Göteborgsstudier i nordisk språkvetenskap 30

Extraction from relative clauses in Swedish

(Swedish summary)

Filippa Lindahl

GÖTEBORGS UNIVERSITET

TITLE: Extraction from relative clauses in Swedish
SWEDISH TITLE: Satsfläta med relativsats i svenskan
LANGUAGE: English (Swedish summary)
AUTHOR: Filippa Lindahl

Abstract

This dissertation presents an empirical study of extraction from relative clauses (ERC) in Swedish, where a phrase outside a relative clause (RC) is related to a gap or a resumptive pronoun inside the RC. The aim of the study is to provide an analysis of Swedish ERC-sentences based on spontaneously produced examples, questionnaire data, and elicitation and to clarify the interplay between information structure, discourse factors, and semantics on the one hand, and syntax on the other, in constraining ERC. The investigation draws on a collection of 270 naturally occurring ERC-sentences from spoken and written Swedish.

The study shows that the syntactic dependency between the extracted phrase and the gap position inside the RC is an A´-movement dependency, and that the RC in many ERC-sentences is a regular restrictive relative clause. From a discourse perspective, preposing in ERC is like preposing in the local clause and from *att*-clauses (*that*-clauses) in that it has the same discourse functions. With respect to the information structural role in the clause, the preposed phrase is often an aboutness topic. It can also be the information focus of the sentence, but not a scene-setter.

An in-depth study of extraction of both adjuncts and arguments from RCs shows that A´-movement from RCs is more restricted than A´-movement in the local clause and from *att*-clauses. Evidence for this is that the semantic type of the extracted phrase affects extractability. Furthermore, *wh*-questions from RCs exhibit a pattern familiar from previous research on long extraction from embedded questions, suggesting that Swedish relative clauses constitute some type of weak island.

The results and their implications for theories of islands are discussed in relation to recent proposals within the Minimalist program.

KEYWORDS: A-bar-movement, extraction, information structure, island constraints, preposing phenomena, relative clauses, Scandinavian, semantics, syntactic dependencies, Swedish, weak islands.

DISTRIBUTION: Institutionen för svenska språket
Box 200
405 30 Göteborg

FOTO PORTRÄTT: Rudolf Rydstedt
ISSN: 1652-3105
ISBN: 978-91-87850-65-3
LÄNK TILL E-PUBLICERING: http://hdl.handle.net/2077/51985
SÄTTNING: Filippa Lindahl och Sven Lindström
TRYCKNING: Ineko AB, 2017

Acknowledgements

This dissertation was written as part of my graduate studies at the Department of Swedish at the University of Gothenburg. Several people have contributed greatly to the work presented here. First and foremost, I want to thank my advisors, Elisabet Engdahl and Maia Andréasson. Elisabet has been my main advisor. Discussing all sorts of problems and examples in syntax and semantics with her during the past few years has been a great privilege, and most of all a lot of fun. Writing this thesis would not have been possible without Elisabet's knowledge and encouragement. Maia has been my secondary advisor, and has valiantly read my longer and longer drafts and provided insightful comments, pushed me to explain my ideas better, and helped me with a number of curious computer- and typography-related problems.

Thanks to a grant from the American-Scandinavian Foundation, I was able to spend the academic year of 2013–2014 at the Linguistics Research Center at University of California at Santa Cruz, which has had a big influence on this thesis. I have benefited greatly both from talking to faculty and graduate students at the Linguistics department at UCSC, and from taking courses. Sandy Chung, Amy Rose Deal, Donka Farkas, Jorge Hankamer, and Jim McCloskey all deserve special thanks for welcoming me to the department and talking to me about my project, as do Karen Duek, Nick Kalivoda, Kelsey Kraus, Deniz Rudin, Erik Zyman, and my adopted cohort. I'd like to especially thank Jim for taking an interest in my project in 2016, when I came back to Santa Cruz for two shorter visits in the spring and fall.

Nick Kalivoda, Jim McCloskey, Johanna Prytz, and Rickard Ramhöj have read and commented on drafts of chapters at various stages, and Johan Brandtler and Henrik Rosenkvist read drafts of articles from which several ideas later became parts of this dissertation. Fredrik Heinat read an earlier manuscript

as opponent for my final seminar, and Nick Kalivoda proofread and checked my English in the final version of the text. Matilda Lindahl read the Swedish summary. Morgan Andréasson and Maia Andréasson proved to be true saviors when it came to fonts in the typesetting process. Special thanks are also due to my editor Inga-Lill Grahn.

I have received valuable feedback from audiences and seminar participants at the UC Santa Cruz Syntax and Semantics Circle, June 2014 and March 2016; at the Grammar Seminar at Lund University, November 2014; at the annual meeting of the Linguistic Society of America in Portland, January 2015; at Grammar in Focus in Lund, February 2015; at Grammatik i Norden in Gothenburg, May 2016; at the UC Berkeley Syntax and Semantics Circle, September 2016; at the Stanford Syntax and Morphology Circle, September 2016; at the Higher Seminar in Scandinavian languages at Stockholm University, December 2016; and at the Grammar Seminar at the University of Gothenburg. The Adlerbert Scholarship Foundation and Stiftelsen Paul och Marie Berghaus donationsfond have made much of my traveling possible.

Numerous people have contributed data for my investigation. Friends, family, and colleagues around dinner and fika tables in Sandviken, Gothenburg, and other parts of Sweden have all had to endure my questionable social skills. I am grateful to all of them for being patient when I ran off to write down what they said instead of keeping up the conversation. Anne Mette Nyvad has helped me with data from Danish.

I have had two office-mates during my time in Gothenburg, Kristian Blensenius and Malin Sandberg, who deserve special thanks for making every workday enjoyable. Thanks also to Marie Rydenvald for sticking it out at Lennart Torstenssonsgatan during long hours in the endless summer of 2016. The Department of Swedish as a whole has provided me with a very stimulating work atmosphere. Tack allihop!

Finally, I want to thank the Lindahl family for moral support and encouragement.

Gothenburg, March 2017
Filippa Lindahl

Contents

Tables

Abbreviations

Glossing abbreviations

SG	singular
PL	plural
MASC	masculine
FEM	feminine
NEUT	neuter
NOM	nominative
ACC	accusative
DEF	definite
PN	pronoun
POSS	possessive
PRT	particle
REFLX	reflexive
SOM	complementizer in embedded question
känna.COG	'know of'
känna.REL	'know' 'be acquainted with'

Other abbreviations

BCI	Backgrounded Constructions are Islands
CED	Condition on Extraction Domain
CNPC	Complex Noun Phrase Constraint
ERC	extraction from relative clause(s)
PIC	Phase Impenetrability Condition
RC	relative clause

Transcription symbols

(.)	short pause
?	the utterance functions as a question
skull-	interrupted word

1. Introduction

In this dissertation, I study a syntactic phenomenon found in the mainland Scandinavian languages: extraction from relative clauses (ERC). Some examples from Swedish are given in (1)–(3).[1]

(1) men ingen av dom är ju varmblodiga (.) det_1 finns det inga insekter [$_{\text{RC}}$
but none of them are PRT *warm-blooded that is there no insects*
som är e_1]
that are
'But none of them are warm-blooded, there are no insects that are.'

(Conversation, Aug. 2015)

(2) ja [ett lodjur]$_1$ har jag inte hört talas om nån [$_{\text{RC}}$ som blivit
yes a lynx have I not heard spoken about someone that has been
uppäten av e_1]
eaten by
'Yes, I've never heard about anyone who was eaten by a lynx.'

(Conversation, Summer 2013)

(3) där$_1$ hade du ju en svartvit flugsnappare [$_{\text{RC}}$ som brukade bo e_1]
there had you PRT *a pied flycatcher that used to live*
'Oh yeah, you had a pied flycatcher that used to live there!'

(Conversation, Aug. 2016)

In each of the examples in (1)–(3) there is a syntactic relation between a phrase outside of a relative clause (RC) and a position inside it, here marked by an *e*.

[1] The (.)-notation signifies a micro pause.

I will refer to phrases like *det* 'that', *ett lodjur* 'a lynx', and *där* 'there' in such examples as **extracted** or **preposed**, and to the position inside the relative clause as a **gap**. The fact that the preposed phrases are related to the gaps is marked by subscript indices. What kind of dependency mediates the relation, and how this kind of dependency is restricted in Swedish, are the overarching questions of the dissertation, together with questions about syntactic and pragmatic properties of the relative clauses that appear in ERC.

Extraction phenomena have been a central topic for linguistic theory since the 1960s, when Chomsky (1964) and Ross (1967) began investigating them. Ross (1967) identified several environments where syntactic dependencies are restricted: syntactic islands. Among them are relative clauses in English, as illustrated by the examples in (4a) and (4b).

(4) a. *The man who I read a statement which was about is sick. (Ross 1967:119)

b. *Who does Phineas know a girl who is working with? (Ross 1967:124)

Relativizing or questioning a position inside a relative clause, as has been done in these examples, is generally not possible in English. This is unexpected given that both question formation and relativization can form dependencies over an unbounded domain. A central question, then, is why examples like (4a) and (4b) are not possible sentences. There are several competing answers to this question, involving syntactic, semantic, or pragmatic constraints, or some combination thereof. Often, these constraints are hypothesized to be universal and innate, since they are presumably not learnable from the input that children receive.

The mainland Scandinavian languages are relevant for theories of extraction in that they appear to exhibit different constraints than other languages in which syntactic islands have been investigated. Erteschik-Shir (1973) shows that in Danish, unlike in English, relative clauses are not always syntactic islands, and subsequent research has shown that Norwegian and Swedish are more like Danish than English in terms of their ability to form dependencies into relative clauses (e.g. Allwood 1976, Andersson 1982, Engdahl 1982, Christensen 1982). The mainland Scandinavian languages can thus give us important clues about the nature of restrictions on extraction, and their place in our mental grammars.

First, any universal theory of extraction needs to be able to account for the fact that sentences such as the ones in (1)–(3) are possible sentences of a language. The Scandinavian languages show a kind of variation that our theories have to permit. A central question from this perspective is how the Scandinavian languages are different from other languages.

Second, since ERC is possible in the mainland Scandinavian languages, we

can answer questions about what kinds of extractions can occur and which cannot. Interestingly, it is not the case that just any phrase can be extracted from a relative clause, even though the Scandinavian languages are more permissive in terms of the domains that allow extraction than languages like English. For example it is not possible to interpret the Swedish example in (5) as a question about the reasons for writing books, which we would expect if the *wh*-expression *varför* 'why' could freely associate with a position inside the relative clause.

(5) Varför känner du många som har skrivit böcker?
why know you many that have written books
'Why do you know many people who have written books?'

a. För att jag är med i Författarförbundets styrelse.
*because I am with in Writer union.*DEF*'s board*
'Because I'm on the board of The Swedish Writer's Union.'

b. #För att chockera sin samtid.
*in order to shock their.*REFLX.SG *contemporaries*
'In order to shock their contemporaries.'

Finding out what constrains extraction from relative clauses in Swedish provides one piece of the puzzle about the nature of extraction phenomena in general.

The structure of relative clauses that can participate in ERC is a point of contention. While Erteschik-Shir sees no reason to assume that there is a structural difference between relative clauses that permit extraction and those that do not (Erteschik-Shir 1973:34), there are proposals that argue precisely this. Platzack (1999, 2014) suggests that the reason we find examples like (1)–(3) in the mainland Scandinavian languages is that they have a special way of deriving subject relative clauses dependent on having both relative complementizers and being a verb second language. Kush (2011) and Kush et al. (2013) propose that the relative clause-like constituents in ERC in the Scandinavian languages are not real relative clauses, but a type of small clause. The latter proposal is cast into doubt by Christensen & Nyvad (2014) and Müller (2015).

