from this that *all v* heads are phases, but does not address the emerging problem for superraising. Moreover, den Dikken (2006) details that none of Legate’s (2003) arguments actually provides evidence for *v*’s phasehood, as all of her empirical points can be accounted for without invoking *v*P phases. Additional evidence against phasal distinctions between different *v* heads is provided by Richards (2004, 2007), further aggravating the conceptual problem just identified.

A third problem concerns the generality of the problem. The *v*_{def} solution is specifically tailored towards superraising in English. While I have illustrated the problem that emerges for *v*P phases using superraising, the problem is much more widespread and extends to virtually all instances of selective opacity. Because the *v*_{def}-based solution is geared towards superraising in English, it does not extend to the various other instances of the problem. I will illustrate this point with two examples.

The first example involves A- and \bar{A} -movement in Hindi. As laid out in detail in chapter 2, the two differ in their locality. In particular, A-movement is possible out of nonfinite clauses, but not out of finite clauses. This contrast is illustrated in (587) using weak crossover as a diagnostic of A-movement. See section 2.2.1 of chapter 2 for more discussion.

(587) *Hindi*

- a. *A-movement possible out of nonfinite clauses* [=(79b)]
 har laṛke-ko_i [us-kii_i bahin] [t_i dekha-naa] caah-tii thii
 every boy-ACC 3SG-GEN sister see-INF.M.SG want-IPFV.F.SG be.PST.F.SG
 ‘For every boy *x*, *x*’s sister wants to see *x*.’

b. *A-movement impossible out of finite clauses* [= (76b)]

*har larke-ko_i [us-kii_i bahin] soc-tii hai [ki raam-ne t_i
 every boy-ACC 3SG-GEN sister think-IPFV.M.SG be.PRES.3SG that Ram-ERG
 dekh-aa]
 see-PFV.M.SG

Intended: ‘For every boy x , x ’s sister thinks that Ram saw x .’

In chapter 2, I argue that nonfinite clauses in Hindi are ambiguous between a vP and a TP structure and that only the vP variant is transparent to A-movement, but this added complication is irrelevant for our concerns here.

I have provided evidence in section 2.5.2 of chapter 2 that A-movement lands in SpecTP in Hindi and \bar{A} -movement in SpecCP, just like their English counterparts (584) and (585). The Spell-Out problem described above for English thus also arises in Hindi: The matrix vP phase forces movement through its specifier. Whether subsequent movement from this Spec vP to the matrix SpecTP position is then possible or not must crucially depend on whether the element had been extracted out of a finite or a nonfinite clause. Just like in the case of English, this information is no longer present, being contained inside an already spelled-out phasal complement. Importantly, it is not possible to appeal to v_{def} to sidestep the problem in Hindi. This is because the matrix predicate in both (587a) and (587b) contains an external argument and must hence contain v^* , which is phasal. The v_{def} -based response to the locality problem for vP phases is thus parochially restricted to superraising in English. It does not apply to Hindi, where the problem arises in full force. I therefore contend that appeal to v_{def} does not provide a real solution.

A second illustrative example comes from German, discussed at length in chapter 4. Among the various instances of selective opacity detailed in chapter 4, I will focus on merely one, but parallel remarks apply to the others as well. Consider once more the distinction between *wh*-movement into a V2 clause and *wh*-movement into a V-final clause in German. As we saw in chapter 4 and again in section 6.2.2 of this chapter, the descriptive constraint in (588) holds:

(588) In German, V-final clauses are transparent to *wh*-movement into V-final and V2 clauses. V2 clauses only allow *wh*-movement into a V2 clause, not into a V-final clause. [= (566)]

The relevant examples are repeated in (589) and (589).

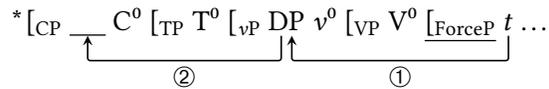
- (589) *Movement out of V-final clause* [= (567)]
- a. *Into V2 clause*
 Wen_i glaubt er [dass die Maria t_i getroffen hat]?
 whom believes he that the Maria met has
- b. *Into V-final clause*
 Ich weiß nicht [wen_i er glaubt [dass die Maria t_i getroffen hat]].
 I know not whom he believes that the Maria met has
- (590) *Movement out of V2 clause* [= (568)]
- a. *Into V2 clause*
 Wen_i glaubt er [t_i hat die Maria t_i getroffen]?
 whom believes he has the Maria met
- b. *Into V-final clause*
 *Ich weiß nicht [wen_i er glaubt [t_i hat die Maria t_i getroffen]]
 I know not whom he believes has the Maria met

This pattern provides another illustration of the *vP* phasehood problem. If *vP* is a phase, all movement dependencies in (589) and (590) must proceed from the edge of the lower clause through Spec*vP* of the higher clause and then to the edge of the higher clause. For the sake of concreteness, I will assume the structural proposal of German in chapter 4, where I argue that V-final clauses are CPs and V2 clauses are ForcePs and that *wh*-movement into a V-final clause lands in SpecCP, whereas *wh*-movement into a V2 clause lands in SpecForceP. The problem brought about by *vP* phases is, however, entirely independent of these assumptions.

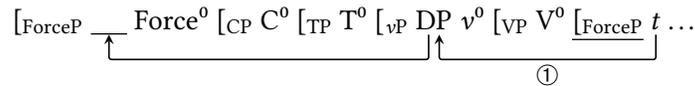
The crucial derivation that needs to be blocked of (590b) is sketched in (591a). The DP first moves from the edge of the embedded V2 clause to Spec*vP* of the higher clause (①), and from there to SpecCP of the matrix clause (②). The problem is that each movement step is independently attested and hence well-formed. Movement from the edge of a V2 clause to Spec*vP* of the higher clause (①) is necessary to derive grammatical movement from a V2 to a V2 clause in (591b). Furthermore, movement from Spec*vP* to SpecForceP (②) is independently required to produce grammatical movement from a V-final into a V-final clause (591c).

(591) *Extraction paths under vP phases*

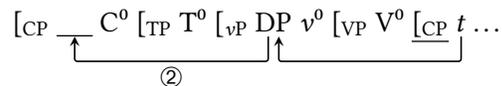
a. *Movement out of V2 clause into V-final clause (590b)*



b. *Movement out of V2 clause into V2 clause (590a)*



c. *Movement out of V-final clause into V-final clause (589b)*

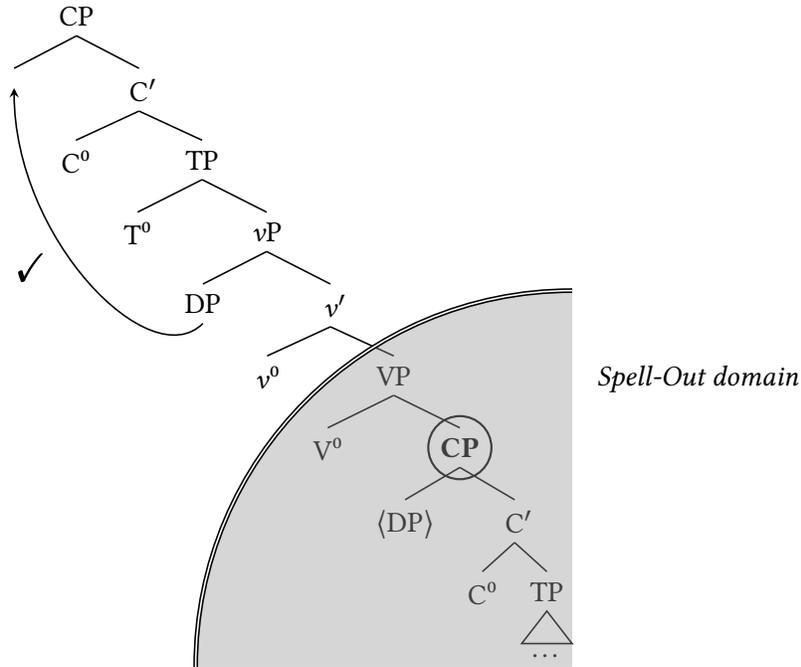


Just as in the case of superraising in English, it is the combination of otherwise well-formed movement steps that needs to be ruled out if *vP* is a phase. Only a nonlocal constraint that has access to information about already spelled-out phases would be able to do so.

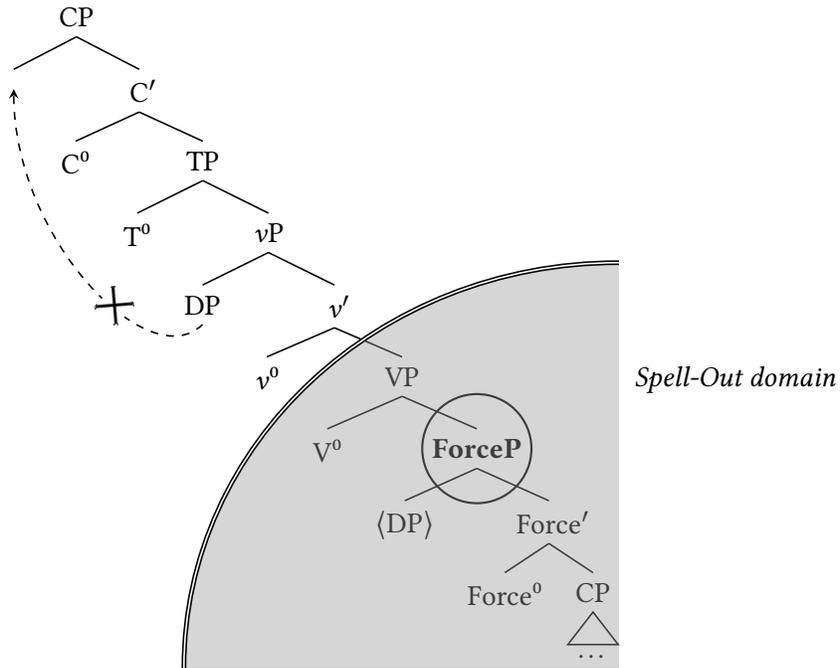
Another way of visualizing the problem is again in strictly derivational terms. Given that movement step ① is possible in (591a), the question of whether DP can move from Spec*vP* to SpecCP (i.e., ②) must depend on information that is no longer present in the workspace. This situation is schematized in (592) and (593).⁶

⁶ Alternatively, one could pursue the view that movement step ① from SpecForceP to Spec*vP* is possible in (592) but impossible in (593). This would not solve the problem either because then this blocking would have to be sensitive to whether the DP ends up in SpecCP or in SpecForceP of the higher clause (compare (591a) and (591b)), hence information about structure that has not yet been built. The choice here is merely between a look ahead problem and a backtracking problem. In either case, information is simultaneously required that should never be simultaneously accessible if *vP* were a phase.

(592) Extraction out of V-final clause → *vP-to-CP* movement possible (591c)



(593) Extraction out of V2 clause → *vP-to-CP* movement impossible (591a)



Whether or not movement from Spec vP to Spec CP is possible in the two structures thus depends on whether the embedded clause is a CP or a $ForceP$. If vP were a phase, this information would be rendered inaccessible by the time the decision has to be made.

Appealing to a special v_{def} head is insufficient here because the matrix clause yet again clearly contains an external argument and hence a phasal v^* head. In sum, invoking a special non-phasal type of v does allow one to circumvent the problem for English, but due to the construction-specific nature of this solution, it does not extend to other instances of precisely the same problem.

The problem is perfectly general and additional examples are easy to come by. For example, relativization in German cannot leave a finite clause, unlike *wh*-movement, but it can leave infinitival clauses (see chapter 4). This would again entail that movement from SpecCP to SpecvP must be possible (because of *wh*-movement), that relativization from SpecvP to SpecCP must be possible (because of relativization that is local or out of an infinitive), but the combination of the two must be ruled out. As in the other examples just discussed, the intermediate touchdown that vP phases enforce requires the relevant constraint to be global in nature.

A second way of addressing this problem induced by vP phases is to invoke case and the Activity Condition to restrict A-movement (Chomsky 2001). As discussed in detail in chapter 5.2 of chapter 5, this account restricts A-movement to elements whose case feature has not yet been valued. Once an element has its case feature valued, it becomes ‘inactive’ for all A-processes, including A-movement. This approach circumvents the problem for English because elements moved out of finite clauses will invariably be case-marked and hence invisible to A-processes.

Yet this second line of solution falls short for exactly the same data as the first. As I have argued in section 5.2 of chapter 5, case-marked elements in Hindi are able to undergo A-movement. In fact, there is compelling evidence that the moving element in both (587a) and (587b) receives its case in the embedded clause. On a case-based account, it remains entirely unclear why A-movement is possible in (587a) but not in (587b). This point is even more pressing for *wh*-movement in German (590). The case-based account does not apply here, not only because the case of the moving element is valued inside the lower clause in both examples, but also because both structures involve \bar{A} -movement, which case does not restrict. Just like the v_{def} solution to the problem, the case-based account is insufficient because it is narrowly geared towards English. The problem created by vP phases is considerably more widespread and ultimately beyond the reach of the two attempts at a solution just discussed.

The problem created by vP phases for selective opacity effects in general and superraising in particular has largely gone unnoticed in the literature. The only direct discussion of the problem

that I know of is Müller (2014a,b), who converges on the conclusion that a system with vP phases makes necessary a constraint that applies to the entire derivational history of the moving element. Müller (2014a,b) cleverly reconciles this requirement with the cyclic Spell-Out inherent in the phase system by invoking a mechanism (so-called ‘buffers’) that essentially records the movement history of an element on that very element. On this proposal, the element carries upward as part of its feature specification a record of the positions it has moved through. Crucial information about the structure of spelled-out phases is thereby retained on the moving element. At the end of the derivation, this record is inspected and unwanted items with illegitimate movement histories are filtered out. A more detailed discussion of Müller’s (2014a,b) is provided in section 5.6 of chapter 5.

While such an account thus provides a technical solution to the problem created by vP phases, it is worth bearing in mind that this solution crucially requires the postulation of ad hoc machinery. There is, to my knowledge, no evidence that an item’s derivational history is recorded on that item apart from the range of facts this proposal is designed to solve. Moreover, as discussed in section 5.6 of chapter 5, once one has access to the derivational history of an element, any number of pathological constraints could be imposed, such as, for instance, that buffers have to have any form *but* the functional sequence, a stipulation that would produce the inverse of the attested facts but is no more or less natural in the buffers account. The proposal thus severely overpredicts the range of movement constraints that could be observed. Finally, it strikes me as conceptually unattractive to devise a system in which portions of the syntactic workspace are removed regularly and then supplement this system with another mechanism that keeps a record of information that has been removed.

There is a very simple solution to the problem that vP phases present for selective opacity. Instead of designating v^0 as a phase head and then introducing a number of additional stipulations to counteract the effects of vP phases, one may question the underlying assumption that vP is a phase. It should be clear from the discussion that the Spell-Out problem ceases to exist if vP is not a phase. No appeal to various v^0 heads with distinct phasal status or buffers is then necessary, because the problem that these concepts were intended to solve does not arise in the first place. This is the approach I have presupposed in the preceding chapters.

I conclude from the considerations in this section that selective opacity effects provide theory-internal evidence against vP phases as they would require interactions between parts of the structure

that could not interact if ν P were phases. Importantly, while this conclusion holds for horizons, it is also true of virtually all other accounts of selective opacity. Two questions immediately arise at this point. First, one may rightly wonder now whether there is independent evidence that ν P is not a phase. Second, we should ask how previous evidence in favor of ν P phases can be satisfactorily accounted for if ν P is not, after all, a phase. The remainder of this chapter will address both of these points.

6.3.2 Nonlocal agreement as an argument against ν P phases

The problem created by selective opacity for ν P phases discussed in the previous section amounts to ν P phases rendering the system *too local*. Converging evidence for an analogous problem arises in the domain of ϕ -agreement. Recall again the definition of the Phase Impenetrability Condition given in (560), repeated here.

- (594) **Phase Impenetrability Condition** (Chomsky 2000: 108) [=(560)]
 In phase α with head H, the domain of H is not accessible to operations outside of α , only H and its edge are accessible to such operations.

If ν P is a phase and the PIC in (594) holds, then the VP complement of a ν^0 head is spelled-out as soon as T^0 , the next higher head, is merged into the structure. This has the effect that T^0 will not be able to access any material c-commanded by ν^0 , as schematized in (595), where the boxed and shaded area represents material rendered inaccessible once T^0 is merged.

- (595) $[_{TP} DP T^0 [_{\nu P} \nu^0 [_{VP} V^0 DP]]]$

It is easy to show that this locality is overly strict. As, e.g., Richards (2011) points out, it is inconsistent with nominative objects in Icelandic, for example. As is well-known, some verbs in Icelandic take a dative subject and a nominative object (see Taraldsen 1995, Sigurðsson 1996, *et seq.*). In these configurations, the object can control verbal agreement in number, with no indication that such agreement is dependent on movement.

- (596) Henni leiddust þeir [Icelandic]
 she.DAT was.bored.by.3PL they.NOM
 ‘She was bored with them.’ (Taraldsen 1995: 307)

The structure in (596) thus requires it to be possible that T^0 agree with the internal argument of a verb, in direct violation of (595).⁷

- (597) [TP DP T^0 [_{vP} v^0 [_{VP} V^0 DP]]]
 [_uφ] [_φ]
- 

This problem is by no means confined to Icelandic. We have in fact already observed it for Hindi in section 2.2.2 of chapter 2. As discussed in detail there, a verb has to agree with its local object if it cannot agree with the subject (i.e., if the subject is case-marked) and if the object can control agreement (i.e., if the object is not case-marked). An example is provided in (598).

- (598) laṛkō-ne yah kitaab paṛh-ii [Hindi]
 boys.M.PL-ERG this book.F.SG read-PFV.F.SG
 ‘The boys read this book.’

Such object agreement is obligatory and also affects weak indefinites, idioms, and complex predicates and elements that demonstrably resist movement. There is furthermore no indication that agreeing objects occupy a structural position different from that of non-agreeing ones. This suggests that object agreement in Hindi is established as in (597).

In both cases, the problem is that vP phases in conjunction with the PIC in (560) yields a system that is too local, as it rules out a direct dependency between a T^0 head and an object across a v projection, because the VP is rendered inaccessible as soon as T^0 is merged. One solution to this problem that was already discussed in section 6.3.1 is to assume that the v^0 head in these cases is defective and hence not a phase. We have already seen on the basis of (587) that this line of account runs afoul of the Hindi facts.

A second common solution to this problem is to slightly delay that Spell-Out of the phase complement in a redefinition of the PIC in Chomsky (2001).

⁷ In both the Icelandic example in (596) and the Hindi example in (598), the clause contains an external argument and hence must involve v^* . The point made here therefore holds irrespective of whether a ‘weak’ v_{def} head is distinguished.

(599) **Phase Impenetrability Condition** (Chomsky 2001:14)

The domain of [phase head] H is not accessible to operations at ZP [= the smallest strong phase]; only H and its edge are accessible to such operations.

The 2001 version of the PIC in (599) differs from the 2000 version in (594) in that the phase complement is not spelled-out as soon as the next higher head is merged, but as soon as the next higher *phase* head is merged. For the case at hand, the VP complement of a v^0 phase head is rendered inaccessible only once C^0 is merged. This enables agreement between T^0 and the object in its base position and thereby accommodates (596) and (598).⁸

There is reason to believe that simply delaying the point at which Spell-Out takes place is an insufficient solution to the problem. First, the revised definition in (599) clearly applies to all phase heads, hence to C^0 as well. This entails that the TP complement of a C^0 head is no longer spelled-out once that CP is completed, but only once the v^0 of the next higher clause is merged. This entails that extraction out of a CP clause no longer has proceed through SpecCP of that clause. It would be equally licit for an element to move to SpecTP of the lower clause, followed by movement to SpecVP of the higher clause, and so on, as depicted in (600), where the boxed and shaded area again designates spelled-out material.

$$(600) \quad [_{VP} DP V^0 [_{CP} C^0 [_{TP} t T^0 \dots]]] \longrightarrow [_{vP} v^0 [_{VP} DP V^0 [_{CP} C^0 [_{TP} t T^0 \dots]]]]$$

This is a significant problem because the account for successive cyclicity through SpecCP that phases were designed for would then be lost. All instances of obligatory successive-cyclic movement through SpecCP would then have to be attributed to a principle *other than phases*, clearly an unsatisfactory result as it would call into question whether phases are necessary at all.

At the heart of this problem lies a fundamental difference between the locality of vPs and CPs . There is compelling empirical evidence that it is possible to have syntactic relations between elements inside a vP and material outside of them, as allowed by the 2001 version of the PIC (599). There is no comparable evidence for CPs . All movement relations across CPs are mediated via its edge, and no agreement past it is possible (see the Polinsky–Bobaljik Generalization (578)). Any

⁸ Interestingly, the empirical ground gained by adopting the 2001 PIC (599) over the 2000 PIC (560) are lost again under the view that the features of T^0 are in fact dependent on C^0 . Chomsky (2007, 2008) proposes that probe features on T^0 originate on C^0 and are handed down onto T^0 (*feature inheritance*). On this account, T^0 will not acquire its ϕ -probe until C^0 is merged. Object agreement is then again incorrectly ruled out even on the 2001 version of the PIC.

account that treats v^0 and C^0 on par with respect to their phase locality status will fail to capture this asymmetry: The 2000 PIC (560) yields adequate locality for CPs, but seriously undergenerates for v P. The 2001 version (599), on the other hand, is more adequate for v P locality (though see the remarks immediately below), but severely overgenerates for CP locality. No account that treats CPs and v P. along the same lines is able to capture this evident empirical difference between the two.

The observed asymmetry between CPs and v P. of course follows without problems if only C^0 is a phase head and v^0 is not. Adopting the 2000 version of the PIC then correctly captures the striking contrast between CPs and v P., precisely because it applies only to CPs.

Long-distance agreement and *wh*-licensing provide additional arguments against v P. phases, irrespective of which version of the PIC is adopted. We have already seen that ϕ -agreement across a v projection is possible, disconfirming the 2000 PIC. The 2001 PIC allows such agreement, but crucially predicts that no more than a single v projection can be crossed by agreement as it delays Spell-Out only until the next higher phase head is merged. This prediction is not borne out. Chapter 2 has discussed in great detail long-distance agreement (LDA) in Hindi. An example is provided in (601), where the embedded object *madad* ‘help’ controls agreement on the matrix verb *caah* ‘want’ and auxiliary.

(601) akbar-ne [merii madad kar-nii] caah-ii thii
 Akbar-ERG my.F help.F do-INF.F.SG want-PFV.F.SG be.PST.F.SG
 ‘Akbar had wanted to help me.’ (Bhatt 2005: 798)

Chapter 2 has provided evidence that LDA does not require movement of the agreement controller as elements that resist movement can nonetheless control LDA (section 2.2.2). Moreover, there are configurations in which LDA becomes obligatory even if the agreement controller has not left its base position (section 2.3.3). Furthermore, there is evidence that even in LDA configurations the embedded clause contains an accusative case licenser and a PRO subject, hence a v P. projection (section 2.3.1). This v^0 head is furthermore able to license the case of the embedded object and introduces an external argument (in the form of PRO). The structure of (601) is thus as in (602). I abstract away from the linear order between the verb and its object for readability. Altogether, two v P. are crossed by ϕ -agreement: one in the matrix clause and another inside the nonfinite clause.

these results is clear: if vP is not a phase head, all the empirical facts just presented fall out without further ado.

I thus conclude that there is empirical evidence from long-distance operations that v^0 is not a phase head, whereas C^0 is. This conclusion is identical to the one reached on the basis of selective opacity effects in section 6.3.1. Supporting evidence for the conclusion that C^0 is phasal but v^0 is not comes from sentence processing and is discussed in detail in chapter 7. One may rightfully raise the question of how the view that v^0 is not a phase is compatible with arguments that extraction over a vP has to proceed through the specifier of that vP , much like successive-cyclic movement over CPs. The next section will address this question.

6.4 A reassessment of previous arguments for vP phases

There are a number of arguments in the literature to the effect that vPs constitute phases on par with CPs. This section will critically examine some of these arguments and demonstrate that they do not, after all, support this conclusion. An important point to be made in this section is methodological. To demonstrate that vP is a phase, it has to be shown (i) that movement through Spec vP is obligatory and (ii) that this movement occurs in order to escape the vP . I will discuss here two often-cited arguments for vP phases and show that they fail to establish these two claims. As a result, these arguments are compatible with the existence of vP phases, but they do not require them. As such, they are orthogonal to the question of whether vP is a phase or not.

Before considering empirical arguments for vP phases, it is worth pointing out that the original motivation for vP phases was conceptual in nature (Chomsky 2000, 2001).⁹ Chomsky (2000, 2001) defines phases as constituents that are ‘propositionally complete’. Because a vP contains a verb and all of its arguments, it counts as propositionally complete and thus constitutes a phase. It has often been noted that this purely conceptual argument faces a number of severe problems (Bošković 2002, Epstein & Seely 2006, Boeckx & Grohmann 2007). For example, Chomsky (2000) argues that finite clauses, but not nonfinite ones, are phases, yet both are propositional (Bošković 2014). Moreover, on the standard view, adopted by Chomsky (2000, 2001) that the external argument is introduced

⁹ The claim that intermediate landing sites are created at VP edges goes back to Chomsky (1986), where this view emerges from purely theory-internal considerations.

by v^0 , VP and vP are both propositions (i.e., type $\langle s, t \rangle$), see Kratzer (1996), yet Chomsky takes only the latter to be a phase head.

It is also worth emphasizing that the original conceptual justification for the existence of phases from Chomsky (2000, 2001, 2004, 2005) does not entail specific heads to be phase heads. Chomsky (2000, 2001, 2004, 2005) contends that the grammar utilizes phases in order to optimize cognitive resource demands, by cyclically transferring built syntactic structure (also see Berwick & Weinberg 1984 for a related proposal). This view is sometimes taken to suggest that phases should be as small as possible in order to minimize the computational burdens to the greatest extent (e.g., Müller 2010b). Yet it is simply not the case that smaller phases invariably lead to a reduction in computational complexity. As the discussion in section 6.3.1 has made clear, accounts of improper movement in particular and selective opacity more generally that invoke vP phases require some form of access to information embedded inside an already spelled-out phase. This information might be obtained by inspecting the contents of a spelled-out phase or preserving this information by encoding it on element that moves out of the phase (Müller 2014a,b). In either case, the account will need to appeal to a mechanism that is otherwise unmotivated and in fact unnecessary if vP is not a phase. Whether or not vP phases outweigh the computational burden imposed by a mechanism is entirely unclear. Moreover, an increase in the number of phases invariably leads to an increase in the number of intermediate landing site that have to be created. It is quite possible, then, that having *too many* phases might in fact incur an undue *increase* in the processing load they incur. From a purely conceptual perspective, one might argue that a larger, CP-based phase system combines the advantages of cyclic Spell-Out with a relatively small number of intermediate landing sites necessary to construct a movement dependency.

In sum, conceptual arguments cut both ways and it is unlikely that considerations of computational complexity will produce reliable conclusions regarding the distribution of phase heads. The matter will have to be decided on empirical grounds. The next sections will take on some previous empirical arguments for vP phases.

6.4.1 Reconstruction

Turning now to empirical arguments for vP phases, one classical consideration comes from Fox (1999) and Sauerland (2003) and revolves around reconstruction. Fox (1999: 174) presents the sentence in (609). There a complex DP containing a proper name (*Ms. Brown*) and a pronoun (*he*) is *wh*-moved. Crucially, this movement path crosses a pronoun coindexed with the proper name and the quantified DP *every student* binding the pronoun. This configuration imposes two constraints: First, due to Principle C, the proper name must not be *c*-commanded by the pronoun *her*. Second, the pronoun *he* contained inside the DP must be *c*-commanded by *every student* in order to be bound. Fox (1999) reasons that the DP can neither be interpreted in its surface position—because *every student* does not *c*-command *he* in this position—, nor in its base position—as the proper name *Ms. Brown* would be *c*-commanded by *her*, creating a Principle C violation. The only position that allows the binding requirement and Principle C to be simultaneously satisfied is situated between *every student* and *her* (indicated by an underline in (609)). Because no CP boundary falls within this range, Fox (1999) concludes that there exists an intermediate trace within the vP region.

(609) [Which of the books that he_i asked Ms. Brown $_j$ for] did every student $_i$ ___ get
from her $_j$ ___* ?

A similar argument is put forth for A-movement by Sauerland (2003) and is based on the example in (610). This example has an interpretation in which the raised subject *every child* takes scope underneath the negation, but at the same time binds the pronoun *his*. As Sauerland (2003) points out, interpreting *every child* in its base position in the lower clause would run afoul of the pronominal binding, and interpreting it in its surface position would give it scope above negation (Sauerland shows that negation cannot scope over subjects). He concludes that *every child* in (610) must have moved through an intermediate position sandwiched between the matrix experiencer PP and the matrix negation, which he proposes is Spec vP . The resulting derivation is shown in (611).

(610) Every child $_i$ doesn't seem to his $_i$ father [t_i to be smart].

(611) Every child_i doesn't [_{vP} t_i seem to his_i father [t_i to be smart]].

Both Fox's (1999) and Sauerland's (2003) arguments show that reconstruction into an intermediate position at the edge of VP/vP region is *possible*.¹⁰ What they demonstrate is that extraction out of vP *can* proceed in two steps. But the claim that vP is a phase is considerably stronger than that, in that it *forces* such movement to proceed in two steps. To demonstrate that vP is a phase, it is necessary to show that extraction out of the vP *has to* proceed through the edge of the vP. The examples in (609) and (610) simply do not establish this stronger claim. Thus, (609) and (610) follow on the standard view that it is possible for one movement step to apply to the output of another, irrespective of phases.¹¹

It seems to me that this is a general property of evidence from reconstruction. The option of reconstructing an element into an intermediate position is uninformative with respect to the question of whether this intermediate position has to be created or merely may be created. Because an argument for phases depends on establishing the former, reconstruction evidence as a matter of principle does not provide evidence for phasehood.

To address this deficiency in evidence from reconstruction, one may consider instances of obligatory reconstruction, e.g., Principle C violations. But Principle C evidence is inherently unable to diagnose obligatory reconstruction into intermediate positions if movement has to land in a position c-commanding the launching site, as is standard. As it turns out, obligatory reconstruction into an intermediate position will be indistinguishable from obligatory reconstruction into the base position. This is because a Principle C violation in an intermediate landing site requires a coindexed pronoun to c-command this intermediate landing site. But it is clear that such a pronoun would

¹⁰ Den Dikken (2006) in fact argues that (609) does not even show that. Adopting Kiss' (1993) account of pair list readings of questions, wherein a universally quantified DP can covertly move above the *wh*-element, den Dikken (2006) proposes the resulting structure in (i):

(i) [every student]_i [which of the books that he_i asked Ms. Brown_j for]_k did t_i get from her_j t_k

In (i), *every student* c-commands *he* and *her* is not c-commanded by *Ms. Brown*, producing the correct interpretation. No reconstruction at all is thus necessary if (i) is the LF for (609).

However, Kyle Johnson (p.c.) points out to me that the binding facts in (609) also obtain if *every student* is changed to *no student*. Because *no student* cannot undergo QR over a *wh*-word as in (i), den Dikken's (2006) account of (609) is not completely general. This does not affect the argument in the main text.

¹¹ Incidentally, Neeleman & van de Koot (2010: 346–347) develop an account on which a moving element can reconstruct into a position even if it has not moved through this position. On their account, reconstruction does not diagnose intermediate landing sites.

automatically also c-command the base position of the moving element. For example, cases like (612) demonstrate that reconstruction of the *wh*-element underneath the subject is required in this case. But because reconstruction into either the base position or the purported intermediate landing site in SpecvP would result in a fatal Principle C violation, (612) does not provide any evidence for reconstruction into an intermediate landing site.

(612) ??/* [Which argument that John_i is a genius] did he_i ___ believe ___? (Fox 1999: 164)

Principle C effects hence provide evidence for obligatory reconstruction, but not for obligatory reconstruction into a position in SpecvP.

6.4.2 Successive cyclicity in Dinka

A second influential argument for vP phases that I would like to discuss here has been presented by van Urk & Richards (2015) and is based on the Nilotic language Dinka (also see van Urk 2015).¹² I will first briefly lay out the data underlying the argument for vP phases and then present van Urk & Richards' (2015) analysis. I will then show that on their account, vP phases do not, after all, play any analytical role. As a result, the Dinka data are uninformative as to whether vP is a phase or not.

Dinka has a preverbal position to the right of the subject that must be filled by exactly one DP in regular declarative clauses. Ditransitive predicates provide the clearest illustration of this requirement. As demonstrated in the paradigm in (613), the preverbal position has to be filled by either the direct object *kitáb* 'book' or the indirect object *Ayén* 'Ayen', but it cannot be filled by both. The other element appears to the right of the verb.

- (613) a. *yêen cé _____ yiên kitáp Ayén
 I PFV give book Ayen
- b. *yêen cé _____ yiên Ayén kitáp
 I PFV give Ayen book
- c. yêen cé Ayén yiên kitáp
 I PFV Ayen give book

¹² I am indebted to Coppe van Urk for very helpful discussion of and help with the Dinka data.

- d. $\gamma\hat{\epsilon}\epsilon\text{n}$ $\text{c}\grave{\epsilon}$ $\text{k}\grave{\text{i}}\text{t}\acute{\text{a}}\text{p}$ $\text{y}\grave{\text{i}}\hat{\epsilon}\text{n}$ $\text{A}\text{y}\acute{\epsilon}\text{n}$
 I PFV book give Ayen
 ‘I gave Ayen a book.’ (van Urk & Richards 2015: 122–123)

Van Urk & Richards (2015) propose that this preverbal position is SpecvP, which bears an EPP feature that requires one vP-internal argument to move to its specifier. Against this background, van Urk & Richards (2015) observe that movement of a vP-internal element to SpecCP must proceed through this preverbal position. In (614), the indirect object *yenà* ‘who’ is *wh*-moved. In this case, the preverbal position must remain empty and cannot be filled by the direct object *kitáp* ‘book’.

- (614) a. $\text{Y}\acute{\epsilon}\text{n}\grave{\text{a}}$ $\text{c}\grave{\text{i}}$ $\text{m}\acute{\text{o}}\text{c}$ --- $\text{y}\grave{\text{i}}\hat{\epsilon}\text{n}$ $\text{k}\grave{\text{i}}\text{t}\acute{\text{a}}\text{p}$?
 who PFV.NS man.GEN give book
 b. * $\text{Y}\acute{\epsilon}\text{n}\grave{\text{a}}$ $\text{c}\grave{\text{i}}$ $\text{m}\acute{\text{o}}\text{c}$ $\text{k}\grave{\text{i}}\text{t}\acute{\text{a}}\text{p}$ $\text{y}\grave{\text{i}}\hat{\epsilon}\text{n}$?
 who PFV.NS man.GEN book give
 ‘Who did the man give the book to?’ (van Urk & Richards 2015: 125)

The same is true of *wh*-movement of the direct object (no example shown here). *Wh*-movement must thus proceed successive-cyclically through SpecvP.¹³ To rule out (614b), a derivation must be blocked in which *kitáp* ‘book’ is moved to SpecvP (as in (613d)), followed by one-fell-swoop movement of *yenà* ‘who’, as schematized in (615). Only movement via SpecvP, as in (616), must be allowed.