Theoretical accounts of ERC in the Scandinavian languages invoke pragmatic or semantic properties of the phenomenon in one way or another (e.g. Andersson 1982, Engdahl 1997, Erteschik-Shir 1973). Andersson (1982) argues that it is important that the resulting sentence "make sense" in the context and that the functions of the involved constructions matter. Engdahl (1997) reports that most cases of extraction from relative clauses involve topicalization,

and proposes that what sets the Scandinavian languages apart might be their way of organizing coherent discourses. However, Engdahl stresses that more research is needed. Erteschik-Shir (1973) argues that a semantic condition rather than a syntactic one is at play in extraction in Danish, but it is unclear whether her account extends to all cases of ERC that are possible in Swedish (see Löwenadler 2015). Furthermore, the proposals differ in the role they give pragmatic or semantic constraints in relation to the syntax. Thus, even though it is clear that ERC is sensitive to context, it is still not obvious what the role of pragmatic or semantic constraints is.

Previous research on ERC has mainly been based on constructed examples without context (but see Engdahl 1997, Lindahl 2010). There are a few reasons that solely relying on judgments of examples without context is not ideal. In particular, it makes it harder to study the role of information structure in constraining extraction. Certain generalizations can be made on the basis of constructed question and answer pairs, but data from spontaneously produced examples is needed to evaluate hypotheses and generalizations that previous research has generated.

There is also a problem relating to acceptability judgements and the context-sensitivity of extractions. Allwood (1976) points out that extraction sentences which sound strange or degraded will often sound better when the right context is imagined. This is probably true of a large class of sentences which are not generally felicitous in an out-of-the-blue context. If we present such a sentence without the context, the acceptability judgement, in addition to saying something about the acceptability of the extraction itself, will reflect the ability of the test participant to come up with a suitable context (Engdahl 1997). When we know more about the contexts in which ERC is acceptable, and what characterizes spontaneously produced ERC-sentences, we can avoid these potential pitfalls.

1.1 Aim

The general aim of the dissertation is to advance our understanding of ERC, and to provide an analysis of Swedish ERC based both on spontaneously occurring examples and acceptability judgements of naturalistic constructed example sentences. More specifically, the aim is to clarify the interplay between information structure, discourse factors, and semantics on the one hand, and syntax on the other, in Swedish ERC. From a broader theoretical perspective, the aim of the dissertation is to contribute to our knowledge about extraction phenomena in natural languages.

I address the following questions:

1. What is the nature of the dependency between the extracted phrase and the empty position inside the relative clause in ERC-sentences?
2. What is the structure of the relative clauses in ERC-sentences, and how do the relative clauses relate to the rest of the sentence?
3. What is the role of discourse and pragmatic constraints in ERC?
4. How is ERC constrained?

The questions will be made more specific in section 2.3, following an introduction of previous research and theoretical concepts.

1.2 Outline

The dissertation consists of seven chapters, including this introduction. Chapter 2 is a background to the study and gives a general introduction to the central theoretical concepts of Ā-dependencies and locality, and a brief overview of research on constraints on extraction. The chapter also provides an overview of the previous research on ERC in the mainland Scandinavian languages.

In chapter 3, I describe the data used, and how they were collected. The main material is a collection of 270 examples from spoken and written Swedish collected between 2011 and 2016. In addition to the collection of spontaneously produced examples, I present a questionnaire study and elicitation interviews which were used to collect additional data.

Chapters 4 to 6 each investigate different aspects of ERC. Broadly speaking, chapter 4 is centered on syntactic issues, and chapter 5 on pragmatics and discourse factors. Chapter 6 relates Swedish ERC to semantic and pragmatic accounts of weak islands.

To be more specific, chapter 4 introduces the view of Swedish clause structure assumed in the dissertation, and investigates various aspects of the syntax of ERC-sentences. I first adopt an analysis of the left periphery of Swedish clauses introduced by Vikner (2017) and Nyvad et al. (forthcoming), and relate it to the preposing phenomena that are central to this study. The remainder of the chapter is devoted to the syntactic dependencies involved in ERC, and to the structure and attachment of the relative clauses that we see in ERC-sentences of various types. Arguments that the dependencies in ERC are created by Ā-movement are presented, as well as arguments that there is extraction from

regular restrictive relative clauses, and not simply from small clauses or other types of relative clause-like constituents.

Chapter 5 investigates ERC in discourse. Based on my collectio of spontaneously produced extraction sentences, I give an overview of the types of relations that hold between the fronted phrase and the context, and the possible information-structural functions that the fronted phrase can have in the utterance. I show that while aboutness topics and information foci can be fronted, the extracted phrase in ERC is never a scene-setter, a function which is otherwise common for clause initial non-subjects in declarative main clauses. Chapter 5 also contains an analysis of the information status of the relative complex in ERC-sentences.

Chapter 6 centers on the question of extraction of adjuncts, and situates Swedish ERC with respect to research on weak islands. I argue on the basis of corpus and questionnaire data that while adjunct extraction is rare, it is possible in certain circumstances. Specifically, it is possible to extract adjuncts that are contrastive or deictic (denoting a specific point in time, for instance), or that can be construed as being D-linked (Pesetsky 1987). In this chapter I also present data about *wh*-question formation from a small interview study. The study shows that *wh*-questions in ERC are acceptable if they can be interpreted as being D-linked, but that *varför*-questions (*why*-questions) out of relative clauses are not acceptable at all. This is a pattern that is familiar from previous research on weak islands. However, Swedish relative clauses are even more transparent than more prototypical weak islands, in that they do not block functional readings of questions (Engdahl 1986).

In chapter 7, I draw together the findings from the previous chapters, and summarize what these findings entail for the analysis of ERC in Swedish and for theories of islands more generally.

2. Ā-dependencies, locality, and extraction in mainland Scandinavian

As a background to my investigation, this chapter introduces some basic concepts in research on extraction, and gives an overview of previous research on extraction in the mainland Scandinavian languages, with a focus on ERC.

The chapter begins with an introduction to the notions of Ā-dependencies and locality, constraints on extraction, and syntactic islands in section 2.1, where I provide a very brief overview of some of the theoretical developments in this domain of research in the last few decades. Section 2.1.1 is an introduction to Ā-dependencies in Phase Theory, section 2.1.2 an overview of properties of Ā-movement dependencies in Swedish, and section 2.1.3 presents a distinction between Ā-movement and Ā-binding due to Cinque (1990). In section 2.1.4, I present some lines of research which argue that constraints on extraction are functional, pragmatic, or semantic and not syntactic, and in section 2.1.5, processing factors are very briefly introduced. A division between strong and weak islands is often assumed in research on extraction, and I introduce this division briefly in section 2.1.6.

In section 2.2, I turn to research on extraction from subordinate clauses in the mainland Scandinavian languages, examining in particular what has been said about ERC. I begin by introducing early accounts and comments, which were often given in style guides or discussions of proper language use, and survey more recent research in section 2.2.1. Accounts that emphasize the role of pragmatics, semantics and information structure are presented in section 2.2.1.1, and mainly structural approaches in section 2.2.1.2. A few recent experimental studies are introduced in section 2.2.1.3.

The last section of the chapter, section 2.3, revisits the four overarching questions of the dissertation introduced in chapter 1 and makes them more specific given the theoretical background and the previous research surveyed.

2.1 Ā-dependencies and locality

In English, relative clauses, constituent questions, clefts and topicalization all involve a link between a position outside, or on the left edge of a clause, and an empty position inside it. Consider the examples in (1).

(1) a. The book which$_1$ I got e_1 from the library last month is overdue.

b. I wonder what$_1$ they will do e_1.

c. It was coffee$_1$ she wanted e_1.

d. Coffee$_1$, I really like e_1.

The dependency between the phrases marked with subscript indices and the empty positions marked with *e* in these examples is often called a ***wh*-dependency**, or in more recent work an **Ā-dependency**.[1] The interpretation of the empty position is in each case dependent on the phrase at the left edge of the clause. This type of dependency is characteristic of the clause types in (1) in many languages.

From a theoretical perspective a particularly interesting property of Ā-dependencies is that they appear to be able to hold over unbounded domains. Consider the examples in (2).

(2) a. The book which$_1$ [$_S$ Anna thought [$_S$ Julia said [$_S$ she'd returned e_1 to the library]]] is overdue.

b. I wonder what$_1$ [$_S$ she thinks [$_S$ they ought to say [$_S$ they will do e_1]]].

c. It was coffee$_1$ [$_{S'}$ that [$_S$ Olle thought [$_S$ she must have said [$_{S'}$ that [$_S$ she wanted e_1]]]]].

d. Coffee$_1$, [$_S$ I think [$_S$ she should have said [$_{S'}$ that [$_S$ he really likes e_1]]]].

There are three clause boundaries between the extracted phrases in these examples and the empty positions which they correspond to, but there is no reason

[1] See e.g. Chomsky (1973, 1977) and Bresnan (1977). The term **Ā-position** was introduced for non-argument positions, as opposed to **A-positions**, which are positions where a theta-role may be assigned (Chomsky 1981). Other names for the group of dependencies include **filler-gap dependencies** (see e.g. Fodor 1978) and **unbounded dependencies** (see e.g. Engdahl & Ejerhed 1982).

to think that there is an upper limit on the number of boundaries that can be crossed. In (3), five clause boundaries are crossed, and it is easy to see that the example could be extended so that the empty position would be contained within a subordinate clause an arbitrary number of clauses down, for instance adding iterations of *X said*.

(3) That's the professor who$_1$ [$_S$ I think [$_S$ Anna said [$_{S'}$ that [$_S$ she thought [$_S$ it's important [$_{S'}$ that [$_S$ you meet with e_1]]]]]]]

This example involves relativization, but clefting, questioning and topcalization also share the property of (apparent) unboundedness.

Examples similar to the ones in English in (1) can be constructed in Swedish as well (4a)–(4d).

(4) a. Jag har en släkting$_1$ som e_1 bor i Ramnäs.
I have a relative that lives in Ramnäs.
'I have a relative who lives in Ramnäs.'

b. Jag undrar *vem*$_1$ som e_1 åt upp min smörgås.
I wonder who SOM *ate* PRT *my sandwich*
'I wonder who ate my sandwich.'

c. Det var Anita$_1$ som e_1 åt upp smörgåsen.
it was Anita that ate PRT *sandwich*-DEF
'It was Anita who ate the sandwich.'

d. Smörgåsen$_1$ åt Anita upp e_1.
sandwich-DEF *ate Anita* PRT
'The sandwich, Anita ate.'

Example (4a) shows a relative clause, example (4b) an embedded constituent question, and example (4c) a *det*-cleft, a construction very similar to the English *it*-cleft. In order to avoid making assumptions about the pragmatic function of the fronted phrase in (4d), I use the term **T-preposing** for fronting to the pre-verbal position in declarative main clauses, instead of **topicalization** (see section 4.1.1 for a more detailed description of how I use the term).

A more articulated analysis of each of the clause types is presented in sections 4.1.4 and 4.1.5 in chapter 4. What will be of immediate interest to us here is that Swedish is one of the languages where these clause types involve forming an Ā-dependency, and that, unsurprisingly, Ā-dependencies in Swedish share many properties with Ā-dependencies in English and other languages. For example, like in English, these constructions seem to be unbounded.

(5) Jag undrar vem$_1$ [$_S$ Anna sa [$_{S'}$ att [$_S$ hon tycker [$_{S'}$ att [$_S$ det är
I wonder who Anna said that she thinks that it is
viktigt [$_{S'}$ att [$_S$ du pratar med e_1]]]]]]].
important that you talk with
'I wonder who Anna said that she thinks it is important that you talk to.'

Vem 'who' in (5) has crossed four clause boundaries, and like in the English case in (3) discussed above, we can imagine adding iterations of *X sa* 'X said'.

Surprisingly, given the apparently unbounded nature of these dependencies, it was discovered that there are nevertheless some limits on when they can be formed. In a discussion on interrogatives and relative clauses Chomsky (1964) provides the set of examples in (6).

(6) a. Mary saw the boy walk towards the railroad station

b. Mary saw the boy who was walking towards the railroad station

c. Mary saw the boy walking towards the railroad station

(Chomsky 1964:44)

Chomsky observes that the sentence in (6c) is ambiguous, and could be interpreted either as meaning the same thing as the example with a small clause structure in (6a), or as the example with the relative clause in (6b).[2] A difference between the two interpretations is that it is only in (6a) that Mary has to see the walking event. Example (6b) could be true even if she just saw the boy. Interestingly, the two sentences in (7) are not ambiguous in the same way.

(7) a. the railroad station that Mary saw the boy walking towards is about to be demolished

b. what did Mary see the boy walking towards?

(Chomsky 1964:45)

In both of these sentences, we must interpret *walking towards* as a small clause, and Mary must be seeing the walking event.

In the transformational grammar of Chomsky (1964), relative clauses and questions were formed by applying a *wh*-transformation to a string. The transformation, which preposed a *wh*-marked NP to the beginning of the string, could apply either as the Relative transformation or as the Interrogative transformation. The two interpretations of *walking towards* in (6c), as a small clause or as a relative clause, are each represented by a different string. What Chomsky

[2] There is also a third interpretation of (6c) where *Mary* is the subject of *walk*, but that interpretation is not relevant to the argument.

points out is that in (7a) and (7b), which are formed by applying the Relative transformation and the Interrogative transformation respectively, only the small clause interpretation of *walking towards* is possible. Chomsky argues that we can understand this if we assume that the *wh*-transformation cannot apply to its own output. Specifically, if the Relative transformation has already applied to form a string, which is embedded as a relative clause, no *wh*-transformation can apply again to the embedded string to prepose an NP to the matrix sentence, neither the Relativization transformation nor the Interrogative transformation. If this is the case, the source string for (7a) and (7b) would have to be the string with the small clause interpretation in each case, and we do not expect an ambiguity similar to that in (6c) to arise.

Similarly, Chomsky also notes the contrast in (8).

(8) a. 'he wondered where John put what'

b. *'what did he wonder where John put'

(Chomsky 1964:44)

Example (8b) is ungrammatical, and according to Chomsky (1964) the explanation for this is the same as the explanation for the lack of ambiguity in (7). Since the Interrogative transformation cannot be applied to an embedded string which was formed by applying a *wh*-transformation, such as the embedded questions in (8), sentences like (8b) cannot be formed.

Ross (1967) found several more environments that were special in not allowing certain types of operations to apply.

(9) a. *The man who I read a statement which was about is sick. (p. 119)

b. *Who does Phineas know a girl who is working with. (p. 124)

c. *The hat which I believe the claim that Otto was wearing is red. (p. 126)

d. *What sofa will he put the chair between some table and. (p. 158)

e. *The hat which that I brought seemed strange to the nurse was a fedora. (p. 246)

Creating an Ā-dependency into a relative clause (9a) and (9b), a noun complement (9c), a co-ordinate structure (9d), or a sentential subject (9e) results in clearly unacceptable sentences in English. To account for this, Ross proposed that there are several constraints restricting formation of certain syntactic dependencies. From the perspective of extraction from relative clauses, the most relevant constraint is the **Complex NP Constraint** (CNPC), given in (10).

(10) **The Complex NP Constraint**
No element contained in a sentence dominated by a noun phrase with a lexical head noun may be moved out of that noun phrase by a transformation. (Ross 1967:127)

The basic idea is that there are structural environments which are **syntactic islands**, inside which phrases are stranded. Complex NPs are one such structural environment. The constraint accounts for the unacceptability of the examples in (9a)–(9c).

Theories about syntactic islands have developed significantly since Ross's proposal. Later research aimed to move away from construction-specific rules such as (10), and to give more general explanations for the unacceptability of sentences like (8b) and (9a)–(9e). Chomsky (1973) proposed that Ā-dependencies are derived by moving a phrase from the empty position inside the clause to the position outside of it. Instead of constraints like the CNPC, Chomsky proposed that movement dependencies are subject to a general condition on locality, **Subjacency**. The definition in (11), where IP stands for "inflection phrase", is from Haegeman (1994).

(11) **Subjacency**
Movement cannot cross more than one bounding node, where bounding nodes are IP (S) and NP. (Haegeman 1994:402)

Subjacency is designed to capture several types of islands, while still allowing unbounded dependencies in certain long extractions, e.g. extractions out of subordinate clauses such as *that*-clauses. Consider (5), where the gap which is dependent on the moved phrase is indicated with a *t* for trace.

(12) Jag undrar vem$_1$ [$_{IP}$ Anna sa att [$_{IP}$ hon tycker att [$_{IP}$ det är viktigt
I wonder who Anna said that she thinks that it is important
att [$_{IP}$ du pratar med t_1]]]].
that you talk with
'I wonder who Anna said that she thinks it is important that you talk to.'

Movement of *vem* 'who' from the gap in the most deeply embedded *att*-clause to its position in the left edge of the embedded question in one long movement would violate subjacency, since it crosses four IP-nodes. However, it is assumed that the left edges of many clauses can function as intermediate landing sites, i.e. that there is a position at the left edge of the clause which can function as an escape hatch in long extraction. This position, COMP, is outside of IP, but still part of the extended clausal projection. The availability of such landing sites means that *vem* can move in short steps from COMP to COMP, in a **successive-cyclic** fashion, crossing only one IP-node each time.

(13) Jag undrar [$_{\text{COMP}}$ vem$_1$ [$_{\text{IP}}$ Anna sa [$_{\text{COMP}}$ t_1 att [$_{\text{IP}}$ hon tycker
I wonder who Anna said that she thinks
[$_{\text{COMP}}$ t_1 att [$_{\text{IP}}$ det är viktigt [$_{\text{COMP}}$ t_1 att [$_{\text{IP}}$ du pratar med t_1]]]]]]]].
that it is important that you talk to
'I wonder who Anna said that she thinks it's important that you talk to.'

As we saw in (8b) extraction out of embedded questions is restricted in English, and the subjacency-based account captures this on the assumption that COMP can only contain one Ā-moved phrase. The COMP in an embedded question is already occupied by a *wh*-phrase, which blocks extraction of another phrase. Similarly, the COMP of relative clauses in languages like Swedish and English is often assumed to contain a relative operator, which also blocks extraction (see e.g. Chomsky 1977, Platzack 2000).[3]

Research on languages like Irish (McCloskey 1979, 2001), Chamorro (Chung 1982, 1994), and many others, provides further evidence that long movement consists of multiple short steps. In these languages, the morphosyntax indicates that information about an Ā-dependency is available in every clause over which the dependency holds, suggesting the presence of intermediate landing sites. In Irish, the relevant data are from the complementizers of finite clauses, which are morphologically marked if Ā-movement has taken place out of the clause that they head (McCloskey 2001:68). In Chamorro, Ā-movement in a clause is signalled by agreement morphology on the verb, which varies with the case of the *wh*-trace (Chung 1994:7).

There are also several types of evidence suggesting that the extracted phrase in an Ā-dependency affects the clause where the gap is situated. An example from Swedish is given in (14).

(14) a. [Vilk-a av student-er-na]$_1$ tror du [t_1 blev missnöjd-a över
which-PL *of student*-PL-DEF *think you* *became displeased*-PL *over*
sina betyg]?
their.REFLX.PL *grades*
'Which of the students do you think were displeased with their grades?'

b. [Vilk-en av student-er-na]$_1$ tror du [t_1 blev missnöjd över
which-SG *of student*-PL-DEF *think you* *became displeased*.SG *over*
sitt betyg]?
their.REFLX.SG *grade*
'Which of the students do you think was displeased with their grade?'

[3] In *som*-introduced relative clauses the operator is silent, but the phrase in COMP can also be pronounced, in the case of relative pronouns and adverbs.

In Swedish, the predicate adjective in a copular clause agrees with its subject in number, and we see that in (14) the adjective *missnöjd* 'displeased' agrees with the *wh*-phrase, so that it is plural when the *wh*-phrase is plural, and singular when the *wh*-phrase is singular. Furthermore, reflexive pronouns like *sina* 'their.PL' and *sitt* 'their.SG' need to have an antecedent in the local clause, but in (14) there is no potential binder in the embedded clause, and the reflexives need to be interpreted as being bound by the *wh*-phrases. Both of these facts indicate that information about the *wh*-phrases is available in the embedded clause in these examples. A common way to model this distribution of information throughout the sentence is to adopt a movement-based account.[4]

Another influential proposal for how to capture certain locality effects is due to Rizzi (1990), who introduced the notion of **Relativized Minimality**, according to which some types of islands are thought of in terms of intervention. Intervention was relativized to the type of movement, such that A-movement blocks A-movement, Ā-movement blocks Ā-movement, and head movement blocks head movement. From this perspective, what makes an example like (15) unacceptable is that it requires *how* to move via Ā-movement over *which problem*, which has also moved via Ā-movement, resulting in a violation of minimality.

(15) *How$_1$ did he wonder which problem$_2$ to fix t_2 t_1?

Rizzi's proposal was framed in the Government and Binding framework. We will look more closely at how examples like (15) can be treated in chapter 6, but informally, *which problem* blocks movement of *how* to the higher CP here because *which problem* is an Ā-specifier which c-commands the trace position of *how*, and *how* would move to a position which *which problem* does not c-command.

The insight of relativized minimality has been captured in later research by principles requiring movement not to skip potential landing sites. What minimality should be relativized to is thought of in a much more fine-grained way, often in terms of features of the moving element. Within the Minimalist program, movement is commonly assumed to be driven by features of a head, which trigger movement of a phrase lower in the structure to the head's specifier. Relativized minimality can then be built into the definition of movement, if the operation has to affect the closest matching phrase (for a brief overview of this development, see Boeckx 2012:19–25).

[4] Other ways include SLASH-features, as in GPSG (Gazdar et al. 1985) and HPSG (Pollard & Sag 1994) (for an overview up until 1988, see McCloskey 1988).

We will return to characteristics of Ā-movement in section 2.1.2, after a brief introduction of how successive cyclicity and Ā-dependencies are treated in Phase Theory.

2.1.1 Ā-dependencies in Phase Theory

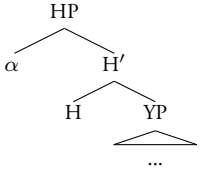
In research within the Minimalist program, successive cyclicity and constraints on locality are derived by the **Phase Impenetrability Condition**, which was introduced as part of **Phase Theory** (Chomsky 1993, 1995, 2000, 2001). According to this view, syntactic structure is not first built and then evaluated with respect to constraints like Subjacency. Instead the syntactic derivation proceeds by **phase**.

In a syntactic derivation, structure is built by the operation **Merge**, which takes two smaller structures A and B and forms a larger structure G={A,B}, as illustrated in (16).

(16)

Merge applies cyclically, building larger and larger structures, until a phase head is merged. Consider the tree in (17).

(17)



If HP is a phase, then H is a **phase head**. Once the phase is fully constructed, the complement of the phase head, here YP, undergoes a process known as **spellout**, by which it is shipped off to the interface levels of LF and PF for semantic and phonetic interpretation, respectively. As a result, the contents of the spellout domain are no longer accessible to syntactic operations as the derivation proceeds. CP and *v*P, and on some accounts DP, are assumed to be phases.

The operation Merge can be either external or internal. **External merge** is when two separate constituents are merged, as in (16). Movement is modelled as **Internal merge**, which is what happens when one of the constituents that Merge applies to is part of the other. In Minimalism, movement is not assumed to create a trace, but rather to leave an unpronounced copy in the gap site (see e.g. Chomsky 1993), as in (18), where the unpronounced copy is within angle brackets.