- (615) * $[\text{C}\text{P } \text{y}\acute{\epsilon}\text{n}\grave{\text{a}}_i \text{ C}^0 [\text{TP } \dots [\text{v}\text{P } \text{k}\grave{\text{i}}\text{t}\acute{\text{a}}\text{p}_j \text{ v}^0 [\text{VP } \text{y}\grave{\text{i}}\hat{\epsilon}\text{n } t_i t_j]]]]$
 who book give
 (616) $[\text{C}\text{P } \text{y}\acute{\epsilon}\text{n}\grave{\text{a}}_i \text{ C}^0 [\text{TP } \dots [\text{v}\text{P } t_i \text{ v}^0 [\text{VP } \text{y}\grave{\text{i}}\hat{\epsilon}\text{n } t_i \text{k}\grave{\text{i}}\text{t}\acute{\text{a}}\text{b}]]]]$
 who give book

The impossibility of the derivation in (615) appears to constitute very strong evidence that extraction out of vP must proceed successive-cyclically and hence that vP is a phase. Upon closer scrutiny, however, it turns out that this is not the case. The empirical picture is complicated by the

¹³ Van Urk & Richards (2015) provide converging evidence from the stranding of a plural marker *ke* in SpecvP whenever a plural element is extracted out of this vP. Because these stranding facts support the same empirical generalizations as those discussed in the text, I will put them aside here in the interest of space. All considerations in the text continue to hold for *ke*-stranding as well.

distribution of PPs. PPs cannot occupy the preverbal position. Thus, in (617), only the DP *Dèŋ* can occupy this position, while the PP *wúut* ‘to the cattle camp’ cannot.

- (617) a. Bòl à-cé Dèŋ tuòc wúut.
 Bol 3SG-PFV Deng send cattle.camp.LOC
 ‘Bol sent Deng to the cattle camp.’
- b. *Bòl à-cé wúut tuòc Dèŋ.
 Bol 3SG-PFV cattle.camp.LOC send Deng (van Urk & Richards 2015:129)

PPs can, however, be extracted to SpecCP. In this case, the preverbal position remains obligatorily filled with the DP argument:

- (618) Yétenô cénne Bòl Dèŋ tòc?
 where PFV.OBL Bol.GEN Deng send
 ‘Where did Bol send Deng?’ (van Urk & Richards 2015: 130)

The sentence in (618) would appear to show that PP movement out of the *vP* can take place in one fell swoop, violating the view that *vPs* are phases. Van Urk & Richards (2015) however argue that *yétenô* ‘where’ first lands in an outer specifier of *vP* and then moves on to SpecCP:

- (619) [CP $yétenô_i$ C⁰ [TP ... [_{vP} t_i Dèŋ_j v^0 [_{VP} tòc t_j t_i]]]]
 where Deng send
-

The analytical challenge that arises from the conjunction of (615,616) and (619) is to ensure that an outer *vP* specifier can be utilized only by PPs. To account for this asymmetry, van Urk & Richards (2015) make two crucial assumptions. First, they propose the featural specification of *v⁰* in (620). According to (620), *v⁰* contains two probes: one Case-related and one *wh*-related. The Case probe attracts a DP to the specifier of *v⁰*, and the *wh*-probe attracts a *wh*-element. If the two probes agree with different elements, and only then, *v⁰* has two specifiers.¹⁴

- (620) *Feature specification of v*
 $v \left[\begin{array}{l} \text{Case} \\ uwh \end{array} \right]$ (van Urk & Richards 2015: 130–131)

¹⁴ Van Urk & Richards (2015) assume that the featural makeup in (620) is invariant, but that unsuccessful probing does not crash the derivation (Preminger 2011). Thus, the *wh*-probe has to agree with a *wh*-element if there is one, but if there is not, the resulting structure is still wellformed.

The second crucial assumption is the *Multitasking* principle in (621).

- (621) *Multitasking* (van Urk & Richards 2015: 132)
 At every step in a derivation, if a probe can trigger two operations A and B, and the features checked by A are a superset of those checked by B, the grammar prefers A.

Applied to (620), Multitasking mandates that if both the Case and the *wh*-probe can agree with the same element (i.e., if there is a *wh*-DP), then they have to. Otherwise, they agree with separate elements.

With these two assumptions in place, van Urk & Richards (2015) are able to account for the data just discussed. In a configuration in which the VP contains two DP arguments, one of which is *wh*, Multitasking requires that both the Case and the *wh* probe on v^0 agree with the *wh*-DP and attract it to v 's specifier. With both probes checked, there is no probe that would license movement of the non-*wh* DP to a second specifier. The *wh*-DP then undergoes further movement from Spec v P to SpecCP. The paradigm in (614) and in particular the contrast between (615) and (616) is thus accounted for. By contrast, in a configuration in which the VP contains a *wh*-PP and a non-*wh* DP, v 's *wh*-probe agrees with the PP and the Case probe agrees with the DP, pulling both into separate specifiers of v^0 . This derives the example in (618) and its structure in (619).

With this description of van Urk & Richards' (2015) system in place, let us now return to the question of whether v P is a phase or not. We have seen that at first glance the impossibility of the derivation in (615), repeated here, in which one-fell-swoop movement out of the v P takes place at first glance constitutes strong evidence for v P phases.

- (622) * $[_{CP} \text{yenà}_i C^0 [_{TP} \dots [_{vP} \text{kítáp}_j v^0 [_{VP} \text{yìèn } t_i t_j]]]]$ [= (615)]
 who book give

As it turns out, however, once Multitasking and the v^0 specification in (620) are added to the account, as van Urk & Richards (2015) propose, (622) is ruled out *regardless of whether v P is a phase or not*. Because v^0 contains a *wh*-probe, the *wh*-element in (622) is attracted to Spec v P simply because Agree is obligatory. Furthermore, a derivation in which both the *wh*-DP and the non-*wh* DP are moved to separate specifiers of v^0 is ruled out by Multitasking. (622) is thus ruled out by Multitasking irrespective of whether v P is a phase or not. In other words, the locality effects of

vP phases are already entailed by Multitasking and (620). vP phases do not make any analytical contribution to the account. Because Multitasking and (620) are analytically required irrespective of whether vP is a phase or not, the Dinka data do not, after all, provide evidence for the existence of vP phases.

It is worth emphasizing that my claim here is not that it is logically possible to reanalyze the Dinka data in a way that does not involve phases. Rather, I have shown that van Urk & Richards (2015) as it stands does not require vP to be a phase. This is because the obligatory movement to SpecvP applies irrespective of whether further movement to SpecCP takes place. To demonstrate that vP is a phase, it is necessary to show movement to SpecvP that applies solely in order to leave the vP, not for independent reasons internal to the vP. The Dinka data simply do not instantiate such a configuration. Therefore, while the Dinka data are clearly *compatible* with vP phases, they do not in any way *require* them, even on van Urk & Richards' (2015) account. The Dinka data thus emerge as orthogonal to the question of whether or not vP is a phase.¹⁵

6.4.3 *meN*-Deletion in Indonesian

An interesting third argument for vP phases to be considered here is presented by Aldridge (2008) and based on Indonesian (also see Georgi 2014).¹⁶ In declarative transitive clauses in Indonesian, the verb may be marked with the prefix *meN*-. It is considered a transitivity marker by Saddy (1991) and as an active marker by Aldridge (2008). Because its precise nature is not relevant for the argument, I will simply gloss it as 'MEN' here.

(623) Ali mem-beli buku.
Ali MEN-buy book
'Ali bought a book.' (Aldridge 2008:1442)

If the object is \bar{A} -moved, *meN*- cannot appear on the verb:

¹⁵ As Coppe van Urk (p.c.) points out to me, the assumption that vP is a phase might provide a meta-theoretical rationale for *why* v^0 in Dinka has the features it does. On the other hand, however, one might invoke Occam's razor to conclude that vPs are not phases, as they render the account of Dinka redundant in that with vP phases a whole class of derivations is simultaneously ruled out by two principles (Multitasking and phase impenetrability). Thus, it is not clear to me that either account is clearly superior on conceptual grounds.

¹⁶ I am indebted to very helpful discussion of these facts with Lisa Travis.

- (624) Siapa_i yang Sally (*men-)cintai t_i.
 who FOC Sally (*MEN-)love
 ‘Who does Sally love?’ (Saddy 1991: 186)

\bar{A} -movement of the subject, on the other hand, does not rule out *meN*-

- (625) Siapa_i yang t_i men-cintai Sally?
 who FOC MEN-love Sally
 ‘Who loves Sally?’ (Saddy 1991: 186)

Although the details of their accounts differ, both Aldridge (2008) and Georgi (2014) analyze *meN*-deletion as a side effect of movement to SpecvP. The fact that such deletion is obligatory if object extraction takes place then argues that the object has to obligatorily move through SpecvP on its way to SpecCP, in support of vP phases. The argument is compelling because it shows that such movement is obligatory and because there is no reason to believe that the object moves to SpecvP in declarative clauses, i.e., in the absence of further movement to SpecCP. These properties render the argument stronger than both the reconstruction evidence and intermediate touchdowns in Dinka.

While I do not have a comprehensive counter-analysis to offer, I would like to note that the argument faces a substantial challenge from in-situ *wh*-questions. As Saddy (1991) notes, an (overt) *wh*-movement is optional in the language and a *wh*-element in object position may induce a regular question interpretation. Crucially, *meN*- may appear in such cases:

- (626) Sally men-cintai siapa?
 Sally MEN-love who
 ‘Who does Sally love?’ (Saddy 1991: 186)

There are two analytical options for *wh*-in-situ questions like (626). First, *siapa* could not undergo any movement whatsoever and instead enter into an Agree relationship with C⁰, much like I have argued for Hindi in section 2.4 of chapter 2. On such an account, (626) would actually entail that vP *cannot* be a phase in Indonesian. Because of the presence of *meN*-, it is clear that *siapa* has not been moved to SpecvP. If vP were a phase, it should therefore be inaccessible to C⁰ and the structure hence uninterpretable (at least as a question). This is clearly incorrect. The second possible line of account is to say that *siapa* has undergone covert movement in (626). Such an account first

of all entails that covert movement cannot be overt movement with the realization of a lower copy, but must be a qualitatively different mode of movement, either in that it need not apply successive-cyclically (in violation of vP phases), or in that for some reason it does not trigger *meN*-deletion (in which case the link between edge movement and *meN*-deletion becomes tenuous). In either case, the account is arguably incompatible with the single-cycle model of syntax that phases are crucially couched within because such model is fundamentally incompatible with the notion that covert movement is a designated movement type that applies after structure building is completed. Consequently, (iv) constitutes a substantial challenge to the view that vP s are phases and current phase-based analyses of (623)–(625) do not extend to (626). I therefore consider the Indonesian evidence inconclusive.

6.4.4 Section summary

In this section, I have reassessed three popular arguments in favor of vP phases. My point here is primarily methodological, using these three arguments as illustrations. The general conclusion that emerges from the considerations in this section are that a successful argument for vP phases has to meet the two requirements in (627).

- (627) a. *Obligatoriness requirement*
 Movement to Spec vP must be obligatory.
- b. *Extraction requirement*
 Movement to Spec vP must take place in order to be extracted out of the vP , not for purely vP -internal requirements.

The reconstruction argument does not establish (627a) because it does not show that the movement is optional. The argument from Dinka fails (627b) because the movement is triggered by purely vP -internal features and hence enforced regardless of whether vP is a phase or not. The Indonesian evidence is of particular interest in that it conforms to both requirements in (627). But as I have shown, the Indonesian evidence itself is inconclusive because there are processes that appear to be able to search into the bowels of a vP even in Indonesian, which would be unexpected if vP were a phase. Consequently, the reconstruction and Dinka arguments are compatible with vP being

phases, but do not require it to be, whereas the Indonesian evidence arguably constitutes evidence against vP phases.

It is likely that other previous arguments for vP phases will also fail to establish (627) and hence be inconclusive, but needless to say, a careful consideration of all previous arguments for vP phases will have to be carried out to establish exactly which ones adhere to the requirements in (627), a task that I cannot undertake here. What I hope to have shown, however, is that the empirical support for vP phases is considerably weaker than it is standardly taken to be. In combination with the arguments in section 6.3 against vP phases, it seems clear that, at the very least, the phase status of vP is much less secure than is commonly assumed.

6.5 Chapter summary

This chapter has investigated the status of phases in a system incorporating horizons as a constraint on the locality of operations. I have argued that there is evidence for CP phases as constraints on operations in addition to horizons. Importantly, there is no tension between the two concepts as they have clearly different empirical signatures. As a result, no analytical redundancy arises in a system that incorporates both. In a nutshell, horizons are operation-specific and determine whether a given operation can access a domain or not. Phases, by contrast, are domain-general and hence apply to all operations alike. Moreover, phases merely induce an edge effect. They hence do not render a domain entirely opaque from the outside, but merely state that movement has to apply successive-cyclically through the edge of CPs. I have presented evidence that corroborated this division of labor.

Next, I have investigated the status of vP phases. We saw that unlike CP phases, vP phases are incompatible with a horizon-based system. It is noteworthy that this incompatibility extends to virtually all previous accounts of selective opacity. The reason is that the formulation of empirically adequate constraints for selective opacity require referral to information that should be properly contained inside a spelled-out vP phase and hence inaccessible. Access to the derivational history is hence required in one way or another, violating at least the spirit of vP phases. I have furthermore presented evidence from the non-locality of ϕ -agreement and *wh*-licensing that provides independent support for these conclusions.

Finally, I have reassessed some previous arguments for vP phases and demonstrated that they do not, after all, establish vP to be a phase because the conclusions they lead to are too weak to demonstrate the presence of v phases. They are, in other words, merely compatible with vP phases, a very weak conclusion. The next chapter will establish that this is more generally the case. More precisely, I will argue that standard diagnostics for phasehood are biased in that they are inherently unable to diagnose the *absence* of a phase head. I will then present new evidence from online sentence processing that converges in a striking way with the conclusions reached on independent grounds in this chapter, i.e., that CPs are phases but vP s are not.

CHAPTER 7

PROCESSING EVIDENCE FOR THE DISTRIBUTION OF PHASES

7.1 Introduction

Chapter 6 has discussed in detail the role of phases in the horizon-based system developed in this dissertation. The key conclusions of this chapter are that (i) CP phases coexist with horizons and that (ii) ν P is not a phase. These conclusions are convergent with the standard view on phases in the domain of CP, but diverge from it in the domain of ν P. Chapter 6 has provided empirical evidence for this disparity between CPs and ν Ps with respect to their status as phases. This chapter is more exploratory in nature and uses evidence from a somewhat untraditional domain – online sentence processing – to further investigate the distribution of phases. The goal of the experiments reported in this chapter is hence not to investigate the processing of selective opacity effects per se,¹ but to investigate a consequence of my proposal for the distribution of phases. I will show that this evidence aligns with the conclusions reached in chapter 6. In doing so, this chapter will first motivate the use of sentence processing as an empirical window into the presence of intermediate landing sites and hence the structure of movement dependencies. I will then present experimental evidence that supports the view that CP hosts an intermediate landing site while ν P does not. This pattern of results is explained if CP is a phase, but ν P is not. The fact that considerations in two

¹ See McKinnon & Osterhout (1996) for an investigation of the processing of superraising in English.

very different domains ('classical' syntactic in chapter 6 and sentence processing evidence in this chapter) lead to convergent conclusions is a heartening result, not only because it corroborates these conclusions, but also because it indicates that sentence processing evidence may provide crucial evidence for or against syntactic proposals and analyses in at least some domains.

This chapter is structured as follows: In section 7.2, I argue that traditional diagnostics for intermediate landing sites are biased in the sense that they are inherently unable to diagnose the absence of an intermediate trace. I then motivate in section 7.3 the use of evidence from online sentence processing to diagnose the presence or absence of intermediate landing sites, following pioneering work by Gibson & Warren (2004), who present processing evidence for intermediate gaps in SpecCP. Two self-paced reading experiments are then presented in sections 7.4 and 7.5, respectively. Experiment 1 provides further evidence for successive cyclicity through SpecCP by ruling out a plausible alternative explanation of Gibson & Warren's (2004) results that does not invoke intermediate landing sites. Experiment 2 extends the experimental methodology to ν P phases. The results of this experiment indicate that ν Ps differ from CPs in a way that follows if they do not host an intermediate landing site and hence are not phases. Section 7.6 discusses the results in the context of the findings of chapter 6 and current theories of online sentence parsing. Section 7.7 concludes.

7.2 Successive cyclicity and a bias problem

In chapter 6, I have argued that CP is a phase, but ν P is not. The view that CP is a phase is standard and relatively straightforward to motivate: Movement dependencies across a CP projection have to proceed successive-cyclically through SpecCP and in-situ relations like ϕ -agreement may have access to SpecCP in some languages, but appear to never be able to reach past a C projection. Motivating that ν P is not a phase is considerably harder and it is instructive to consider the reasons for this. Chapter 6 has provided two types of arguments that ν P is not a phase. The first argument (section 6.3.1) concerned the formulation of syntactic constraints that rule out selective opacity effects like superraising. I have shown that if ν P were a phase, such a constraint (be it horizons or something else) would have to make reference to material that has already undergone Spell-Out and should hence no longer be accessible. In other words, ν P phases make necessary reference to

the derivational history of an element, arguably undermining the very notion of cyclic Spell-Out that phases encapsulate. The second argument (section 6.3.2) was based on in-situ relations like ϕ -agreement and *wh*-licensing. I have shown that both may proceed over an arbitrary number of vP projections, a situation strictly ruled out if vP were a phase. In all of these cases, vP phases render the system too local and rule out a number of attested syntactic dependencies. As also shown in chapter 6, standard caveats like delaying phasal Spell-Out, distinguishing between phasal and non-phasal flavors of v etc. are unsuccessful for a variety of reasons.

The question of whether or not vP makes very different predictions with regard to movement dependencies. If vP is a phase, movement out of vP has to proceed through Spec vP . If vP is not a phase, such an intermediate landing site is not necessary, although it may of course be possible if triggered by independent factors. Despite their clearly different predictions, assessing the empirical state of affairs is very difficult, for principled reasons. I have shown in section 6.4 of chapter 6 that several arguments in favor of successive-cyclic movement through Spec vP are ultimately unsuccessful in this regard. This is because they either establish that an intermediate landing site in the vP region is merely possible, a point too weak to support the existence of a phase head, or the intermediate landing site is created due to aspects of the system entirely unconnected to the vP status. In the latter case, the assumption that vP is a phase has no analytical contribution and may equally well be dispensed with. These purported arguments, then, fail to constitute evidence that vP is a phase. Crucially, however, merely showing that a particular argument in favor for vP phases does not go through is not the same as showing that vP is not a phase. In the same vein, Dayal (to appear) argues that Manetta's (2010) argument for vP phases in Hindi is faulty, but she is careful to emphasize that her refutation of Manetta's (2010) argumentation does not entail that vP is not a phase. The same is true of den Dikken's (2006) reply to Legate's (2003) arguments for vP phases.

To provide an argument from movement in favor of the view vP is not a phase, one would have to demonstrate that extraction out of a vP does not need to proceed through the edge of that vP . The standard diagnostics for movement paths are arguably inherently unable to show that, even if it is factually true. Consider as an example morphological reflexes of successive-cyclic movement. As is well-known and illustrated in section 6.2.1, in Irish the form of a complementizer changes if movement into its specifier has taken place. Crucially, because all complementizers that are crossed by a movement dependency undergo this change, this change constitutes evidence for

intermediate landing sites in SpecCP. In English, on the other hand, the form of the complementizer is not affected by extraction over it. But this of course does not entail that long extraction in English proceeds in one fell swoop. It merely shows that English lacks an overt reflex like Irish. As a consequence, the absence of an overt reflex of successive cyclicity is uninformative with respect to the presence of absence of intermediate landing sites.²

This point generalizes to other diagnostics for intermediate landing sites as well. Since Sportiche (1988), floating quantifiers have often been taken as diagnostics for intermediate landing sites (e.g., McCloskey 2000), on the view that stranding is possible in intermediate positions. As is well-known, however, *wh*-movement in English may strand a preposition in the base position, but it may not do so in an intermediate SpecCP (Postal 1972: 213).

- (628) a. Who_i do you believe [_{CP} Mary thinks [_{CP} Joan talked [_{PP} to *t_i*]]]?
 b. *Who_i do you believe [_{CP} Mary thinks [_{CP} [_{PP} to *t_i*]_j Joan talked *t_j*]]?
 c. *Who_i do you believe [_{CP} [_{PP} to *t_i*]_j Mary thinks [_{CP} Joan talked *t_j*]]?

While Postal (1972) and Perlmutter & Soames (1979: 512–514) argue based on this restriction that long *wh*-movement does *not* proceed successive-cyclically, the more common response to this fact is that movement *is* successive-cyclic, but that there exists an independent constraint that bans P-stranding in SpecCP (e.g., McCloskey 2000: 64n9). Consequently, if stranding is impossible in a given position in a given language, it again tells us nothing about the presence or absence of an intermediate landing site in this position.

The same reasoning applies to other diagnostics standardly invoked to motivate the existence of intermediate landing sites. If reconstruction into a particular position is impossible, this can be due to an independent constraint on reconstruction; if copying in a position is impossible, a constraint on copying can be invoked, etc. For standard diagnostics, the failure of a positive result is uninformative.

² A strong version of this view is adopted by Georgi (2014). She contends that even if a language generally has an overt morpho-syntactic reflexes of movement into a position, the absence of such reflexes in intermediate positions does not disqualify the claim that there are no intermediate traces in these positions. This is of course a reasonable position, but it underscores the analytical asymmetry noted in the text: The absence of a reflex is entirely uninformative with respect to the existence of intermediate landing sites.

A consequence of this state of affairs is that standard diagnostics are *inherently biased* in that they may never provide evidence for the absence of an intermediate landing site and hence of phases. This is a significant result as it entails that traditional diagnostics for intermediate landing sites are by their very nature unable to detect the *absence* of intermediate landing sites. Put differently, the claim that a given head is a phase makes no falsifiable predictions with respect to traditional diagnostics. Evidence supporting the claim in chapter 6 that *vP* is not a phase will hence not come from these traditional diagnostics. To fairly assess this hypothesis in the domain of movement requires an unbiased diagnostic, i.e., one that is able to detect the presence as well as the absence of intermediate landing sites. The next section will argue that real-time sentence processing provides such a diagnostic.

7.3 Successive cyclicity in online sentence parsing

Despite the immense impact of successive cyclicity on theoretical syntax, there are very few direct investigations into the existence of successive-cyclic movement in sentence comprehension. The only designated attempts to address this question are Frazier & Clifton (1989) and Gibson & Warren (2004), both of which argue in favor of it. However, while the experimental results of these studies are compatible with successive-cyclic movement, they may also be attributed to independently motivated properties of the parser that are unrelated to successive-cyclic movement. This section will review previous results and discuss their limitations. These limitations will form the basis for Experiment 1, reported in section 7.4.

7.3.1 Previous results

The first study that explicitly argues for successive cyclicity in online sentence processing is Frazier & Clifton (1989), who argue that an unassigned filler remains active across clause boundaries. In addition, the results of two end-of-sentence acceptability rating experiments indicate that crossclausal movement dependencies are harder than monoclausal ones, an effect also observed by Phillips, Kazanina & Abada (2005). Frazier & Clifton (1989) argue that these findings can be explained if (i) an intermediate landing site of the filler is responsible for carrying it across clause

boundaries, and (ii) that the construal of the intermediate landing site necessary in crossclausal dependencies increases the demand such dependencies pose onto the parser.

While these interpretations are consistent with the observed results, they are not necessary. First, that a filler is held active across a clause boundary does not, in and of itself, require that filler to be linked to an intermediate trace at the edge of the lower clause. Second, that movement over a clause boundary incurs more difficulty than intraclausal movement may be attributed to a variety of factors unrelated to successive cyclicity. One plausible alternative is that it is the increased distance between the filler and its gap in the two-clause condition that delays the establishment of the movement dependency and thereby increases the difficulty of parsing this structure.

A more recent investigation into processing reflexes of successive-cyclic movement is Gibson & Warren (2004) (henceforth G&W). The rationale underlying G&W's experiment is the following: There is independent evidence that the distance between the moved element and its trace is positively correlated with the difficulty with which the postulation of the gap and its semantic integration take place. In other words, the greater the distance between the moved element and its trace, the harder it is to integrate that trace (King & Just 1991, Gibson 1998, 2000, Gordon, Hendrick & Johnson 2001, Warren & Gibson 2002, Grodner & Gibson 2005, Lewis & Vasishth 2005, Vasishth & Lewis 2006, Staub 2010, Bartek et al. 2011). It is, for instance, well-known that object relative clauses are harder to process than subject relative clauses (King & Just 1991, Gordon et al. 2001, Traxler, Morris & Seely 2002). One line of account attributes this contrast to the fact that the distance between the moved element and the trace is greater in object than in subject relative clauses (e.g., Just & Carpenter 1992). That the distance between the moved element and its trace should affect processing speed at the trace position is commonly assumed to follow from the fact that the semantic and morpho-syntactic features of the moved element have to be retrieved in order to successfully construe and interpret the trace. The greater the distance to this antecedent, the harder this retrieval process will be. How exactly distance is measured is subject to considerable debate in the literature, but the choice is inconsequential for the discussion in this chapter.³

³ Metrics that have been proposed include the number of intervening words (Hawkins 1994), the number of intervening similar elements (Gordon et al. 2001, Lewis 1996), the number of intervening discourse referents (Gibson 1998, 2000, Warren & Gibson 2002, Grodner & Gibson 2005) and the time elapsed between the processing of the filler and the gap in combination with interference from similar constituent encodings (Lewis & Vasishth 2005, Vasishth & Lewis 2006).

Against this general background, G&W's investigated the processing of structures like the ones in (629).

- (629) a. *CP condition*
- i. *Movement*
The manager *who_i* the consultant claimed *t_i* that the new proposal has pleased *t_i* will hire five workers tomorrow.
 - ii. *Control*
The consultant claimed that the new proposal had pleased the new manager who will hire five workers tomorrow.
- b. *DP condition*
- i. *Movement*
The manager *who_i* the consultant's claim about the new proposal had pleased *t_i* will hire five workers tomorrow.
 - ii. *Control*
The consultant's claim about the new proposal had pleased the manager who will hire five workers tomorrow.

In (629a.i) the relative pronoun *who* is moved across a finite clause boundary; in (629b.i) the same element is moved over a complex subject, but this movement does not cross a clause boundary. The sentences in (b) constitute the respective control structures, in which the object of the embedded verb *pleased* is not moved. Of interest are the reading times at the verb that hosts the trace of the moved element – *pleased* in (629). To obtain a measure of the difficulty of integrating the trace, G&W compared the reading times at *pleased* in the movement condition in (i) to the reading times in the non-movement structure in (ii). Because the lexical content and the immediately preceding syntactic context are identical between the two versions, an increase from the non-movement structure to the movement structure must reflect the cost of integrating the trace.

G&W reason as follows: If successive-cyclic movement proceeds through the specifier of each CP that it crosses, there will be an intermediate trace in the CP structure (boxed in (629a.i)). Because no CP is crossed in the DP condition, no intermediate trace exists in this structure. Recall that the difficulty of integrating a gap depends on the distance to its antecedent. If there exists an intermediate trace in the CP condition but not the DP condition, the distance to the antecedent should be smaller in the CP structure than in the DP structure. This in turn predicts that the integration of the trace should be easier in the CP condition than in the DP condition. On the other

hand, if no intermediate trace exists in the CP condition, the integration of the trace should either be equally hard in the two conditions (if it is the linear distance between the trace and its antecedent that matters) or the integration should be harder in the CP structure than the DP structure (if structural distance is the decisive factor). Of interest, then, is whether the integration of the traces is facilitated by a silent intermediate trace in the CP condition. A noteworthy difference between the logic in Frazier & Clifton (1989) and Gibson & Warren (2004) is that Frazier & Clifton (1989) reason that the presence of an intermediate landing site makes processing of a structure harder, where Gibson & Warren (2004) reason that it makes processing at the final gap site easier.

G&W's results confirm the prediction of successive cyclicity. In addition to an overall reading time increase in the movement conditions compared to the non-movement conditions, this increase was reliably smaller in the CP condition (629a) than in the DP condition (629b). G&W take this result to support the existence of an intermediate landing site in Spec,CP: The reactivation of the filler at the intermediate gap site in the CP structure aids subsequent retrieval of this filler at the ultimate gap site. This effect has subsequently been replicated by Marinis, Roberts, Felser & Clahsen (2005).⁴

7.3.2 Limitations

While suggestive and entirely novel, G&W's findings are not completely conclusive. As it turns out, their results are amenable to an analysis in terms of independently motivated parsing strategies that are unrelated to successive-cyclic movement. This alternative account takes the following form: A crucial feature of the design of G&W's experiment is that the movement dependency crosses a verb (*claim* in (629a)) in the CP condition (629a.i), but not in the DP condition (629b.i). One very general concern is that the parser might have initially postulated the trace of *who* as the object of the higher verb *claim* in the CP condition. Note that clear evidence that the filler originates from a lower clause becomes available only when the complementizer *that* is encountered. While thus ultimately incorrect, it is possible that the parser construes *who* as the object of the verb *claim* in its initial parse of this region. It is then furthermore plausible that this intermediate, though

⁴ Marinis et al. (2005) also tested L2 speakers of English (L1: Chinese, Japanese, German or Greek) in addition to native speakers. Interestingly, they replicated G&W's crucial effect for L1 speakers, but not for L2 speakers. Marinis et al. (2005) argue from this result that L2 speakers underuse syntactic information in online sentence processing.

incorrect, reactivation of *who* facilitates its retrieval when its actual trace position is encountered. This facilitation would then be reflected as a smaller reading time increase compared to the DP condition, where no such erroneous intermediate reactivation takes place.

On this alternative account, the effect observed by G&W is entirely due to the particular mechanisms and decision procedures underlying the parser, and have nothing to do with successive-cyclic movement. Rather, they reflect the effect of incorrect structural analyses temporarily pursued by the parser on the construction of a subsequent, and correct, representation. I will subsequently refer to this account as the **PREMATURE GAP FILLING ACCOUNT**, as it crucially involves the premature postulation of a gap that later turns out to be incorrect.

What makes this account a viable alternative to successive cyclicity is that the pieces it comprises are independently motivated in the literature on sentence parsing. First, it is well-known that the parser pursues an **ACTIVE FILLER STRATEGY** when scanning an input string for a gap position (also see de Vincenzi's (1991) *Minimal Chain Principle*):

(630) **Active Filler Strategy** (Frazier & Clifton 1989: 95)

When a filler has been identified, rank the option of assigning a gap above all other options.

The active filler strategy states that the parser is extremely eager to terminate an open movement dependency by postulating a trace at the earliest grammatically licit position. Crucially, it does so even in the absence of direct evidence from the input that a trace is present. As a result, the parser may in some cases postulate a trace prematurely, i.e., in a position that will subsequently turn out to be filled by a lexical element, giving rise to 'filled gap effects', first observed by Stowe (1986). Filled gap effects arise when a position that could in principle contain a trace turns out not to contain one. In (631), for instance, Stowe (1986) observed an increase in reading times at *us* relative to a control condition in which no movement took place. The rationale behind this effect is that the parser postulates a trace after *bring* as a first resort. Upon encountering *us*, this choice turns out to be incorrect and thus requires structural revision, which in turn manifests itself in increased reading times.⁵

⁵ Filled gap effects are remarkably robust. They have likewise been observed in other constructions (see, e.g., Crain & Fodor 1985, Frazier & Clifton 1989, Frazier & Flores D'Arcais 1989, Lee 2004, Wagers & Phillips 2013) and in language other than English: Frazier (1987) and Kaan (1997) present evidence from Dutch; Schlesewsky, Fanselow, Kliegl & Krems (2000) from German, de Vincenzi (1991) from Italian, Aoshima, Phillips & Weinberg (2004) from Japanese and Sekerina (2003) from Russian.

(631) My brother wanted to know *who_i* Ruth will bring us home to *t_i* at Christmas. (Stowe 1986)

In addition to filled gap effects, a variety of other experimental paradigms have yielded support for the active filler strategy.⁶

With regard to the structures in (629), these findings are fully in line with the alternative account just sketched. Because the movement dependency crosses a higher verb (e.g. *claim*) in the CP structure, it is likely that the parser initially postulates the trace of *who* as the object of this higher verb. In the DP condition, on the other hand, the complex subject is an island. In light of evidence that the parser does not postulate traces in islands (Stowe 1986, Traxler & Pickering 1996, McElree & Griffith 1998), it is expected that the parser does not construe a premature trace in the DP structure. The ultimately incorrect but temporarily entertained trace in the CP condition would then require a reactivation of *who* and thereby facilitate its subsequent retrieval at the actual site of the trace.

G&W briefly address this concern (p. 64), noting that the verbs they used were strongly biased towards a clausal complement and that, to the extent that these verbs were compatible with a nominal object, required their object to be inanimate. As the fillers in all of their experimental items were animate, G&W conclude that these properties provide sufficient cues to prevent the parser from temporarily postulating a trace of the moved element at the higher verb.

It is, however, not altogether clear that considerations of frequency, animacy and real world plausibility can completely suspend the parser's active filler strategy, and thus prevent the parser from postulating an otherwise preferred gap. First, while subcategorization constraints have been found to affect initial parsing decisions (e.g., Staub 2007), it is less clear whether the frequency of a subcategorization frame has the same effect. The frequency of a subcategorization frame has been argued not to affect whether a gap is postulated as the argument of an incoming verb (Frazier & Clifton 1989, Staub 2007), or whether a gap is postulated for an optionally transitive verb whose argument could in principle have been extraposed (Staub, Clifton & Frazier 2006).

⁶ Traxler & Pickering (1996), Phillips (2006), Staub (2007) and Wagers & Phillips (2009), for instance, find reading time increases if a verb semantically mismatches an unassigned filler even if that filler is not syntactically related to the verb in the ultimately correct parse of the input string. There is moreover converging evidence from electroencephalography (Garnsey, Tanenhaus & Chapman 1989, Kaan, Harris, Gibson & Holcomb 2000, Phillips et al. 2005), the stops-making-sense task (Boland, Tanenhaus, Garnsey & Carlson 1995), cross-modal lexical priming (Nicol & Swinney 1989, Nicol, Fodor & Swinney 1994) and anticipatory eye movement (Sussman & Sedivy 2003).

Second, it is likewise controversial whether animacy and plausibility constraints have enough of an impact on initial parsing decisions to prevent the parser from construing an otherwise licit trace. Pickering & Traxler (2003), for example, crossed optionally transitive verbs that occur more or less frequently with a direct object with a filler that is either semantically compatible with the verb or not, and observed elevated reading times on a verb if the moved element was an implausible object of that verb, indicating that a trace was postulated even if plausibility considerations disfavor such a choice. Moreover, these elevated reading times arose regardless of the verb's subcategorization frequency, thus corroborating Staub's (2007) findings.

In light of previous evidence, it is thus possible that the parser postulates a trace of the moved element when it encounters the higher verb even if frequency and plausibility disfavor such a decision.⁷ While G&W's results are thus fully compatible with the interpretation that the facilitation at the ultimate gap site in the CP condition is due to successive-cyclic movement, their results could also be attributed to premature gap filling at the intermediate verb. To distinguish between these competing explanations, Experiment 1 investigates the role of the higher verb by systematically manipulating its subcategorization restrictions. The results favor the successive cyclicity account over the premature gap filling one.

7.4 Experiment 1

To distinguish whether the effect observed by Gibson & Warren (2004) is due to the integration of the filler into the intermediate Spec,CP – i.e., successive cyclicity – or to a temporary incorrect parse in which the filler is construed as the object argument of the higher verb – i.e., premature gap filling –, Experiment 1 manipulates the subcategorization frame of the higher verb. As reviewed above, there is evidence that the parser postulates an object trace only for verbs whose subcategorization requirement allows the verb to take a DP object (Staub 2007). Under the premature gap filling account, the moved element should hence only be reactivated at the higher verb if this verb can

⁷ In addition to the general concerns just mentioned, the particular properties of G&W's stimuli also deserve some remarks. First, seven of the nine verbs used by G&W are similar to *claim* in that they productively take nominal objects (*predict, claim, conclude, imply, confirm, realize, state*). Of these seven verbs, at least four appear compatible with animate objects, as a cursory search on Google reveals: *predict* ('Jesus predicted the prophet Mohammed'), *claim* ('The queen claimed the slaves as her own property'), *confirm* ('The Senate confirmed Robert Hanna as Superior Court judge') and *realize* ('He created a trade that reached to all parts of the Union, and realized him a large fortune' [NY Times]).

in principle take an object DP. Consequently, the premature gap filling hypothesis predicts that the G&W effect should be modulated by properties of the intervening verb. By contrast, if G&W's effect is the result of successive cyclicity, the reactivation of the filler is independent of the higher verb and the verb type manipulation should hence not have an impact on the effect.