(18) Which book did you like <which book>?

In this dissertation I discuss proposals from several theoretical frameworks, and I will continue to use **movement** as a descriptive term. In order to save space, I will also use the trace notation for the gap created by movement/Internal merge in example sentences and trees.

Since Ā-movement is potentially unbounded, there must be a way to extract certain elements from within the phase. This necessitates introducing the concept of a **phase edge**, which serves as an escape hatch through which a moving element may escape before spellout of the complement of the phase head. In (17) the specifier of H is the phase edge, and it is occupied by α.

Successive-cyclic movement through phase edges is enforced by the Phase Impenetrability Condition (PIC) (19), which holds for a phase HP with a head H.

(19) **Phase Impenetrability Condition**
The domain of H is not accessible to operations outside HP; only H and its *edge* are accessible to such operations. (Chomsky 2001:13)

The complement of the phase head becomes invisible for further operations after the phase is spelled out. If a phrase is to be extracted from the phase, it has to move to its edge before spellout.

To take a concrete example, consider (20), which is a slightly adapted version of (9a) above, and involves extraction from a relative clause.

(20) *What$_1$ does Phineas know a girl [$_{\text{CP}}$ who$_2$ [$_{\text{TP}}$ t_2 is working with t_1]]?

Here, the relative pronoun *who* has moved to the edge of the embedded CP-phase, and since C is a phase head, the TP *<who> is working with what* is spelled out, and therefore inaccessible to further operations. This means that the *wh*-phrase *what* cannot be extracted, since it is no longer visible for the higher C[wh].[5]

The theory that spellout occurs cyclically phase-by-phase, which is known as **multiple spellout**, differs from previous conceptions of syntactic derivation such as that in Government and Binding, in which every part of the syntactic tree is in principle accessible to syntactic operations at any point in the

[5] Since relative clauses are generally assumed to be adjuncts, it is common to refer not only to PIC but also to the **Condition on Extraction Domain** (CED) (Huang 1982) in order to explain why ERC, as in (20), is impossible. We will return to the status of the CED in Swedish in section 4.5.

derivation, unless additional principles intervene to block the application of a syntactic process.

2.1.2 Properties of Ā-movement

As we have seen in the chapter so far, Ā-movement dependencies in Swedish share some basic properties with English Ā-movement dependencies. In both languages, Ā-movement creates a gap, and is (apparently) unbounded. It is nevertheless constrained in both languages. Even if (1)–(3) in chapter 1 raise the question whether a constraint which blocks movement out of relative clauses holds in Swedish, there are some constructions where constraints on movement seem to apply. Consider the Swedish versions of (9d) and (9e) in (21) and (22).

(21) *[Vilken soffa]$_1$ tänker han ställa stolen mellan ett bord och t_1?
*which sofa thinks he put chair.*DEF *between a table and*

(22) *Hatten$_1$ som [att jag tog med t_1] verkade konstigt för sjuksköterskan var
*hat.*DEF *that that I took with seemed strange for nurse.*DEF *was*
en stråhatt.
a straw hat

These examples illustrate that it is not possible to extract a phrase from a coordinate structure or from a sentential subject, which means that Swedish does exhibit some types of island effects.

The past decades of research have shown that Ā-movement dependencies share many other properties as well. This is particularly well described for English (for overviews, see for example Chomsky 1977, McCloskey 1988, Haegeman 1994, Pesetsky 2013), but has also been investigated in many other languages. In this section, I will show that many of these cross-linguistically common characteristics are shared by Swedish Ā-movement dependencies.[6]

In (14) above we saw that the Ā-moved phrase can affect the clause where it originated in various ways. Another characteristic property of Ā-movement is that facts about the form of the fronted phrase or its interpretation often indicate that the phrase originates in the gap site. One example of this is **case connectivity**, which is illustrated in (23) and (24).

[6] For other accounts of Ā-movement in Swedish, see Engdahl & Ejerhed (1982), Engdahl (1986), Platzack (1998, 2011) and Teleman et al. (1999 4:405–437).

(23) a. Dig$_1$/*Du$_1$ vill de gärna träffa t_1.
*you.*ACC/*you.*NOM *want they gladly meet*
'They would love to meet you.'

b. Dig$_1$/*Du$_1$ sa de [att de gärna ville träffa t_1].
*you.*ACC/*you.*NOM *said they that they gladly wanted meet*
'They said that they would love to meet you.'

(24) a. Det var dig$_1$/*du$_1$ de ville träffa t_1.
*it was you.*ACC/*you.*NOM *they wanted meet*
'You are the one that they wanted to meet.'

b. Det var dig$_1$/*du$_1$ de sa [att de gärna ville träffa t_1].
*it was you.*ACC/*you.*NOM *they said that they gladly wanted meet*
'You are the one that they said that they would love to meet.'

Example (23) shows local T-preposing (23a), and T-preposing from an *att*-clause (23b). Example (24) exhibits the case pattern for the pivot of a *det*-cleft, with local clefting (24a), and clefting of an argument in an *att*-clause (24b). In all of the examples, the fronted phrase is in the accusative, which is expected if it originates as a complement of e.g. *träffa* 'meet'. Since the only phrases that are overtly case marked in Swedish are personal pronouns, we can only see case connectivity in examples with T-preposing and clefts; we cannot test for case connectivity in constituent questions or relative clauses. Where it is possible to test, as in (23), and (24), however, the examples show that the fronted pronoun will have the case it would be assigned if it were realized in the gap position.[7]

Furthermore, Ā-movement typically exhibits connectivity with respect to binding, i.e. a certain pattern with respect to the possible interpretations of reflexives, pronouns, and full DPs. Consider the set of examples in (25), where co-reference is indicated with a subscript *x*.

(25) a. [Sina$_x$ sista ord]$_1$ yttrade [Julius Caesar]$_x$ t_1 år 44 f.kr.
*his.*REFLX *last words uttered Julius Caesar year 44 BC*
'Julius Caesar uttered his last words in 44 BC.'

[7] The default case in Swedish is nominative, unlike in English, which is revealed by examples like (i).

(i) a. Vem vill följa med på bio?
who wants follow with on movie
'Who wants to go to the movies?'

b. Jag!/*Mig!
I me

b. *[Hans$_x$ sista ord]$_1$ yttrade [Julius Caesar]$_x$ t_1 år 44 f.kr.
*his.*PN *last words uttered Julius Caesar year 44 BC*

In (25a), the initial phrase contains a reflexive pronoun *sin* 'his.REFLX', which is co-referent with the subject of the clause, *Julius Caesar*. In (25a), the initial phrase contains a pronoun *hans* 'his.PN', and the sentence is not acceptable when this pronoun is co-referent with the subject. A reflexive must be bound by a sufficiently local subject which c-commands it. Pronouns cannot be locally bound.[8] The available and unavailable interpretations of the nominal expressions in (25) can be explained if we assume that the fronted phrase in each case originates as a complement of *yttrade* 'uttered', where it is c-commanded by the subject, and then Ā-moves to the clause-initial position.

The possible interpretations of the personal pronouns in (26) illustrate another characteristic pattern.

(26) a. Gunnar$_x$ sa att Julia inte hade ringt honom$_{x/y}$ ännu.
Gunnar said that Julia not had called him yet
'Gunnar$_x$ said that Julia hadn't called him$_{x/y}$ yet.'

b. [Gunnar$_x$]$_1$ sa han$_{*x/y}$ att Julia inte hade ringt t_1 ännu.
Gunnar said he that Julia not had called yet.
'Gunnar$_x$, he$_{*x/y}$ said that Julia hadn't called yet.'

Example (26a) is acceptable both when the proper noun *Gunnar* is co-referential with the pronoun *honom* 'him', and when the proper noun and the pronoun are not co-referential. Example (26b) is only acceptable on one interpretation, namely when the pronoun and the proper noun are not co-referential. The example can not be interpreted as *Gunnar* and *han* 'he' referring to the same person. If we assume that the phrase *Gunnar* originated as an object of *ringa* 'call' in the embedded clause, and reached its position by Ā-movement, we can connect these facts to the restriction that proper nouns can not be co-referential with a phrase that c-commands them.[9] In the position of the trace, *Gunnar* would be c-commanded by *han* 'he', and it is thus expected that the co-referential interpretation is impossible.

In constituent questions, **crossover effects** (Postal 1971), as in (27a) and (27b) are also characteristic of Ā-movement dependencies.

[8] Condition A and Condition B of the Binding Theory (Chomsky 1981).

[9] Condition C of Binding Theory (Chomsky 1981).

(27) a. *[Vilken student]$_{x1}$ sa han$_x$ att Julia inte hade ringt t_1 ännu?
which student said he that Julia not had called yet
*'[Which student]$_x$ did he$_x$ say that Julia hadn't called yet?

b. *[Vilken student]$_{x1}$ brukar hans$_x$ mamma ringa till t_1?
which student tends his mom call to
*'[Which student]$_x$ does his$_x$ mom often call?'

The sentence in (27a) exhibits strong crossover, and (27b) weak crossover, where the co-referent 'crossed over' pronoun is the possessor inside a DP. In neither case is co-reference between the *wh*-phrase and the pronoun possible.

In English, weak crossover effects are absent in relative clauses, but in Swedish, relativization induces weak crossover effects. The example in (28) is adapted from Engdahl (1985b:9).

(28) *[mannen_x]$_1$ som hans$_x$ mor tyckte bäst om t_1
man.DEF *that his mother thought best about*
'the man who his mother liked best'

The final characteristic for Ā-movement which I will review here is its ability to license parasitic gaps, i.e. gaps that are dependent on other gaps (Engdahl 1983, 1985b). We see an example in (29a), which is a adapted from Engdahl (1983:5). The real gap, which is the complement of *file*, licenses the parasitic gap which is the complement of *reading*, in the adjunct. The example in (29b) shows that the parasitic gap is not acceptable when there is no licensing gap.

(29) a. [Which articles]$_1$ did John file t_1 [without reading $\_\__{pg}$]?
b. *John filed the articles without reading $\_\__{pg}$.

There is some variation in speakers' intuitions about the acceptability of parasitic gaps, but the variation is not random. Rather, there is an implicational hierarchy of domains in which parasitic gaps are allowed. One factor that influences their acceptability is tense. If a speaker allows a parasitic gap in a tensed clause, the same speaker will generally also allow a parasitic gap in an untensed domain (Engdahl 1983).

For speakers who accept parasitic gaps, these can be used to diagnose Ā-movement dependencies, since the real, licensing gap is created by precisely the types of syntactic dependencies which we have identified as created by Ā-movement, i.e. constituent questions, topicalization, clefting and relativization.

Engdahl (1983) also investigates parasitic gaps in Swedish, and the general pattern seems to be that they are more acceptable in Swedish than in English. The examples in (30) show that all of the clause types that we have described as involving Ā-movement dependencies license parasitic gaps in Swedish.

(30) a. Det där är en sorts bönor$_1$ som man inte kan äta t_1 [utan att koka $\_\__{pg}$].
the there is a kind beans that one not can eat w/o to cook
'That is a kind of bean which you cannot eat without cooking.'

b. [Vilken sorts bönor]$_1$ äter du inte t_1 [utan att koka $\_\__{pg}$]?
which kind beans eat you not w/o to cook
'What kind of beans won't you eat without cooking?'

c. Det är [den sortens bönor]$_1$ som man inte kan äta t_1 [utan att koka $\_\__{pg}$].
it is that kind beans that one not can eat w/o to cook
'It is that kind of beans that you can't eat without cooking.

d. Bönorna$_1$ tror jag att hon åt t_1 [utan att koka $\_\__{pg}$]
*beans.*DEF *think I that she ate w/o to cook*
'The beans, I think she ate without cooking.'

e. *Jag åt bönorna [utan att koka $\_\__{pg}$].
*I ate beans.*DEF *w/o to cook*

The parasitic gap in the adjunct clause *utan att koka* 'without cooking' is licensed by the Ā-movement dependency in a relative clause in (30a), in a constituent question in (30b), in a *det*-cleft in (30c), and in T-preposing in (30d). Again, the last example shows that no gap is licensed in the adjunct clause when there is no Ā-movement dependency (30e).

An overview of the characteristics of Ā-movement is given in (31).

(31) *Ā-movement*
- creates a gap
- is (apparently) unbounded
- respects island constraints
- exhibits crossover effects
- exhibits connectivity effects
- licenses parasitic gaps

When I discuss the nature of the Ā-dependencies in ERC in the following chapters, I will use these characteristics as diagnostics.

2.1.3 Two types of Ā-dependencies

Cinque (1990) develops a more fine-grained understanding of Ā-dependencies, arguing that they are not all created the same way. Investigating English and Italian Ā-dependency constructions, he finds indications that the gaps in these constructions are not all of the same type. The gap left by movement is assumed to be a trace, or in more recent analyses, as we saw in the previous section, a copy

of the moved phrase. Cinque observes that the gap in some constructions does not conform to all of the characteristics of traces (or copies). He investigates examples like the ones in (32).

(32) a. The article was too long for us to read *e* ...

b. (?) The article that we filed without reading *e* ...

c. (?) The article that we went to England without reading *e* ...

(Cinque 1990:98)

In (32a), the gap is in a complement object deletion construction (Lasnik & Fiengo 1974). The gap in (32b) is a parasitic gap, similar to those we saw in the previous section, and in (32c) the gap is inside an adjunct island.

Even though these constructions share several properties with regular Ā-movement constructions, there are also differences. An important difference is that, at least in English and Italian, they are restricted to DP-gaps. This would be surprising if they were derived by regular Ā-movement, because Ā-movement is not restricted in such a way, and displaces many phrasal categories.

Cinque's (1990) proposal is that the empty elements in such constructions are not the same as in regular Ā-movement constructions. In regular Ā-movement, successive-cyclic movement forms a chain of traces (or copies), headed by the Ā-moved phrase. In the case of the constructions in (32), Cinque proposes that the head of the Ā-chain is base generated in its surface position, and is related to an empty pronominal element *pro* in the gap site via what he calls **Ā-binding**, which, unlike Ā-movement, is proposed to be a long-distance relation. This is shown in (33) for (32a).

(33) $[\text{The article}]_x$ was too long for us to read pro_x.

The pronominal element can be thought of as a silent resumptive pronoun, i.e. it has to be interpreted as co-referent with an antecedent.

In what follows, I will use the trace notation *t* to indicate that I analyze a dependency as created by Ā-movement, and *pro* when I analyze it as created by binding. In cases where it is not clear what kind of dependency it is, I will use the *e* for an unspecified empty category.

In order to argue that there really is extraction from relative clauses in Swedish, we need to rule out that ERC-sentences involve binding of a silent *pro* rather than Ā-movement. I will investigate this in detail in section 4.2, and I will mark the dependency with an *e* in the meantime. Examples cited from other authors are often rendered in my notation in order to increase readability.

2.1.4 Functional approaches

The constraints on extraction proposed by Chomsky, Ross, and others were syntactic, but early on it was noted that there is a need for a functional, pragmatic component in accounts of extraction. The argument was first made by Erteschik-Shir (1973, 1982, 2007). Erteschik-Shir discusses extraction in Danish, which permits extraction from a wider variety of embedded clauses than English. She observes that in Danish, there are certain relative clauses which are not islands for extraction, as illustrated in (34).

(34) Det$_1$ kender jeg mange [der har gjort e_1] (Da.)
that know I many who have done
'I know many people who have done that.'
(Erteschik-Shir 1973:63)

However, other relative clauses in Danish are islands for extraction.

(35) *Det$_1$ har jeg peget paa mange [der har gjort e_1] (Da.)
that have I pointed at many who have done
(Erteschik-Shir 1973:64)

Since, according to Erteschik-Shir (1973, 1982, 2007), there is no structural difference between the relative clauses which allow extraction and those which do not, the difference has to lie elsewhere. She argues that the possibility of extraction is governed by the pragmatics and semantics of the utterance. Specifically, it is only possible to extract out of those embedded constituents which are **dominant**, i.e. those which are prominent in that they can be commented on in the upcoming discourse.

The idea that the discourse function, information impact, or semantics of an embedded constituent is what decides whether it will be an island or not is taken up by several other researchers (e.g. Van Valin 1994, 1996, Van Valin & LaPolla 1997, Goldberg 2006, 2013). An idea that these proposals have in common, and which they also share with Erteschik Shir's approach, is that a constituent is an island if it is not discourse-prominent in some respect.

Goldberg (2006, 2013) takes a construction-based approach, and suggests that constraints on extraction can in most, if not all, cases be explained by the information-structural properties of the constructions involved. On this view, constraints on extraction should be conceptualized as pragmatic clashes which come about when a long-distance dependency construction, e.g. a topicalization construction or a question construction, is combined with another construction, which it is not compatible with. According to Goldberg (2006), extracted phrases are "positioned in discourse-prominent slots" (p. 135), and the gap in such constructions cannot be in a part of the sentence which is **backgrounded**,

because this would require us to treat the extracted phrase as both backgrounded and discourse-prominent at once. Goldberg's approach is inspired by Van Valin (1994, 1996) and Van Valin & LaPolla (1997), who propose that restrictions on extraction should be expressed in terms of **focus domains**. On Van Valin and Van Valin & LaPolla's view, the extracted phrase needs to be in the potential focus domain of the sentence, i.e. in the part of the sentence where focal elements can occur, in order for extraction to be possible. We will look more closely at such approaches in relation to Swedish ERC in section 5.5.

2.1.5 Processing factors

In the previous sections, I have introduced structural and functional conditions on extraction. Many researchers have also observed processing constraints. For instance, island effects may be connected to difficulties in processing certain types of dependencies (see Fodor's (1978) Nested Dependency Constraint). It has also been proposed that extraction phenomena can be explained in terms of limits on attention span (Deane 1991), or in terms of a combination of factors related to sentences processing, such as working memory and lexical semantic processing (Kluender & Kutas 1993).

Investigating island effects in online processing has become an important field in psycholinguistics. There have been attempts, for example, to reduce island effects to processing (Hofmeister & Sag 2010). However, it has also been argued that such a reduction is not possible (Phillips 2006, Sprouse et al. 2012). See Sprouse & Hornstein (2013) and the papers in that volume for recent discussion of various issues in relation to island effects and processing.

With respect to processing of extraction sentences in the mainland Scandinavian languages, there have been some recent experimental investigations which I will bring up in sections 2.2.1.2 and 2.2.1.3.

2.1.6 Strong and weak islands

In recent work on extraction, it is common to distinguish between two classes of domains, **strong islands** and **weak islands** (for an overview, see Szabolcsi 2006). A basic characteristic that can be taken as a starting point is that strong islands are those domains from which no extraction is possible, whereas weak islands are domains which allow some phrases to be extracted but not others. These are also sometimes referred to as **absolute** and **selective islands**, respectively.

Several of the types of islands we have seen in the previous sections are considered strong islands. Co-ordinate structures, subjects (being left branches),

adjuncts, and complex DPs, such as the relative clause complex, are usually counted as belonging to this class. The prototypical weak islands are embedded questions, but later research has connected many other phenomena to the class, such as negative islands (also known as **inner islands** Ross 1984) and complements of factive verbs (Rizzi 1990, see also Szabolcsi 2006 and references therein).

Given the distinction between Ā-movement and Ā-bound *pro* introduced in section 2.1.3 (Cinque 1990), things get slightly more complicated. One of the places where Ā-bound *pro* is assumed to occur is precisely in constructions which look like they are derived by Ā-movement, but where movement would violate island constraints. This would mean that we can expect to find some cases of apparent extraction even in strong islands. However, we can expect strong islands to be able to contain a DP-gap at most, given the assumption that *pro* can only be bound by DPs. As soon as a domain can contain a gap of e.g. a PP, the domain should be classified as a weak island (Szabolcsi 2006:482).

Szabolcsi (2006) extends this reasoning to question formation. Consider the two sets of examples in (36) and (37), which show a typical pattern used to illustrate the distinction between strong and weak islands. The %-sign indicates that some speakers accept the sentence while others do not.

(36) a. %About which topic did John ask <whether to talk ___ >?
b. Which topic did John ask <whether to talk about ___ >?
c. *How did John ask <whether to behave ___ >?

(37) a. *About which topic did you leave <without talking ___ >?
b. Which topic did you leave <without talking about ___ >?
c. *How did you leave <without behaving ___ >?

(Szabolcsi 2006:481)

Examples (36a)–(36c) show extraction from infinitival questions. Extracting a PP is accepted by some speakers (36a), and DP-extraction is generally accepted (36b). However, some phrases are not acceptable to extract at all, as shown by extraction of the manner *wh*-phrase *how* (36c). This is the kind of pattern expected for weak islands. In extraction from adjuncts, as in (37), only DP-extraction is reported to be possible, and adjuncts are consequently taken to be strong islands.

Theories of weak islands have developed significantly in the last decades, covering more and more phenomena and making finer and finer distinctions in what types of phrases can be extracted. Since Swedish apparently permits extraction from relative clauses, their status as strong islands in the language

is in doubt, and an interesting question is whether they fit into the class of weak islands. In chapter 6, I return to this question, and aim to situate Swedish relative clauses with respect to research on weak islands.

2.2 Extraction in the mainland Scandinavian languages

The awareness that phrases which have a grammatical function inside a subordinate clause or phrase can sometimes be realized outside of that phrase predates the theoretical discussions detailed in section 2.1. In the Scandinavian context, early references include Mikkelsen (1894), Cederschiöld (1897), and Lindstedt (1926). The Danish grammarian Mikkelsen refers to the phenomenon as *"sammenslyngning af en overordnet og en underordnet sætning"* 'intertwining of a superordinate and a subordinate clause'. A later Danish term is *sætningsknude* 'sentence knot'. Norwegian has a similar term, *setningsknute* 'sentence knot', and the Swedish traditional term is *satsfläta* 'sentence braid'.

Some of the first examples of long extractions mentioned in the scholarly literature are found in prescriptive discussions of proper language use, where they are sometimes described as "unkempt" or illogical, but are not always entirely condemned (for an overview of the early stylistic advice on extraction see Teleman 1991). Cederschiöld (1897) studies written Swedish, and argues that the written language has to be what he considers to be logical and correct. In a passage comparing written and spoken language, he notes that it is not uncommon to hear sentences like (38) in spoken, informal language. The translation and glosses are mine in this example, as in several other examples in this section where the sources are in Swedish or Danish.

(38) I morgon tror ja' säkert (att) ja' ska bli färdig.
tomorrow think I surely that I will become ready
'I think that I'll be ready tomorrow, for sure.'
(Cederschiöld 1897:153)

The sentence involves fronting of an adverbial which is interpreted as modifying the event in the embedded *att*-clause to the preverbal position in the matrix clause. I have left the fronted phrase in situ in the English paraphrase in order to make it more natural sounding, a practice I will maintain in the dissertation when it is possible. According to Cederschiöld (1897), the word order in (38) is reckless, and the sentence is mentioned together with what he calls errors in coherence, order, and clarity. Ljunggren (1917) evaluates a similar example as

illogical, but permissible in informal style (§46, p. 36).

As far back as I have been able to trace references, Mikkelsen (1894) is the first to give examples of long extraction where the gap is inside a relative clause, namely the examples in (39).