7.4.1 Method

Materials

Fifty-six sets of sentences like the one in (632) were constructed, incorporating materials adapted from Gibson & Warren (2004) and Marinis et al. (2005). Due to the fact that the semantic relations in the sentences differ by condition, a plausibility norming study was carried out to ensure that the eight conditions matched each other in plausibility. The details of this norming study are described in Appendix A. Forty-eight sets of sentences with closely matching overall plausibility ratings were selected and the remaining eight discarded.

The experiment manipulated INTERVENER TYPE (*CP* vs. *DP*) and MOVEMENT ([+*move*] vs. [-*move*]) in a way parallel to G&W's original study. In addition, the TYPE OF THE VERB preceding the complementizer was manipulated. One set of verbs productively took nominal direct objects (like, e.g., *claim*). Verbs in this class will be referred to as *CP/DP-verbs* as their subcategorization allows for either a *CP* or a *DP* object. The second class of verbs was incompatible with a *DP* object and only allowed a clausal object (e.g., *think*). This second class will be referred to as *CP-verbs*. Membership in the two classes was determined by the sentence completion results reported in Trueswell, Tanenhaus & Kello (1993) and a subcategorization database developed by Sabine Schulte im Walde, based on the British National Corpus and containing data for over 3000 verbs (Schulte im Walde 1998).⁸ Verb type was crossed with the other two factors, yielding a total of eight conditions.

⁸ The 15 *CP/DP-verbs* used in the stimuli are *assert, assume, claim, conclude, confirm, decide, declare, demonstrate, guess, illustrate, imply, predict, prove, recall* and *state*. These verbs had a average *DP* object rate of .21 in the British National Corpus. The 14 *CP-verbs* used are *agree, argue, boast, comment, dream, hint, hope, hypothesize, insist, pretend, remark, speculate, theorize* and *think*. Of these, *agree, boast, dream, hint, hope, insist, pretend, remark* and *think* were selected because their *DP* object rate in Trueswell et al.'s (1993) sentence completion experiment was 0%. Five additional verbs (*argue, comment, hypothesize, speculate, theorize*) were included to decrease the amount repetition in the stimuli. Overall, these *CP-verbs* had a average *DP* object rate of .09 in the British National Corpus. It should be noted that this number is likely inflated because two verbs (*boast* and *theorize*) had a very high *DP* object rate (.41 and .30, respectively), which is due to the highly literary character of some of the materials in this corpus. That *boast* did not elicit a single *DP* object completion in Trueswell et al. (1993) makes it very unlikely that these usages of these two verbs have a significant impact on parsing decisions in the population of interest. To

Because of the relative difficulty of the stimuli, the target sentences were preceded by a theme-setting context sentence. This context sentence gave lexical information about the lower verb in the relative clause in the target sentences (e.g., *hurt* in (632)). The rationale for including this context was to decrease the processing and integration load in the critical region. Informal consultation of native speakers of American English confirmed that this context sentence reduced the perceived difficulty of the target stimuli. All target sentences were followed by a comprehension question about the semantic relations in the sentence the participant had just seen. No question targeted the context sentence. All questions were multiple choice, with both answers being referents present in the sentence. Answer options were presented in random order.

(632) *Context:* Groundless allegations really could hurt people in our company.

a. *CP intervener*

(i) [+move]

The secretary who_i the lawyer $\left\{ \begin{array}{l} CP/DP: \text{ claimed} \\ CP: \text{ boasted} \end{array} \right\}$ that the accusations had hurt t_i was fired from her job.

(ii) [-move]

The lawyer {claimed/boasted} that the accusations had hurt the secretary who was fired from her job.

b. *DP intervener*

(i) [+move]

The secretary who_i the lawyer's $\left\{ \begin{array}{l} CP/DP: \text{ claim} \\ CP: \text{ boast} \end{array} \right\}$ about the accusations had hurt t_i was fired from her job.

(ii) [-move]

The lawyer's {claim/boast} about the accusations had hurt the secretary who was fired from her job.

Comprehension question: Who made a {claim/boast}?
the lawyer – the secretary

In addition to the 48 target items, 48 additional items were created that matched the target sentences in syntactic complexity and length. 24 of these were part of unrelated experiments and the additional 24 were a haphazard collection of sentences.

confirm this conclusion, all analyses reported here were also conducted with all items containing these two verb eliminated. The critical effects remained unchanged in these analyses. With these two verbs excluded, the average DP completion rate of the CP-verbs in the Corpus was .05.

Participants

The experiment reported here involved 130 participants, recruited via Amazon Mechanical Turk. All were native speakers of American English and naïve to the purpose of the experiment. Each received a compensation of USD 1.

Procedure

The experiment was conducted using the online experiment platform Ibex (Drummond 2013) and employed a region-by-region self-paced noncumulative moving-window task (Just, Carpenter & Woolley 1982).⁹ The regioning of the target sentences followed the general schema in Table 7.1.¹⁰ At the beginning of each trial participants saw the theme-setting context sentence, which was displayed in its entirety. Upon pressing the space bar, the context sentence was replaced by dashes masking the regions of the target sentences. Pressing the space bar caused the dashes in the first region to be replaced by the actual content of the region. When the space bar was pressed again this region reverted back to dashes and the next region appeared. Participants traversed through the entire sentence by repeatedly pressing the space bar. Pressing the space bar after the final region had been displayed caused the dashes to disappear and a comprehension question accompanied by two possible answers appeared on the screen. Participants selected the answer on the left by pressing the ‘f’ key and the one on the right by pressing the ‘j’ key. No feedback on answer accuracy was given. After every twelve trials the participant had to take a ten-second break and could rest for longer if they so desired. Altogether, the experiment contained seven of these mandatory breaks. The experimental trials were preceded by a screen collecting general demographic data, three screens

⁹ One might wonder about the reliability of reading data gathered online. Use of online platforms for this type of study is becoming increasingly mainstream among researchers (see Wagers & Phillips 2013 for a recent example). There is so far no indication that this methodology produces results that are qualitatively different from data elicited in a more traditional lab setting (e.g., Wagers & Phillips 2013). Moreover, with regard to the two experiments reported here, it is noteworthy that both extend the experimental design used by Gibson & Warren (2004) and Marinis et al. (2005) in a traditional lab setting. A measure of the reliability of the online data reported here is thus whether they replicate the key effects observed in these two studies. As emphasized in the discussion of the respective results, both Experiment 1 and 2 do replicate these previous findings. It is therefore very unlikely that the effects reported here are merely an artifact of the experimental methodology.

¹⁰ This regioning is identical to the one employed by G&W in their analysis with one difference. In G&W materials, relative pronouns were grouped inconsistently. In [+move] conditions the pronoun was grouped with the head noun phrase (*the secretary* in Table 7.1), while it was grouped with the remainder of the relative clause in the [-move] condition. This inconsistency was irrelevant for G&W as they used a word-by-word presentation and the words were grouped into regions only for the purposes of the analysis. Since, by contrast, the segmentation affects the units of presentation in the present experiment, a consistent regioning was employed and relative pronouns were presented with the head nominal throughout.

Table 7.1. Regioning of stimuli in Experiment 1 ([±mv] = [±move])

		Region							
		1	2	3	4	5	6	7	
CP-verb	CP		The lawyer boasted	that	the accusations	had hurt	the secretary who	was fired from her job	
		[−mv]							
		[+mv]	The secretary who	the lawyer boasted	that	the accusations	had hurt	was fired from her job	
	DP			The lawyer's boast	about	the accusations	had hurt	the secretary who	was fired from her job
		[−mv]							
		[+mv]	The secretary who	the lawyer's boast	about	the accusations	had hurt	was fired from her job	
CP/DP-verb	CP		The secretary who	the lawyer claimed	that	the accusations	had hurt	was fired from her job	
		[−mv]							
		[+mv]	The secretary who	the lawyer's claim	about	the accusations	had hurt	was fired from her job	
	DP			The lawyer's claim	about	the accusations	had hurt	the secretary who	was fired from her job
		[−mv]							
		[+mv]	The secretary who	the lawyer's claim	about	the accusations	had hurt	the secretary who	was fired from her job

of instructions and six practice trials. The experiment took about forty minutes. Throughout the entire experiment a progress bar was displayed. Items were arranged on eight different lists such that each list contained one instance of every item and all eight conditions of each item appeared on one list. Participants were randomly assigned to one of the eight lists. The 48 target sentences were interspersed with the 48 fillers and the order of presentation was randomized for each participant.

Analysis

The data analysis procedures were identical for both experiments reported here and are discussed here in the context of Experiment 1.

All data analysis was carried out in the R software environment (R Core Team 2014). All reading times were logarithmically transformed and then entered into linear mixed effects (LME) models; answer accuracy was analyzed by means of logistic mixed effects modeling. Models were fit using the lme4 package (Bates, Maechler, Bolker & Walker 2014). All models were maximal in the sense that they incorporated random intercepts for participants and items and random slopes for all fixed effects and their interactions for both participants and items, following the suggestions

in Barr, Levy, Scheepers & Tily (2013). The estimate of the regression slope β and the corresponding t/z -statistics will be reported. To obtain p -values for the model coefficients, the Satterthwaite approximation to the degrees of freedom associated with a coefficient's t -value was employed, using the `lmerTest` package (Kuznetsova, Brockhoff & Christensen 2014). All contrasts used in the analyses were orthogonal and will be specified for each experiment individually.

The following general exclusion procedure was applied: Participants who indicated that their native language was not American English were excluded from analysis. If a participant took the experiment multiple times, only the first time was included in the data analysis. Because of the difficulty of the materials and the subtlety of the effect of interest, only subjects whose overall accuracy over the entire experiment (including all target and filler items) was at least 75% were included in the data analysis. Reading time data less than 200 ms or greater than 5000 ms were taken as not reflecting the process of interest and hence rejected as outliers. Finally, log reading times that were more than 3 standard deviations away from the condition mean in that region were discarded.

To adjust for difference in participants' reading speed and the substantial differences in the length of the various regions, residual reading times were calculated. Using all target and filler items used in the experiment, an LME model was fit that predicted raw reading times from the number of characters in a region and included random intercepts and slopes for participants (see Ferreira & Clifton 1986 for discussion of this procedure). At every region the reading time predicted by the regression for a given participant and condition was subtracted from the actual measured reading time. This difference constitutes the residual reading time. A positive value hence indicates slower reading times than were predicted by the model while negative values predict faster reading times. Residual reading times were submitted to LME model analyses only for regions whose length differed across conditions.

7.4.2 Results

In order to streamline the presentation and discussion of the results, the main text will focus attention on the critical regions of the relevant experiment. A comprehensive overview of the analyses of all regions and a discussion thereof is provided in Appendix B.

Of the 130 participants who took part in Experiment 1, 6 were excluded for taking the experiment twice. An additional 30 were discarded for falling below the accuracy threshold.¹¹ Outlier elimination based on the absolute thresholds excluded less than 0.4% of the data. The *z*-score-based rejection of outliers resulted in the exclusion of 0.8% of the observations. For the analysis of the reading time data the results were subjected to a $2 \times 2 \times 2$ -factorial LME model crossing the factor MOVEMENT ([+move] vs. [-move]), INTERVENER (*CP* vs. *DP*) and VERB TYPE (*CP-verbs* vs. *CP/DP-verbs*). Accuracy data were analyzed using a parallel logistic mixed effects model. All covariates were coded numerically using sum-coding (MOVEMENT: [-move] = -.5, [+move] = .5; INTERVENER: *CP* = -.5, *DP* = .5; VERB TYPE: *CP-verb* = -.5, *CP/DP-verb* = .5). Because one item was regioned inconsistently across conditions, it was excluded from all analyses.¹²

Comprehension question response accuracy

The overall answer accuracy over all items and conditions was 85.3%. The proportion of correct responses by condition is given in Table 7.2. Logistic LME modeling revealed a main effect of *Movement* such that accuracy in the [+move] condition was lower than in the [-move] condition ($\hat{\beta} = -1.25, z = -11, p < .001$). No other effect was significant.

Table 7.2. Mean answer accuracy in Experiment 1

	CP-verbs		CP/DP-verbs	
	CP	DP	CP	DP
[-move]	.92	.93	.92	.91
[+move]	.78	.76	.83	.78

Reading times

The mean reading times by region and condition computed across participants as well as residual reading times for region which differed in length between conditions are given in Table 7.3. A plot of the residual reading times by condition and region is given in Figure 7.1. An overview of the by-condition standard error as a measure of the response variability is provided in Appendix B.

¹¹ The critical statistical results reported below also hold if no accuracy-based exclusion of participants is applied.

¹² Eliminating this item did not introduce a plausibility confound (all *ps* > .5, cf. Appendix B).

Table 7.3. Mean raw reading times by condition and region in ms computed over participants for Experiment 1. For regions that differed in length between conditions, residual reading times are additionally provided in parentheses. ('Vt' = Verb type, 'Intv' = Intervener, 'Mvmt' = Movement, [\pm mv] = [\pm move])

Vt	Intv	Mvmt	Region						
			1	2	3	4	5	6	7
CP	CP	[-mv]	—	911	536 (-37)	775	708	812 (-48)	1288 (68)
		[+mv]	731	1032	598 (25)	797	836	801 (59)	868 (-64)
	DP	[-mv]	—	964	561 (-29)	824	761	796 (-62)	1213 (-10)
		[+mv]	735	1216	679 (89)	921	860	863 (120)	866 (-64)
CP/DP	CP	[-mv]	—	896	526 (-48)	765	738	840 (-19)	1279 (58)
		[+mv]	736	1066	631 (57)	790	893	820 (79)	888 (-44)
	DP	[-mv]	—	1018	562 (-20)	824	736	815 (-47)	1219 (-2)
		[+mv]	710	1215	657 (74)	892	903	845 (102)	887 (-43)

Table 7.4 provides an overview of the LME analyses of the log-transformed reading times for the regions that are crucial in evaluating the two contrasting hypotheses. Of central importance are the region containing the trace-hosting verb (region 5) and the spillover region (region 6). In addition, because the manipulation of the type of the higher verb should affect whether the parser prematurely posits a trace or not, the reading times in the region following this verb (region 3) should be affected by the type of the preceding verb. In the interest of space, Table 7.4 only reports the analyses for these three regions. A comprehensive analysis of all regions is provided in Appendix B.

Analysis of region 3 reveals main effects of *movement* and *intervener*. In addition, there was a significant three-way interaction between all factors. Because the length of the region differed between condition, an additional analysis of the residual reading times was performed, which replicated the two main effects ($p < .001$) and the three-way interaction ($\hat{\beta} = -65, t = 2.7, p < .01$). In addition, this analysis indicated an interaction between *movement* and *intervener* ($\hat{\beta} = 25, t = 2.0, p < .05$), which is, however, uninterpretable given the higher-level interaction.

To further investigate the three-way interaction in region 3, the predictors *movement* and *verb type* were nested under the levels of *intervener* in the LME analysis of log reading times, with the full random effects structure being preserved. The rationale for doing so is the question whether CP and DP structures differ in whether or not the reading time increase induced by movement is modulated by the type of the verb (in the CP structure) or noun (in the DP structure).

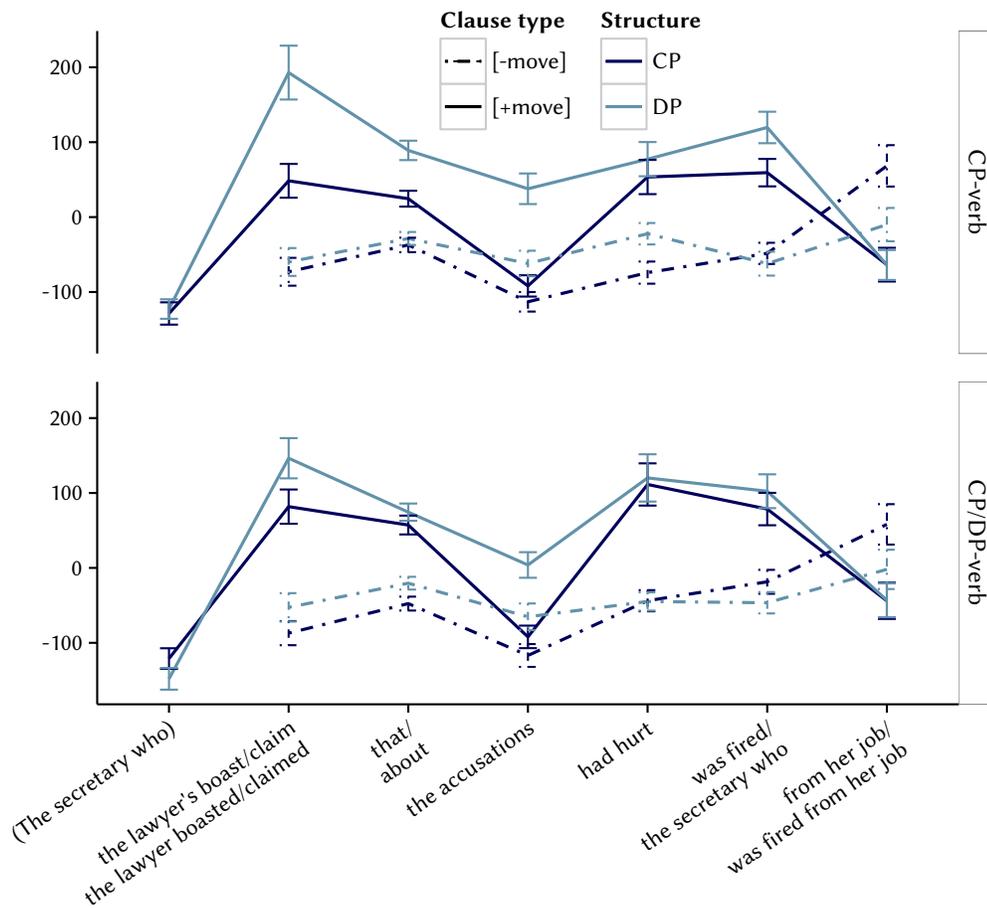


Figure 7.1. Residual reading times by condition and region in Experiment 1 in ms with by-participant standard errors

For DP structures, the model detected a main effect of *movement* ($\hat{\beta} = .14, t = 10, p < .001$) but neither a main effect of *verb type* ($\hat{\beta} = -.01, t = -1.0, p = .3$) nor an interaction between the two ($\hat{\beta} = -.02, t = -1.1, p = .3$). For the CP structure, on the other hand, the model revealed a main effect of *movement* ($\hat{\beta} = .11, t = 8, p < .001$), no main effect of *verb type* ($\hat{\beta} = .01, t = 1.0, p = .3$), but an interaction between both factors ($\hat{\beta} = .05, t = 2.1, p < .05$) such that the reading time increase was higher for CP/DP-verbs than for CP-verbs. The type of the verb/noun in region 2 thus affects the reading times increase in region 3 in the CP structures but not the DP structures. The same pattern emerged in the analogous analysis of the residual reading times. It will be considered in greater detail in the discussion section.

In the gap region (region 5), the DP condition was read more slowly than the CP condition and the [+move] condition was read more slowly than [-move]. There is, in addition, a main effect of

Table 7.4. Coefficient estimates and corresponding t -value for linear mixed effects model analyses of log reading times in critical regions of Experiment 1. *Mvmt:Intv* refers to the interaction between *Movement* and *Intervener*, *Mvmt:Vt* to the interaction between *Movement* and *Verb type* and *Intv:Vt* refers to the interaction between *Intervener* and *Verb type*. *Mvmt:Intv:Vt* refers to the three-way interaction of all predictors. Cells with $p < .05$ are shaded. For statistical analyses of all regions, see Appendix B.

	Region					
	3		5		6	
	$\hat{\beta}$	t	$\hat{\beta}$	t	$\hat{\beta}$	t
Movement	0.127	11.33	0.112	5.03	-0.009	-0.38
Intervener	0.063	6.23	0.031	2.66	0.015	1.17
Verb type	0.000	0.03	0.023	1.99	0.012	1.11
Mvmt:Intv	0.029	1.78	-0.004	-0.17	0.060	2.55
Mvmt:Vt	0.013	0.75	0.018	0.80	-0.026	-1.27
Intv:Vt	-0.024	-1.47	-0.049	-2.37	-0.017	-0.79
Mvmt:Intv:Vt	-0.075	-2.31	0.031	0.70	-0.018	-0.43

verb type such that CP/DP-verb conditions were read more slowly than CP-verb conditions. Finally, there was an interaction between *verb type* and *intervener* such that the reading time difference between CP and DP structures was smaller for CP/DP-verbs than for CP-verbs. There was no interaction between *movement* and *intervener*.

The spillover region (region 6) shows a significant interaction between *intervener* and *movement* such that the DP [+move] structure leads to higher reading times relative to its [-move] control than the CP structure. There is no interaction with verb type. Because the length of this region differed between the [+move] and [-move] conditions and to confirm the reliability of these critical findings, an analysis of the residual reading times was carried out. This model showed a main effect of *movement* ($\hat{\beta} = 132, t = 7, p < .001$) and an interaction between *movement* and *intervener* ($\hat{\beta} = 62, t = 2.4, p < .05$). There was no three-way interaction ($\hat{\beta} = -26, t = -.6, p > .5$). That there was a main effect of movement in only the residual reading times is unsurprising because region 6 was systematically longer in the [-move] than in the [+move] condition. Higher proportional reading times in the movement condition hence are only detectable when the length of the region is factored out.

Figure 7.2 plots the crucial increase in the residual reading times between [-move] and [+move] structures in the spillover region. The analysis just presented reveals the following key components:

- (i) There is a positive reading time increase incurred by the presence of a movement dependency

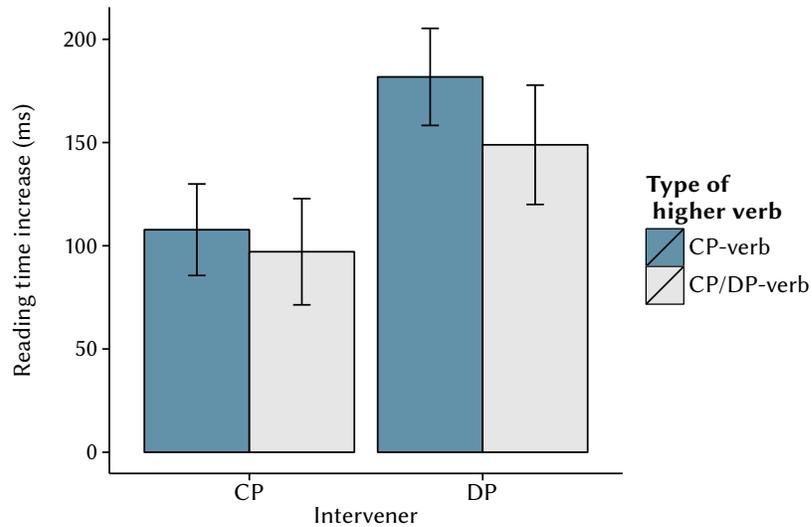


Figure 7.2. Mean increase in residual reading times between [-move] and [+move] conditions and by-participant standard errors in region 6 of Experiment 1 by intervener and verb type

in all conditions, (ii) this increase is greater in DP structures than in CP structures, and (iii) the pattern of this increase does not differ across the two verb types.

7.4.3 Discussion

The results of Experiment 1 replicate G&W's crucial finding: The difference between [-move] and [+move] structures was greater in DP structures than in CP structures when the position of the trace was encountered. While the region that the effect surfaced in differed (region 6 in the present experiment but region 5 in Gibson & Warren 2004 and Marinis et al. 2005), the shape of the effect is identical. Because it is common in self-paced reading experiments for an effect to surface in the spillover region, this discrepancy is not surprising.

The central finding of Experiment 1 is that this effect is not modulated by the type of the higher verb. Because the type of the higher verb is crucially implicated in the premature gap filling account but not the successive cyclicity account, this result provides evidence for successive cyclicity. It is worth noting that the advantage of the CP condition relative to the DP condition was numerically, but not significantly, greater for CP-verbs than for CP/DP-verbs, the opposite of what the premature gap filling account predicts.

The three-way interaction observed in region 3 – the region containing the complementizer/preposition – is particularly instructive. The pattern of this effect is that the reading times increase between [–move] and [+move] conditions is sensitive to the type of the verb in the preceding region, but only in the CP condition. More specifically, the reading time increase is greater for CP/DP-verbs like *claim* than for CP-verbs like *boast*. In the DP condition, on the other hand, this increase is not affected by whether the preceding noun was *claim* or *boast*. This pattern is very plausibly a filled gap effect: If the verb is of the CP/DP-kind, the moved element is initially construed as its object and the parser incorrectly postulates a trace of the moved element when the higher verb is encountered. The complementizer in the following region makes it clear that this parse is incorrect and that reanalysis is required. This reanalysis amplifies the reading time increase compared to the non-movement baseline. The fact that this increase is less pronounced if the higher verb is of the CP-type demonstrates that either no incorrect trace is postulated or that it is postulated on fewer trials. No such difference was observed in the DP condition: Regardless of whether the head noun is *claim* or *boast*, no gap can be postulated in either region 2 or 3. As such, no filled gap effect results as the type of the higher noun has no impact on whether or not a trace is postulated. This accounts for the observed three-way interaction. This finding has important repercussions: It demonstrates that the two groups of verbs (*CP/DP* vs. *CP*) indeed differ in their subcategorization frames and that premature gap filling is affected by these frames. Yet crucially, the critical effect in region 6 is entirely independent of these subcategorization frames. This corroborates the conclusion that the effect at the trace-hosting verb cannot be attributed to premature gap filling, as it is stable regardless of whether or not premature gap filling takes place.

There was, in addition, an interaction between *intervener* and *verb type* in the gap region such that the reading time difference between CP and DP structures is smaller for CP/DP-verbs than for CP-verbs. Because this pattern does not involve the reading time increase between [–move] and [+move] structured, it is orthogonal to the crucial effects just discussed. It is plausible that this effect is a spillover from the preceding region and can be accounted for with reference to the definitional properties of the two verb classes: While CP/DP-verbs allow for a DP object, CP-verbs do not and it is natural to assume that this property of the verb is preserved under nominalization. In the DP structures, the noun in region 4 has to be conceptually integrated with the head noun (*claim/boast*). Due to their conceptual properties, this integration is likely easier for nouns based

on CP/DP-verbs than for noun based on CP-verbs. In the CP condition, by contrast, the noun in region 4 is not semantically construed as an argument of the higher verb and verb type should hence not matter. As a result, the reading time increase from the CP to the DP condition is greater for CP-verbs than for CP/DP-verbs.

That the presence of a movement dependency incurred an increase in reading times in regions prior to the trace is expected under the assumption that an unassigned filler has to be held active in working memory (Wanner & Maratsos 1978, King & Kutas 1995, Gibson 1998, 2000, Fiebach, Schlesewsky & Friederici 2002, Grodner, Gibson & Tunstall 2002, Chen, Gibson & Wolf 2005), a requirement that leaves fewer resources available for the processing of incoming material.

Before proceeding, it is worth considering the role of *structural distance* in the results of Experiment 1. There is some indication that the linear distance between the filler and the gap is an insufficient predictor of integration difficulty. Investigating Chinese relative clauses, Lin (2006) and Lin & Bever (2006) found an advantage for subject relatives over object relatives despite the fact that the object position is linearly closer to the head noun in Chinese. They conclude that structural distance must also have an impact on integration cost. Taking into account structural distance does not affect the interpretation of the present results. First, suppose that no successive-cyclic movement took place in the CP condition. Because movement is crossclausal in the CP condition but intra-clausal in the DP condition, the structural distance between the filler and the gap, and hence the reading time increase, would be predicted to be larger in the CP condition, the opposite of the observed pattern. Appeal to successive cyclicity is hence necessary even if structural distance is factored in. Moreover, if successive-cyclic movement through Spec,CP takes place, the ultimate trace is equally far away from the closest antecedent position in both conditions. Structural distance hence leaves unaffected the conclusion that successive cyclicity takes place in the CP condition but not the DP condition.¹³

¹³ It is noteworthy that there is a second parsing-based account of the facilitation effect in the CP condition, not considered here so far. According to this account, the disproportional difficulty in the DP condition is due to a super-additive interaction between the presence of a complex subject and the existence of a movement dependency across it, the combination of which creates a computational bottleneck. Under this view, the crucial effect does not reflect particularly fast reading times in the CP extraction condition but unexpectedly long reading times in the DP extraction condition. In order to evaluate this account, a post hoc test was performed. Because, by hypothesis, the interaction is the super-additive combination of the relative complexity of the DP structure and the movement dependency, the size of the critical effect at region 6 should positively correlate with the complexity of the DP structure relative to the CP structure in the absence of movement. In other words, for each item, the greater the relative difficulty of the DP structure in the absence of extraction the greater the size of the interaction at region 6

How exactly does the presence of an intermediate landing site in SpecCP confer a processing advantage at the ultimate gap site? The linking hypothesis between properties of the grammar and the parser presupposed here is that the parser needs to construct a syntactic representation that obeys the well-formedness requirements of the grammar. If the grammar requires the existence of an intermediate trace in SpecCP, then the representation constructed by the parser needs to provide such a trace. Integrating a representation of the filler into this intermediate position makes necessary a retrieval of the filler from memory and thereby reactivates it. One way of conceiving this reactivation is as an *anti-locality effect*, whereby a boost in the activation of an element strengthens this element's representation in memory and consequently facilitates subsequent access to it (e.g., Lewis & Vasishth 2005, Vasishth & Lewis 2006, Nicenboim, Vasishth, Gattei, Sigman & Kliegl 2015). The term anti-locality effect stems from the findings that an increased distance between the initial encoding of an element in memory and its subsequent retrieval can in fact facilitate the retrieval if the intervening structure reactivates the element and thereby boosts its memory representation. The effect of intermediate landing sites can be conceived of as parallel in nature. The reactivation boost conferred to the memory representation of a filler by an intermediate trace encoding of this filler aids and facilitates its subsequent retrieval at the ultimate gap site. In Nicenboim et al.'s (2015) terminology, encoding of the intermediate trace *preactivates* the filler. This manifests itself as a relatively fast response times relative to the DP control structures, which lack an intermediate representation of the filler and in which therefore no reactivation takes place. The facilitation in the CP structures is hence a consequence of both properties of the grammar as well as the parser. First, the grammar enforces the existence of an intermediate representation of the filler in the embedded

is predicted to be, assuming that the complexity contributed by the movement dependency is roughly equal for all items. To evaluate this prediction, the residual reading times per structure type for regions 2 through 4 for the no-move conditions of each item were summed up. This yields a measure of the respective difficulty of the CP and DP structures in the regions preceding the critical verb in the absence of a movement dependency. To arrive at the difficulty of the DP condition relative to the CP condition, the difference of the two sums was calculated for each item. In addition, to estimate the size of the critical interaction between extraction and intervener type at region 6, the increase in log reading times from the non-extraction to the extraction condition for CP structures was subtracted from the same difference of the DP structures for each item. According to the alternative account outlined above, the size of the critical effect in region 6 should be positively correlated with the relative difficulty of the DP condition in regions 2 through 4. To assess this prediction, a linear model was devised that regressed the structural difficulty scores against the interaction scores. The model did not revealed any reliable relationship between the two measures ($r = -0.07$, $adjusted R^2 = -0.02$, $F(1, 45) = 0.22$, $p > .5$). It is worth noting that the numerical trend was in the direction opposite of what the structural complexity account predicts. The results of Experiment 1 therefore do not support the view that the G&W effect is simply a non-cumulative combination of the two main effects.

SpecCP. Second, given general mechanisms governing fluctuating activation levels in memory during sentence parsing, this intermediate representation boosts the activation level of the filler in memory and thus aids its retrieval at the final gap site.

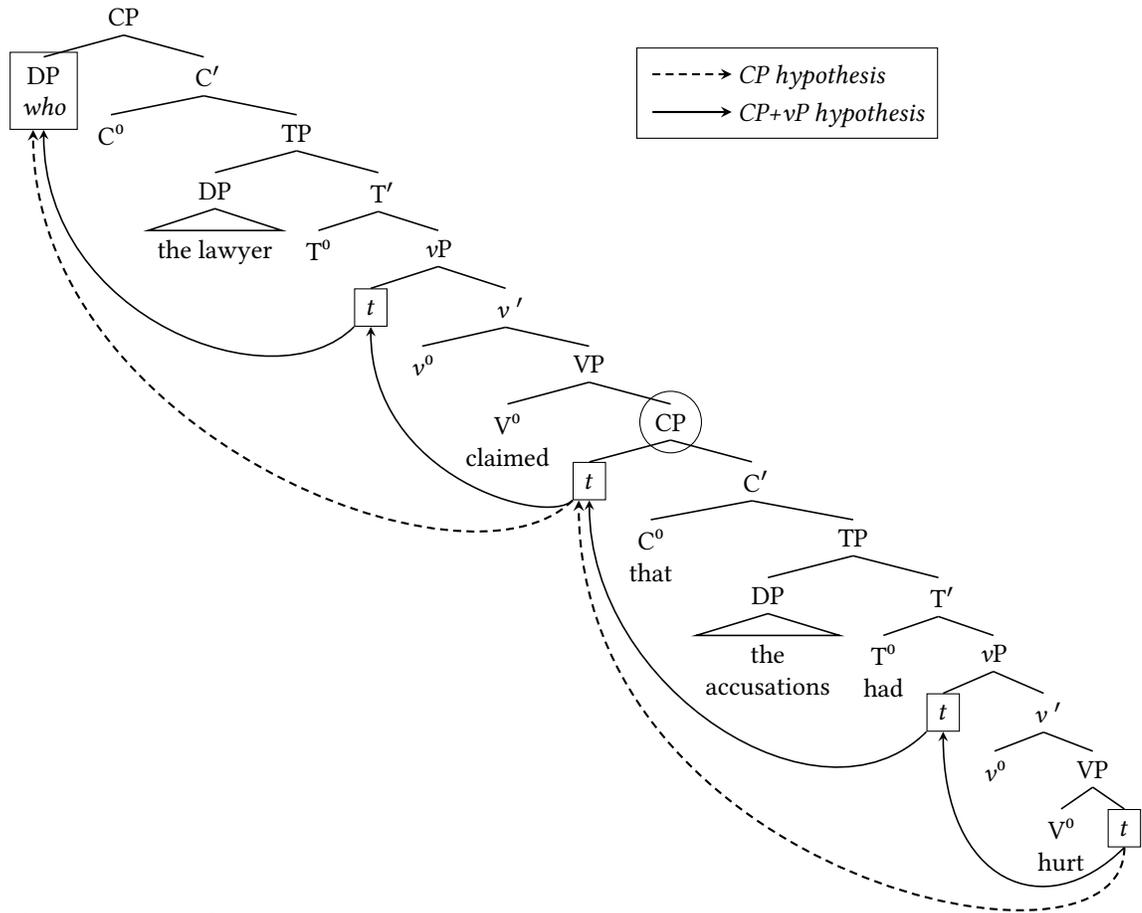
7.5 Experiment 2

G&W's results as well as the ones of Experiment 1 provide evidence for successive-cyclic movement through Spec,CP. Experiment 2 extends their basic experimental design to ask whether intermediate traces are limited to SpecCP (as argued in chapter 6) or also occur in SpecvP (the standard phase-based view). Let us refer to the former view as the **CP HYPOTHESIS** and to the latter position as the **CP+vP HYPOTHESIS**.

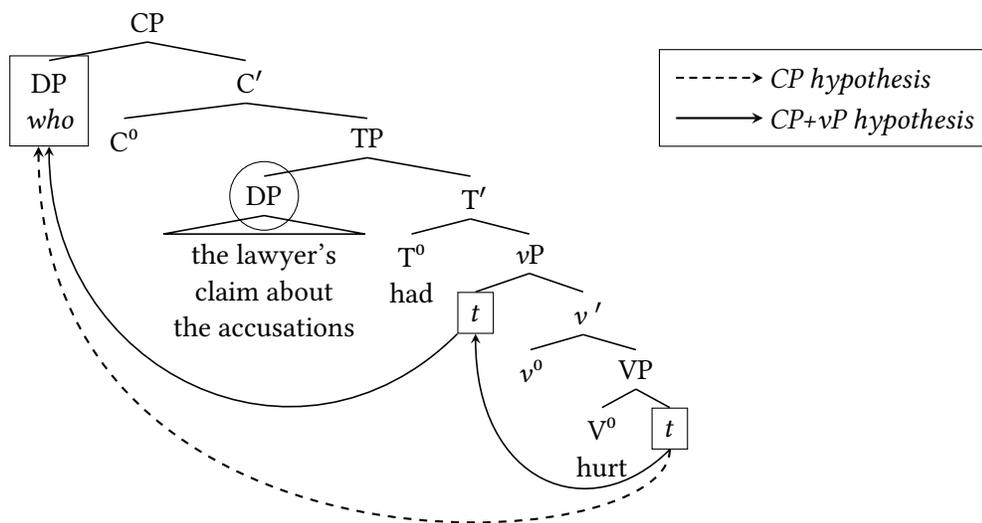
- (633) a. **CP Hypothesis**
Only C^0 is a phase head and intermediate landing sites are therefore created only at CP edges.
- b. **CP+vP Hypothesis**
 C^0 as well as v^0 are phase heads and consequently successive-cyclic movement targets SpecCP and SpecvP.

Returning to the structures investigated by G&W and in Experiment 1, the two hypotheses predict a different number of intermediate landing sites in the CP and the DP structures. (634) and (635) give the schematic structures and the movement steps necessary to derive them under each hypothesis. The dashed lines indicate the movement path under the CP hypothesis; There is a single intermediate gap in the CP condition and none in the DP condition. The solid lines indicate movement under the CP+vP hypothesis: There are three intermediate gaps in the CP structure, and one in the DP structure.

(634) Structure of CP condition in Experiment 1



(635) Structure of DP condition in Experiment 1



The results of Experiment 1 do not necessarily allow us to distinguish between the two hypotheses. They are clearly compatible with the predictions of the CP hypothesis: Because there exists an intermediate trace in the CP structure but not the DP structure, retrieval of the moved

element at the final trace position is easier in the CP structure, thus giving rise to a smaller reading time increase than in the DP structure.

Whether or not they are also compatible with the CP+*v*P hypothesis depends on the mechanics of how the moved element is reactivated over the course of the dependency. For the sake of the argument, consider the simplest version: When the ultimate trace site is encountered, the parser has to retreat to its closest antecedent – either the moved element itself, or an intermediate trace. On the CP+*v*P hypothesis, the most recent intermediate trace in both the CP and the DP structures lies in the *v*P edge immediately above the verb that hosts the trace. In other words, the distance between the ultimate trace and the closest intermediate trace is identical in the two structures. If retrieving the antecedent amounted to a search for the most recent trace, there should be no difference between the two structures, contrary to fact. On this simple view of antecedent retrieval, the results of Experiment 1 favor the CP hypothesis.

There is, however, reason to believe that this picture of how the antecedent is retrieved is overly simplistic. Based on anti-locality effects in processing, Vasishth & Lewis (2006) argue that it is not merely the distance to the closest antecedent that affects processing speed, but also the number of times an element has been previously activated. Applied to the structures at hand, it is plausible in light of Vasishth & Lewis' (2006) findings that the retrieval of a moved element is inversely related to the number of intermediate traces of this element. In other words, the more intermediate traces exist between the overt position of the moved item and its ultimate trace, the faster the postulation and interpretation of this trace proceeds. This view thus contrasts with the simpler one just considered in that intermediate landing sites have a *cumulative* effect on the processing speed at the trace position.¹⁴

Against this view of antecedent retrieval, let us consider again the predictions of the CP and the CP+*v*P hypothesis for the structures in (634) and (635). The predictions of the CP hypothesis remain unchanged: Because there is only a single intermediate trace in the CP structure and none

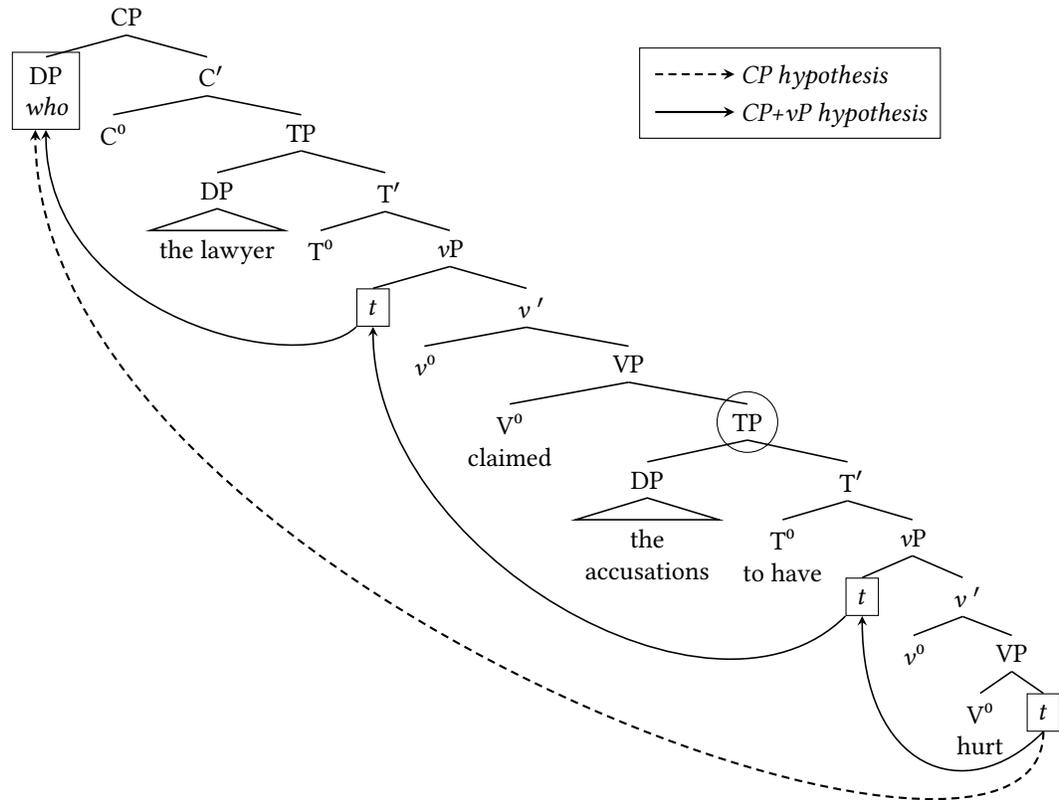
¹⁴ In addition to the empirical evidence by Vasishth & Lewis (2006), this view is very plausible in light of current accounts of the mechanisms underlying retrieval processes in online parsing (see, among others, Stevenson 1994, Gordon et al. 2001, McElree et al. 2003, Lewis et al. 2006, Wagers & Phillips 2009, 2013, Bartek et al. 2011). According to this family of theories, the ease with which an item is retrieved from memory is affected by this element's activation level and interference from similar encodings. An element's activation level is subject to decay over time. Retrieval of an element boosts its activation level and thereby counteracts the effects of decay and interference. Crucially, reactivation is taken to be cumulative so that a sequence of reactivations boost an element's activation level to a greater extent than a single reactivation, thus aiding subsequent retrieval.

in the DP structure, cumulative and non-cumulative reactivation are indistinguishable. The CP+ ν P hypothesis, by contrast, predicts that the antecedent is intermediately reactivated three times in the CP structure (corresponding to the three intermediate landing sites), but only once in the DP structure. If reactivation is cumulative, the antecedent's activation level will be greater in the CP than in the DP structure at the point at which the ultimate trace is encountered, and integration of this trace will as a result be faster in the CP structure than the DP structure. As both hypotheses thus derive the critical effect, there is so far no evidence to distinguish between them if reactivation is assumed to be cumulative.

Experiment 2 is intended to distinguish between the two hypotheses by comparing the CP and DP structure to a third structure which contains an additional ν P layer, but *no* additional CP layer. The CP hypothesis and the CP+ ν P hypothesis differ in the effect an additional ν P layer is predicted to have on retrieval times. Consider first the CP hypothesis: Because no intermediate gap is created at ν P edges on this account, a ν P layer will not lead to filler reactivation, and hence should not facilitate filler retrieval at the gap site. By contrast, the CP+ ν P hypothesis requires Spec, ν P to contain an intermediate trace. This trace should reactivate the filler and facilitate subsequent filler retrieval at the ultimate gap site. On the CP+ ν P hypothesis, then, a ν layer should have effects entirely parallel to those of a C layer. To assess these predictions, Experiment 2 investigates the processing of nonfinite-clause embedding structures. In particular, it focuses on extraction out of *exceptional case marking* (ECM) constructions like (636), which is modeled after the sample stimuli in (634) and (635) above. ECM constructions were chosen because it is fairly uncontroversial that they lack a CP layer. Other infinitival complementation structures, like control, are commonly taken to contain a CP (e.g. Landau 2004), and are hence unsuitable to distinguish between the two hypotheses.

The central characteristic of ECM constructions is that the embedded clause is not a CP, but a TP (circled in (636)), and structures of this type will subsequently be referred to as the **TP CONDITION**. Due to the removal of the CP, the CP hypothesis predicts no intermediate landing site in (636). As two ν Ps are crossed, the CP+ ν P hypothesis predicts two intermediate landing sites in the TP condition.

(636) *Structure of TP condition*



The number of intermediate gaps postulated by the two hypotheses for each of the three structures are summarized in (637). On the null assumption that all intermediate gaps have identical effects on filler retrieval, the two hypotheses make different predictions about the difficulty of integrating the filler in the three constructions. On the CP+*vP* hypothesis, the additional *vP* layer should have the same effect as a CP layer: It should reactivate the filler and thereby facilitate its retrieval at the ultimate gap site. On the CP hypothesis, on the other hand, *vP* layers should critically differ from CP layer in that only the latter facilitates filler retrieval downstream. Given cumulative reactivation, the more intermediate landing sites there are, the smaller the reading time increase at the gap site is predicted to be. According to the CP hypothesis, successive-cyclic movement is limited to the CP structure and the trace should consequently be more easily integrated in the CP structure than in the other two. This pattern should manifest itself in a smaller reading time increase in the CP structure. The CP+*vP* hypothesis, on the other hand, predicts three intermediate traces in the CP structure, two in the TP structure and one in the DP structure. The ease of integrating the

trace should mirror this order: The reading time increase should be smallest in the CP structure, larger in the TP structure and largest in the DP structure:

(637) *Predicted number of intermediate traces per structure*

	CP hypothesis	CP+vP hypothesis
CP structure (634)	1	3
DP structure (635)	0	1
TP structure (636)	0	2

Prediction:
 Reading time increase: {DP, TP} > CP DP > TP > CP

To summarize, the critical prediction of the CP hypothesis is that movement in the CP condition is easier than in the other two, as reflected in the reading time increase. The critical prediction of the CP+vP hypothesis is that the movement in the TP structure is easier than in the DP condition.

An important feature of (637) is that both the CP hypothesis and the CP+vP hypothesis make testable predictions. As I have discussed in section 7.2 above, this is not the case of traditional diagnostics for intermediate landing sites, as those are inherently biased in the sense that they are unable to detect the absence of an intermediate landing site. In the experimental design just outlined, on the other hand, the two views makes distinct predictions, both of which are not null results. Both hypotheses are therefore testable and falsifiable. This feature of the design allows us to overcome the bias problem that arises with the more standard diagnostics commonly employed in the literature.

7.5.1 Method

Materials

Thirty-six items like the one in (638) were constructed, out of which a subset of thirty with closely matching plausibility ratings between conditions was selected, the remaining six being discarded. The details of the norming procedure are described in Appendix A. The experiment manipulated INTERVENER TYPE (*CP* vs. *DP* vs. *TP*) and MOVEMENT ([+*move*] vs. [−*move*]). To decrease the processing load at the critical verb and to increase comparability to Experiment 1, a general scene-setting context sentence preceded each target sentence. As in Experiment 1, this context sentence

gave lexical information about the embedded verbal region in the corresponding target sentence. All sentences were followed by a comprehension question about the semantic relations in the sentence. All questions were multiple choice and both possible answers occurred in the target sentence. The order of the answer options was randomized.

- (638) *Context:* At the press conference last Monday several people became very agitated.
- a. *CP intervener*
 - (i) [+move]
The journalist who the union member believed that the tax policy had intensely agitated was planning a series of articles.
 - (ii) [-move]
The union member believed that the tax policy had intensely agitated the journalist who was planning a series of articles.
 - b. *DP intervener*
 - (i) [+move]
The journalist who the union member's beliefs about the tax policy had intensely agitated was planning a series of articles.
 - (ii) [-move]
The union member's beliefs about the tax policy had intensely agitated the journalist who was planning a series of articles.
 - c. *TP intervener*
 - (i) [+move]
The journalist who the union member believed the tax policy to have intensely agitated was planning a series of articles
 - (ii) [-move]
The union member believed the tax policy to have intensely agitated the journalist who was planning a series of articles.

Comprehension question: Who believed something regarding the tax policy?
the union member – the journalist

Some general remarks about the design of the stimuli are in order. Most items contained the auxiliary *have* (*to have* in the TP condition). Other items contained the future auxiliary *will* (*to* in the TP condition). Because Experiment 2 seeks to test whether there is an intermediate landing site at the vP edge, it is crucial to dissociate the region containing the ultimate trace from the one containing the left edge of the vP . It is uncontroversial that finite auxiliaries and the infinitival *to* occupy T. An intermediate landing site in the vP edge can hence be reliably postulated when the

parser reaches this segment of the clause. An adverb was inserted between the auxiliary and the verb hosting the trace to separate the hypothetical landing site in *vP* from the final trace position.

In addition to the 30 target sentences, the experiment contained 24 sentences of an experiment not reported here and 36 distractor sentences. All stimuli matched the target sentences in length and complexity.

Participants

The experiment reported here involved 162 participants, recruited via Amazon Mechanical Turk. All were naïve to the purpose of the experiment and each received a compensation of USD 1.

Procedure

The experiment was conducted using the online experiment platform Ibex (Drummond 2013) and employed a region-by-region self-paced noncumulative moving-window task (Just et al. 1982). Stimuli were regioned as shown in Table 7.5. At the beginning of each trial, participants saw the theme-setting context sentence, which was displayed in its entirety. Upon pressing the space bar, the context sentence was replaced by dashes masking the regions of the target sentences. Pressing the space bar caused the dashes in the first region to be replaced by the actual content of the region. When the space bar was pressed again this region reverted back to dashes and the next region appeared. Participants traversed through the entire sentence by repeatedly pressing the space bar. Pressing the space bar after the final region had been displayed caused the dashes to disappear and a comprehension question accompanied by two possible answers appeared on the screen. Participants selected the answer on the left by pressing the ‘f’ key and the one on the right by pressing the ‘j’ key. No feedback on answer accuracy was given. After every twelve trials the participant had to take a ten-second break and could rest for longer if they so desired. The total experiment thus contained seven of these breaks. The experiment took about thirty-five minutes.

Analysis

Preprocessing and outlier rejection followed the same procedure as Experiment 1.

Table 7.5. Regioning of stimuli in Experiment 2 ($[\pm mv] = [\pm move]$)

		Region								
		1	2	3	4	5	6	7	8	9
CP	$[-mv]$		The union member believed	that	the tax policy	had	intensely	agitated	the journalist who	was planning a series of articles
	$[+mv]$	The journalist who	the union member believed	that	the tax policy	had	intensely	agitated	was planning	a series of articles
DP	$[-mv]$		The union member's beliefs	about	the tax policy	had	intensely	agitated	the journalist who	was planning a series of articles
	$[+mv]$	The journalist who	the union member's beliefs	about	the tax policy	had	intensely	agitated	was planning	a series of articles
TP	$[-mv]$		The union member believed		the tax policy	to have	intensely	agitated	the journalist who	was planning a series of articles
	$[+mv]$	The journalist who	the union member believed		the tax policy	to have	intensely	agitated	was planning	a series of articles

7.5.2 Results

Of the 162 participants, one was excluded for not being a native speaker of English. No participant took the experiment twice. One item was coded incorrectly, and seen twice by seven participants. This item was therefore excluded from all analyses.¹⁵ Moreover, in order to be maximally conservative, and safeguard against the possibility that encountering the same item twice might have affected the response pattern of the affected subjects, these seven subjects were also excluded from analysis.¹⁶ 53 of the remaining participants had to be excluded from analysis for falling below the accuracy threshold.¹⁷

Outlier rejection based on absolute reading time thresholds eliminated less than 0.4% of the data. The z -score based outlier rejected excluded less than 1% of the data. Reading time data were analyzed using a 2×3 -factorial LME model crossing the factors MOVEMENT ($[-move]$ vs. $[+move]$) and INTERVENER (CP vs. DP vs. TP). Accuracy data were analyzed using a parallel logistic LME model. All covariates were numerically coded. *Movement* was sum-coded ($[-move] = -.5$, $[+move] = .5$). The predictor *intervener* used Helmert coding: The first contrast compared the CP condition to the mean of the DP and TP conditions ($CP = -2/3$, $DP = 1/3$, $TP = 1/3$). The second contrast compared the DP and the TP condition ($CP = 0$, $TP = -.5$, $DP = .5$). This coding makes sense in light of the critical predictions of the two hypotheses. The first contrast tests whether the CP condition was read faster than the other two (the critical prediction of the CP hypothesis) and is reported as *IntervCP-TPDP*. The second contrast tests whether the TP condition was faster than the DP condition (the critical prediction of the CP+vP hypothesis) and is reported as *IntervTP-DP*. Because the TP condition had no observations region 3, the factor *intervener* was sum-coded in this region ($CP = -.5$, $DP = .5$).

¹⁵ Eliminating this item did not introduce a plausibility confound (p 's $> .5$), cf. Appendix A.

¹⁶ The inclusion or exclusion of these subjects did not affect the statistical patterns reported here. The coding error mentioned in the text also led to one item being somewhat less balanced across conditions. To be conservative, analyses were conducted both with and without this item and the patterns of significance were identical in all regions. I report here the results with this item included.

¹⁷ As in Experiment 1, a relatively stringent accuracy criterion was chosen as the most reliable method of ensuring that the results reflect genuine properties of the parsing of the syntactic structures in questions, and not properties of shallow parsing. The critical statistical results reported here also obtain if no accuracy-based exclusion of participants is applied.

Comprehension question accuracy

The mean accuracy on the comprehension questions was 81%. Accuracy by condition is given in Table 7.6. Logistic LME modeling revealed that (i) accuracy was lower in the [+move] conditions ($\hat{\beta} = -1.6, z = -12, p < .001$); (ii) accuracy in the CP conditions was higher than the mean of the DP and TP conditions ($\hat{\beta} = -.7, z = -4.8, p < .001$); and (iii) accuracy in the TP conditions was higher than in the DP conditions ($\hat{\beta} = -.6, z = -3.3, p < .001$). Moreover, while the accuracy drop in the [+move] conditions did not differ between CP structures and the mean of TP and DP structures ($\hat{\beta} = .3, z = 1.2, p > .2$), this drop was greater in the TP condition than in the DP condition ($\hat{\beta} = .6, z = 2.1, p < .05$). This finding will be picked up in the discussion section.

Table 7.6. Mean answer accuracy by condition in Experiment 2

	CP	DP	TP
[-move]	.95	.84	.92
[+move]	.78	.67	.72

Reading times

The mean reading times per region and condition as well as the residual reading times for regions that differed in length are given in Table 7.7. A plot of the residual reading times by condition and region is given in Figure 7.3. The standard error by condition and region as a measure of data variability is provided in Appendix B.

To distinguish between the CP hypothesis and the CP+vP hypothesis, the crucial evidence is the pattern at the region containing the trace (region 7) and the spillover region (region 8). Of primary interest is the reading time difference between the movement condition and the no-movement controls for each syntactic structure. Table 7.8 provides the results of the analyses of these two regions. In the interest of space, only the crucial gap and spillover regions are discussed in the main text. Analyses and discussions all other regions are provided in Appendix B.

In region 7 (the gap region), reading times in the TP condition were greater than in the DP condition while the CP condition did not differ from the mean of the other two. Crucially, there was an interaction between the reading time increase due to movement noted above and the intervener type: The reading time increase was greater in the TP condition than in the DP condition. Pairwise comparisons made it clear that this pattern was driven by an exceptionally high reading time increase

Table 7.7. Mean raw reading times by condition and region in ms computed over participants for Experiment 2. For regions that differed in length between conditions, residual reading times are additionally provided in parentheses ('Intv' = Intervener, 'Mvmt' = Movement)

Intv	Mvmt	Region								
		1	2	3	4	5	6	7	8	9
CP	[-move]	—	844	494(10)	657	493(29)	495	522	732(-46)	1270(90)
	[+move]	699	949	559(75)	717	495(31)	502	548	677(-2)	877(-1)
DP	[-move]	—	1032	520(12)	675	495(30)	499	514	705(-66)	1215(38)
	[+move]	706	1190	618(111)	770	515(50)	507	536	714(39)	903(36)
TP	[-move]	—	852	—	775	543(9)	523	525	701(-73)	1275(99)
	[+move]	691	1015	—	833	579(44)	527	593	698(26)	888(10)

in the TP condition: The increase was greater than in the CP condition, marginally significant by participants ($t_1(100) = 1.89, p = .06$) and fully significant by items ($t_2(28) = 2.8, p < .01$) but did not differ between the CP and DP conditions ($t_1(100) = .2, p > .5$; $t_2(28) = .3, p > .5$).

In region 8 (the spillover region), the only significant effect was an interaction such that the reading time difference between movement and non-movement structures was significantly smaller

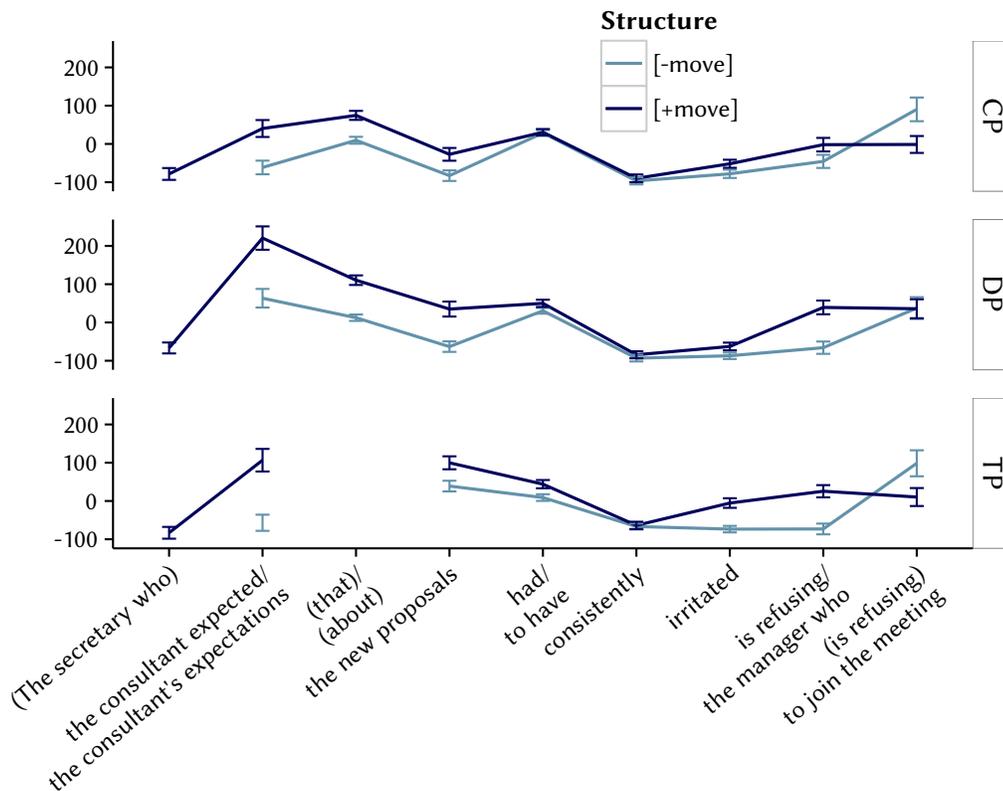


Figure 7.3. Residual reading times by condition and region in ms with by-participant standard errors

Table 7.8. Coefficient estimates and corresponding t -values for the linear mixed effects model analyses of log reading times in critical regions of Experiment 2. *Mvmt* refers to the sum-coded factor *movement*; *IntervCP-TPDP* compares the CP condition to the mean of the TP and DP conditions and *IntervTP-DP* compares the TP condition to the DP condition. *Mvmt:IntCP-TPDP* and *Mvmt:IntTP-DP* refer to the interaction between *Mvmt* and *IntervCP-TPDP* and *IntervTP-DP*, respectively. Cells with $p < .05$ are shaded. For statistical analyses of all regions, see Appendix B.

	Region			
	7		8	
	$\hat{\beta}$	t	$\hat{\beta}$	t
Movement	0.053	4.00	-0.039	-1.11
IntervCP-TPDP	0.013	1.13	-0.000	-0.00
IntervTP-DP	-0.042	-3.65	-0.003	-0.21
Mvmt:IntervCP-TPDP	0.024	1.09	0.059	2.08
Mvmt:IntvTP-DP	-0.049	-2.10	-0.018	-0.61

in the CP condition than in the mean of the other two. The latter two did not differ from each other. Because the length of the region differed between movement and non-movement structures, an additional LME analysis of the residual reading times was carried out to validate this effect. This analysis corroborated the interaction ($\hat{\beta} = 59, t = 2.5, p < .05$). It also showed that the [+move] conditions were read more slowly than the [-move] ones ($\hat{\beta} = 81, t = 5.2, p < .001$). Just like in Experiment 1, the fact that the region was longer in the [-move] condition than in the [+move] one renders the effect of movement visible only in the residual reading times.¹⁸

Figure 7.4 plots the crucial reading time increase between the [-move] and [+move] conditions in the gap and spillover regions. As the analyses just described reveal, (i) there is a positive increase in all conditions, (ii) the increase is substantially larger for TP structures in the gap region and (iii) the increase is substantially smaller for CP structures in the spillover region.

¹⁸ Although not critical for the evaluation of the two hypotheses considered here, a post hoc analysis was carried out to compare the reading time increase in the CP condition to that in the DP condition in the spillover region. The rationale for this test was to replicate G&W's key finding, replicated in Experiment 1 above, that the reading time increase is greater in the DP condition than in the CP condition. Pairwise t -tests comparing the reading time difference between [+move] and [-move] conditions in CP and DP structures were carried out on the log-transformed and residual reading times. Analysis of the former revealed that the difference was marginally smaller in the CP condition than in the DP condition ($t_1(100) = 1.62, p = .11; t_2(28) = 2.11, p < .05$). This difference was fully significant in the residual reading times ($t_1(100) = 2.23, p < .05; t_2(28) = 2.61, p < .05$). The results of Experiment 2 are thus fully consistent with those of Experiment 1.

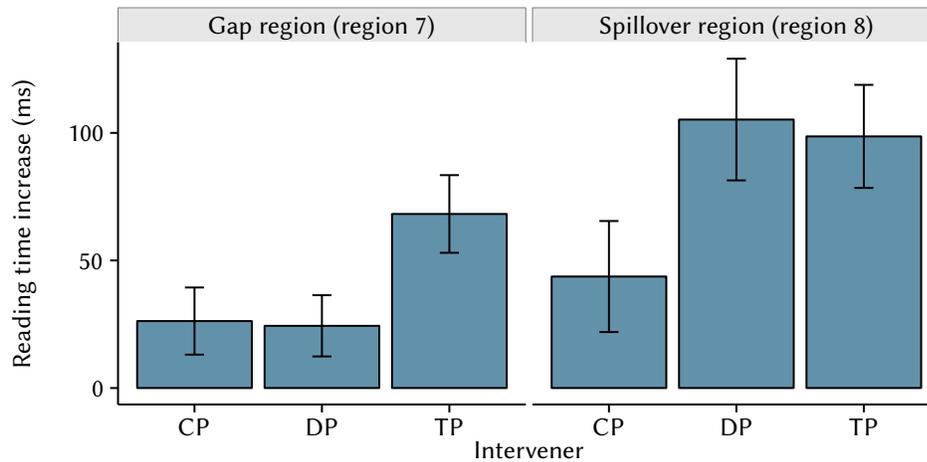


Figure 7.4. Mean increase in residual reading times between [-move] and [+move] conditions in gap and spillover regions and by-participant standard errors in Experiment 2 by intervener

7.5.3 Discussion

Recall from (637) the predictions of the two hypotheses: The critical prediction of the CP hypothesis is that a CP crossed by movement will lead to an intermediate reactivation of the filler, while a ν P has no such effect. On the CP+ ν P hypothesis, on the other hand, both CPs and ν Ps are expected to reactivate the filler. The CP hypothesis thus predicts that the reading increase in the CP condition is smaller than in the DP and the TP condition, whereas the CP+ ν P hypothesis predicts that the reading time increase is greater in the DP condition than in the TP condition.

The results of Experiment 2 show facilitation in the CP structure relative to the DP and TP structures. In line with the findings of Gibson & Warren (2004) and Experiment 1 above, this result provides evidence that crossing a CP layer reactivates the filler, indicating the presence of an intermediate landing site in Spec,CP. Crucially, however, there was no indication that crossing a ν P has a similar effect. Not only was the retrieval speed in the TP condition not faster than in the DP control structures, retrieval times were in fact *slower* in the TP structure than in the DP structure. This indicates that the addition of a ν P layer in the TP structure does *not* aid filler retrieval at the gap site, but in fact incurs a greater difficulty of filler retrieval, a result to which we will return shortly. The crucial finding of the experiment is thus that CPs and ν P differ substantially with respect to the processing of a filler-gap dependency: While a CP layer leads to reactivation of the filler, a ν P layer does not.

These findings are in line with predictions of the CP hypothesis. Because on this hypothesis only CPs host intermediate landing sites, the filler is intermediately reactivated only in the CP structure, but not in the TP or DP structures, as in the latter two structures no CP layer is crossed by movement. On the other hand, the CP+vP hypothesis postulates intermediate gaps in both CP and vP. If intermediate gaps in vP lead to reactivation of the filler just like intermediate gaps in CP do, then filler integration should be easier in the TP than in the DP condition (recall the reasoning in (637)). This prediction was not borne out. In sum, the evidence for intermediate reactivation in Spec,CP and the simultaneous absence of parallel reactivation in Spec,vP thus supports the CP hypothesis.

While the results thus bear out the critical prediction of the CP hypothesis because only an embedded CP layer correlated with facilitated filler retrieval, the overall difficulty of the TP condition was unexpected under both hypotheses and therefore deserves consideration. The specific difficulty of movement in the TP structures relative to their nonmovement controls is not limited to the reading time patterns, but was also evident in the accuracy analysis. Note that this increased difficulty of the TP structures relative to the DP structures directly contradicts the predictions of the CP+vP hypothesis, which after all predicts filler retrieval to be easier in TP structures than in DP structures. By contrast, the CP hypothesis does not by itself make a commitment with regard to the relation between the TP and DP structures. The failure of the CP hypothesis to predict the TP–DP contrast therefore does not put it on a par with the CP+vP hypothesis: Whereas the former merely fails to predict an observed contrast, the latter critically predicts the opposite of the observed results. Yet the question remains as to why filler retrieval should be exceptionally hard in TP structures. One plausible source of this difficulty is *structural distance*.¹⁹ Although the materials used in Experiment 2 controlled the *linear* distance between the overt filler and its gap, the *structural* distance between the two differed across conditions. Concretely, in the DP condition, the moved element was in the same clause as its trace (as not clause boundary was crossed in this condition). In the TP condition, on the other hand, the moved element resided in a higher clause than its trace. Retrieval of the filler thus has to traverse a clause boundary in the TP condition. That

¹⁹ There is evidence that increased structural distance between a moved element and its trace is associated with greater retrieval difficulty at the gap site (e.g. Lin 2006, Lin & Bever 2006).

the reading time increase was greater in the TP than in the DP condition is likely a reflex of this difference in the structural length of the dependency.

Importantly, this structural distance-based account of the TP–DP difference does not affect the argument for the CP hypothesis, nor is it able to reconcile the CP+*v*P hypothesis with the results. While structural difference induces a penalty in the TP condition relative to the DP condition, it must not do so in the CP condition (or else the CP and TP condition would pattern alike). If movement in the CP condition were crossclausal in the same way that movement in the TP condition is, they should pattern alike in being harder than the DP condition, contrary to fact. The ease of postulating the trace in the CP condition therefore shows that the structural distance is shorter in this condition than in the TP condition. This is, of course, precisely what successive cyclicity through Spec,CP claims. In other words, the penalty for crossclausal movement is absent in the CP condition precisely because there exists an intermediate trace within the same clause. An appeal to structural distance hence does not obviate the need for successive cyclicity through Spec,CP.

Equally importantly, a structural distance account in conjunction with the CP+*v*P hypothesis does not alter the incorrect prediction of this hypothesis. To see why, consider again the structures in (634–636). On the CP+*v*P hypothesis, the intermediate trace closest to the ultimate trace site is at the edge of the embedded *v*P in all three structures. As a consequence, the structural distance between the trace and its closest antecedent is *identical* in all three structures. Thus, integrating structural distance into the analysis does not affect the predictions of the CP+*v*P hypothesis. In particular, it has no impact on the incorrect prediction that the integration of the trace should be easier in the TP structure than in the DP structure. In sum, while reference to structural distance does provide an account for the difficulty in the TP condition, it does not modify the conclusion drawn above: The results of Experiment 2 support the CP hypothesis because the application of structural distance is successful only on this hypothesis.

Three potential limitations of Experiment 2 deserve to be discussed. First, while the results of Experiment 2 show that only CP layers lead to facilitated retrieval of the filler at the gap site, it is a priori possible that both CPs and Ps host intermediate landing sites, but that only traces in Spec,CP reactivate the filler. A difference between CP and P traces along this line would reconcile the CP+P hypothesis with the results of Experiment 2 as it would effectively render P traces irrelevant for the purposes of filler retrieval. Whether this account constitutes a viable alternative to the CP

hypothesis depends on whether there is independent evidence that intermediate traces in CP and P differ in whether or not they lead to filler reactivation. It is not evident that there is. First, note that such a difference cannot plausibly emerge from the *syntax* of successive cyclicity because successive-cyclic movement through CP and through P are formed by the same syntactic mechanism. Second, it is not clear that there is independent *processing* motivation for such a contrast. In fact, standard parsing accounts that incorporate fluctuating activation explicitly assume that retrieval of an item invariably leads to reactivation of this item (e.g., Lewis & Vasishth 2005, Vasishth & Lewis 2006) and it is not clear on this view how the filler could be retrieved at P without concomitantly reactivating it.²⁰ Thus, although a full evaluation of this alternative hypothesis is beyond the scope of the experiments and results reported here, this hypothesis requires independent evidence for the purported difference between CP and P traces or their interaction with general principles of sentence parsing. Only concrete alternative proposals can be evaluated with respect to this question. In the absence of such evidence, I will adopt here the null assumption that retrieval induces reactivation regardless of the context it applies in, and that CP and P traces are therefore identical in the relevant respects. On this null assumption, the lack of reactivation in the TP condition does provide evidence against the presence of an intermediate gap in Spec,P, and hence in favor of the CP hypothesis.

A second potential limitation of the results of Experiment 2 is that the higher verb is a CP/DP-verb (using the terminology of Experiment 1), a direct consequence of the experimental design. Given the results of Experiment 1, it is expected that the parser initially postulates a trace of the moved element as the object of these verbs in the TP and CP conditions, but not in the DP condition. It is a priori possible that the filled-gap effect in the TP condition has affected the results at the actual gap site, a possibility that warrants consideration. While the results of Experiment 2 do not allow us to directly investigate the impact of this confound, there is reason to doubt that the crucial pattern at the gap and spillover regions are an artifact of premature gap filling at the higher verb.