(39) a. Det$_1$ er der ingen [der har sagt mig e_1]. (Da.)
that is there no one who have said me
'No one has told me this.' (p. 441)

b. Her er en Person$_1$, som du næppe kan nævne nogen [der kan maale
here is a person that you hardly can mention someone who can measure
sig med e_1]. (Da.)
REFLX *with*
'Here is a person such that you can hardly find anyone who can be compared to them.' (p. 442)

c. Han fortalte mange Historier, hvorav$_1$ han selv var den [der lo
he told many stories of which he REFLX *was the one who laughed*
mest e_1]. (Da.)
most
'He told many stories such that he was the one who laughed the most at them.' (p. 442)

Det 'that' in (39a) is T-preposed. In (39b) and (39c) the long extraction consists of relativization. It is necessary to insert resumptive pronouns in the gap sites in the English translations of (39b) and (39c) to make them acceptable, but there are no resumptive pronouns in the Danish ERC-sentences.

Mikkelsen's (1894) examples are given in a broader overview of preposing in Danish, and there is no stylistic discussion or remark about their correctness or logic. The paragraph at hand is specifically about the intertwining of superordinate and subordinate clauses, and the ERC-examples are listed among examples with extraction from other clause types, without any special comment.

For Swedish, an early source for stylistic advice about long extraction is Wellander (1939:503ff). Wellander's view is a little more nuanced than Cederschiöld's. On the one hand, he states that long extraction is not suitable for polished prose, and that it amounts to a "strange disturbance of the syntax" [my translation]. This is especially the case, according to Wellander, if many long words come in between the fronted phrase and the gap. Extraction sentences, which are by their nature illogical in Wellander's opinion, should be avoided if they can be reformulated in a natural way.

On the other hand, Wellander points out that extraction is very common in Swedish, and that it has an important role to play in spoken language

and colloquial prose. He also argues that a natural but more informal way of expressing oneself is preferable to a more logical way of speaking if the latter is forced and "un-Swedish" (Wellander 1939:513). A few of the ERC-examples Wellander (1939) provides are given in (40).

(40) a. Gustav$_1$ är det ingen människa [som har hört ett ord ifrån e_1].
Gustav is there no human that has heard a word from
'There is no one who has heard a word from Gustav.' (p.504)

b. [Mot myggor]$_1$ finns det folk [som använder eukalyptusolja e_1].
against mosquitos exist there people that use eucalyptus oil
'There are people who use eucalyptus oil against mosquitos.' (p. 507)

c. Det$_1$ hade jag aldrig träffat någon [som hade gjort e_1].
that had I never met someone that had done
'I had never met anyone who had done that.' (p. 507)

d. Framför altaret stod en dopfunt av silver, [[i vilken]$_1$ Karl XII
*in front of altar.*DEF *stood a baptismal font of silver in which Charles XII*
var den förste [som döptes e_1]].
was the first that was christened
'In front of the altar was a baptismal font made of silver, in which Charles XII was the first to be christened.' (p. 510)

As in Mikkelsen's grammar, ERC is not singled out among the other types of long extraction. Wellander (1939) does point out, however, that one kind of ERC is easy to avoid, namely the kind in (41a), which could be simplified as in (41b).

(41) a. Mig$_1$ är det ingen [$_{RC}$ som tänker på att hjälpa e_1].
me is there no one that thinks on to help
'There is no one who thinks about helping me.' (p. 511)

b. Mig tänker ingen på att hjälpa.
me thinks no one about to help
'No one thinks about helping me.' (p. 511)

With respect to present day research about ERC, these early references are interesting for two reasons. First, they show that extraction from relative clauses is not a new phenomenon. Second, the fact that extraction from relative clauses has been discussed from a stylistic point of view, and has sometimes been advised against, is interesting, since it indicates that these sentences were spontaneously used in everyday speech (for a similar argument, see Teleman 1991).

A more modern, descriptive source is the *Swedish Academy grammar* (Teleman et al. 1999). According to Teleman et al. (1999 4:419ff) long extraction is rather common in Swedish, especially in spoken language. Several factors are reported to affect the acceptability and use of extraction: the type of the embedded clause which contains the gap, the meaning of the matrix predicate, the grammatical function of the extracted element, the position of the gap, and the complexity of the clause. Speakers differ in how willing they are to use different kinds of extraction, and formality level and medium (spoken vs. written) are important factors in this.

The Swedish Academy grammar discusses extraction from three types of subordinate clauses: complement clauses, relative clauses and adverbial clauses. Extraction from complement clauses is generally accepted by all speakers according to Teleman et al. (1999), whereas extraction from adverbial clauses is described as marginal. An exception is extraction from conditional and temporal adverbial clauses, which is accepted by many speakers in informal spoken language.

Teleman et al. (1999) write that ERC is probably accepted by most speakers, as long as a number of conditions are fulfilled. Taken together, these conditions are said to be stricter than in extraction from complement clauses.

An important condition is that the relative clause has to be restrictive. Extraction from non-restrictive relative clauses is impossible, as illustrated by the following pair from Teleman et al. (1999 4:423).

(42) a. Biljard$_1$ fanns där många [som spelade e_1].
billiards existed there many that played
'There were many people who played billiards.'

b. *Biljard$_1$ fanns där väldigt många människor [som alla spelade e_1].
billiards existed there very many people that all played

Teleman et al. (1999) base their description on a sentence schema, in which the clause has three fields, an initial field, a middle field and an end field (4:5ff).[10] In main clauses, the middle field corresponds to the area between the finite and the non-finite verbs. There is a clear contrast between extraction from a relative clause which is part of a DP in the end field, as in (43a), and extraction from a relative clause in a DP in the middle field, as in (43b).[11]

[10] See also Diderichsen 1946.

[11] This could be understood as a **freezing effect** (Wexler & Culicover 1980).

(43) a. Akupunktur$_1$ brukar det delta en läkare [som kan e_1] vid våra
acupuncture tends there participate a doctor that can at our
seminarier.
seminars
'There is a doctor who knows how to do acupuncture who usually participates at our seminars.' (p. 423)

b. *Akupunktur$_1$ brukar en läkare [som kan e_1] delta vid våra
acupuncture tends a doctor that can participate at our
seminarier.
seminars (p. 423)

Extraction is also said to be considered more acceptable if the gap comes last in the relative clause, or is at least in the end field of the relative clause, and not, e.g. in the middle field.

With respect to the matrix predicate, Teleman et al. (1999) report that there is a preference for it to express either existence or presence (*finnas* 'exist', *komma* 'come'), knowledge (*känna* 'know.REL' [12], *veta* 'know.COG'), observation/perception (*se* 'see', *få syn på* 'spot', *hitta* 'find'), or belonging (*ha* 'have', *förlora* 'lose', *längta efter* 'long for').

(44) a. [Överblivna biljetter]$_1$ fanns/kom det en [som ville sälja e_1].
leftover tickets existed/came there one that wanted sell
'There was a guy who wanted to sell leftover tickets.'
'A guy who wanted to sell leftover tickets came.' (p. 423)

b. Där$_1$ känner/vet/såg/har jag en flicka [som bor e_1].
*there know.*REL*/know.*COG*/saw/have I a girl that lives*
'I know/know of/saw/have a girl who lives there.' (p. 423)

ERC is also more acceptable if the head of the relative clause is indefinite and non-specific.

(45) a. Johan$_1$ känner jag många [som skulle vilja gifta sig med e_1].
*Johan know.*REL *I many that would want marry* REFLX *with*
'I know many people who would like to marry Johan.' (p. 424)

b. *Johan$_1$ har jag inte träffat den flickan [som vill gifta sig med e_1].
*Johan have I not met the girl.*DEF *that wants marry* REFLX *with*
(p. 424)

[12] Both *känna* and *veta* translate to *know* in English. I will translate *känna*, which is used for a relation of acquaintance between people, as 'know.REL' and *veta*, which means 'to have knowledge about or be aware of something', as 'know.COG.

In (45b) the head of the relative clause is definite, and the example is unacceptable.

As mentioned above, the *Swedish Academy grammar* is a descriptive grammar, which aims to give a characterization of what is common in usage and what most people will accept as natural sounding Swedish. In chapter 4, we will look more closely at how well Teleman et al.'s characterization matches spontaneously produced data.

2.2.1 Recent approaches to Scandinavian extraction

Ever since ERC in Danish was brought to the attention of the international research community by Erteschik-Shir (1973), the phenomenon has engaged many researchers. In order to give the general overview in this section some structure, I will organize it around three lines of investigation in recent research: accounts that place an emphasis on the pragmatics, information structure, and semantics of extraction; accounts that center on structural explanations; and some recent experimental studies. This division of the chapter is mainly for expository reasons. Several accounts assume that both pragmatic, information-structural, and semantic factors on the one hand, and syntactic factors on the other, are important. I will discuss each analysis in the section where it makes its largest contribution. While this section focuses on ERC, I also include discussion of several relevant studies that investigate extraction more broadly.

2.2.1.1 Pragmatics, information structure, and semantics

Allwood (1976, 1982) demonstrates that there are sentences in Swedish, just as in Danish, where ERC is acceptable, and argues that this shows that the Complex NP Constraint (Ross 1967) does not hold in the mainland Scandinavian languages. Instead, he argues that semantic and pragmatic factors, and context, are important. The most important factor in determining the acceptability of an extraction sentence, according to Allwood, is that the sentence should form a coherent predication about the fronted constituent, i.e. that the sentence as a whole should have a topic-comment structure.

Even though pragmatics and semantic constraints are considered most important in Allwood's (1976, 1982) account, he observes examples where he argues that the unacceptability could have a syntactic explanation, namely examples like (46a) and (46b).

(46) a. *Salome$_1$ levde Herodes i [hoppet om [att e_1 skulle förföra den
*Salome lived Herod in hope.*DEF *about that would seduce that*
mannen]].
*man.*DEF (p. 5)

b. *[En flicka]$_1$ känner jag till [en bok [som e_1 gav en pojke]].
a girl know I of a book that gave a boy (p. 9)

Both (46a) and (46b) involve extracting a subject out of a complex NP, in (46a) an *att*-clause complement, in (46b) a relative clause. Neither of the examples is grammatical. Allwood (1976, 1982) proposes that what blocks such extractions is a less restrictive version of the Complex NP Constraint, which he calls the **Complex NP Subject Constraint** (47).

(47) **The Complex NP Subject Constraint**
The subject constituent cannot be moved out of a sentence which is embedded in a complex NP. (Allwood 1976:10)

In fact this constraint could be a lot more general, because extraction of a subject next to a complementizer can never leave a gap in standard Swedish.[13] Consider for example (48).

(48) a. Jag tror [att Anna kommer hem från Paris nästa vecka].
I think that Anna comes home from Paris next week
'I think that Anna will be back from Paris next week.'

b. *Anna tror jag [att e_1 kommer hem från Paris nästa vecka].
Anna think I that comes home from Paris next week

It is possible to extract a subject from this position, but only if the trace is spelled out as a resumptive pronoun (Zaenen et al. 1981, Engdahl 1986). Extraction of subjects from relative clauses has not been investigated in detail, but Teleman et al. (1999 4:429) provide the two examples in (49), where a subject has been extracted from a relative clause, and there is a resumptive pronoun in the gap site.

(49) a. [Vilken läkare]$_1$ sa han till dig att han har träffat en [som hon$_1$ har
which doctor said he to you that he has met one that she has
botat]?
cured
'Which doctor did he tell you that he met someone she cured?'

[13] In *finlandssvenska*, the variety of Swedish spoken in Finland, this generalization does not hold (Engdahl 1986:124, Teleman et al. 1999 4:426).

b. [Vilken sonat]$_1$ sa du att du känner flöjtisten [som den$_1$ är skriven för]?
*which sonata said you that you know flutist.*DEF *that it is written for*
'Which sonata did you say that you know the flutist that it's written for?'