²⁰ One line of analysis could be that intermediate traces give rise to reactivation only if they reside at a clause edge, a line of account that would liken intermediate reactivation to clausal wrap-up effects. While a priori possible, it remains unclear why intermediate reactivation should be principally restricted to clause boundaries. Notice, for instance, that in contrast to wrap-up effects at the end of sentences, thematic processing in the higher clause is not completed by the time the CP of the lower clause is postulated, simply because this CP is itself a thematic argument of the higher verb. Moreover, this account would appear to be at variance with evidence for retrieval-induced activation of a non-moved constituents occurring in the middle of a clause (e.g., Vasishth & Lewis 2006).

In Experiment 1, verb type had an impact on the reading time pattern in the region following the higher verb, but not in the trace region. This indicates that structural revision is fast and a parse temporarily pursued in the higher clause has no discernable impact on the regions of interest, in concordance with the previous literature on structural reanalysis. Moreover, the crucial reading time increase is not modulated by the intervener structure in any region prior to the gap region (see Section B.2 in the supplementary materials), which all else equal makes it implausible that the pattern at the gap region is a sustained effect of gap filling at the higher verb. While there is thus no compelling reason to believe that the critical results in the TP condition are an artifact of premature gap filling, this property of the design does impose a certain limitation on the interpretation of the results.

A third limitation imposed by the experimental design is that the TP structure contained a sustained ambiguity in the regions following the ECM verb. A DP following an ECM verb could be either a direct object to this verb, the subject of an ECM infinitive, or the subject of an embedded finite clause (without a complementizer). In the CP condition, on the other hand, the complementizer and the following DP disambiguate the input towards the correct structure earlier. One might wonder whether the increased difficulty in the TP structure could be attributed to this sustained ambiguity. As mentioned, the experimental design makes it impossible to directly assess this alternative. But due the factorial design employed in the experiment, effects of baseline difficulty of the three structures are ‘filtered out’ in the crucial interactions. That is, the critical dependent measures are reading time increases in a structure relative to a non-movement baseline of that structure. Because these structural properties of ECM constructions also arise in the non-movement baselines, there is no a priori expectation that these properties should affect the reading time increase and hence the interaction terms. Moreover, the time course of the processing difficulty is unexpected under this account. As just mentioned, the difficulty in the TP structure is limited to the gap and the spillover regions (regions 7 and 8). On the ambiguity account, the ambiguous regions of the TP condition are the ones immediately following the embedding verb (region 4). The ambiguity account thus faces the challenge of explaining why the difficulty purportedly created by the structural indeterminacy does not arise until after this indeterminacy has been eliminated.

Notice also that the experimental results do not allow us to determine with precision when the intermediate trace is postulated during the course of parsing the sentence. The fact that the

presence of an intermediate trace confers an advantage when the ultimate trace is constructed entails that this intermediate trace is postulated before the final trace is encountered, hence not ‘retroactively’. Any more fine-grained determination of the precise time course is beyond the scope of the experiments reported here.

Finally, the relevance of the notion of cumulative reactivation deserves some remarks. As discussed above, cumulative reactivation refers to the hypothesis that a sequence of reactivations eases an element’s subsequent retrieval to a greater extent than a single reactivation. As also discussed above, the CP+P hypothesis is compatible with the results of G&W and Experiment 1 only if cumulative reactivation is adopted. However, even under cumulative reactivation the CP+P hypothesis fails to account for the results of Experiment 2, adopting, as before, the null assumption that all intermediate traces behave similarly. Consequently, the CP+P hypothesis makes incorrect predictions regardless of whether or not cumulative reactivation is adopted. The CP hypothesis, by contrast, makes correct predictions for both experiments reported here regardless of whether cumulative reactivation is adopted or not. This is trivially the case because on this hypothesis no structure investigated here involves more than a single intermediate landing site. Cumulative and non-cumulative reactivation are hence indistinguishable. This ensures maximal generality of the results obtained here: No stance on whether intermediate reactivation is cumulative or not is necessary for the experiments here to be interpretable in the context of the hypotheses.

7.6 General discussion

7.6.1 The distribution of phases

The preceding section have brought a novel type of empirical evidence to bear on the question of successive-cyclic movement and the distribution of phases. The processing evidence presented here corroborates the general conclusion that long movement dependencies are constructed successive-cyclically, in line with the overwhelming support for this view from traditional syntactic diagnostics and in line with the results of chapter 6. In addition to providing evidence for successive cyclicity in sentence processing, the results presented here shed light on the question of *where* successive-cyclic movement takes place, and hence on the distribution and location of the phase heads. On the

one hand, we have seen evidence that the edges of finite clauses are targeted by successive-cyclic movement. On the other hand, such evidence was altogether absent for vP s. This likewise aligns with the conclusions of chapter 6.

These results support the view arrived at on entirely independent grounds in chapter 6, where I have argued that C^0 is a phase head, but that v^0 is not. As discussed in the previous section, unless a principled reason can be motivated for why intermediate traces in Spec, vP should systematically differ from traces in Spec,CP, this result calls into question the existence of an intermediate landing site in Spec, vP . As a consequence, these findings challenge the view that vP edges act just like CP edges in enforcing intermediate touchdowns.²¹

An important limitation of traditional diagnostics for intermediate landing sites discussed in section 7.2 is that they are *biased*. Standard diagnostics of successive cyclicity like morpho-syntactic alternations, stranding of material, and so on, are inherently unable to provide evidence for the absence of an intermediate landing site. Quite generally, nothing can be concluded from the absence of evidence for successive cyclicity. But this in turn entails that claims that certain heads are phase heads are effectively unfalsifiable using these standard diagnostics. Sentence processing contributes a highly valuable tool to our inventory of phase diagnostics because it provides an unbiased test for intermediate landing sites. In the experimental design employed in Experiment 2, both the hypothesis that vP is a phase and the hypothesis that it is not made distinct and falsifiable predictions, thus offering us a fairer test for adjudicating between the two. As we saw, the evidence is consistent with the view that CP is a phase, but vP is not.

These results converge with the conclusions reached in section 6.3 on the basis of evidence from selective opacity effects and long-distance relationships like ϕ -agreement and *wh*-licensing. I have argued there that the assumption that vP constitutes a phase leads to syntactic domains that are too local for a variety of attested patterns. In particular, I have shown that the formulation of any reasonably comprehensive account of selective opacity effects becomes impossible unless access to information inside an already spelled-out phase is allowed, at the risk of rendering the

²¹ The results of Experiment 2 also provide evidence against successive cyclicity in ECM constructions more generally. In other words, these results not only cast doubts on successive cyclicity through Spec, vP , they also suggest that there is no clause-level cyclicity in TP clauses. This finding is inconsistent with recent claims that the highest projection of a clause is phasal, regardless of its category (e.g., Bošković 2014). To the extent that these conclusions are on the right track, they converge with the conclusion above that only C^0 is a phase head (at least within the verbal domain).

very notion of cyclic Spell-Out vacuous. Moreover, I have demonstrated that in-situ ϕ -agreement and *wh*-licensing is possible across an arbitrary number of ν P projections, clearly demonstrating that ν Ps are not designated locality domains. Finally, I have shown in section 6.4 that at least several previous arguments in favor of ν P phases are faulty and do not in fact support the conclusion that ν Ps are phasal. The processing evidence presented in this chapter provides independent evidence that converges in a rather striking way with these conclusions because it shows that intermediate landing sites are created only in SpecCP and not in Spec ν P. Thus, evidence from a variety of unrelated phenomena converges on the conclusion that ν P is not phasal, whereas CP is.

7.6.2 Implications for theories of parsing

The main focus of this chapter was the existence and role of successive cyclicity in sentence parsing. The evidence gathered in the course of investigating this question additionally sheds light on questions regarding general parsing strategies. This section will discuss some of these implications.

Experiments 1 and 2 investigated how the type of the verb that is crossed by a crossclausal movement dependency affects the processing of that dependency at the gap site. This manipulation provided evidence against a premature gap filling account of G&W's effect. In addition, the manipulations of these factors sheds light on how the parser employs them to decide on where to postulate traces.

We have seen evidence that subcategorization frequency and real world plausibility have less of an impact on trace postulation. The results of Experiment 1 provide evidence that the parser postulates a DP trace after DP-taking verbs like *claim* even if considerations of frequency and plausibility disfavor such a trace. This result does not, of course, entail that frequency and plausibility information is ignored in initial parsing (see, e.g., Gibson & Pearlmuter 1998). But it indicates that frequency and plausibility does not invariably prevent the parser from postulating an otherwise licit trace.

It is also worth emphasizing once more that the experiments reported here do not offer evidence for or against cumulative reactivation of a moved element. As discussed on p. 443, reactivation is cumulative if multiple reactivations of some item increase this item's activation level to a greater degree than a single reactivation. As already discussed in section 7.5.3, the results here are

compatible with both cumulative and non-cumulative reactivation and so are the conclusions with respect to the distribution successive cyclicity. Under non-cumulative reactivation, Experiment 1 provides evidence against the CP+vP hypothesis because the closest reactivation site in Spec,vP is equally far away from the actual trace position in both the CP and the DP condition. If cumulative reactivation is assumed, the results of Experiment 2 constitute evidence against this hypothesis. The CP hypothesis, on the other hand, makes the correct predictions for both experiments here regardless of whether reactivation is cumulative or not, simply because in every structure there is at most one intermediate reactivation site. Because cumulative and non-cumulative reactivation make different predictions only if there is more than one reactivation site, the difference between the two is irrelevant to the predictions of the CP hypothesis. As a consequence, the experiments do not offer evidence for or against either view of reactivation and the CP hypothesis emerges as superior under either view.

Lastly, the results here have implications for current theories of the retrieval processes involved in the online computation of long-distance dependencies. Recall from Experiment 2 that the integration of the trace incurred greater processing cost in the TP condition than in the DP condition. Because no CP was crossed in either condition, this contrast cannot be attributed to successive cyclicity. In other words, in both conditions the closest antecedent is the overt instance of the moved element. This differences in the retrieval difficulty in the TP and the DP condition is particularly illuminating against the background of current theorizing of the factors regulating retrieval processes in sentence comprehension. Traditional accounts of retrieval processes assume that elements within a local syntactic domain are easier to access than elements outside of this domain (Kimball 1973, Frazier 1978, Berwick & Weinberg 1984, Frazier & Clifton 1989, Gibson 1998, Sturt, Pickering & Crocker 1999). Alternative accounts of retrieval processes eschew a syntactic distinction between local and non-local domains and characterize retrieval difficulties solely in terms of an element's *activation decay* and *similarity-based interference* from other elements (e.g., McElree et al. 2003, Van Dyke & Lewis 2003, Lewis & Vasishth 2005, Lewis et al. 2006, Bartek et al. 2011). This family of accounts treats the availability of an element for retrieval as a function of its activation level, which is subject to decay over time, and the interference from similar elements that match a retrieval cue. In contrast to the former type of analysis, syntactic locality and domains do not enter directly into the computation of the ease of retrieving an element. Instead, recency

effects are recast as the indirect result of decay and interference: The greater the structural or linear distance between the encoding of an element and the point at which it needs to be retrieved the greater the effect of decay on this element's activation level, thus increasing retrieval difficulty. Similarly, greater distance corresponds to a greater number of similar elements that interfere with retrieval.

Against this background, the aforementioned contrast between the ease of retrieving the moved element in the TP and DP condition poses an interesting problem. Crucially, filler retrieval was harder in the TP condition than in the CP condition. This difference does not follow from similarity-based interference or decay alone. Because the DP and TP structures do not differ with respect to the linear distance between the moved element and its trace, the effect of decay should be identical. Whether or not the two structures are expected with respect to interference depends on the exact retrieval cues. If the additional noun in the DP condition (*beliefs* in (638)) matches these retrieval cues, the DP condition is expected to be harder than the TP condition. If it does not, no difference between the two structures is predicted. Yet the results of Experiment 2 showed that retrieval was harder in the TP condition than it was in the DP condition. Neither decay nor similarity-based interference alone derive this pattern. This indicates that there must be some other factor at play that is responsible for this difference.

As discussed at length in section 7.5.3, one factor that might plausibly underlie the difficulty in the TP condition is syntactic locality. While the moved element is equally distant from its trace linearly in the TP and DP structures, their structural distance differs considerably. Whereas the moved element is situated in the same clause as its trace in the latter structure, it is located in a higher clause in the former. If leaving the current clause increases the cost of retrieving an element, the disadvantage of the TP structure is accounted for. This analysis involves crucial reference to the syntactic structure separating the element to be retrieved and its trace. It requires that the retrieval process be affected by properties of syntactic domains and is hence not available to theories that deny a direct impact of syntactic structures on retrieval processes.

These considerations converge in a rather striking way with evidence from the processing of reflexives reported in Dillon, Chow, Wagers, Guo, Liu & Phillips (2014). Based on Mandarin Chinese, Dillon et al. (2014) investigated whether the retrieval of an antecedent for a reflexive is affected above and beyond the impacts of decay and interference by whether or not this antecedent

is contained within the same clause as the reflexive or a higher one. Their findings support the view that clause affiliation does affect retrieval difficulty and conclude that syntactic domains have a direct impact on retrieval processes. The results here corroborate this conclusion and suggest that this conclusion is not limited to the processing of reflexives, but likewise holds for movement dependencies.

7.6.3 Limitations of the experiments

I started out this chapter by pointing out a limitation inherent in traditional diagnostics for successive-cyclicity. These limitations have led me to explore evidence in a somewhat less orthodox domain. It is instructive to now consider the limitations of evidence from this domain, both with respect to the particular experiments reported here as well as for sentence processing more generally.

One important limitation of the two experiments here stems from the design that they involve. Both experiments contrast constructions with a differing number of hypothesized intermediate landing sites. But of course these structures differed in more than just the number of intermediate landing sites because they involve constructions with distinct properties above and beyond the mere number of hypothesized landing sites. This opens up the possibility that the reading time differences observed here are not a result of the number of intermediate landing sites, but due to some other distinctive feature of these constructions, like temporal ambiguity, frequency, etc.

Both experiments were not designed to directly address this question and therefore a consideration of these alternatives has to be indirect. The crucial dependent variable in both experiments was the reading time increase in these constructions relative to a non-movement baseline. All else equal, baseline effects of these constructions like frequency, temporal ambiguity, etc., will manifest themselves in both the movement version and the baseline structure and therefore not show up in the interaction term. In other words, the crucial interactions filter out baseline complexity effects of these constructions and thereby isolate the relative difficulty of the movement dependency. Second, the hypothesis that the critical interaction pattern is a result of baseline complexity can be tested indirectly by assessing whether differences in the baseline complexity of these constructions across items was correlated with the critical interaction in the gap and spillover regions. This was

done in fn. 13 on p. 439 for Experiment 1 and failed to find evidence for such a correlation. Third, because self-paced reading provides measures of processing complexity over the entire course of the dependency, the timing of this complexity can be used to assess different hypotheses about their underlying source (see the Discussion sections above). In particular in Experiment 2, the fact that the critical reading time increase was confined to the gap and spillover regions is consonant with the interpretation that this processing difficulty arose from processes of gap construal. A simple appeal to the baseline complexity of the structures fails to appropriately localize the timing of the effect. While these indirect considerations thus support the analytical conclusions drawn here, a direct test of these alternative accounts would require additional experiments. One potential way of avoiding these concerns would be to attempt to find evidence for successive cyclicity by comparing version of the same structure. This would be a suitable direction for future research.

Another crucial component that enters the interpretation of these results is that it requires detailed considerations of both properties of the grammar as well as mechanisms employed by the parser. That is, testable predictions emerge from the interplay of assumptions about the grammar as well as the parser. The central assumptions that have entered the discussion here are that (i) the parser constructs a phrase-structure tree that obeys the wellformedness requirements of the grammar, (ii) the integration of an intermediate trace reactivates the filler in memory and thereby strengthens its representation in memory (Lewis & Vasishth 2005, Vasishth & Lewis 2006, Nicenboim et al. 2015), (iii) reactivation of an item facilitates subsequent retrieval of this item. Assumptions about the structural representation of a movement dependency thus give rise to reading time predictions only in conjunction with a commitment to processing mechanisms. Consequently, there is no a priori method of determining whether a given response time pattern is due to properties of the grammar or the parser and adjudicating between the two options is necessarily indirect.

To illustrate, one assumption that is crucial for the evaluation of the results of Experiment 2 is that an intermediate trace has the same effect on the memory representation of the filler irrespective of its location. While the results of Experiment 2 show that only CP layers lead to facilitated retrieval of the filler at the gap site, it is a priori possible that both CPs and vPs host intermediate landing sites, but that only traces in Spec,CP reactivate the filler. A difference between CP and vP traces along this line would reconcile the CP+vP hypothesis with the results of Experiment 2 as it would effectively

render *vP* traces irrelevant for the purposes of filler retrieval. As a result, it is less straightforward to reason about grammatical representations on the basis of response time data than it is on the basis of more traditional linguistic evidence. This does not, of course, entail that response time data is uninformative or uninterpretable because claims about parsing mechanisms can be independently tested. The alternative account just sketched would have to provide independent evidence that reactivation of an element facilitates its subsequent retrievals in some cases but not in others. In the absence of such evidence, this alternative account is unconvincing.

The fact that reasoning from online processing data to grammatical structures involves two domains the relationship of which is not always clear has sometimes been taken to question the validity of grammatical inferences based on processing evidence more generally (Boland 2005, Phillips & Wagers 2007, Phillips & Parker 2014). These authors focus on domains in which there is universal agreement for the existence of grammatical dependency (such as between a displaced element and its gap) and theories diverge with respect to the implementation of these dependencies (e.g., movement vs. slash features). Evidence from processing has not succeeded in providing evidence for one account over another, but this does not, of course, preclude the possibility of processing evidence being relevant elsewhere. Moreover, I would like to draw attention to the fact that the experiments reported here did not ask about the nature of dependency whose existence is beyond doubt. Rather, these experiments have probed for syntactic dependencies whose existence was under dispute. Thus, in a configuration where an object is locally *wh*-moved, the view that *vP* is not a phase amounts to the claim that there is a direct grammatical dependency between the filler and the gap. By contrast, the view that *vP* hosts an intermediate landing site states there is a syntactic dependency between the gap site and Spec,*vP* and another dependency between Spec,*vP* and the overt filler position. Unlike the types of theoretical questions addressed by Boland (2005), Phillips & Wagers (2007), and Phillips & Parker (2014), the present study involved a very different kind of question, one that does not ask for the correct way of representing a dependency, but investigated the presence or absence of a dependency.

I therefore would like to emphasize that I do not wish to give the impression that processing evidence is in some sense inherently superior to more traditional diagnostics. Rather, different sources of evidence bring to the table different strengths and limitations. As I have argued at the outset of this chapter, a limitation of traditional diagnostics of successive cyclicity is that they

are biased in the sense that they are inherently unable to detect the absence of an intermediate landing site. I have furthermore argued that reading time evidence has the potential to overcome this limitation. Yet reading time evidence comes with its own limitations arguably not shared by traditional diagnostics, like the crucial role of parsing assumptions as a mediator between structural hypotheses and processing measures.

These inherent limitations in the various diagnostics underscore the need to evaluate theoretical hypotheses using a variety of types of empirical evidence. With regard to the distribution of successive cyclicity, I believe that the strongest argument for the view that CPs are phases but vPs are not is the convergence between the processing evidence in this chapter with the evidence from selective opacity and agreement in chapter 6. As such, evidence from language processing constitutes a valuable addition to the range of evidence used to assess theoretical proposals because its strengths and limitations complement those of more standard types of evidence.

7.7 Chapter summary

In this chapter, I have used a novel type of empirical evidence to elucidate the conclusions reached in chapter 6 regarding the distribution of phase heads. Chapter 6 has concluded on the basis of traditional successive-cyclicity effects, selective opacity, and in-situ dependencies that CP is a phase, hence inducing a locality boundary and requiring movement through its specifier, whereas vP crucially is not. In this chapter, I have argued that evidence from online sentence parsing provides a novel window into the distribution of successive cyclicity and hence phases. Following the pioneering work of Gibson & Warren (2004), I have presented experimental evidence that movement dependencies that cross a CP are easier to process at their gap position than dependencies that do not cross a CP. Following the reasoning in Gibson & Warren (2004), this finding is accounted for if crossing a CP requires an intermediate landing site of the moved element, which leads to reactivation of the filler, facilitating subsequent retrieval at the ultimate gap site. Experiment 1 has contrasted Gibson & Warren's (2004) successive-cyclicity account with an alternative premature gap filling account, which does not resort to successive cyclicity. The results of Experiment 1 support the successive-cyclicity account. Reading time evidence, then, supports the standard syntactic view that CPs are cyclic domains, i.e., phases in common parlance.

Experiment 2 has then extended this basic experimental methodology to assess whether *v*P_s have a similar facilitatory effect on the creation of the gap. If *v*P_s host an intermediate landing site analogously to CP_s, they too should induce facilitation. By contrast, if *v*P_s are not cyclic node, they should not aid filler retrieval at the gap site and hence not induce facilitation. Experiment 2 directly contrasted these two hypotheses. The results indicate that a *v*P crossed by movement has no facilitatory effect on dependency completion, in direct contrast to CP_s. This striking asymmetry between CP_s and *v*P_s is straightforwardly accounted for if CP_s host an intermediate landing site, but *v*P_s do not. This is, of course, precisely what is predicted if CP_s are phases and *v*P_s are not. The experimental results in this chapter thus corroborate the key conclusions of chapter 6.

An important virtue of processing evidence for intermediate landing sites is that it allows us to overcome the bias problem noted in section 7.2 above. There I showed that traditional diagnostics of successive cyclicity are biased in that they are inherently unable to detect the absence of an intermediate landing site. In other words, a claim that some projection X is a phase is virtually unfalsifiable using traditional tests for intermediate landing sites. Processing evidence helps overcome this bias problem because both the claim that X is a phase and the claim that it is not make testable and falsifiable predictions. For this reason, sentence processing provides an important addition to our battery of empirical diagnostics, worthy of continued exploration.

Synthesizing chapters 6 and 7, I have provided evidence from traditional diagnostics for successive cyclicity, selective opacity, in-situ dependencies, and sentence processing that CP_s are phases. By contrast, evidence from selective opacity, in-situ dependencies, and sentence processing converged on the conclusion that *v*P_s are not phases. Finally, I have argued that previous arguments in favor of *v*P_s are flawed and do not in fact support the existence of *v*P phases. Needless to say, these findings have considerable repercussions for the theory of phases and syntactic locality more generally, which give rise to fruitful avenues for future work.

Appendix A. Plausibility norming

Experiment 1

As in G&W's experiment, the semantic relations in the DP and CP condition differed in Experiment 1. Thus, in (632a) it is the accusations that hurt the secretary, while it is the claim or boast about such accusations in (632b). The verb type manipulation introduces a second potential source for

plausibility differences between the conditions. Because plausibility considerations are well-known to affect parsing decisions, a plausibility norming study on the experimental items was conducted. Because the semantic relations within any given item in Experiment 1 did not change between the [+move] and [-move] conditions, only the latter were included in the norming study. The norming experiment hence contained four versions of each item, corresponding to (632a.ii) and (632b.ii) above, thus crossing the factors *intervener* and *verb type*. All sentences were accompanied by the respective context-sentence but were not followed by a comprehension question. Participants were asked to judge how natural the meaning of the sentence was, i.e., how likely the events depicted were to occur in the real world. Answers were recorded on a 7-point scale with 1 corresponding to ‘extremely unnatural’ and 7 to ‘very natural.’ The target sentence, but not the context sentence, remained on screen while the participant made her choice.

Thirty native speakers of American English were recruited via personal communication. All were naïve to the purpose of the experiment and did not receive compensation for their participation.

A set of fifty-six items with four conditions each was initially constructed and used in the norming study. These items were arranged in a Latin Square with four lists so that each condition of each item appeared on exactly one list. Participants were randomly assigned to one list. The order of presentation was randomized for each participant.

The results were subjected to a 2×2 -factorial LME model analysis. Predictors were sum-coded (*INTERVENER*: $CP = -.5$, $DP = .5$; *VERB TYPE*: $CP-verb = -.5$; $CP/DP-verb = .5$). For the initial set of 56 items, the model revealed a lower plausibility rating in the DP condition compared to the CP condition ($p < .05$), making it clear that the experiment has the ability to detect differences in plausibility between the conditions. No other effect approached significance. To avoid a confounding of clause structure and plausibility, 48 items were selected so that the plausibility differences between the conditions was minimized. Table 7.9 provides the mean plausibility ratings for these 48 items. LME modeling did not detect even a hint of a plausibility difference between the DP and the CP condition within the set of these 48 items ($p > .5$). As before, no other contrast approached significance. This set of 48 items was used in the main experiment, the remaining eight were discarded. The 48 items are listed in Appendix C.

Table 7.9. Mean plausibility rating (and by-participant standard errors) by condition for stimuli used in Experiment 1

<i>Intervener</i>	<i>Verb type</i>	
	CP-verbs	CP/DP-verbs
CP	4.9 (.1)	4.9 (.1)
DP	4.8 (.1)	4.8 (.1)

Experiment 2

As in Experiment 1, the semantic relations differed between the conditions in Experiment 2. To ensure that *intervener* type is not confounded with plausibility, a plausibility norming study was conducted. Because the semantic relations remain constant between the [+move] and [-move] manipulation, only the latter was included in the rating study. The experiment thus comprised a single three-level factor *INTERVENER*. The norming experiment hence contained three versions of each item, corresponding to (638a.ii), (638b.ii) and (638c.ii) above. All sentences were accompanied by the respective context-sentence but were not followed by a comprehension question. Participants were asked to judge how natural the meaning of the sentence was, i.e., how likely the events depicted were to occur in the real world. Answers were recorded on a 7-point scale with 1 corresponding to

‘extremely unnatural’ and 7 to ‘very natural.’ The target sentence, but not the context sentence, remained on screen while the participant made her choice.

Twenty-four native speakers of American English were recruited via Amazon Mechanical Turk. Each received USD 0.50 as compensation.

A set of thirty-six items with three conditions each was constructed. These items were arranged in a Latin Square with three lists. Each item appeared in exactly one list. Participants were randomly assigned to one list and the order of presentation of the stimuli was randomized for each participant.

To maximize comparability, the LME model used to analyze the plausibility ratings is identical to the one used to analyze the reading time data, modulo the absence of the predictor *MOVEMENT*. In particular, the factor *INTERVENER* was Helmert coded (first contrast: CP structures versus the mean of TP and DP structures; second contrast: TP versus DP structures) to detect whether patterns in the reading time data are potentially due to plausibility confounds.

The analysis of the original set of 36 items revealed a somewhat higher plausibility of the CP condition compared to the other two, which did however not reach significance ($p > .1$) and no substantial difference between the DP and TP conditions ($t < 1$). To closely match the plausibility of the three structures across the experimental items, thirty items were extracted. The mean plausibility ratings for these thirty items by condition are provided in Table 7.10. In this set there was no hint of a plausibility difference (all p 's $> .5$). This set of 30 sentences was then used in the self-paced reading experiment, the remaining 6 being discarded. These 30 items are listed in Appendix C.

Table 7.10. Mean plausibility rating (and by-participant standard errors) by condition for stimuli used in Experiment 2

CP	<i>Intervener</i>	
	DP	TP
4.7 (.2)	4.8 (.2)	4.7 (.2)

Appendix B. Statistical analysis of non-crucial regions

Experiment 1

As a measure of the response variability by condition and region, the by-participant standard errors computed over raw reading times in ms is provided in Table 7.11. A comprehensive overview of the coefficients obtained from the LME modeling for all regions in Experiment 1 is provided in Table 7.12. This appendix will discuss the results of the regions not considered in the main text.

In region 1, there is a significant interaction between *intervener* and *verb type*. This must be a type 1 error as the materials were identical in all conditions.

In region 2, there were main effects of *intervener* and *structure* such that DP conditions were read significantly slower than CP conditions and [+move] conditions slower than [-move] conditions.

In region 4, there likewise were main effects of *intervener* type and *movement*. In addition, there was a significant interaction between the two factors such that the reading time increase was greater in the DP condition than the CP condition. While the interpretation of this effect is not completely clear, it is possibly related to the hypothesis that the moved element is reactivated by the intermediate trace in the embedded Spec,CP, which the parser can postulate as soon as it encounters the complementizer. As noted above, the results here show a reading time increase due to movement that is plausibly due to the fact that an unassigned element has to be held in

Table 7.11. By-participant standard error of raw reading times in ms by condition and region for Experiment 1

Verb type	Intervener	Movement	Region						
			1	2	3	4	5	6	7
CP	CP	[-move]	—	32	13	28	25	27	50
		[+move]	19	36	17	26	37	28	34
	DP	[-move]	—	30	14	33	25	28	41
		[+move]	22	52	22	37	37	31	30
CP/DP	CP	[-move]	—	26	11	26	21	29	46
		[+move]	21	35	20	27	41	33	35
	DP	[-move]	—	34	14	32	24	26	49
		[+move]	20	45	19	33	45	31	36

memory. If this element is integrated and hence re-activated when the complementizer is processed, the smaller reading time difference incurred by movement in the CP condition in region 4 can be viewed as the result of this reactivation.²² If this interpretation is correct, it provides additional evidence for an intermediate landing site in the embedded Spec,CP.

In the final region, the [-move] conditions were read more slowly than the extraction conditions. There was furthermore a significant interaction between intervener and movement such that the reading time increase between extraction and non-extraction structures is greater for CP structures than for DP structures. Because the length of the region differed between conditions, an additional analysis of the residual reading times was performed. This analysis replicated the effect of *movement* ($\beta = -83, t = -4.1, p < .001$) and the interaction between movement and intervener ($\beta = 70, t = 2.4, p < .05$). In addition, there was a main effect of *intervener* ($\beta = -34, t = -2.3, p < .05$). This pattern of results is plausibly a wrap-up effect (see, e.g., King & Just 1991): The reading time distribution mirrors the level of embedding in the final region. In the [+move] conditions, the material in region 7 is part of the matrix clause. In the [-move] DP condition, the material in the final region is part of a relative clause and hence belongs to a first level of embedding. In the [-move] CP condition the material in this region is a relative clause inside a complement clause and thus at the second level of embedding. The reading times increase with the level of embedding in this region.²³

Two effects are evident in a number of consecutive regions and deserve some comment. First, in regions 2 through 5, DP structures were generally read more slowly than CP structures. Because the syntactic structure in regions 2 through 4 and the lexical material in regions 2 and 3 differed between these conditions, this contrast is not surprising. That the pattern persisted into region 5 is likely either a spillover from the preceding region or an artifact of the interaction of *intervener* and *verb type* in this region. Second, that the reading times are higher in the [+move] condition than in the [-move] one throughout the length of the movement dependency mirrors the results of Experiment 1 and is plausibly due to the fact that an unassigned filler has to be held active in working memory (Wanner & Maratsos 1978, King & Kutas 1995, Gibson 1998, 2000, Fiebach et al. 2002, Grodner et al. 2002, Chen et al. 2005), under the assumption that this cost for storing the filler leaves fewer resources available for the processing of incoming material.

²² This type of interaction is consistent with Gibson's (1998, 2000) length-dependent notion of storage cost, according to which the processing load incurred by holding a filler in memory grows with the distance to the filler.

²³ A very similar pattern of results is observed by Frazier & Clifton (1989).

Table 7.12. Coefficient estimates and corresponding t -value for linear mixed effects model analyses of log reading times in Experiment 1 for all regions. *Mvmt:intv* refers to the interaction between *Movement* and *Intervener*, *Mvmt:Vt* to the interaction between *Movement* and *Verb type* and *Intv:Vt* refers to the interaction between *Intervener* and *Verbttype*. *Mvmt:Intv:Vt* refers to the three-way interaction of all predictors. Cells with $p < .05$ are shaded.

Region	Movement		Intervener		Verb type		Mvmt:Intv		Mvmt:Vt		Intv:Vt		Mvmt:Intv:Vt	
	$\hat{\beta}$	t	$\hat{\beta}$	t	$\hat{\beta}$	t	$\hat{\beta}$	t	$\hat{\beta}$	t	$\hat{\beta}$	t	$\hat{\beta}$	t
1	—	—	-0.016	-1.53	-0.010	-0.89	—	—	—	—	-0.048	-2.31	—	—
2	0.138	7.70	0.101	7.46	0.021	1.53	0.033	1.59	-0.000	-0.01	0.012	0.44	-0.062	-1.35
3	0.127	11.33	0.063	6.23	0.000	0.03	0.029	1.78	0.013	0.75	-0.024	-1.47	-0.075	-2.31
4	0.081	5.31	0.077	5.78	-0.008	-0.82	0.055	2.67	-0.014	-0.61	-0.014	-0.59	-0.034	-0.74
5	0.112	5.03	0.031	2.66	0.023	1.99	-0.004	-0.17	0.018	0.80	-0.049	-2.37	0.031	0.70
6	-0.009	-0.38	0.015	1.17	0.012	1.11	0.060	2.55	-0.026	-1.27	-0.017	-0.79	-0.018	-0.43
7	-0.350	-15.85	-0.021	-1.90	0.001	0.12	0.058	2.62	0.010	0.51	0.003	0.14	0.015	0.34

Experiment 2

The standard error of the raw reading times in ms per condition and region is given in Table 7.13. A comprehensive overview of the coefficients obtained from the LME modeling for all regions in Experiment 2 is provided in Table 7.14. This appendix will discuss the results of the regions not considered in the main text.

Most regions exhibited a main effect of *movement*: Reading times in the [+move] conditions were greater than in the [-move] conditions in regions 2 through 5 and in region 7. Region 6 followed the same trend but without reaching significance. In the final region, reading times were significantly lower in the [+move] conditions than the [-move] conditions.

In region 2, reading times in the CP condition were reliably smaller than the mean of the DP and TP conditions and reading times in the DP conditions were significantly higher than in the TP condition. A post-hoc pairwise comparisons of the means showed that the DP condition had significantly higher reading times than the CP condition ($t_1(100) = 10, p < .001$; $t_2(28) = 7, p < .001$) and the TP condition ($t_1(100) = 8, p < .001$; $t_2(28) = 6, p < .001$) while the latter two did not differ ($t_1(100) = 1.63, p > .1$; $t_2(28) = 1.93, p = .06$). This pattern is unsurprising given that the lexical material in the DP condition differed from the other two in this region.

Because region 3 did not contain any material in the TP condition, the CP and DP condition were compared directly (see the discussion in the main text). The DP condition elicited greater reading times than the CP condition. This is again unsurprising as the lexical content and the preceding region differed between the two structures. Analysis of the residual reading times indicated a parallel effect of *movement* ($\hat{\beta} = 81, t = 8, p < .001$) but no main effect of *intervener* nor an interaction (p 's $> .1$).

In region 4, reading times in the CP condition were smaller than in the combination of the other two. At the same time, reading times in the DP condition were faster than in the TP condition. Post-hoc pairwise comparisons revealed that this pattern was produced by faster reading times in the CP condition compared to the DP condition ($t_1(100) = 1.94, p = .06$; $t_2(28) = 2.71, p < .011$). In addition, the CP condition was read faster than the TP condition ($t_1(100) = 9, p < .001$; $t_2(28) = 9, p < .001$). These differences are not interpretable because the lexical content of the preceding region as well as the syntactic context of that region differed between the three structures.

In region 5, the model analysis produced a pattern similar to region 4. Pairwise comparisons showed that the CP condition was read faster than the DP condition ($t_1(100) = 1.94, p = .055$; $t_2(28) = 2.14, p < .05$) and the TP condition ($t_1(100) = 10, p < .001$; $t_2(28) = 8, p < .001$). Analysis of the residual reading times indicated only a main effect of *movement* ($\hat{\beta} = 18, t = 2.5, p < .05$; all other p 's $> .1$). This pattern is plausibly related to the design of the experiment. For the majority of items, region 5 was longer in the TP condition (*to have* vs. *had*). The increased reading in the

Table 7.13. By-participant standard error of raw reading times in ms by condition and region for Experiment 2

Intervener	Movement	Region								
		1	2	3	4	5	6	7	8	9
CP	[-move]	—	31	12	22	12	14	16	30	54
	[+move]	22	39	17	28	11	15	19	29	35
DP	[-move]	—	43	11	26	11	13	15	30	52
	[+move]	23	53	18	33	13	14	16	31	38
TP	[-move]	—	29	—	28	14	16	16	28	53
	[+move]	21	43	—	29	15	14	24	27	36

Table 7.14. Coefficient estimates and corresponding *t*-values for the linear mixed effects model analyses of log reading times in Experiment 2 for all regions. *Mvmt* refers to the sum-coded factor *movement*; *IntervCP-TPDP* compares the CP condition to the mean of the TP and DP conditions and *IntervTP-DP* compares the TP condition to the DP condition. *Mvmt:IntCP-TPDP* and *Mvmt:IntTP-DP* refer to the interaction between *Mvmt* and *IntervCP-TPDP* and *IntervTP-DP*, respectively. Cells with *p* < .05 are shaded.

Region	<i>Mvmt</i>		<i>IntervCP-TPDP</i>		<i>IntervTP-DP</i>		<i>Mvmt:IntCP-TPDP</i>		<i>Mvmt:IntTP-DP</i>	
	$\hat{\beta}$	<i>t</i>	$\hat{\beta}$	<i>t</i>	$\hat{\beta}$	<i>t</i>	$\hat{\beta}$	<i>t</i>	$\hat{\beta}$	<i>t</i>
1	—	—	0.002	0.12	0.012	0.69	—	—	—	—
2	0.108	4.87	0.099	6.17	0.141	6.77	0.046	1.68	-0.012	-0.35
3*	0.113	7.84	0.073	4.50	—	—	0.034	1.18	—	—
4	0.089	5.10	0.101	7.14	-0.126	-7.66	0.035	1.12	0.053	1.68
5	0.032	2.61	0.069	6.56	-0.092	-5.55	0.037	1.65	-0.030	-1.32
6	0.015	1.22	0.029	2.73	-0.037	-3.41	0.006	0.29	-0.005	-0.23
7	0.053	4.00	0.013	1.13	-0.042	-3.65	0.024	1.09	-0.049	-2.10
8	-0.039	-1.11	-0.000	-0.00	-0.003	-0.21	0.059	2.08	-0.018	-0.61
9	-0.318	-10.63	0.003	0.26	-0.007	-0.39	0.019	0.73	0.085	2.62

* Because the TP condition did not contain region 3, *IntervCP-TPDP* effectively corresponds to a direct comparison between the CP and the DP condition and *IntervTP-DP* is undefined. The same holds for their interactions with *Movement* (*Mvmt:IntCP-TPDP* and *Mvmt:IntTP-DP*), respectively. Contrasts were redefined accordingly for that region. See the main text for details.

DP condition are likely due to the fact that the element that had to be construed in the subject position is of considerably greater syntactic and semantic complexity in the DP condition than the CP condition.

In region 6, reading times in the CP condition were faster than the mean of the other two. Additionally, reading times in the TP conditions were higher than in the DP condition. Pairwise comparisons indicated that this pattern was produced by particularly high reading times in the TP condition: The reading times in the TP condition were higher than in the CP condition ($t_1(100) = 3.7, p < .001$; $t_2(28) = 3.0, p < .01$) and there was no difference between the CP and the DP conditions ($t_1(100) = .9, p = .4$; $t_2(28) = .7, p > .5$). This pattern is plausibly a result of the fact that the lexical content in the preceding region differed between the TP and the other two conditions.

Finally, region 9 exhibited significantly greater reading times in the [-move] conditions than the [+move] ones and an interaction such that this reading times increase was significantly greater in the TP condition than the DP condition. In the analysis of the residual reading times, both effects are marginally significant ($.05 < p < .1$). This effects mirrors the pattern in Experiment 1 and like the latter is likely a wrap-up effect reflecting the differing depths of clausal embedding of the material in the final region: Not only is the final region longer in the [-move] condition than in the [+move] condition, this regions is also part of a matrix clause in the latter but inside an embedded clause in the former. Furthermore, the region is embedded within an embedded clause in the CP and TP region. This accounts for the particularly high reading time increase in the CP and TP conditions compared to the DP conditions.

Appendix C. Materials used

Experiment 1

This appendix provides the materials in the extraction conditions of each item in both the CP and the DP condition (a. and b., respectively). The verb type manipulation is given between curly brackets. The first verb/noun shown falls into the CP/DP-class, the second one into the CP-class. The non-extraction conditions can be constructed by moving the clause-initial noun phrase and the relative pronoun (e.g., *the manager who* in item 1) immediately after the end of the relative clause (e.g., between *pleased* and *will hire* in item 1). The first sentence of each item constitutes the context sentence for that item. The question following the actual target sentences in a. and b are the comprehension questions. Possible answers are given in italics. The order of presentation of these answers was randomized for each subject.

- | | |
|--|---|
| <p>1. People tend to make favorable decisions when they are pleased by something.</p> <p>a. The manager who the consultant {claimed / hinted} that the new proposal had pleased will hire five workers tomorrow.</p> <p>b. The manager who the consultant's {claim / hint} about the new proposal had pleased will hire five workers tomorrow.</p> <p>Who was (allegedly) pleased?
<i>the manager – the consultant</i></p> | <p>b. The actress who the agent's {implications / remark} about the controversial rumor had distressed lobbied to play Evita.</p> <p>Who felt distressed? <i>the actress – the agent</i></p> |
| <p>2. At the rehearsal some felt distressed.</p> <p>a. The actress who the agent {implied / remarked} that the controversial rumor had distressed lobbied to play Evita.</p> | <p>3. Thrilling experiences are the spice of life.</p> <p>a. The woman who the man {confirmed / dreamed} that the recent events had thrilled was hoping to get married.</p> <p>b. The woman who the man's {confirmation of / dream about} the recent events had thrilled was hoping to get married.</p> <p>Who was thrilled? <i>the woman – the man</i></p> |
| | <p>4. Sometimes people are embarrassed by what others have to say.</p> |

- a. The freshman who the sophomore {implied / insisted} that the late night confession had embarrassed was new to the partying scene.
- b. The freshman who the sophomore's {implication / insistence} about the late night confession had embarrassed was new to the partying scene.
- Who was the embarrassed one?
the freshman – the sophomore
5. Some things are really annoying.
- a. The victim who the counselor {concluded / agreed} that the mean-spirited comment had annoyed has switched to another psychoanalyst.
- b. The victim who the counselor's {conclusion about / agreement with} the mean-spirited comment had annoyed has switched to another psychoanalyst.
- Who was annoyed? *the victim – the counselor*
6. Reassuring someone who is in a difficult situation in life is often of utmost importance.
- a. The patient who the doctor {predicted / hoped} that the test results will reassure had been afraid she had cancer.
- b. The patient who the doctor's {prediction / hope} about the test results will reassure had been afraid she had cancer.
- Who will be reassured? *the patient – the doctor*
7. People generally do not deal well with being embarrassed.
- a. The judge who the reporter {implied / speculated} that the controversial decision had embarrassed has decided to sue the paper.
- b. The judge who the reporter's {implication / speculation} about the controversial decision had embarrassed has decided to sue the paper.
- Who was embarrassed? *the judge – the reporter*
8. To produce a masterpiece artists need to be inspired by something.
- a. The author who the secret lover {declared / insisted} that his love had inspired unexpectedly won the Nobel Prize.
- b. The author who the secret lover's {declaration / insistence} of his love had inspired unexpectedly won the Nobel Prize.
- Who was inspired? *the author – the lover*
9. (*The regioning in this item was incorrect and it was discarded from all analyses*)
In the trial yesterday someone got implicated in the crime.
- a. The defendant who the witness {confirmed / agreed} that the convincing testimony had implicated admitted the truth.
- b. The defendant who the witness's {confirmation of / agreement with} the convincing testimony had implicated admitted the truth.
- Who was it that got implicated in the crime?
the defendant – the witness
10. At the hearing last month some participant supposedly got disconcerted.
- a. The senator who the committee {stated / insisted} that the charges had disconcerted will nonetheless remain on the ethics panel.
- b. The senator who the committee's {statement of / insistence on} the charges had disconcerted will nonetheless remain on the ethics panel.
- Who was it that got disconcerted?
the senator – the committee
11. Occasionally comfort in rough times comes from unexpected places.
- a. The daughter who the guidance counselor {predicted / thought} that the court decision will comfort has left home for good.
- b. The daughter who the guidance counselor's {prediction / thoughts} about the court decision will comfort has left home for good.
- Who {might / will} receive some comfort?
the daughter – the guidance counselor
12. Working in politics can be bothersome.
- a. The politician who the journalist {predicted / pretended} that the government announcement will bother is calling a press conference.
- b. The politician who the journalist's {prediction / pretense} about the government announcement will bother is calling a press conference.
- Who did the {announcement / prediction / pretense} bother?
the politician – the journalist
13. Young people are easy to frighten.
- a. The girl who the teacher {concluded / remarked} that the nasty threat had frightened has stopped going to school.
- b. The girl who the teacher's {conclusions / remark} about the nasty threat had frightened has stopped going to school.
- Who got frightened by {(the conclusion / remark about) the threat}?
the girl – the teacher
14. Many things in life can be upsetting.
- a. The orphan who the social worker {decided / thought} that the problem had upset was unhappy with her foster parents.
- b. The orphan who the social worker's {decision / thoughts} about the problem had upset was unhappy with her foster parents.
- Who was upset by something?
the orphan – the social worker
15. The court hearing the other day seems to have shocked many people.
- a. The witness who the lawyer {proved / pretended} that the crime had shocked does not want to testify.

- b. The witness who the lawyer's {proof / pretense} about the crime had shocked does not want to testify.
Who did the {crime / proof / pretense} shock?
the witness – the lawyer
16. People inspired by something tend to choose artistic professions.
- a. The actress who the journalist {guessed / hoped} that the talented writer had inspired will go on stage tonight.
- b. The actress who the journalist's {guess about / hopes for} the talented writer had inspired will go on stage tonight.
Who was inspired? *the actress – the journalist*
17. It is important to be vocal if one feels distressed.
- a. The schoolboy who the teacher {proved / theorized} that the bully's aggression had distressed will complain at the meeting.
- b. The schoolboy who the teacher's {proof of / theory about} the bully's aggression had distressed will complain at the meeting.
Who was feeling distressed?
the schoolboy – the teacher
18. Many people pursue grad school because they are fascinated by their subject.
- a. The student who the professor {concluded / thought} that the Ancient Greek had fascinated is planning to go on an archaeological excursion.
- b. The student who the professor's {conclusions / thoughts} about the Ancient Greek had fascinated is planning to go on an archaeological excursion.
Who was fascinated? *the student – the professor*
19. Getting inspired by something sometimes leads to new discoveries.
- a. The student who the teacher {predicted / boasted} that the new idea will inspire is studying artificial intelligence.
- b. The student who the teacher's {prediction / boast} about the new idea will inspire is studying artificial intelligence.
Who was inspired? *the student – the teacher*
20. If one is bothered by an aspect of one's life it is a good idea to look for alternatives.
- a. The chef who the cook {assumed / speculated} that the head waitress had bothered wants to find another job.
- b. The chef who the cook's {assumption / speculation} about the head waitress had bothered wants to find another job.
Who did {the head waitress / the assumption about the head waitress / the speculation about the head waitress} bother?
the chef – the cook
21. Every so often even grumpy people are delighted by something.
- a. The CEO who the secretary {claimed / insisted} that the new project had delighted is normally very hard to please.
- b. The CEO who the secretary's {claim / insistence} about the new project had delighted is normally very hard to please.
Who was delighted? *the CEO – the secretary*
22. Some people act irresponsibly when they get annoyed.
- a. The customer who the receptionist {implied / remarked} that the lazy cleaner had annoyed will not pay his bill.
- b. The customer who the receptionist's {implications / remark} about the lazy cleaner had annoyed will not pay his bill.
Who got annoyed by something?
the customer – the receptionist
23. Making someone feel alarmed by something can often sell products.
- a. The man who the detective {concluded / remarked} that the dangerous thief had alarmed will buy a new alarm.
- b. The man who the detective's {conclusion / remark} about the dangerous thief had alarmed will buy a new alarm.
Who was alarmed? *the man – the detective*
24. It's advisable not to anger a person you are working for.
- a. The tourist who the tour guide {claimed / thought} that the hotel manager had angered wants to return home now.
- b. The tourist who the tour guide's {claim / thoughts} about the hotel manager had angered wants to return home now.
Who was (allegedly) angered by something?
the tourist – the tour guide
25. At the annual convention some attendees allegedly were infuriated.
- a. The activists who the reporter {asserted / remarked} that human-rights violations had infuriated wrote a strong letter of complaint.
- b. The activists who the reporter's {assertion / remark} about human-rights violations had infuriated wrote a strong letter of complaint.
Who got infuriated at the convention?
the activists – the reporter
26. Some people work relentlessly if they are fascinated by something.
- a. The therapist who the patient {assumed / dreamed} that the strange woman had fascinated is writing a new book.
- b. The therapist who the patient's {assumption / dream} about the strange woman had fascinated is writing a new book.
Who is fascinated by something?
the therapist – the patient

27. Even appalling events in one's life can teach one a lesson.
- A lot of people who the veteran {recalled / hoped} that the war had appalled told their sons not to join the Navy.
 - A lot of people who the veteran's {recollections / hope} about the war had appalled told their sons not to join the Navy.
- Who was appalled? *a lot of people – the veteran*
28. Being agitated by some problem can be very motivational.
- Many engineers who the scientist {demonstrated / hypothesized} that the effects of global warming agitated invented greener technologies.
 - Many engineers who the scientist's {demonstration of / hypothesis about} the effects of global warming agitated invented greener technologies.
- Who was agitated and used it for a good cause?
the engineers – the scientist
29. There are people who are very sensitive and easy to displease.
- The heiress who the lawyer {confirmed / insisted} that the recently proposed plan had displeased will attend an important meeting this afternoon.
 - The heiress who the lawyer's {confirmation of / insistence on} the recently proposed plan had displeased will attend an important meeting this afternoon.
- Who did {the plan / the confirmation / the insistence} displease?
the heiress – the lawyer
30. I have been informed that some person was unsettled by what happened in court yesterday.
- The lawyer who the spectator {recalled / boasted} that the incident had unsettled made his own declaration.
 - The lawyer who the spectator's {recollection of / boast about} the incident had unsettled made his own declaration.
- Who was unsettled in court yesterday?
the lawyer – the spectator
31. Because many political situations are complicated different sources often contradict each other.
- The journalist who the editor {decided / argued} that the new report had contradicted was planning a series of articles.
 - The journalist who the editor's {decision / argument} about the new report had contradicted was planning a series of articles.
- Who did {the report / the decision / the argument} contradict?
the journalist – the editor
32. I heard that a visitor at the art gallery had gotten emotionally affected.
- The grumpy old man who the guide {recalled / speculated} that the painting had deeply affected turned out to be a wealthy collector.
 - The grumpy old man who the guide's {recollections / speculations} about the painting had deeply affected turned out to be a wealthy collector.
- Who was it that got emotionally affected?
the old man – the guide
33. From time to time people who are angered do irresponsible things.
- The general who the sergeant {assumed / thought} that the message had angered complained to the lieutenant.
 - The general who the sergeant's {assumption / thoughts} about the message had angered complained to the lieutenant.
- Who did the general complain to?
the lieutenant – the sergeant
34. Groundless allegations really could hurt people in our company.
- The secretary who the lawyer {claimed / boasted} that the accusation had hurt was fired from her job.
 - The secretary who the lawyer's {claim / boast} about the accusation had hurt was fired from her job.
- Who made a {claim / boast}?
the lawyer – the secretary
35. At the movie set the other day someone had the blues and needed to be cheered up.
- The make-up artist who the producer {claimed / commented} that the hair stylist had cheered up was enthusiastic to get to work again.
 - The make-up artist who the producer's {claim / comment} about the hair stylist had cheered up was enthusiastic to get to work again.
- Who made a {claim / comment}?
the producer – the hair stylist
36. In some cases conclusions which are supposed to be impartial are actually unfairly influenced.
- The committee members who the professor {assumed / speculated} that the dean's recommendation had influenced finally reached a decision.
 - The committee members who the professor's {assumption / speculation} about the dean's recommendation had influenced finally reached a decision.
- Who made a recommendation?
the dean – the professor
37. It frequently happens that people who are angered won't cooperate.
- The nurse who the doctor {assumed / agreed} that the new patient had angered is refusing to work late.

- b. The nurse who the doctor's {assumptions about / agreement with} the new patient had angered is refusing to work late.
Who was new to the hospital?
the patient – the nurse
38. At the convention last Thursday someone who got startled made a surprising decision.
a. The farmer who the builder {implied / thought} that the dedicated worker had startled will give everybody extra money.
b. The farmer who the builder's {implication / thoughts} about the dedicated worker had startled will give everybody extra money.
Who is dedicated?
the worker – the builder
39. Yesterday someone got offended and a drama happened.
a. The singer who the musician {asserted / hinted} that the drunken guitarist had offended will not perform this evening.
b. The singer who the musician's {assertion / hint} about the drunken guitarist had offended will not perform this evening.
Who asserted something?
the musician – the guitarist
40. Annoying someone can have drastic consequences.
a. The coach who the manager {decided / remarked} that the violent boxer had annoyed will cancel the match today.
b. The coach who the manager's {decision / remark} about the violent boxer had annoyed will cancel the match today.
Who made a {decision / remark}?
the manager – the boxer
41. Embarrassing someone is rarely a good idea.
a. The film star who the interviewer {implied / hinted} that the horrible photographer had embarrassed will not answer any questions.
b. The film star who the interviewer's {implication / hint} about the horrible photographer had embarrassed will not answer any questions.
Who {implied / hinted at} something?
the interviewer – the photographer
42. In the headquarters there was a dispute this morning that displeased various workers.
a. The captain who the officer {decided / agreed} that the young soldier had displeased will write a formal report.
b. The captain who the officer's {decision about / agreement with} the young soldier had displeased will write a formal report.
Which person is young?
the soldier – the officer
43. Recent advancements in science could help many patients.
a. The schizophrenic who the psychologist {concluded / hypothesized} that the new theory could help has stopped taking his pills.
b. The schizophrenic who the psychologist's {conclusion / hypothesis} about the new theory could help has stopped taking his pills.
Who {drew a conclusion / made a hypothesis}?
the psychologist – the schizophrenic
44. Running for office can be disturbing.
a. The candidate who the senator {proved / speculated} that the allegations had disturbed might retract his candidacy.
b. The candidate who the senator's {proof of / speculation about} the allegations had disturbed might retract his candidacy.
Who {proved / speculated about} something?
the senator – the candidate
45. Even if you are properly prepared certain things you learn in school can be confusing.
a. The smart student who the teacher {illustrated / remarked} that the problem had confused did very poorly on the exam.
b. The smart student who the teacher's {illustration of / remark about} the problem had confused did very poorly on the exam.
Who is smart?
the student – the teacher
46. Magic sure is impressive.
a. The magician who the apprentice {guessed / boasted} that the spell had impressed suddenly vanished from the stage.
b. The magician who the apprentice's {guess / boast} about the spell had impressed suddenly vanished from the stage.
Who made a {guess / boast}?
the apprentice – the magician
47. Owning property can be disconcerting.
a. The wary landlord who the tenant {stated / commented} that the burglary has disconcerted is planning to have an alarm installed.
b. The wary landlord who the tenant's {statement / comment} about the burglary has disconcerted is planning to have an alarm installed.
Who is wary?
the landlord – the tenant
48. Having an idol can be thrilling indeed.
a. The fan who the rock star {confirmed / commented} that the rumors had thrilled immediately told his crazy uncle about it.
b. The fan who the rock star's {confirmation of / comment about} the rumors had thrilled immediately told his crazy uncle about it.
Who is crazy?
the uncle – the rock star

Experiment 2

This appendix provides the materials in the extraction conditions of each item in the CP, TP and DP conditions (a., b. and c., respectively). The non-extraction conditions can be constructed analogously to Experiment 1. The first sentence of each item constitutes the context sentence for that item. The question following the actual target sentences in a., b. and c. are the comprehension questions. Possible answers are given in italics. The order of presentation of these answers was randomized for each subject.

1. Thrilling experiences are the spice of life.
 - a. The actress who the agent believed that the recent events had secretly thrilled lobbied to play Evita.
 - b. The actress who the agent believed the recent events to have secretly thrilled lobbied to play Evita.
 - c. The actress who the agent's belief about the recent events had secretly thrilled lobbied to play Evita.Who was thrilled? *the actress – the agent*
2. In the trial today someone surprising was incriminated.
 - a. The witness who the prosecutor proved that the bloody footprint had conclusively incriminated admitted the truth.
 - b. The witness who the prosecutor proved the bloody footprint to have conclusively incriminated admitted the truth.
 - c. The witness who the prosecutor's proof about the bloody footprint had conclusively incriminated admitted the truth.Who had been incriminated?
the witness – the prosecutor
3. It is important to be vocal if one feels distressed.
 - a. The schoolboy who the teacher believes that the bully's aggression had obviously distressed will complain at the meeting.
 - b. The schoolboy who the teacher believes the bully's aggression to have obviously distressed will complain at the meeting.
 - c. The schoolboy who the teacher's belief about the bully's aggression had obviously distressed will complain at the meeting.Who got distressed by the belief about the aggression?
the schoolboy – the teacher
4. Many things you have to learn in college are very boring.
 - a. The student who the professor proved that the Ancient Greeks had terribly bored is forced to go on an archaeological excursion.
 - b. The student who the professor proved the Ancient Greeks to have terribly bored is forced to go on an archaeological excursion.
 - c. The student who the professor's proof about the Ancient Greeks had terribly bored is forced to go on an archaeological excursion.
5. Who was really bored by the proof?
the student – the professor
5. Being agitated by some problem can be very motivational.
 - a. Many engineers who the scientist believed that global warming had seriously agitated invented greener technologies.
 - b. Many engineers who the scientist believed global warming to have seriously agitated invented greener technologies.
 - c. Many engineers who the scientist's beliefs about global warming had seriously agitated invented greener technologies.Who was agitated by something?
the engineers – the scientist
6. I have been informed that some person was unsettled by what happened in court yesterday.
 - a. The spectator who the detective proved that the incident had noticeably unsettled decided to make his own declaration.
 - b. The spectator who the detective proved the incident to have noticeably unsettled decided to make his own declaration.
 - c. The spectator who the detective's proof about the incident had noticeably unsettled decided to make his own declaration.Who did the proof unsettle?
the spectator – the detective
7. Everyday things can be astonishing to some people.
 - a. The woman who the man expected that the park will totally astonish had to tell her friends about it.
 - b. The woman who the man expected the park to totally astonish had to tell her friends about it.
 - c. The woman who the man's expectations about the park will totally astonish had to tell her friends about it.Who might the expectations about the park astonish?
the woman – the man
8. (*This item was coded incorrectly and eliminated from all analyses.*)
Even appalling events in one's life can teach one a lesson.
 - a. A lot of people who the veteran suspects that the war had completely appalled told their sons not to join the Navy.

- b. A lot of people who the veteran suspects the war to have completely appalled told their sons not to join the Navy.
- c. A lot of people who the veteran's suspicion about the war had completely appalled told their sons not to join the Navy.
- Who felt appalled? *a lot of people – the veteran*
9. Providing help for individuals in need should be an objective for politicians.
- a. Countless families who the government proved that the new legislation had profoundly helped would have gone bankrupt otherwise.
- b. Countless families who the government proved the new legislation to have profoundly helped would have gone bankrupt otherwise.
- c. Countless families who the government's proof about the new legislation had profoundly helped would have gone bankrupt otherwise.
- Who was helped by the proof?
many families – the government
10. Being aggravated by a problem can make people persistent.
- a. The political candidate who the inspector revealed that the Senate had immensely aggravated is lacking rich donors to back him up.
- b. The political candidate who the inspector revealed the Senate to have immensely aggravated is lacking rich donors to back him up.
- c. The political candidate who the inspector's revelations about the Senate had immensely aggravated is lacking rich donors to back him up.
- Who was aggravated?
the candidate – the inspector
11. People caught up in the recent political affair were greatly disconcerted by it.
- a. The senator who the political consultant suspected that the charges had greatly disconcerted will nonetheless remain on the panel.
- b. The senator who the political consultant suspected the charges to have greatly disconcerted will nonetheless remain on the panel.
- c. The senator who the political consultant's suspicions about the charges had greatly disconcerted will nonetheless remain on the panel.
- Who was disconcerted?
the senator – the consultant
12. There are people who are very sensitive and easy to displease.
- a. The heiress who the lawyer believed that the recently proposed plan had noticeably displeased will attend an important meeting this afternoon.
- b. The heiress who the lawyer believed the recently proposed plan to have noticeably displeased will attend an important meeting this afternoon.
- c. The heiress who the lawyer's beliefs about the recently proposed plan had noticeably displeased will attend an important meeting this afternoon.
- Who was it that was displeased?
the heiress – the lawyer
13. At the meeting last week everyone had to say what they were currently interested in.
- a. The newswriter who the legal correspondent discovered that the trial had greatly interested is having serious doubts about the verdict.
- b. The newswriter who the legal correspondent discovered the trial to have greatly interested is having serious doubts about the verdict.
- c. The newswriter who the legal correspondent's discovery about the trial had greatly interested is having serious doubts about the verdict.
- Who was interested by the discovery?
the newswriter – the correspondent
14. Excessive drinking can lead to worrisome results.
- a. The freshman who the sophomore suspected that the scandalous party had needlessly worried was absent from class this morning.
- b. The freshman who the sophomore suspected the scandalous party to have needlessly worried was absent from class this morning.
- c. The freshman who the sophomore's suspicion about the scandalous party had needlessly worried was absent from class this morning.
- Who was it that was needlessly worried?
the freshman – the sophomore
15. I heard that a visitor at the garden exhibition was delighted.
- a. The elderly man who the guide believed that the beautiful flower had immensely delighted turned out to be a wealthy collector.
- b. The elderly man who the guide believed the beautiful flower to have immensely delighted turned out to be a wealthy collector.
- c. The elderly man who the guide's beliefs about the beautiful flower had immensely delighted turned out to be a wealthy collector.
- Who was it that felt delighted?
the elderly – the guide
16. In some parts of the world the weather is often a cause for concern.
- a. The teacher who the little boy expected that the sky will seriously concern is planning to leave as soon as possible.
- b. The teacher who the little boy expected the sky to seriously concern is planning to leave as soon as possible.
- c. The teacher who the little boy's expectation about the sky will seriously concern is planning to leave as soon as possible.

- Who did the expectation about the sky concern?
the teacher – the boy
17. Some people are quite easy to impress.
- The girl who the teacher expected that the old building will thoroughly impress is telling all her friends about it.
 - The girl who the teacher expected the old building to thoroughly impress is telling all her friends about it.
 - The girl who the teacher's expectation about the old building will thoroughly impress is telling all her friends about it.
- Who will be impressed? *the girl – the teacher*
18. Trying to change things that infuriate you can be a driving force in one's work.
- The author who the agent revealed that the media had frequently infuriated is writing a new book.
 - The author who the agent revealed the media to have frequently infuriated is writing a new book.
 - The author who the agent's revelations about the media had frequently infuriated is writing a new book.
- Who got infuriated? *the author – the agent*
19. Sometimes even very simple things can be pleasing.
- The boy who the mother expected that the weather will easily please can't wait to go outside.
 - The boy who the mother expected the weather to easily please can't wait to go outside.
 - The boy who the mother's expectation about the weather will easily please can't wait to go outside.
- Who expected something? *the mother – the boy*
20. Taking courses outside of one's field of study can be overwhelming.
- The undergrad who the math professor suspected that the problem had really overwhelmed has already given up on that class.
 - The undergrad who the math professor suspected the problem to have really overwhelmed has already given up on that class.
 - The undergrad who the math professor's suspicion about the problem had really overwhelmed has already given up on that class.
- Who had a suspicion about something?
the professor – the undergrad
21. If you are famous, reporters often will want to embarrass you with shady affairs.
- The movie star who the spokesperson revealed that social media had thoroughly embarrassed is strangely enough not going to sue.
 - The movie star who the spokesperson revealed social media to have thoroughly embarrassed is strangely enough not going to sue.
 - The movie star who the spokesperson's revelation about social media had thoroughly embarrassed is strangely enough not going to sue.
- Who revealed something?
the spokesperson – the movie star
22. At the family reunion last month someone seemed distressed.
- The aunt who the girl suspected that the controversial rumor had seriously distressed was hoping to get married.
 - The aunt who the girl suspected the controversial rumor to have seriously distressed was hoping to get married.
 - The aunt who the girl's suspicion about the controversial rumor had seriously distressed was hoping to get married.
- Who suspected something regarding the rumor?
the girl – the aunt
23. Alarming news arrived today.
- The daughter who the guidance counselor suspected that the court decision had probably alarmed might leave home for good.
 - The daughter who the guidance counselor suspected the court decision to have probably alarmed might leave home for good.
 - The daughter who the guidance counselor's suspicion about the court decision had probably alarmed might leave home for good.
- Who suspected something?
the guidance counselor – the daughter
24. At the press conference last Monday several people became very agitated.
- The journalist who the union member believed that the tax policy had intensely agitated was planning a series of articles.
 - The journalist who the union member believed the tax policy to have intensely agitated was planning a series of articles.
 - The journalist who the union member's beliefs about the tax policy had intensely agitated was planning a series of articles.
- Who believed something regarding the tax policy?
the union member – the journalist
25. Some employees at this hospital are very sensitive and easy to offend.
- The nurse who the doctor believed that the admitted patient had gravely offended is well-known for not getting along with people.
 - The nurse who the doctor believed the admitted patient to have gravely offended is well-known for not getting along with people.
 - The nurse who the doctor's beliefs about the admitted patient had gravely offended is well-known for not getting along with people.

Who believed something about the patient?
the doctor – the nurse

26. At the recent hearing, everyone looks calm and composed.
- The minister who the attorney believed that the charges had curiously calmed down is secluding himself from the public.
 - The minister who the attorney believed the charges to have curiously calmed down is secluding himself from the public.
 - The minister who the attorney's beliefs about the charges had curiously calmed down is secluding himself from the public.

Who had a belief about the charges?
the attorney – the minister

27. Many students pursue grad school because they are fascinated by something.
- The history major who the professor discovered that the Middle Ages had deeply fascinated is especially interested in the crusades.
 - The history major who the professor discovered the Middle Ages to have deeply fascinated is especially interested in the crusades.
 - The history major who the professor's discoveries about the Middle Ages had deeply fascinated is especially interested in the crusades.

Who had discovered something?
the professor – the history major

28. Foreign places can make for an unsettling experience.
- The tourists who the local zealot expected that the ruins had deeply unsettled were looking for an adventure.
 - The tourists who the local zealot expected the ruins to have deeply unsettled were looking for an adventure.
 - The tourists who the local zealot's expectations about the ruins had deeply unsettled were looking for an adventure.

Who expected something? *the zealot – the tourists*

29. Children can be intrigued by very simple things.
- The girl who the uncle expects that the puzzle will endlessly intrigue does not like to play outside.
 - The girl who the uncle expects the puzzle to endlessly intrigue does not like to play outside.
 - The girl who the uncle's expectation about the puzzle will endlessly intrigue does not like to play outside.

Who expects something? *the uncle – the girl*

30. The work in a lab is full of surprises.
- The lab technician who the researcher proved that the new supercomputer had genuinely surprised is working late again.
 - The lab technician who the researcher proved the supercomputer to have genuinely surprised is working late again.

- The lab technician who the researcher's proof about the supercomputer had genuinely surprised is working late again.

Who proved something?
the researcher – the lab technician

BIBLIOGRAPHY

- Abels, Klaus (2003). Successive cyclicity, anti-locality, and adposition stranding. Ph.D. dissertation, University of Connecticut, Storrs, Connecticut.
- Abels, Klaus (2007). Towards a restrictive theory of (remnant) movement. In: *Linguistic Variation Yearbook 7*, ed. by Jeroen van Craenenbroeck & Johan Rooryck, Amsterdam: John Benjamins, pp. 53–120.
- Abels, Klaus (2009). Some implications of improper movement for cartography. In: *Alternatives to Cartography*, ed. by Jeroen van Craenenbroeck, Berlin: de Gruyter, pp. 325–359.
- Abels, Klaus (2012a). The Italian left periphery: A view from locality. *Linguistic Inquiry* 43: 229–254.
- Abels, Klaus (2012b). *Phases: An Essay on Cyclicity in Syntax*. Berlin: de Gruyter.
- Adger, David (2003). *Core Syntax: A Minimalist Approach*. Oxford: Oxford University Press.
- Adger, David (2013). *A Syntax of Substance*. Cambridge, MA: MIT Press.
- Aissen, Judith & David Perlmutter (1976). Clause reduction in Spanish. In: *Proceedings of the 2nd Annual Meeting of the Berkeley Linguistic Society*, University of California, Berkeley: BLS, pp. 1–30.
- Aldridge, Edith (2008). Phase-based account of extraction in Indonesian. *Lingua* 118: 1440–1469.
- Alexiadou, Artemis & Elena Anagnostopoulou (2002). Raising without infinitives and the nature of agreement. In: *Dimensions of Movement: From Features to Remnants*, ed. by Artemis Alexiadou, Elena Anagnostopoulou, Sjef Barbiers & Hans-Martin Gärtner, Amsterdam: Benjamins, pp. 17–30.
- Anand, Pranav & Andrew Nevins (2006). The locus of ergative case assignment: Evidence from scope. In: *Ergativity: Emerging Issues*, ed. by Alana Johns, Diane Massam & Juvenal Ndayiragije, Dordrecht: Springer, pp. 3–25.
- Antomo, Mailin (2012). Interpreting embedded verb second. In: *Proceedings of ConSOLE XVII*, ed. by Camelia Constantinescu, Bert Le Bruyn & Kathrin Linke, Leiden University, pp. 27–51.

- Aoshima, Sachiko, Colin Phillips & Amy Weinberg (2004). Processing filler-gap dependencies in a head-final language. *Journal of Memory and Language* 51: 23–54.
- Arregi, Karlos (2003). Clausal pied-piping. *Natural Language Semantics* 11: 115–143.
- Baltin, Mark (1978). Toward a theory of movement rules. Ph.D. dissertation, MIT, Cambridge, MA.
- Baltin, Mark (1982). A landing site theory of movement rules. *Linguistic Inquiry* 13: 1–38.
- Barr, Dale, Roger Levy, Christoph Scheepers & Harry Tily (2013). Random effects structure for confirmatory hypothesis testing: Keep it maximal. *Journal of Memory and Language* 68: 255–278.
- Barss, Andrew (1986). Chains and anaphoric dependence. Ph.d. thesis, MIT, Cambridge, MA.
- Bartek, Brian, Richards Lewis, Shravan Vasishth & Mason Smith (2011). In search of on-line locality effects in sentence comprehension. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 37: 1178–1198.
- Bates, Douglas, Martin Maechler, Ben Bolker & Steven Walker (2014). lme4: Linear mixed-effects models using Eigen and S4, R package version 1.1-7.
- Bayer, Josef (1989). Notes on the ECP in English and German. *Groninger Arbeiten zur Germanistischen Linguistik* 30: 1–55.
- Bayer, Josef (1996). *Directionality and Logical Form: On the Scope of Focussing Particles and Wh-in-situ*. Dordrecht: Kluwer.
- Bayer, Josef & Jaklin Kornfilt (1990). Restructuring effects in German. In: *Parametric Variation in Germanic and Romance: Proceedings from a DYANA Workshop*, ed. by Elisabeth Engdahl, Mike Reape, Martin Mellor & Richard Cooper, Edinburgh: University of Edinburgh, Centre for Cognitive Science, pp. 21–42.
- Bayer, Josef & Jaklin Kornfilt (1994). Against scrambling as an instance of move-alpha. In: *Studies on Scrambling: Movement and Non-Movement Approaches to Free Word-Order Phenomena*, ed. by Norbert Corver & Henk van Riemsdijk, Berlin: Mouton de Gruyter, pp. 17–60.
- Bayer, Josef & Martin Salzmann (2013). That-trace effects and resumption: How Improper Movement can be repaired. In: *Repairs: The Added Value of Being Wrong*, ed. by Patrick Brandt & Eric Fuß, Berlin: de Gruyter, pp. 275–333.
- Bech, Gunnar (1955/1957). *Studien über das deutsche Verbum Infinitum*. Copenhagen: Munksgaard, 2 vols.
- Beck, Sigrd (1996). Quantified structures as barriers for LF movement. *Natural Language Semantics* 4: 1–56.
- Beck, Sigrd (2006). Intervention effects follow from focus interpretation. *Natural Language Semantics* 14: 1–56.
- Beck, Sigrd & Shin-Sook Kim (1997). On wh- and operator scope in Korean. *Journal of East Asian Linguistics* 6: 339–384.