It is possible that the examples in (46) are best described as an instantiation of a more general **Comp-trace effect** (Chomsky & Lasnik 1977). In any case, on Allwood's view, there is a role for both syntactic and pragmatic constraints on ERC.

In a similar spirit as Allwood (1976, 1982), Andersson (1982) also argues for a diminished role for syntactic constraints on extraction, and proposes that general semantic and pragmatic restrictions explain why some extractions are acceptable and others are not. He points out counterexamples to some generalizations previously thought to hold. For example, he observes that it is possible to construct examples in which the head of the relative clause may be definite, and in which a head like *ingen* 'no one' is impossible, despite the fact that *ingen* as a head is very common in ERC-sentences otherwise (50). The example is adapted from Andersson (1982:41).

(50) [Den här tavlan]$_1$ känner jag [killen/en kille/*ingen [som har målat e_1]].
*the here painting.*DEF *know I guy.*DEF/*a guy/no one that has painted*
'I know the guy/a guy/*no one who painted this painting.'

The version of (50) with a definite head is more natural than the one with an indefinite head, and as already mentioned, *ingen* 'no one' seems entirely impossible. The point that Andersson wants to make is that there are restrictions on the form of the head, but that they are not syntactic in nature; they are semantic and pragmatic. For example, encyclopedic knowledge tells us that it is most common for a painting to have only one painter and this is why it is natural to use a definite head here.

Another factor which is discussed by Andersson (1982) is the discourse function of the constructions involved. The extraction sentences which he discusses all involve T-preposing or left dislocation, which he describes as "text-linguistically" motivated rules. What this means is that there has to be a contextual motivation for using the resulting extraction sentence, just like for using a sentence with local T-preposing and left dislocation. There are similarities between Andersson's points and the approach taken later by Goldberg (2006) which was briefly discussed above, in that he aims to get the acceptability of ERC to fall out from the pragmatic and semantic function of the different constructions which are involved in extraction sentences. More generally, Andersson (1982)

argues that it is important to make sure that the perceived unacceptability is not due to semantic or pragmatic factors. To do this, we often have to find a context where the sentence in question makes sense.

There are also a few studies that investigate long extraction in spontaneous language use. Jensen (2002) studies extraction in spoken Danish. Her study is restricted to extraction in declarative main clauses, i.e. what I call T-preposing, but she investigates extraction from all types of clauses. The empirical basis of the study is 18 hours of spoken Danish.[14] In these 18 hours of speech, Jensen found 230 sentences with extraction from an embedded clause. From the point of view of ERC, the study is particularly interesting in that it gives a hint about the frequency of that construction. In the 18 hours of spoken Danish, Jensen found ten sentences with extraction out of a relative clause or relative clause-like constituent. All of the spontaneously occurring examples from her study have *være* 'be' as the matrix predicate. Two of them are given in (51). The annotation of the examples, which are from Jensen (2002:107), is slightly adapted.

(51) a. og det$_1$ var der sgu nogen [der ikke forstod e_1] (Da.)
and that was there PRT *some who not understood*
'... and there were some who didn't understand that.'

b. det$_1$ er jeg stort set den eneste [der synes e_1 i andelsforeningen her]
*that am I almost the only who thinks in cooperative.*DEF *here*
'I am almost the only one who thinks so in the cooperative here.'

The T-preposed phrase in each of these examples is the pronoun *det* 'that', which Jensen reports was typical for the examples in her study.

An article by Engdahl (1997) is the only previous study of ERC based on a larger collection of naturally occurring examples. Engdahl argues that syntactic constraints like the Complex NP Constraint or Subjacency are not very useful in distinguishing between acceptable and unacceptable extractions in Swedish. The pattern that she finds is instead that many of the naturally occurring examples involve constructions where the relative clause is the information center of the utterance. This is in line with the approach taken by Erteschik-Shir (1973), and with Allwood's (1976, 1982) and Andersson's (1982) accounts, as well as Jensen's (2002).

Engdahl (1997) also studies the function of the fronted constituent in ERC, and finds that the extractions are mainly topicalizations (T-preposing in my

[14] 13.2 hours from the corpus BySoc, four hours from conversation, and one hour from a television-broadcast from the Danish election in 1998.

terminology), and that there are two different information-structural functions that the fronted phrase can have: a contrastive one and a cohesive one. Two pragmatic factors are singled out as especially important for the acceptability of ERC: that the fronting is motivated in the context, and that the information structure of the clause fits the information states of the participants.

Engdahl (1997:74) also presents a hypothesis as to why the mainland Scandinavian languages are different from Icelandic, Faroese, German, Dutch, English, and Romance languages when it comes to ERC. Since she finds no structural differences between the languages that would explain why only the mainland Scandinavian languages permit ERC, she proposes that we might find a more relevant difference in how the languages use fronting for discourse purposes. The mainland Scandinavian languages frequently employ topic fronting, both in the local clause and in long extraction, and they use both contrastive and cohesive fronting. English mostly uses contrastive fronting, and even though German and Dutch use cohesive fronting, this is not possible out of subordinate clauses. This might be an important difference. For Faroese and Icelandic, more research is needed.

A recent article by Löwenadler (2015) takes as its starting point pragmatic approaches to ERC, and compares ERC in English an Swedish. While he argues that a dominance- and processing-based approach is on the right track, he also notes that there are examples which are not predicted to be acceptable on such an approach, but which several speakers nevertheless accept. An example is given in (52).

(52) a. Ser du det stora huset bakom träden?
'Do you see the big house behind the trees?'

b. Ja, det huset avundas jag dom som bor i.
yes that house envy I those who live in
'Yes, I envy those who live in that house.' (= because they live in that house)

(Löwenadler 2015:44)

In an acceptability study, this example, where there is a causal relation between the emotion expressed by the matrix verb and the state or event expressed by the relative clause, was accepted by 11 out of 13 informants (Löwenadler 2015:61, footnote 8).

Löwenadler takes a Construction Grammar approach (Croft 2001, Verhagen 2009) to ERC, and proposes that there are three constructions that license ERC in Swedish: a Presentational complex NP extraction construction, a Focused relative complex NP extraction construction, and a Cause-related complex NP extraction construction. The first two of these are much in line with the dominance-based approach, in that the constructions specify that the

relative clause is the information center of the utterance. The last, cause-related construction licenses examples like (52). In such examples, the relative complex is not the information center of the utterance, and they are unexpected on pragmatic approaches such as Erteschik-Shir's (1973) and Goldberg's (2006).

On Löwenadler's view, constructions are conventionalized schemas with slots that are specified for what types of phrases can occur in them. The reason that ERC is judged as more acceptable by Scandinavian speakers on this view is that the conventionalized ERC constructions are "more deeply entrenched among the speakers" (Löwenadler 2015:59). Extraction appears to be more restricted in English simply because there are more subtypes of ERC conventionalized in Swedish. We return to Löwenadler's account in section 5.5.

2.2.1.2 Structural approaches

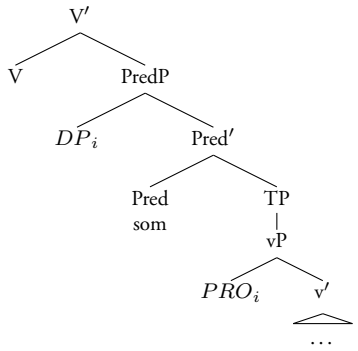
While Erteschik-Shir (1973) found no structural differences between relative clauses which permit extraction and those which do not, there are recent analyses that argue that structure matters both when it comes to which relative clauses permit extraction, and when it comes to accounting for cross-linguistic variation. Platzack (1999, 2000, 2014) takes such an approach, and proposes a structural explanation for why extraction is possible in the mainland Scandinavian languages but not in Icelandic, English, or German. The account takes as a starting point the idea that an escape hatch is needed for the extracted phrase in ERC to be able to move out of the relative clause successive-cyclically, thereby satisfying locality requirements. In relative clauses in languages like German, English, and Icelandic, no such escape hatch is available, because the specifier in the complementizer domain is needed to form the Ā-dependency in the relative clause.

According to Platzack, the reason that ERC is permitted in the mainland Scandinavian languages is that they have a way of forming subject relative clauses which does not involve Ā-movement to the complementizer domain (for a similar idea based on observations about English, see Chung & McCloskey 1983). This possibility is connected to the verb second (V2) property and to the fact that the relativizers in the mainland Scandinavian languages are complementizers and not relative pronouns. It is the combination of these two features that matters; having only one of them is not enough. This means that English does not have the same opportunities for extraction; even though it has a relative complementizer (*that*), it is not a V2 language. German is V2, but has relative pronouns and no relative complementizer. Icelandic would seem to be a counterexample, since it is V2 and has the relative complementizer *sem* 'that', but Platzack argues that *sem* is not a complementizer but a pronoun.

A prediction made by Platzack's account is that extraction from non-subject relative clauses should not be possible.[15] I discuss further aspects of Platzack's analysis of relative clauses and ERC in section 4.5.

Another structural approach to ERC in the mainland Scandinavian languages is the **Small Clause Hypothesis** put forth by Kush (2011) and Kush et al. (2013). On this approach, ERC is only illusory, and all cases which purportedly involve extraction from a relative clause should actually be analyzed as extraction from a small clause (cf. Chomsky's (1964) observations about (7) referred to in section 2.1). The proposal builds on the fact that the relative complementizer *som* 'that' is homophonous with a predicational *som*, similar to the English particle 'as', which has been argued to head small clauses (Eide & Åfarli 1999), and on the idea that the preference for certain matrix predicates in extraction is really syntactically encoded: only verbs that select small clauses allow extraction from what look like relative clauses. As an extraction sentence is parsed, there will be an amelioration effect if it is possible to reanalyze the relative clause as a small clause. On this account, the structure of the relative clause-like constituent in ERC should really look like (53).[16]

(53)



(Kush 2011:21)

Som is the head of a predicational phrase, and the rest of the clause is its complement. This complement is predicated of the DP, which is in the specifier of PredP. It should be noted that the small clause in this analysis differs from other small clauses in Swedish in that it is tensed.

Like in the proposal put forth by Platzack (1999), it is assumed that ex-

[15] In the later version of the analysis Platzack reports on an apparent counterexample to this prediction, and suggests that it might follow given the **Principle of Minimal Compliance** (Richards 1998, Platzack 2014).

[16] This structure is from Kush (2011), and seems to be the structure which Kush et al. (2013) assume as well, even though it is not explicitly given in their article.

traction is only possible from relative clauses formed on the subject. On the Small Clause Hypothesis, this follows because small clauses are subject-oriented (Kush et al. 2013:254, footnote 9), i.e. the DP in the specifier of PredP will correspond to the subject argument in a clause, and can never be interpreted as corresponding to a non-subject argument.

The analysis gets some support from an acceptability study, where extraction from the complement of a small clause-selecting verb received higher acceptability ratings than extraction from the complement of a verb that does not select a small clause (Kush 2011). A weakness of the study is that only three matrix predicates were tested, namely *vara* 'be', *se* 'see', and *träffa* 'meet'. This means that the differences in acceptability which Kush (2011) and Kush et al. (2013) interpret as evidence in favor of the Small Clause Hypothesis could have other explanations. Two studies, Christensen & Nyvad (2014) and Müller (2015), pursue this matter further.

Christensen & Nyvad (2014) investigated the acceptability of ERC with 16 matrix verbs in Danish. Unlike Kush (2011), they found no significant effect of the ability of the verb to appear with a small clause.[17] There is variation in acceptability between the verbs in the study, but Christensen & Nyvad conclude that it does not seem to stem from structural factors. In their study, verb frequency has a significant ameliorating effect on the acceptability of extraction sentences, together with a significant effect of trial, i.e. a repetition effect, both of which they argue are compatible with a processing explanation rather than a structural explanation.