- Béjar, Susana & Milan Rezac (2009). Cyclic agree. *Linguistic Inquiry* 40: 35–73.
- Berwick, Robert & Amy Weinberg (1984). *The Grammatical Basis of Linguistic Performance: Language Use and Acquisition*. Cambridge, MA: MIT Press.
- Bhatia, Sakshi & Ethan Poole (2015). Deriving subject and antisubject orientation, Presentation at *Formal Approaches to South Asian Linguistics 6* (FASAL 6).
- Bhatt, Rajesh (2005). Long distance agreement in Hindi-Urdu. *Natural Language and Linguistic Theory* 23: 757–807.
- Bhatt, Rajesh (2006). *Covert Modality in Non-Finite Contexts*. Berlin: de Gruyter.
- Bhatt, Rajesh (2016). Minimalist approaches to South Asian syntax. In: *The Languages and Linguistics of South Asia: A Comprehensive Guide*, ed. by Hans Henrich Hock & Elena Bashir, Berlin: de Gruyter, pp. 506–529.
- Bhatt, Rajesh & Elena Anagnostopoulou (1996). Object shift and specificity: Evidence from *ko*-phrases in Hindi. In: *Papers from the 32nd Chicago Linguistics Society*, ed. by Lise M. Dobrin, Kora Singer & Lisa McNair, Chicago: Chicago Linguistics Society, pp. 11–22.
- Bhatt, Rajesh & Veneeta Dayal (2007). Rightward scrambling as rightward remnant movement. *Linguistic Inquiry* 38: 287–301.
- Bhatt, Rajesh & Vincent Homer (2014). The interpretation of infinitival negation in Hindi-Urdu, Handout of talk at FASAL 4, Rutgers University.
- Bhatt, Rajesh & Stefan Keine (to appear). Long-distance agreement. In: *The Companion to Syntax: Second Edition*, ed. by Martin Everaert & Henk van Riemsdijk, Hoboken, NJ: Wiley-Blackwell.
- Bhattacharya, Tanmoy (2002). Peripheral and clause-internal complementizers in Bangla: A case of remnant movement. In: *Proceedings of the 12th Western Conference On Linguistics (WECOL 12)*, ed. by Vida Samiian, Fresno: California State University, Department of Linguistics, pp. 100–112.
- Bickel, Balthasar & Yogendra Yadava (2000). A fresh look at grammatical relations in Indo-Aryan. *Lingua* 110: 343–373.
- Bierwisch, Manfred (1963). *Grammatik des deutschen Verbs*. Studia Grammatica II, Berlin: Akademie Verlag.
- Bobaljik, Jonathan (2008). Where's Phi? Agreement as a postsyntactic operation. In: *Phi Theory: Phi-Features Across Interfaces and Modules*, ed. by Daniel Harbour, David Adger & Susana Béjar, Oxford: Oxford University Press, pp. 295–328.
- Bobaljik, Jonathan & Susanne Wurmbrand (2005). The domain of agreement. *Natural Language and Linguistic Theory* 23: 809–865.
- Boeckx, Cedric (2003). *Islands and Chains: Resumption as Stranding*. Amsterdam: Benjamins.
- Boeckx, Cedric (2004). Long-distance agreement in Hindi: Some theoretical implications. *Studia Linguistica* 58: 23–36.

- Boeckx, Cedric & Kleanthes Grohmann (2007). Putting phases in perspective. *Syntax* 10: 204–222.
- Boland, Julie (2005). Cognitive mechanisms and syntactic theory. In: *Twenty-First Century Psycholinguistics: Four Cornerstones*, ed. by Anne Cutler, Mahwah, NJ: Lawrence Erlbaum, pp. 23–42.
- Boland, Julie, Michael Tanenhaus, Susan Garnsey & Greg Carlson (1995). Verb argument structure in parsing and interpretation. *Journal of Memory and Language* 34: 774–806.
- Bošković, Željko (2002). A-movement and the EPP. *Syntax* 5: 167–218.
- Bošković, Željko (2003). Agree, phases, and intervention effects. *Linguistic Analysis* 33: 54–96.
- Bošković, Željko (2005). On the locality of left branch extraction and the structure of NP. *Studia Linguistica* 59: 1–45.
- Bošković, Željko (2014). Now I'm a phase, now I'm not a phase: On the variability of phases with extraction and ellipsis. *Linguistic Inquiry* 45: 27–89.
- Brandner, Ellen (2004). Head-movement in minimalism, and V2 as FORCE-marking. In: *The Syntax and Semantics of the Left Periphery*, ed. by Horst Lohnstein & Susanne Trissler, Berlin: de Gruyter, pp. 97–138.
- Branigan, Phil & Marguerita MacKenzie (2002). Altruism, \bar{A} -movement and object agreement in Innu-aimûn. *Linguistic Inquiry* 33: 385–407.
- Bresnan, Joan (1976). On the form and functioning of transformations. *Linguistic Inquiry* 7: 3–40.
- Bruening, Benjamin (2001). Syntax at the edge: Cross-clausal phenomena and the syntax of Passamaquoddy. Ph.D. dissertation, MIT, Cambridge, MA.
- Bruening, Benjamin (2006). Differences between the wh-scope-marking and wh-copy constructions in Passamaquoddy. *Linguistic Inquiry* 37: 25–49.
- Büring, Daniel & Katharina Hartmann (1994). The dark side of wh-movement. *Linguistische Berichte* 149: 56–74.
- Butt, Miriam (1993). The structure of complex predicates in Urdu. Ph.D. dissertation, Stanford University.
- Butt, Miriam (1995). *The Structure of Complex Predicates in Urdu*. Stanford, California: CSLI Publications.
- Butt, Miriam (2014). Control vs. complex predication: Identifying non-finite complements. *Natural Language and Linguistic Theory* 32: 165–190.
- Butt, Miriam & Tracy King (2004). The status of case. In: *Clause Structure in South Asian Languages*, ed. by Veneeta Dayal & Anoop Mahajan, Boston: Kluwer, pp. 153–198.
- Cable, Seth (2007). The grammar of Q: Q-particles and the nature of wh-fronting, as revealed by the wh-questions of Tlingit. Ph.D. dissertation, MIT, Cambridge, MA.

- Cable, Seth (2010). *The Grammar of Q: Q-Particles and Pied-Piping*. Oxford: Oxford University Press.
- Carstens, Vicki (2010). Grammatical gender and the theory of uninterpretable features. In: *Exploring Crash-Proof Grammars*, ed. by Michael Putnam, Amsterdam: John Benjamins, pp. 31–57.
- Carstens, Vicki (2011). Hyperactivity and hyperagreement in Bantu. *Lingua* 121: 721–741.
- Carstens, Vicki (2016). Delayed valuation: A reanalysis of goal features, “upward” complementizer agreement, and the mechanics of case. *Syntax* 19: 1–42.
- Carstens, Vicki & Michael Diercks (2013). Parametrizing case and activity: Hyperraising in Bantu. In: *Proceedings of the 40th North East Linguistic Society (NELS 40)*, ed. by Seda Kan, Claire Moore-Cantwell & Robert Staubs, Amherst, MA: GLSA, pp. 99–115.
- Chandra, Pritha (2007). (Dis)Agree: Movement and agreement reconsidered. Ph.D. dissertation, University of Maryland, College Park.
- Chen, Evan, Edward Gibson & Florian Wolf (2005). Online syntactic storage costs in sentence comprehension. *Journal of Memory and Language* 52: 144–169.
- Chomsky, Noam (1957). *Syntactic Structures*. The Hague and Paris: Mouton.
- Chomsky, Noam (1964). The logical basis of linguistic theory. In: *Proceedings of the Ninth International Congress of Linguists, Cambridge, Mass., 1962*, ed. by Horace Lunt, The Hague: Mouton, pp. 914–978.
- Chomsky, Noam (1965). *Aspects of the Theory of Syntax*. Cambridge, MA: MIT Press.
- Chomsky, Noam (1973). Conditions on transformations. In: *A Festschrift for Morris Halle*, ed. by Stephen Anderson & Paul Kiparsky, New York: Academic Press, pp. 232–286.
- Chomsky, Noam (1977). On wh-movement. In: *Formal Syntax*, ed. by Peter Culicover, Tom Wasow & Adrian Akmajian, New York: Academic Press, pp. 71–132.
- Chomsky, Noam (1981). *Lectures on Government and Binding*. Dordrecht: Foris.
- Chomsky, Noam (1986). *Barriers*. Cambridge, MA: MIT Press.
- Chomsky, Noam (1993). A minimalist program for syntactic theory. In: *The View from Building 20*, ed. by Ken Hale & Samuel Jay Keyser, Cambridge, MA: MIT Press, pp. 1–52.
- Chomsky, Noam (1995a). Bare phrase structure. In: *Government and Binding Theory and the Minimalist Program*, ed. by Gert Webelhuth, Oxford: Blackwell, pp. 383–440.
- Chomsky, Noam (1995b). *The Minimalist Program*. Cambridge, MA: MIT Press.
- Chomsky, Noam (2000). Minimalist inquiries: The framework. In: *Step by Step: Essays in Syntax in Honor of Howard Lasnik*, ed. by Roger Martin, David Michaels & Juan Uriagereka, Cambridge, MA: MIT Press, pp. 89–155.

- Chomsky, Noam (2001). Derivation by phase. In: *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, Cambridge, MA: MIT Press, pp. 1–52.
- Chomsky, Noam (2004). Beyond explanatory adequacy. In: *Structures and Beyond. The Cartography of Syntactic Structures, Vol. 3*, ed. by Adriana Belletti, Oxford: Oxford University Press, pp. 104–131.
- Chomsky, Noam (2005). Three factors in language design. *Linguistic Inquiry* 36: 1–22.
- Chomsky, Noam (2007). Approaching UG from below. In: *Interfaces + Recursion = Language?*, ed. by Uli Sauerland & Hans-Martin Gärtner, Berlin: de Gruyter, pp. 1–31.
- Chomsky, Noam (2008). On phases. In: *Foundational Issues in Linguistic Theory: Essays in Honour of Jean-Roger Vergnaud*, ed. by Robert Freidin, Carlos Otero & Maria Luisa Zubizarreta, Cambridge, MA: MIT Press, pp. 89–155.
- Chomsky, Noam (2012). Foreword. In: *Phases: Developing the Framework*, ed. by Ángel Gallego, Berlin: de Gruyter, pp. 1–7.
- Chomsky, Noam (2013). Problems of projection. *Lingua* 130: 33–49.
- Chung, Sandra (1994). Wh-agreement and ‘referentiality’ in Chamorro. *Linguistic Inquiry* 25: 1–44.
- Chung, Sandra (1998). *The Design of Agreement: Evidence from Chamorro*. Chicago: Chicago University Press.
- Chung, Sandra (2004). Restructuring and verb-initial order in Chamorro. *Syntax* 7: 199–233.
- Cinque, Guglielmo (1989). On embedded verb second clauses and ergativity in German. In: *Sentential Complementation and the Lexicon: Studies in Honour of Wim de Geest*, ed. by Dany Jaspers, Wim Klooster, Yvan Putseys & Pieter Seuren, Dordrecht: Foris, pp. 77–96.
- Cinque, Guglielmo (1990). *Types of A-bar Dependencies*. Cambridge, MA: MIT Press.
- Cinque, Guglielmo (1999). *Adverbs and Functional Heads: A Cross-Linguistic Perspective*. Oxford: Oxford University Press.
- Clements, George (1994). Principles of tone assignment in Kikuyu. In: *Autosegmental Studies in Bantu Tone*, ed. by George Clements & John Goldsmith, Dordrecht: Foris, pp. 281–339.
- Clements, George & Kevin Ford (1979). Kikuyu tone shift and its synchronic consequences. *Linguistic Inquiry* 10: 179–210.
- Cole, Peter (1982). *Imbabura Quechua*. Amsterdam: North-Holland Publishing.
- Collins, Chris (2005a). A smuggling approach to raising in English. *Linguistic Inquiry* 36: 289–298.
- Collins, Chris (2005b). A smuggling approach to the passive in English. *Syntax* 8: 81–120.
- Coon, Jessica, Pedro Mateo Pedro & Omer Preminger (2014). The role of case in A-bar extraction asymmetries. *Linguistic Variation* 14: 179–242.

- Corver, Norbert & Henk van Riemsdijk (1994). Introduction: Approaches to and properties of scrambling. In: *Studies on Scrambling: Movement and Non-Movement Approaches to Free Word-Order Phenomena*, ed. by Norbert Corver & Henk van Riemsdijk, Berlin: de Gruyter, pp. 1–15.
- Crain, Stephen & Janet Fodor (1985). How can grammars help parsers? In: *Natural Language Parsing: Psychological, Computational, and Theoretical Perspectives*, ed. by David Dowty, Lauri Karttunen & Arnold Zwicky, Cambridge: Cambridge University Press, pp. 94–128.
- Cremers, Crit (1983). On two types of infinitival complementation. In: *Linguistic Categories: Auxiliaries and Related Puzzles, Vol. 1: Categories*, ed. by Frank Heny & Barry Richards, Dordrecht: Reidel, pp. 169–221.
- Cresti, Diana (1995). Extraction and reconstruction. *Natural Language Semantics* 3: 79–122.
- Culicover, Peter (1996). On distinguishing \bar{A} -movements. *Linguistic Inquiry* 27: 445–463.
- d’Avis, Franz-Josef (1996). On *wh*-islands in German. In: *On Extraction and Extraposition in German*, ed. by Uli Lutz & Jürgen Pafel, Amsterdam: John Benjamins, pp. 89–120.
- Davison, Alice (1991). Feature percolation and agreement in Hindi/Urdu, Ms., University of Iowa.
- Davison, Alice (1999). Ergativity: Functional and formal issues. In: *Functionalism and Formalism in Linguistics, Vol. 1*, ed. by Michael Darnell, Edith Moravcsik, Federick Newmeyer, Michael Noonan & Kathleen Wheatley, Amsterdam: John Benjamins, pp. 177–208.
- Davison, Alice (2004a). Non-nominative subjects in Hindi/Urdu: VP structure and case parameters. In: *Non-Nominative Subjects, Vol. 1*, ed. by Peri Bhaskararao & Karumuri Venkata Subbarao, Amsterdam: Benjamins, pp. 141–168.
- Davison, Alice (2004b). Structural case, lexical case, and the verbal projection. In: *Clause Structure in South Asian Languages*, ed. by Veneeta Dayal & Anoop Mahajan, Dordrecht: Kluwer, pp. 199–225.
- Davison, Alice (2008). A case restriction on control: Implications for movement. *Journal of South Asian Languages* 1: 29–54.
- Davison, Alice (2010). Long-distance agreement and restructuring, Ms., University of Iowa.
- Davison, Alice (2014). Non-finite complements and modality in *de-na* ‘allow’ in Hindi-Urdu. *Natural Language and Linguistic Theory* 32: 137–164.
- Dayal, Veneeta (1994a). Binding facts in Hindi and the scrambling phenomenon. In: *Theoretical Perspectives on Word Order in South Asian Languages*, ed. by Miriam Butt, Tracy Holloway King & Gillian Ramchand, Stanford: CSLI, pp. 237–262.
- Dayal, Veneeta (1994b). Scope marking as indirect *wh*-dependency. *Natural Language Semantics* 2: 137–170.
- Dayal, Veneeta (1996). *Locality in Wh-Quantification: Questions and Relative Clauses in Hindi*. Dordrecht: Kluwer.

- Dayal, Veneeta (2000). Scope marking: Cross-linguistic variation in indirect dependency. In: *Wh-Scope Marking*, ed. by Uli Lutz, Gereon Müller & Arnim von Stechow, Amsterdam: John Benjamins, pp. 157–193.
- Dayal, Veneeta (to appear). Does Hindi-Urdu have feature-driven wh movement to Spec vP?, *Linguistic Inquiry*.
- de Hoop, Helen & Bhuvana Narasimhan (2005). Differential case-marking in Hindi. In: *Competition and Variation in Natural Languages: The Case For Case*, ed. by Mengistu Amberber & Helen de Hoop, Amsterdam: Elsevier, pp. 321–345.
- de Vincenzi, Marica (1991). *Syntactic Parsing Strategies in Italian*. Dordrecht: Kluwer.
- den Besten, Hans (1983). On the interaction of root transformations and lexical deletive rules. In: *On the Formal Syntax of the Westgermania*, ed. by Werner Abraham, Amsterdam: John Benjamins, pp. 47–131.
- den Besten, Hans & Jean Rutten (1989). On verb raising, extraposition, and free word order in Dutch. In: *Sentential Complementation and the Lexicon*, ed. by Dany Jaspers, Wim Klooster, Yvan Putseys & Pieter Seuren, Dordrecht: Foris, pp. 77–92.
- den Dikken, Marcel (2006). A reappraisal of vP being phasal: A reply to Legate, Ms., CUNY.
- den Dikken, Marcel (2007). Phase extension: Contours of a theory of the role of head movement in phrasal extraction. *Theoretical Linguistics* 33: 1–41.
- Déprez, Viviane (1989). On the typology of syntactic positions and the nature of chains: Move α to the specifier of functional projections. Ph.D. dissertation, MIT, Cambridge, MA.
- Diercks, Michael (2012). Parametrizing Case: Evidence from Bantu. *Syntax* 15: 253–286.
- Diesing, Molly (1992). *Indefinites*. Cambridge, MA: MIT Press.
- Dillon, Brian, Wing-Yee Chow, Matthew Wagers, Taomei Guo, Fengqin Liu & Colin Phillips (2014). The structure-sensitivity of memory access: Evidence from Mandarin Chinese. *Frontiers in Psychology* 5: 1025.
- Drummond, Alex (2013). *Ibex Farm* (version 0.3) [software]. (Retrieved from <https://github.com/addrummond/ibexfarm>).
- Dukes, Michael (1992). On the status of Chamorro wh-agreement. In: *Proceedings of the Eleventh West Coast Conference on Formal Linguistics (WCCFL 11)*, Stanford: CSLI Publications, pp. 177–190.
- Epstein, Samuel David & T. Daniel Seely (2002). Rule applications as cycles in a level-free syntax. In: *Derivation and Explanation in the Minimalist Program*, ed. by Samuel David Epstein & T. Daniel Seely, Oxford: Blackwell, pp. 65–89.
- Epstein, Samuel David & T. Daniel Seely (2006). *Derivations in Minimalism: Exploring the Elimination of A-Chains and the EPP*. Cambridge: Cambridge University Press.

- Fanselow, Gisbert (1987). *Konfiguralität: Untersuchungen zur Universalgrammatik am Beispiel des Deutschen*. Tübingen: Narr.
- Fanselow, Gisbert (1989). Coherent infinitives in German: Restructuring vs. IP-complementation. In: *Syntactic Phrase Structure Phenomena in Noun Phrases and Sentences*, ed. by Christa Bhatt, Elisabeth Löbel & Claudia Schmidt, Amsterdam: John Benjamins, pp. 1–16.
- Fanselow, Gisbert (1990). Scrambling as NP-movement. In: *Scrambling and Barriers*, ed. by Günther Grewendorf & Wolfgang Sternefeld, Amsterdam: Benjamins, pp. 113–140.
- Fanselow, Gisbert (1991). Minimale Syntax. *Groninger Arbeiten zur Germanistischen Linguistik* 32: 1–402, Published version of Habilitation thesis, Universität Passau.
- Fanselow, Gisbert (2001). Features, θ -roles, and free constituent order. *Linguistic Inquiry* 32: 405–436.
- Fanselow, Gisbert (2004). The MLC and derivational economy. In: *Minimality Effects in Syntax*, ed. by Arthur Stepanov, Gisbert Fanselow & Ralf Vogel, Berlin: de Gruyter, pp. 73–123.
- Felser, Claudia (2004). Wh-copying, phases and successive cyclicity. *Lingua* 114: 543–574.
- Ferreira, Fernanda & Charles Clifton (1986). The independence of syntactic processing. *Journal of Memory and Language* 25: 348–368.
- Fiebach, Christian, Matthias Schlesewsky & Angela Friederici (2002). Separating syntactic memory costs and syntactic integration costs during parsing: The processing of German wh-questions. *Journal of Memory and Language* 47: 250–272.
- Fitzpatrick, Justin (2002). On minimalist approaches to the locality of movement. *Linguistic Inquiry* 33: 443–463.
- Fox, Danny (1999). Reconstruction, Binding Theory, and the interpretation of chains. *Linguistic Inquiry* 30: 157–196.
- Fox, Danny (2000). *Economy and Semantic Interpretation*. Cambridge, MA: MIT Press.
- Fox, Danny & Howard Lasnik (2003). Successive-cyclic movement and island repair: The difference between sluicing and VP-ellipsis. *Linguistic Inquiry* 34: 143–154.
- Frank, Robert (2004). Restricting grammatical complexity. *Cognitive Science* 28: 669–697.
- Frazier, Lyn (1978). On comprehending sentences: Syntactic parsing strategies. Ph.D. dissertation, University of Connecticut, Storrs.
- Frazier, Lyn (1987). Syntactic processing: Evidence from Dutch. *Natural Language and Linguistic Theory* 5: 519–560.
- Frazier, Lyn & Charles Clifton (1989). Successive cyclicity in the grammar and the parser. *Language and Cognitive Processes* 4: 93–126.
- Frazier, Lyn & Giovanni Flores D'Arcais (1989). Filler-driven parsing: A study of gap filling in Dutch. *Journal of Memory and Language* 28: 331–344.

- Frey, Werner (1993). *Syntaktische Bedingungen für die semantische Interpretation: Über Binding, implizite Argumente und Skopus*. Berlin: Akademie Verlag.
- Frey, Werner (2004). A medial topic position for German. *Linguistische Berichte* 198: 153–190.
- Frey, Werner (2011). Peripheral adverbial clauses, their licensing and the prefield in German. In: *Satzverknüpfungen: Zur Interaktion von Form, Bedeutung und Diskursfunktion*, ed. by Eva Breindl, Gisella Ferraresi & Anna Volodina, Berlin: de Gruyter, pp. 41–77.
- Fukui, Naoki (1993). A note on improper movement. *The Linguistic Review* 10: 111–126.
- Gallego, Ángel & Juan Uriagereka (2007). Sub-extraction from subjects: A phase theory account. In: *Romance Linguistics 2006*, ed. by José Camacho, Nydia Flores-Ferrán, Liliana Sánchez, Viviane Déprez & María José Cabrera, Amsterdam: John Benjamins, pp. 155–168.
- Gambhir, Vijay (1981). Syntactic restrictions and discourse functions of word order in Standard Hindi. Ph.D. dissertation, University of Pennsylvania, Philadelphia, PA.
- Garnsey, Susan, Michael Tanenhaus & Robert Chapman (1989). Evoked potentials and the study of sentence comprehension. *Journal of Psycholinguistic Research* 18: 51–60.
- Gärtner, Hans-Martin (2001). Are there V2 relative clauses in German? *Journal of Comparative Germanic Linguistics* 3: 97–141.
- Gärtner, Hans-Martin (2002). On the force of V2 declaratives. *Theoretical Linguistics* 28: 33–42.
- Gazdar, Gerald, Ewan Klein, Geoffrey Pullum & Ivan Sag (1985). *Generalized Phrase Structure Grammar*. Oxford: Blackwell.
- Georgi, Doreen (2014). Opaque interactions of Merge and Agree: On the nature and order of elementary operations. Ph.D. dissertation, Universität Leipzig.
- Gibson, Edward (1998). Linguistic complexity: Locality of syntactic dependencies. *Cognition* 68: 1–76.
- Gibson, Edward (2000). The dependency locality theory: A distance-based theory of linguistic complexity. In: *Image, Language, Brain*, ed. by Alec Marantz, Yasushi Miyashita & Wayne O’Neill, Cambridge, MA: MIT Press, pp. 95–126.
- Gibson, Edward & Neal Pearlmuter (1998). Constraints on sentence comprehension. *Trends in Cognitive Science* 2: 262–268.
- Gibson, Edward & Tessa Warren (2004). Reading-time evidence for intermediate linguistic structure in long-distance dependencies. *Syntax* 7: 55–78.
- Gordon, Peter, Randall Hendrick & Marcus Johnson (2001). Memory interference during language processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 27: 1411–1423.
- Grewendorf, Günther (1988). *Aspekte der deutschen Syntax: Eine Rektions-Bindungs-Analyse*. Tübingen: Narr.

- Grewendorf, Günther (2001). Multiple wh-fronting. *Linguistic Inquiry* 32: 87–122.
- Grewendorf, Günther (2003). Improper remnant movement. *Gengo Kenkyu: The Journal of the Linguistic Society of Japan* 123: 47–94.
- Grewendorf, Günther (2005). The discourse configurationality of scrambling. In: *The Free Word Order Phenomenon: Its Syntactic Sources and Diversity*, ed. by Joachim Sabel & Mamoru Saito, Berlin: Mouton de Gruyter, pp. 75–135.
- Grewendorf, Günther (2013). Problems of remnant movement, Ms., Universität Frankfurt/Main.
- Grewendorf, Günther (2015). Problems of remnant movement. In: *Remnant Movement*, ed. by Günther Grewendorf, Berlin: de Gruyter, pp. 3–31.
- Grewendorf, Günther & Joachim Sabel (1994). Long scrambling and incorporation. *Linguistic Inquiry* 25: 263–308.
- Grewendorf, Günther & Joachim Sabel (1999). Scrambling in German and Japanese. *Natural Language and Linguistic Theory* 17: 1–65.
- Grewendorf, Günther & Wolfgang Sternefeld (1990). Scrambling theories. In: *Scrambling and Barriers*, ed. by Günther Grewendorf & Wolfgang Sternefeld, Amsterdam: John Benjamins, pp. 3–37.
- Grimshaw, Jane (1991). Extended projections, Ms., Brandeis University.
- Grimshaw, Jane (2000). Extended projection and locality. In: *Lexical Specification and Insertion*, ed. by Peter Coopmans, Martin Everaert & Jane Grimshaw, Amsterdam: Benjamins, pp. 115–133.
- Grodner, Daniel & Edward Gibson (2005). Consequences of the serial nature of linguistic input for sentential complexity. *Cognitive Science* 29: 261–291.
- Grodner, Daniel, Edward Gibson & Susanne Tunstall (2002). Syntactic complexity in ambiguity resolution. *Journal of Memory and Language* 46: 267–295.
- Grohmann, Kleanthes (1997). German superiority. *Groninger Arbeiten zur Germanistischen Linguistik* 40: 97–107.
- Grosu, Alexander & Julia Horvath (1984). The GB theory and raising in Rumanian. *Linguistic Inquiry* 15: 348–353.
- Gurtu, Madhu (1985). Anaphoric relations in Hindi and English. Ph.D. dissertation, CIEFL, Hyderabad.
- Gurtu, Madhu (1992). *Anaphoric Relations in Hindi and English*. New Delhi: Munshiram Manoharlal.
- Haegeman, Liliane (2006). Conditionals, factives and the left periphery. *Lingua* 116: 1651–1669.
- Haegeman, Liliane & Henk van Riemsdijk (1986). Verb projection raising, scope, and the typology of rules affecting verbs. *Linguistic Inquiry* 17: 417–466.
- Hagstrom, Paul (1998). Decomposing questions. Ph.D. dissertation, MIT, Cambridge, MA.

- Haider, Hubert (1983). Connectedness effects in German. *Groninger Arbeiten zur Germanistischen Linguistik* 23: 82–119.
- Haider, Hubert (1984). Topic, focus, and V-second. *Groninger Arbeiten zur Germanistischen Linguistik* 25: 72–120.
- Haider, Hubert (1986). Affect α : A reply to Lasnik and Saito, “On the nature of Proper Government”. *Linguistic Inquiry* 17: 113–126.
- Haider, Hubert (1993). *Deutsche Syntax – generativ*. Tübingen: Narr.
- Haider, Hubert (2003). V-clustering and clause union: Causes and effects. In: *Verb Constructions in German and Dutch*, ed. by Pieter Seuren & Gerard Kempen, Amsterdam: John Benjamins, pp. 91–126.
- Haider, Hubert (2010). *The Syntax of German*. Cambridge: Cambridge University Press.
- Halpert, Claire (2012). Argument licensing and agreement in Zulu. Ph.D. dissertation, MIT, Cambridge, MA.
- Halpert, Claire (2015). Raising parameters, Handout of talk at GLOW 38, Paris.
- Hamblin, C. L. (1973). Questions in Montague English. *Foundations of Language* 10: 41–53, Reprinted in 1976 in: *Montague Grammar*, ed. by Barbara Partee, New York: Academic Press: pp. 247–259.
- Hawkins, John (1994). *A Performance Theory of Order and Constituency*. Cambridge: Cambridge University Press.
- Heck, Fabian & Gereon Müller (2000). Successive cyclicity, long-distance superiority, and local optimization. In: *Proceedings of the 19th West Coast Conference on Formal Linguistics (WCCFL 19)*, ed. by Roger Billerey & Brook D. Lillehaugen, Somerville, MA: Cascadilla Press, pp. 218–231.
- Henry, Alison (1995). *Belfast English and Standard English: Dialect Variation and Parameter Setting*. Oxford: Oxford University Press.
- Hicks, Glyn (2009). *Tough*-constructions and their derivation. *Linguistic Inquiry* 40: 535–566.
- Hinterhölzl, Roland (2006). *Scrambling, Remnant Movement, and Restructuring in West Germanic*. Oxford: Oxford University Press.
- Hoji, Hajime (1985). Logical form constraints and configurational structures in Japanese. Ph.D. dissertation, University of Washington, Seattle, WA.
- Hornstein, Norbert & Andrew Nevins (2005). Review of “Representation Theory” by Edwin Williams. *Language* 81: 757–761.
- Horvath, Julia (1986). *FOCUS in the Theory of Grammar and the Syntax of Hungarian*. Dordrecht: Foris.

- Horvath, Julia (1995). Structural focus, structural case, and the notion of feature-assignment. In: *Discourse Configurational Languages*, ed. by Katalin É. Kiss, Oxford: Oxford University Press, pp. 28–64.
- Horvath, Julia (1997). The status of wh-expletives and the partial wh-movement construction of Hungarian. *Natural Language and Linguistic Theory* 15: 509–572.
- Hsu, Brian (2015). Variation in Bangla complementizer order at the syntax-prosody interface. In: *Proceedings of the 45th North East Linguistic Society (NELS 45)*, ed. by Thuy Bui & Deniz Özyıldız, Amherst, MA: GLSA, pp. 35–44.
- Huang, Cheng-Teh James (1982). Logical relations in Chinese and the theory of grammar. Ph.D. dissertation, MIT, Cambridge, MA.
- Iatridou, Sabine & David Embick (1997). Apropos *pro*. *Language* 73: 58–78.
- Jones, Douglas Arnold (1993). Binding as an interface condition: An investigation of Hindi scrambling. Ph.D. dissertation, MIT, Cambridge, MA.
- Just, Marcel Adam & Patricia Carpenter (1992). A capacity theory of comprehension: Individual differences in working memory. *Psychological Review* 99: 122–149.
- Just, Marcel Adam, Patricia Carpenter & Jaqueline Woolley (1982). Paradigms and processes in reading comprehension. *Journal of Experimental Psychology* 111: 228–238.
- Kaan, Edith (1997). Processing subject–object ambiguities in Dutch. Ph.D. dissertation, University of Groningen.
- Kaan, Edith, Anthony Harris, Edward Gibson & Phillip Holcomb (2000). The P600 as an index of syntactic integration difficulty. *Language and Cognitive Processes* 15: 159–201.
- Kaplan, Aaron (2005). Long-distance wh-movement in Chamorro. In: *Proceedings of the Twelfth Meeting of the Austronesian Formal Linguistics Association (AFLA XII)*, ed. by Jeffrey Heinz & Dimitris Ntelitheos, Los Angeles: UCLA, no. 12 in UCLA Working Papers in Linguistics, pp. 117–129.
- Karttunen, Lauri (1977). Syntax and semantics of questions. *Linguistics and Philosophy* 1: 3–44, Reprinted in 2002 in: *Formal Semantics: The Essential Readings*, ed. by Paul Portner & Barbara Partee, Oxford: Blackwell: pp. 382–420.
- Kayne, Richard (1994). *The Antisymmetry of Syntax*. Cambridge, MA: MIT Press.
- Kayne, Richard & Jean Yves Pollock (1978). Stylistic inversion, successive cyclicity, and Move NP in French. *Linguistic Inquiry* 9: 595–621.
- Keine, Stefan (2010). *Case and Agreement from Fringe to Core: A Minimalist Approach*. Berlin: de Gruyter.
- Keine, Stefan (2013). On the role of movement in Hindi/Urdu long-distance agreement. In: *Proceedings of the 42nd North East Linguistic Society (NELS 42)*, ed. by Stefan Keine & Shayne Sloggett, Amherst, MA: GLSA, pp. 273–284.

- Keine, Stefan & Rajesh Bhatt (to appear). Interpreting verb clusters, accepted for publication in *Natural Language and Linguistic Theory*.
- Kidwai, Ayesha (2000). *XP-Adjunction in Universal Grammar: Scrambling and Binding in Hindi-Urdu*. Oxford: Oxford University Press.
- Kim, Shin-Sook (2002). Intervention effects are focus effects. In: *Japanese/Korean Linguistics 10*, ed. by Noriko Akatsuka & Susan Strauss, Stanford, CA: CSLI, pp. 615–628.
- Kimball, John (1973). Seven principles of surface structure parsing in natural language. *Cognition* 2: 15–47.
- King, Jonathan & Marcel Adam Just (1991). Individual differences in syntactic processing: The role of working memory. *Journal of Memory and Language* 30: 580–602.
- King, Jonathan & Marta Kutas (1995). Who did what and when? Using word- and clause-level ERPs to monitor working memory usage in reading. *Journal of Cognitive Neuroscience* 7: 376–395.
- Kiss, Katalin É. (1987). *Configurationality in Hungarian*. Dordrecht: Reidel.
- Kiss, Katalin É. (1993). Wh-movement and specificity. *Natural Language and Linguistic Theory* 11: 85–120.
- Kitahara, Hisatsugu (1997). *Elementary Operations and Optimal Derivations*. Cambridge, MA: MIT Press.
- Koizumi, Masatoshi (1995). Phrase structure in minimalist syntax. Ph.D. dissertation, MIT, Cambridge, MA.
- Koopman, Hilda (2006). Agreement configurations: In defense of “Spec head”. In: *Agreement Systems*, ed. by Cedric Boeckx, Amsterdam: John Benjamins, pp. 159–199.
- Kotek, Hadas (2014). Composing questions. Ph.D. dissertation, MIT, Cambridge, MA.
- Kratzer, Angelika (1991). The representation of focus. In: *Handbuch Semantik/Handbook Semantics*, ed. by Arnim von Stechow & Dieter Wunderlich, Berlin: Berlin, pp. 825–834.
- Kratzer, Angelika (1996). Severing the external argument from its verb. In: *Phrase Structure and the Lexicon*, ed. by Johan Rooryck & Laurie Zaring, Dordrecht: Kluwer, pp. 109–137.
- Kratzer, Angelika & Junko Shimoyama (2002). Indeterminate pronouns: The view from Japanese. In: *Proceedings of the Third Tokyo Conference on Psycholinguistics*, ed. by Yukio Otsu, Tokyo: Hituzi Publishing Company, pp. 1–25.
- Krifka, Manfred (1998). Scope inversion under the rise-fall contour in German. *Linguistic Inquiry* 29: 75–112.
- Kuznetsova, Alexandra, Per Bruun Brockhoff & Rune Haubo Bojesen Christensen (2014). *lmerTest: Tests for Random and Fixed Effects for Linear Mixed Effect Models (lmer Objects of lme4 Package)*. R package version 2.0-11.

- Lahiri, Utpal (2002). On the proper treatment of “expletive *Wh*” in Hindi. *Lingua* 112: 501–540.
- Lahne, Antje (2009). Where there is fire there is smoke: Local modelling of successive-cyclic movement. Ph.D. dissertation, Universität Leipzig.
- Landau, Idan (2004). The scale of finiteness and the calculus of control. *Natural Language and Linguistic Theory* 22: 811–877.
- Lasnik, Howard & Cedric Boeckx (2006). Long NP-movement. In: *The Blackwell Companion to Syntax, Vol. III*, ed. by Martin Everaert & Henk van Riemsdijk, Oxford: Blackwell, pp. 109–130.
- Lasnik, Howard & Mamoru Saito (1992). *Move α* . Cambridge, MA: MIT Press.
- Lebeaux, David (1988). Language acquisition and the form of the grammar. Ph.D. dissertation, University of Massachusetts, Amherst, MA.
- Lebeaux, David (2009). *Where Does Binding Theory Apply?* Cambridge, MA: MIT Press.
- Lee, Ming-Wei (2004). Another look at the role of empty categories in sentence processing (and grammar). *Journal of Psycholinguistic Research* 33: 51–73.
- Lee-Schoenfeld, Vera (2007). *Beyond Coherence: The Syntax of Opacity in German*. Amsterdam: John Benjamins.
- Legate, Julie Anne (2003). Some interface properties of the phase. *Linguistic Inquiry* 34: 506–516.
- Legate, Julie Anne (2012). The size of phases. In: *Phases: Developing the Framework*, ed. by Ángel Gallego, Berlin: de Gruyter, pp. 233–250.
- Lewis, Richard (1996). Interference in short-term memory: The magical number two (or three) in sentence processing. *Journal of Psycholinguistic Research* 25: 93–115.
- Lewis, Richard & Shravan Vasishth (2005). An activation-based model of sentence processing as skilled memory retrieval. *Cognitive Science* 29: 1–45.
- Lewis, Richard, Shravan Vasishth & Julie Van Dyke (2006). Computational principles of working memory in sentence comprehension. *Trends in Cognitive Science* 10: 447–454.
- Li, Yafei (1990). Conditions on X^0 -movement. Ph.D. dissertation, MIT, Cambridge, MA.
- Lin, Chien-Jer Charles (2006). Grammar and parsing: A typological investigation of relative-clause processing. Ph.D. dissertation, University of Arizona, Tuscon, AZ.
- Lin, Chien-Jer Charles & Thomas Bever (2006). Subject preference in the processing of relative clauses in Chinese. In: *Proceedings of the 25th West Coast Conference on Formal Linguistics (WC-CFL 25)*, ed. by Donald Baumer, David Montero & Michael Scanlon, Somerville, MA: Cascadilla Press, pp. 254–260.
- Lohnstein, Horst & Ursula Bredel (2004). Inflectional morphology and sentence mood in German. In: *The Syntax and Semantics of the Left Periphery*, ed. by Horst Lohnstein & Susanne Trissler, Berlin: de Gruyter, pp. 235–264.

- Lühr, Rosemarie (1988). Zur Satzverschränkung im heutigen Deutsch. *Groninger Arbeiten zur Germanistischen Linguistik* 29: 74–87.
- Mahajan, Anoop (1989). Agreement and agreement phrases. In: *Functional Heads and Clause Structure*, ed. by Itziar Laka & Anoop Mahajan, Cambridge, MA: MIT, MIT Working Papers in Linguistics 10, pp. 217–252.
- Mahajan, Anoop (1990). The A/A-bar distinction and movement theory. Ph.D. dissertation, MIT, Cambridge, MA.
- Mahajan, Anoop (1994). Towards a unified theory of scrambling. In: *Studies on Scrambling: Movement and Non-Movement Approaches to Free Word-Order Phenomena*, ed. by Norbert Corver & Henk van Riemsdijk, Berlin: de Gruyter, pp. 301–330.
- Mahajan, Anoop (1997). Rightward scrambling. In: *Rightward Movement*, ed. by Dorothee Beerman, David LeBlanc & Henk van Riemsdijk, Amsterdam: John Benjamins, pp. 185–213.
- Mahajan, Anoop (2000). Towards a unified treatment of wh-expletives in Hindi and German. In: *Wh-Scope Marking*, ed. by Uli Lutz, Gereon Müller & Arnim von Stechow, Amsterdam: John Benjamins, pp. 317–332.
- Manetta, Emily (2010). Wh-expletives in Hindi-Urdu: The vP phase. *Linguistic Inquiry* 41: 1–34.
- Manetta, Emily (2011). *Peripheries in Kashmiri and Hindi-Urdu: The Syntax of Discourse-Driven Movement*. Amsterdam: John Benjamins.
- Manetta, Emily (2012). Reconsidering rightward scrambling: Postverbal constituents in Hindi-Urdu. *Linguistic Inquiry* 43: 43–74.
- Manetta, Emily (2016). Parasitic gaps redux: PGs in Hindi-Urdu and Kashmiri, Ms., University of Vermont.
- Manzini, Rita (1994). Locality, Minimalism, and parasitic gaps. *Linguistic Inquiry* 25: 481–508.
- Marantz, Alec (1984). *On the Nature of Grammatical Relations*. Cambridge, MA: MIT Press.
- Marinis, Theodore, Leah Roberts, Claudia Felser & Harald Clahsen (2005). Gaps in second language sentence processing. *Studies in Second Language Acquisition* 27: 53–78.
- May, Robert (1979). Must Comp-to-Comp movement be stipulated? *Linguistic Inquiry* 10: 719–725.
- McCloskey, James (1979). *Transformational Syntax and Model Theoretic Semantics: A Case Study in Modern Irish*. Dordrecht: Reidel.
- McCloskey, James (2000). Quantifier float and wh-movement in an Irish English. *Linguistic Inquiry* 31: 57–84.
- McCloskey, James (2002). Resumptives, successive cyclicity, and the locality of operations. In: *Derivation and Explanation in the Minimalist Program*, ed. by Samuel David Epstein & T. Daniel Seely, Oxford: Blackwell, pp. 184–226.

- McDaniel, Dana (1989). Partial and multiple wh-movement. *Natural Language and Linguistic Theory* 7: 565–604.
- McElree, Brian, Stephanie Foraker & Lisbeth Dyer (2003). Memory structures that subserve sentence comprehension. *Journal of Memory and Language* 48: 67–91.
- McElree, Brian & Teresa Griffith (1998). Structural and lexical constraints on filling gaps during sentence comprehension: A time-course analysis. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 24: 432–460.
- McFadden, Thomas (2004). The position of morphological case in the derivation: A study on the syntax-morphology interface. Ph.D. dissertation, University of Pennsylvania.
- McKinnon, Richard & Lee Osterhout (1996). Constraints on movement phenomena in sentence processing: Evidence from event-related brain potentials. *Language and Cognitive Processes* 11: 495–523.
- Meinunger, André (2004). Verb position, verbal moodle and the anchoring (potential) of sentences. In: *The Syntax and Semantics of the Left Periphery*, ed. by Horst Lohnstein & Susanne Trissler, Berlin: de Gruyter, pp. 313–341.
- Miyagawa, Shigeru (2010). *Why Agree? Why Move? Unifying Agreement-Based and Discourse-Configurational Languages*. Cambridge, MA: MIT Press.
- Mohanan, Tara (1994). *Argument Structure in Hindi*. Stanford: CSLI Publications.
- Moore, John (1989). Spanish restructuring and psych verbs: A case for VP-complementation. In: *Proceedings of the Eighth West Coast Conference on Formal Linguistics (WCCFL 8)*, ed. by E. Jane Fee & Katherine Hunt, Stanford: CSLI, pp. 262–275.
- Moore, John (1996). *Reduced Constructions in Spanish*. New York, NY: Garland.
- Moro, Andrea (2000). *Dynamic Antisymmetry*. Cambridge, MA: MIT Press.
- Müller, Gereon (1995). *A-bar Syntax: A Study in Movement Types*. Berlin: de Gruyter.
- Müller, Gereon (1996). A constraint on remnant movement. *Natural Language and Linguistic Theory* 14: 355–407.
- Müller, Gereon (1998). *Incomplete Category Fronting: A Derivational Approach to Remnant Movement in German*. Dordrecht: Kluwer.
- Müller, Gereon (1999a). Imperfect checking. *The Linguistic Review* 16: 359–404.
- Müller, Gereon (1999b). Optimality, markedness, and word order in German. *Linguistics* 37: 777–818.
- Müller, Gereon (2010a). Movement from verb-second clauses revisited. In: *Language and Logos: A Festschrift for Peter Stauchacher*, ed. by Thomas Hanneforth & Gisbert Fanselow, Berlin: Akademieverlag, pp. 97–128.
- Müller, Gereon (2010b). On deriving CED effects from the PIC. *Linguistic Inquiry* 41: 35–82.

- Müller, Gereon (2011). *Constraints on Displacement: A Phase-Based Approach*. Amsterdam: John Benjamins.
- Müller, Gereon (2014a). A local approach to the Williams Cycle. *Lingua* 140: 117–136.
- Müller, Gereon (2014b). *Syntactic Buffers: A Local-Derivational Approach to Improper Movement, Remnant Movement, and Resumptive Movement*, Linguistische Arbeitsberichte 91. Universität Leipzig: Institut für Linguistik.
- Müller, Gereon & Wolfgang Sternefeld (1993). Improper movement and unambiguous binding. *Linguistic Inquiry* 24: 461–507.
- Müller, Gereon & Wolfgang Sternefeld (1994). Scrambling as A-bar movement. In: *Studies on Scrambling*, ed. by Norbert Corver & Henk van Riemsdijk, Mouton de Gruyter, pp. 331–385.
- Müller, Gereon & Wolfgang Sternefeld (1996). A-bar chain formation and economy of derivation. *Linguistic Inquiry* 27: 480–511.
- Neeleman, Ad & Hans van de Koot (2010). A local encoding of syntactic dependencies and its consequences for the theory of movement. *Syntax* 13: 331–372.
- Nevins, Andrew (2005). Derivations without the activity condition. In: *Perspectives on Phases*, ed. by Martha McGinnis & Norvin Richards, Cambridge, MA: MITWPL, no. 49 in MIT Working Papers in Linguistics, pp. 283–306.
- Nicenboim, Bruno, Shravan Vasishth, Carolina Gattei, Mariano Sigman & Reihnhold Kliegl (2015). Working memory differences in long-distance dependency resolution. *Frontiers in Psychology* 6: 312.
- Nicol, Janet, Janet Fodor & David Swinney (1994). Using cross-modal lexical decision tasks to investigate sentence processing. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 20: 1229–1238.
- Nicol, Janet & David Swinney (1989). The role of structure in coreference assignment during sentence comprehension. *Journal of Psycholinguistic Research* 18: 5–19.
- Nunes, Jairo (2008). Inherent case as a licensing condition for A-movement: The case of hyper-raising constructions in Brazilian Portuguese. *Journal of Portuguese Linguistics* 7: 83–108.
- Nunes, Jairo (2010). Relativizing minimality for A-movement: ϕ - and θ -relations. *Probus* 22: 1–25.
- Obata, Miki (2010). Root, successive-cyclic and feature-splitting Internal Merge: Implications for Feature-Inheritance and Transfer. Ph.D. dissertation, University of Michigan, Ann Arbor.
- Obata, Miki & Samuel David Epstein (2011). Feature-splitting Internal Merge: Improper movement, intervention, and the A/A' distinction. *Syntax* 14: 122–147.
- Ott, Dennis (2012). *Local Instability: Split Topicalization and Quantifier Float in German*. Berlin: de Gruyter.

- Ott, Dennis (2015). Symmetric Merge and local instability: Evidence from split topics. *Syntax* 18: 157–200.
- Pandharipande, Rajeshwari & Yamuna Kachru (1977). Relational grammar, ergativity, and Hindi-Urdu. *Lingua* 41: 217–238.
- Perlmutter, David & Scott Soames (1979). *Syntactic Argumentation and the Structure of English*. Berkeley: The University of California Press.
- Pesetsky, David (2000). *Phrasal Movement and Its Kin*. Cambridge, MA: MIT Press.
- Pesetsky, David & Esther Torrego (2001). T-to-C movement: Causes and consequences. In: *Ken Hale: A Life in Language*, ed. by Michael Kenstowicz, Cambridge, MA: MIT Press, pp. 355–426.
- Phillips, Colin (2006). The real-time status of island phenomena. *Language* 82: 795–823.
- Phillips, Colin, Nina Kazanina & Shani Abada (2005). ERP effects of the processing and syntactic long-distance dependencies. *Cognitive Brain Research* 22: 407–428.
- Phillips, Colin & Dan Parker (2014). The psycholinguistics of ellipsis. *Lingua* 151: 78–95.
- Phillips, Colin & Matthew Wagers (2007). Relating structure and time in linguistics and psycholinguistics. In: *The Oxford Handbook of Psycholinguistics*, ed. by Gareth Gaskell, Oxford: Oxford University Press, pp. 739–756.
- Pickering, Martin & Matthew Traxler (2003). Evidence against the use of subcategorization frequencies in the processing of unbounded dependencies. *Language and Cognitive Processes* 18: 469–503.
- Plank, Frans (1983). Transparent versus functional encoding of grammatical relations: A parameter for syntactic change and typology. *Linguistische Berichte* 86: 1–13.
- Polinsky, Maria (2003). Non-canonical agreement is canonical. *Transactions of the Philological Society* 101: 279–312.
- Polinsky, Maria & Eric Potsdam (2001). Long-distance agreement and topic in Tsez. *Natural Language and Linguistic Theory* 19: 583–646.
- Pollard, Carl J. & Ivan A. Sag (1994). *Head-Driven Phrase Structure Grammar*. Chicago: University of Chicago Press.
- Poole, Ethan (2015). Deconstructing quirky subjects. In: *Proceedings of the 45th Meeting of the North East Linguistic Society (NELS 45)*, ed. by Thuy Buy & Deniz Özyıldız, Amherst, MA: GLSA, pp. 247–256.
- Postal, Paul (1971). *Cross-Over Phenomena*. New York: Holt, Rinehart and Winston.
- Postal, Paul (1972). Some rules that are not successive cyclic. *Linguistic Inquiry* : 211–222.
- Preminger, Omer (2009). Breaking agreements: Distinguishing agreement and clitic doubling by their failures. *Linguistic Inquiry* 40: 619–666.

- Preminger, Omer (2011). Agreement as a fallible operation. Ph.D. dissertation, MIT, Cambridge, MA.
- Preminger, Omer (2014). *Agreement and Its Failures*. Cambridge, MA: MIT Press.
- Pullum, Geoffrey (1979). *Rule Interaction and the Organization of a Grammar*. New York: Garland.
- Pullum, Geoffrey (1982). Syncategorematicity and English infinitival *to*. *Glossa* 16: 181–215.
- Putnam, Michael (2007). *Scrambling and the Survive Principle*. Amsterdam: John Benjamins.
- R Core Team (2014). R: A language and environment for statistical computing.
- Rackowski, Andrea & Norvin Richards (2005). Phase edge and extraction. *Linguistic Inquiry* 36: 565–599.
- Ramchand, Gillian & Peter Svenonius (2014). Deriving the functional hierarchy. *Language Sciences* 46: 152–174.
- Reinhart, Tanya (1998). *Wh*-in-situ in the framework of the minimalist program. *Natural Language Semantics* 6: 29–56.
- Reis, Marga (1985). Satzeinleitende Strukturen im Deutschen. In: *Erklärende Syntax des Deutschen*, ed. by Werner Abraham, Tübingen: Narr, pp. 271–311.
- Reis, Marga (1996). Extractions from verb-second clauses in German? In: *On Extraction and Extraposition in German*, ed. by Uli Lutz & Jürgen Pafel, Amsterdam: Benjamins, pp. 45–88.
- Richards, Marc (2004). Object shift and scrambling in North and West Germanic: A case study in symmetrical syntax. Ph.D. dissertation, University of Cambridge, Cambridge, UK.
- Richards, Marc (2007). Dynamic linearization and the shape of phases. *Linguistic Analysis* 33: 209–237.
- Richards, Marc (2011). Deriving the edge: What’s in a phase? *Syntax* 14: 74–95.
- Rizzi, Luigi (1982). *Issues in Italian Syntax*. Dordrecht: Foris.
- Rizzi, Luigi (1990). *Relativized Minimality*. Cambridge, MA: MIT Press.
- Rizzi, Luigi (1997). The fine structure of the left periphery. In: *Elements of Grammar*, ed. by Liliane Haegeman, Dordrecht: Kluwer, pp. 281–337.
- Rizzi, Luigi (2001). On the position of Interrogative in the left periphery of the clause. In: *Current Studies in Italian Syntax: Essays Offered to Lorenzo Renzi*, ed. by Guglielmo Cinque & Giampaolo Salvi, Oxford: Elsevier North-Holland, pp. 287–296.
- Rizzi, Luigi (2004). Locality and left periphery. In: *Structures and Beyond: The Cartography of Syntactic Structures, Vol. 3*, ed. by Adriana Belletti, Oxford: Oxford University Press, pp. 223–251.
- Rizzi, Luigi (2006). On the form of chains: Criterial positions and ECP effects. In: *Wh-Movement: Moving On*, ed. by Lisa Cheng & Norbert Corver, Cambridge, MA: MIT Press, pp. 97–133.

- Rochemont, Michael (1989). Topic islands and the subjacency parameter. *Canadian Journal of Linguistics* 34: 145–170.
- Rochette, Anne (1988). Semantic and syntactic aspects of Romance sentential complementation. Ph.D. dissertation, MIT, Cambridge, MA.
- Romero, Maribel (1997). The correlation between scope reconstruction and connectivity effects. In: *Proceedings of the Sixteenth West Coast Conference on Formal Linguistics (WCCFL 16)*, ed. by Emily Curtis, James Lyle & Gabriel Webster, Stanford: CSLI Publications, pp. 351–366.
- Rooryck, Johan (1994). Against optional movement for clitic climbing. In: *Issues and Theory in Romance Linguistics: Selected Papers from the Linguistics Symposium on Romance Languages XXIII*, ed. by Mike Mazzola, Washington, DC: Georgetown University Press, pp. 417–443.
- Rooth, Mats (1985). Association with focus. Ph.D. dissertation, University of Massachusetts, Amherst, MA.
- Rooth, Mats (1992). A theory of focus interpretation. *Natural Language Semantics* 1: 75–116.
- Rosen, Sara (1990). Restructuring verbs are light verbs. In: *Proceedings of the Ninth West Coast Conference on Formal Linguistics (WCCFL 9)*, ed. by Aaron Halpern, Stanford: CSLI, pp. 477–491.
- Ross, John Robert (1967). Constraints on variables in syntax. Ph.D. dissertation, MIT, Cambridge, MA, Published in 1986 as *Infinite Syntax!*, Ablex, Norwood.
- Ross, John Robert (1970). On declarative sentences. In: *Readings in English Transformational Grammar*, ed. by Roderick Jacobs & Peter Rosenbaum, Waltham, MA: Ginn, pp. 222–272.
- Sabel, Joachim (1996). *Restrukturierung und Lokalität: Universelle Beschränkungen für Wortstellungsvarianten*. Berlin: Akademie-Verlag.
- Saddy, Douglas (1991). Wh-scope mechanisms in Bahasa Indonesia. In: *More Papers on Wh-Movement*, ed. by Lisa Cheng & Hamida Demirdache, Cambridge, MA: MITWPL, MITWPL 15, pp. 183–218.
- Sakai, Hiromu (1994). Derivational economy in long distance scrambling. In: *Formal Approaches to Japanese Linguistics*, ed. by Hiroyuki Ura & Masatoshi Koizumi, Cambridge, MA: MITWPL, MITWPL 24, pp. 295–314.
- Salzmann, Martin (2006). *Resumptive Prolepsis: A Study in Indirect \bar{A} -Dependencies*. Utrecht: LOT.
- Sauerland, Uli (1998). On the making and meaning of chains. Ph.D. dissertation, MIT, Cambridge, MA.
- Sauerland, Uli (1999). Erasability and interpretation. *Syntax* 3: 161–188.
- Sauerland, Uli (2003). Intermediate adjunction with A-movement. *Linguistic Inquiry* 34: 308–314.
- Schlesewsky, Matthias, Gisbert Fanselow, Reinhold Kliegl & Josef Krems (2000). The subject preference in the processing of locally ambiguous *wh*-questions in German. In: *German Sentence Processing*, ed. by Barbara Hemforth & Lars Konieczny, Dordrecht: Kluwer, pp. 65–93.

- Schmid, Tanja, Markus Bader & Josef Bayer (2005). Coherence: An experimental approach. In: *Linguistic Evidence: Empirical, Theoretical, and Computational Perspectives*, ed. by Stephan Kepser & Marga Reis, Berlin: de Gruyter, pp. 435–456.
- Schulte im Walde, Sabine (1998). Automatic semantic classification of verbs according to their alternation behavior, MA thesis, Universität Stuttgart.
- Schwarz, Florian (2007). *Ex-situ* focus in Kikuyu. In: *Focus Strategies in African Languages: The Interaction of Focus and Grammar in Niger-Congo and Afro-Asiatic*, ed. by Enoch Oladé Aboh, Katharina Hartmann & Malte Zimmermann, Berlin: de Gruyter, pp. 139–159.
- Sekerina, Irina (2003). Scrambling and processing: Dependencies, complexity, and constraints. In: *Word Order and Scrambling*, ed. by Simin Karimi, Malden, MA: Blackwell, pp. 301–324.
- Shlonsky, Ur (2006). Extended projections and CP cartography. *Nouveaux cahiers de linguistique française* 27: 87–93.
- Sigurðsson, Halldór Ármann (1996). Icelandic finite verb agreement. *Working Papers in Scandinavian Syntax* 57: 1–46.
- Sigurðsson, Halldór Ármann (2004). Icelandic non-nominative subjects: Facts and implications. In: *Non-nominative Subjects, Vol. 2*, ed. by Peri Bhaskararao & Karumuri Venkata Subbarao, Amsterdam: John Benjamins, pp. 137–159.
- Simpson, Andrew (2000). *Wh-Movement and the Theory of Feature-Checking*. Amsterdam: John Benjamins.
- Simpson, Andrew & Tanmoy Bhattacharya (2003). Obligatory overt *wh*-movement in a *wh*-in-situ language. *Linguistic Inquiry* 34: 127–142.
- Spencer, Andrew (2005). Case in Hindi. In: *Proceedings of the LFG05 Conference*, ed. by Miriam Butt & Tracy Holloway King, Stanford: CSLI, pp. 429–446.
- Sportiche, Dominique (1988). A theory of floating quantifiers and its corollaries for constituent structure. *Linguistic Inquiry* 19: 425–450.
- Sportiche, Dominique (2006). Reconstruction, binding, and scope. In: *The Blackwell Companion to Syntax, Vol. IV*, ed. by Martin Everaert & Henk van Riemsdijk, Oxford: Blackwell, pp. 35–93.
- Sprouse, Jon, Ivano Caponigro, Ciro Greco & Carlo Cecchetto (2016). Experimental syntax and the variation of island effects in English and Italian. *Natural Language and Linguistic Theory* 34: 307–344.
- Srivastav, Veneeta (1991a). Subjacency effects at LF: The case of Hindi *Wh*. *Linguistic Inquiry* 22: 762–769.
- Srivastav, Veneeta (1991b). *Wh* dependencies in Hindi and the theory of grammar. Ph.D. dissertation, Cornell University, Ithaca, NY.
- Starke, Michal (2001). Move dissolves into merge: A theory of locality. Ph.D. dissertation, University of Geneva.

- Staub, Adrian (2007). The parser doesn't ignore intransitivity, after all. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 33: 550–569.
- Staub, Adrian (2010). Eye movement and processing difficulty in object relative clauses. *Cognition* 116: 71–86.
- Staub, Adrian, Charles Clifton & Lyn Frazier (2006). Heavy NP shift is the parser's last resort: Evidence from eye movements. *Journal of Memory and Language* 54: 389–406.
- Staudacher, Peter (1990). Long movement from verb-second-complements in German. In: *Scrambling and Barriers*, ed. by Günther Grewendorf & Wolfgang Sternefeld, Amsterdam: Benjamins, pp. 319–339.
- Stepanov, Artur (2007). On the absence of long-distance A-movement in Russian. *Journal of Slavic Linguistics* 15: 81–108.
- Sternefeld, Wolfgang (1989). Extraction from verb-second clauses in German. *Working Papers in Scandinavian Syntax* 44: 119–140.
- Sternefeld, Wolfgang (1992). Transformationstypologie und strukturelle Hierarchie, Ms., Universität Tübingen.
- Stevenson, Suzanne (1994). Competition and recency in a hybrid network model of syntactic disambiguation. *Journal of Psycholinguistic Research* 23: 295–322.
- Stowe, Laurie (1986). Parsing wh-constructions: Evidence for on-line gap location. *Language and Cognitive Processes* 1: 227–245.
- Sturt, Patrick, Martin Pickering & Matthew Crocker (1999). Structural change and reanalysis difficulty in language comprehension. *Journal of Memory and Language* 40: 136–150.
- Sussman, Rachel & Julie Sedivy (2003). The time course of processing syntactic dependencies. *Language and Cognitive Processes* 18: 143–163.
- Szabolcsi, Anna (1997). Strategies for scope taking. In: *Ways of Scope Taking*, ed. by Anna Szabolcsi, Dordrecht: Kluwer, pp. 109–155.
- Szabolcsi, Anna (2006). Strong vs. weak islands. In: *The Blackwell Companion to Syntax, Vol. IV*, ed. by Martin Everaert & Henk van Riemsdijk, Oxford: Blackwell, pp. 479–531.
- Takahashi, Daiko (1990). Negative polarity, phrase structure, and the ECP. *English Linguistics* 7: 129–146.
- Takahashi, Daiko (1993). Movement of wh-phrases in Japanese. *Natural Language and Linguistic Theory* 11: 655–678.
- Takahashi, Masahiko (2011). Some consequences of Case-marking in Japanese. Ph.D. dissertation, University of Connecticut, Storrs.
- Takahashi, Masahiko (2012). On restructuring infinitives in Japanese: Adjunction, clausal architecture, and phases. *Lingua* 122: 1569–1595.

- Takahashi, Shoichi & Sarah Hulse (2009). Wholesale late merger: Beyond the A/A' distinction. *Linguistic Inquiry* 40: 387–426.
- Takano, Yuji (1994). Unbound traces and indeterminacy of derivation. In: *Current Topics in English and Japanese*, ed. by Masaru Nakamura, Tokyo: Hituzi Syobo, pp. 229–253.
- Tappe, Hans Thilo (1981). Wer glaubst du hat recht? In: *Sprache: Formen und Strukturen*, ed. by Manfred Kohrt & Jürgen Lenerz, Tübingen: Niemeyer, pp. 203–212.
- Tappe, Hans Thilo (1984). On infinitival clauses without COMP. In: *Sentential Complementation: Proceedings of the International Conference held at UFSAL, Brussels, June 1983*, ed. by Wim de Geest & Yvan Putseys, Dordrecht: Foris, pp. 227–237.
- Taraldsen, Knut Tarald (1995). On agreement and nominative objects in Icelandic. In: *Studies in Comparative Germanic Syntax*, ed. by Hubert Haider, Susan Olsen & Sten Vikner, Dordrecht: Kluwer, pp. 307–327.
- Thiersch, Craig (1978). Topics in German syntax. Ph.D. dissertation, MIT, Cambridge, MA.
- Thráinsson, Höskuldur (1984). Different types of infinitival complements in Icelandic. In: *Sentential Complementation: Proceedings of the International Conference held at UFSAL, Brussels, June 1983*, ed. by Wim de Geest & Yvan Putseys, Dordrecht: Foris, pp. 247–255.
- Tomioka, Satoshi (2007). Intervention effects in focus: From the Japanese point of view. In: *Interdisciplinary Studies on Information Structure, Vol. 9: Proceedings of the Second Workshop on Prosody, Syntax, and Information Structure*, ed. by Shinichiro Ishihara, Potsdam University Press, pp. 97–118.
- Torrego, Esther (1984). On inversion in Spanish and some of its effects. *Linguistic Inquiry* 15: 103–129.
- Travis, Lisa (1984). Parameters and effects of word order variation. Ph.D. dissertation, MIT, Cambridge, MA.
- Traxler, Matthew, Robin Morris & Rachel Seely (2002). Processing subject and object relative clauses: Evidence from eye movements. *Journal of Memory and Language* 47: 69–90.
- Traxler, Matthew & Martin Pickering (1996). Plausibility and the processing of unbounded dependencies: An eye-tracking study. *Journal of Memory and Language* 35: 454–475.
- Truckenbrodt, Hubert (2006). On the semantic motivation of syntactic verb movement to C in German. *Theoretical Linguistics* 32: 257–306.
- Trueswell, John, Michael Tanenhaus & Christopher Kello (1993). Verb-specific constraints in sentence processing: Separating effects of lexical preference from garden-paths. *Journal of Experimental Psychology: Learning, Memory, and Cognition* 19: 528–553.
- Ura, Hiroyuki (1994). *Varieties of Raising and Their Implications for the Theory of Case and Agreement*. No. 7 in MIT Occasional Papers in Linguistics, Cambridge, MA: MIT.

- Van Dyke, Julie & Richards Lewis (2003). Distinguishing effects of structure and decay on attachment and repair: A cue-based parsing account of recovery from misanalyzed ambiguities. *Journal of Memory and Language* 49: 285–316.
- van Riemsdijk, Henk (1988). The representation of syntactic categories. In: *Proceedings of the Conference on the Basque Language, Basque World Congress, Vol. 1*, pp. 104–116.
- van Riemsdijk, Henk (1998). Categorial feature magnetism: The endocentricity and distribution of projections. *Journal of Comparative Germanic Linguistics* 2: 1–48.
- van Riemsdijk, Henk & Edwin Williams (1981). NP structure. *The Linguistic Review* 1: 171–218.
- van Riemsdijk, Henk & Edwin Williams (1986). *Introduction to the Theory of Grammar*. Cambridge, MA: MIT Press.
- van Urk, Coppe (2015). A uniform syntax for phrasal movement: A case study of Dinka Bor. Ph.D. dissertation, MIT, Cambridge, MA.
- van Urk, Coppe & Norvin Richards (2015). Two components of long-distance extraction: Successive cyclicity in Dinka. *Linguistic Inquiry* 46: 113–155.
- Vasishth, Shrvan & Richard Lewis (2006). Argument–head distance and processing complexity: Explaining both locality and antilocality effects. *Language* 82: 767–794.
- von Stechow, Arnim (1990). Status government and coherence in German. In: *Scrambling and Barriers*, ed. by Günther Grewendorf & Wolfgang Sternefeld, Amsterdam: John Benjamins, pp. 143–198.
- von Stechow, Arnim (1992). Kompositionsprinzipien und grammatische Struktur. In: *Biologische und soziale Grundlagen der Sprache*, ed. by Peter Suchsland, Tübingen: Niemeyer, pp. 175–248.
- von Stechow, Arnim & Wolfgang Sternefeld (1988). *Bausteine syntaktischen Wissens: Ein Lehrbuch der generativen Grammatik*. Opladen: Westdeutscher Verlag.
- Wagers, Matthew & Colin Phillips (2009). Multiple dependencies and the role of the grammar in real-time comprehension. *Journal of Linguistics* 45: 395–433.
- Wagers, Matthew & Colin Phillips (2013). Going the distance: Memory and control processes in active dependency construction. *The Quarterly Journal of Experimental Psychology* : 1–31.
- Wanner, Eric & Michael Maratsos (1978). An ATN approach to comprehension. In: *Linguistic Theory and Psychological Reality*, ed. by Morris Halle, Joan Bresnan & George Miller, Cambridge, MA: MIT Press, pp. 119–161.
- Warren, Tessa & Edward Gibson (2002). The influence of referential processing on sentence complexity. *Cognition* 85: 79–112.
- Wasow, Thomas (1972). Anaphoric relations in English. Ph.D. dissertation, MIT, Cambridge, MA.
- Webelhuth, Gert (1989). Syntactic saturation phenomena and the modern Germanic language. Ph.D. dissertation, University of Massachusetts, Amherst.

- Webelhuth, Gert (1992). *Principles and Parameters of Syntactic Saturation*. Oxford: Oxford University Press.
- Wechsler, Stephen (1991). Verb second and illocutionary force. In: *Views on Phrase Structure*, ed. by Katherine Leffel & Denis Bouchard, Dordrecht: Kluwer, pp. 177–191.
- Wexler, Ken & Peter Culicover (1980). *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press.
- Williams, Edwin (1974). Rule ordering in syntax. Ph.D. dissertation, MIT, Cambridge, MA.
- Williams, Edwin (1978). Across the board rule application. *Linguistic Inquiry* 9: 31–43.
- Williams, Edwin (2003). *Representation Theory*. Cambridge, MA: MIT Press.
- Williams, Edwin (2009). There is no alternative to cartography. In: *Alternatives to Cartography*, ed. by Jeroen van Craenenbroeck, Amsterdam: Mouton de Gruyter, pp. 361–373.
- Williams, Edwin (2011). *Regimes of Derivation in Syntax and Morphology*. New York: Routledge.
- Williams, Edwin (2013). Generative semantics, generative morphosyntax. *Syntax* 16: 77–108.
- Wiltschko, Martina (1997). Scrambling, D-linking and superiority in German. *Groninger Arbeiten zur Germanistischen Linguistik* 41: 107–142.
- Wiltschko, Martina (1998). Superiority in German. In: *Proceedings of the 16th West Coast Conference on Formal Linguistics (WCCFL 16)*, ed. by Emily Curtis, James Lyle & Gabriel Webster, Stanford: CSLI, pp. 431–446.
- Wold, Dag (1996). Long distance selective binding. In: *Proceedings of Semantics and Linguistic Theory (SALT) VI*, ed. by Teresa Galloway & Justin Spence, Ithaca, NY: Cornell University, pp. 311–328.
- Wöllstein-Leisten, Angelika (2001). *Die Syntax der Dritten Konstruktion: Eine repräsentationelle Analyse zur Monosententialität von 'zu'-Infinitives im Deutschen*. Tübingen: Stauffenberg.
- Woolford, Ellen (2001). Case patterns. In: *Optimality-Theoretic Syntax*, ed. by Géraldine Legendre, Jane Grimshaw & Sten Vikner, Cambridge, MA: MIT Press, pp. 509–543.
- Wurmbrand, Susanne (2001). *Infinitives: Restructuring and Clause Structure*. Berlin: Mouton de Gruyter.
- Wurmbrand, Susanne (2007). How complex are complex predicates? *Syntax* 10: 243–288.
- Wurmbrand, Susi (2014). Tense and aspect in English infinitives. *Linguistic Inquiry* 45: 403–447.
- Wurmbrand, Susi (2015). Restructuring cross-linguistically. In: *Proceedings of the 45th North East Linguistic Society (NELS 45)*, ed. by Thuy Bui & Deniz Özyıldız, Amherst, MA: GLSA, pp. 227–240.