Like in Swedish, Danish relative clauses can be introduced by *som*, but Danish diverges from Swedish in also allowing relative clauses introduced by *der* 'who', which cannot head a small clause. In Christensen & Nyvad's study, half of the experimental sentences included *der* instead of *som*. On the Small Clause Hypothesis, reanalysis of the relative clause as a small clause should only be possible with *som*, since *som*, but not *der*, is also a predicational head. The prediction is that ERC with *der* should be less acceptable than ERC with *som*. This is not the case, however. Christensen & Nyvad found no significant difference between extraction from *der*-relatives and extraction from *som*-relatives. Both this result and the fact that there is no significant amelioration for small clause verbs provide evidence against the Small Clause Hypothesis.

[17] The verbs were classified according to their ability to appear with a small clause based on a norming study in this experiment. It is not clear that all of the verbs that were classed as able to appear with a small clause would be considered small clause-selecting verbs by Kush (2011) and Kush et al. (2013).

Müller (2015) makes a similar point, reporting on an acceptability study of ERC in Swedish. She compares extraction from relative clauses embedded under three types of verbs, small clause-selecting verbs compatible with a small clause headed by *som*, small clause-selecting verbs incompatible with a small clause headed by *som*, and verbs which do not select small clauses. In the study, there is no significant difference in acceptability between these three groups of verbs. This means that the Small Clause Hypothesis is weakened in light of experimental studies in both Danish and Swedish.

Turning to extraction from *att*-clauses, recent research by Bentzen et al. (2007a) and Hrafnbjargarson et al. (2010) reveals an interaction between the structure of the subordinate clause and availability of extraction. In Swedish, main clauses have V2 structure, while subordinate clauses typically do not. However, in certain circumstances, it is possible to get embedded V2 clauses, as in (54) (Bentzen et al. 2007b).

(54) a. Han sa att Lisa **hade inte** läst boken.
he said that Lisa had not read book-the

b. Han sa att **den här boken** hade Lisa läst.
he said that this here book-the had Lisa read

(Bentzen et al. 2007b:95)

We can see that the subordinate clauses have V2 structure because the verb precedes negation in (54a), and because it precedes the subject in (54b). Interestingly, in Swedish, extraction of any phrase from embedded V2 clauses is impossible, as (55) illustrates.

(55) a. Den där boken sa han att Lisa **inte hade** läst.
the there book.DEF *said he that Lisa not had read*
'That book, he said that Lisa hadn't read.'

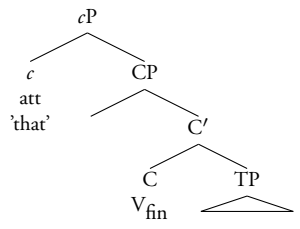
b. *Den där boken sa han att Lisa **hade inte** läst.
the there book.DEF *said he that Lisa had not read*

Example (55a) shows extraction from a subordinate clause with regular embedded word order, which is acceptable, whereas (55b) shows that extraction from an embedded V2 clause is blocked.[18]

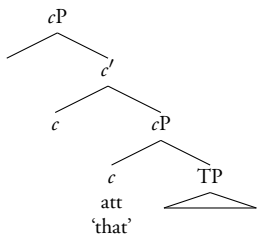
[18] There is some variation regarding extraction from V2 clauses in the Scandinavian languages. Bentzen et al. (2007a) and Hrafnbjargarson et al. (2010) investigate extraction from embedded V2 clauses in Danish, Faroese, Icelandic, Norwegian, and Swedish. Danish and Swedish disallow extraction from all V2 clauses. Extraction from V2 clauses with a fronted non-subject is ungrammatical in all of the languages, whereas Faroese and Icelandic allow extraction of

Danish is like Swedish in this respect; extraction from a subordinate clause with embedded V2 is impossible. A recent proposal by Vikner (2017) and Nyvad et al. (forthcoming) connects this fact to the observations about extraction from relative clauses reported above (Christensen & Nyvad 2014), as well as to findings about extraction from embedded questions (Christensen et al. 2013a,b). The proposal introduces a distinction between clauses with main clause word order, which are CPs, and have verb movement into the complementizer domain, as in (56), and clauses with subordinate clause word order, which are *c*Ps, and where there is no verb movement, as in (57).

(56)



(57)



Acceptable extraction is connected to the availability of an escape hatch which can occur only in clauses where there is no CP. I will introduce this model in much greater detail in section 4.1.4.

2.2.1.3 Other recent experimental studies

In addition to the acceptability studies related to the Small Clause Hypothesis, there are a few other recent experimental studies about extraction in the mainland Scandinavian languages. Tutunjian et al. (2015, under review) report on an on-line processing experiment in Swedish, investigating the status of extraction from restrictive relative clauses as compared to extraction from *att*-clauses (non-islands), non-restrictive relative clauses (which are assumed to be islands), and relative clauses where the verb in the relative clause was part of a pseudo-coordination structure.[19] Tutunjian et al. use eye-tracking to determine whether there are any differences in processing between the different types of extraction.

---

both arguments and adjuncts from subject-initial V2 clauses, and Norwegian allows extraction of arguments from subject-initial V2 clauses. It is suggested that the variation is connected to whether embedded V2 is a root phenomenon in the language.

[19] The experiment is part of the ongoing project *Universality and domain-specificity: processing relative clause extractions in Swedish* at Lund University.

An example of a test sentence from the restrictive relative clause condition is given in (58).

(58) Såna där gamla skottkärror såg jag en man som alltid tvättade [-] på
such old wheelbarrows saw I a man that always washed [-] at
bensinmacken när …
gas-station-the when

(Tutunjian et al. 2015)

Measures from two regions are reported: Region 1, which is the verb in the relative clause (*tvättade* in (58)), and Region 2, which is the PP following the gap site (*på bensinmacken* in (58)). The results show that ERC with restrictive relative clauses is processed more like extraction from *att*-clauses than like extraction from non-restrictive relative clauses in both Region 1 and Region 2. This is interpreted as non-island like behavior on the part of restrictive relative clauses. More generally, it seems as if it is easier to process extraction from restrictive relative clauses than extraction from non-restrictive relative clauses. The results also indicate that non-structural factors, like working memory span and pragmatic fit, affect processing. The results are taken as evidence, confirming off-line intuitions, that ERC is acceptable in Swedish.

Two recent experimental studies (Christensen et al. 2013a,b) about extraction from embedded questions in Danish are also relevant here. Christensen et al. (2013a) report on two acceptability judgement experiments which test two assumptions from the syntactic literature: that a syntactic constraint blocks extraction of *wh*-phrases from embedded questions, i.e that embedded questions are syntactic islands, and that there is an adjunct/argument asymmetry in extraction from embedded questions. Christensen et al. (2013a) find that extraction from an embedded question is degraded in comparison to extraction from an *at*-clause (the Danish equivalent of a Swedish *att*-clause, i.e. a non-island *that*-clause), which is in turn reduced in acceptability compared to local *wh*-movement. They find no evidence for an adjunct/argument asymmetry in extraction from embedded questions, but interestingly, the studies show that there are training effects, which Christensen et al. (2013a) take as evidence that extraction from embedded questions is grammatical, even though it is degraded. The reasoning relies on Sprouse (2007), who argues that training effects are only possible with grammatical strings, since the grammar does not generate full syntactic representations for ungrammatical strings.

In the second of the two studies, Christensen et al. (2013b) report on an fMRI study where the same stimulus materials as in Christensen et al. (2013a) were investigated. Here the objective was to investigate potential neural correlates of the differences in acceptability between different kinds of extraction

that were found in Christensen et al. (2013a). The study investigated activation of the left inferior frontal gyrus in reading sentences with local fronting, extraction from an *at*-clause and extraction from an embedded question, and also included an acceptability judgement task. With respect to the acceptability judgements, the judgement task replicated the findings from Christensen et al. (2013a): local fronting is most acceptable, and extraction from an *at*-clause is more acceptable than extraction from an embedded question. The activation of the left inferior frontal gyrus was predicted to correlate negatively with acceptability, but this prediction was not borne out. Instead, extraction from an *at*-clause and extraction from an embedded question were not significantly different with respect to the imaging data. The activation of the left inferior frontal gyrus thus correlated with crossing a clause boundary, and not with decreased acceptability.

2.3 Research questions revisited

In chapter 1, I introduced four overarching questions, which I am now in a position to make more specific. The first question concerns the nature of the dependency between the extracted phrase and the empty position inside the relative clause in ERC-sentences. Given the properties of Ā-movement dependencies which were introduced in section 2.1.2, we may ask whether the extraction dependency in ERC is in fact an Ā-movement dependency, as opposed to base generation with binding of *pro* in the gap site. I address this question in chapter 4.

The second question concerns the structure of the relative clause. As we have seen, there are several proposals for what structure to assign to the relative clause in ERC-sentences. Specifically, we want to find out to what extent relative clauses in ERC are like other restrictive relative clauses in Swedish. For example, we want to know whether there is extraction from attributive relative clauses, or whether there is evidence that they instead behave as structures predicated of the head DP, as in a small clause structure. We also want to know where the relative clause attaches in the DP; and whether it is necessary to form an Ā-dependency to derive the relative clause. The structure of the relative clause is discussed in sections 4.3 and 4.5.

The third question has to do with the role of discourse and pragmatic constraints in ERC. As we have seen, much of the previous research puts an emphasis on pragmatic and semantic factors in constraining extraction, and some accounts aim to replace syntactic constraints with pragmatic and/or semantic constraints entirely. Given how the problem has been approached in

previous research, it can be broken down into two parts: the function of the fronted phrase in ERC, and the information impact of the relative complex. With respect to the fronted phrase, we want to compare fronting in ERC to local fronting, and fronting from other types of embedded constituents. Specifically, we are interested in finding out whether there are stronger restrictions on fronting out of relative clauses than on fronting in general, or whether they are the same. As for the information impact of the relative complex, we have seen that several of the accounts which argue for functional/pragmatic constraints rather than syntactic ones in constraining extraction, build on the idea that a domain has to be prominent in some sense if extraction is to be licit. We may ask then, if such theories find support from Swedish ERC. These questions are addressed in chapter 5, and the broader question about the role of discourse and pragmatics in ERC is also considered in chapter 7 against the backdrop of the findings from the entire dissertation.

The fourth question concerns how ERC is constrained. While the previous literature has established that extraction is freer in Swedish and the other mainland Scandinavian languages than in English, there are nevertheless extractions that are impossible. An example was given in (5) in chapter 1, which shows that extraction of *varför* 'why' from a relative clause is not acceptable. To my knowledge, this type of example has not been noticed in the previous literature on ERC, and I discuss it in chapter 6, where I investigate adjunct and argument extraction in both T-preposing and *wh*-questions in light of previous research on domains that are considered to be weak islands. The investigation leads into a discussion about how constraints on ERC are best captured: in the syntax, in the semantics, or in the pragmatics. In chapter 7, I approach the question from a broader perspective, widening the discussion to other types of constraints on Ā-movement.

Before moving on to answering these questions, I give an overview of the data used in the investigation in chapter 3.

3. Data and methodology

3.1 A collection of naturally occurring examples

The main material that the thesis is based on is a collection of 270 spontaneously produced examples which were gathered between 2011 and 2016. The collection contains ERC-sentences from both spoken and written Swedish, where ERC-sentence is taken to mean a sentence with some type of RC-like constituent which contains a gap or a resumptive pronoun bound by a phrase outside the constituent.

The examples from spoken language are from radio and television and from everyday conversation. The examples from written language are from fiction, social media, and newspapers. The spoken language part of the collection consists of 161 examples and the written language part of 109 examples.

Collecting spoken data has been my main objective, since such data have not been investigated systematically before, with the exception of Engdahl (1997). I have collected all spoken examples that I have heard during the collection period, except if the circumstances of the conversation did not permit me to write the example down. With respect to the written data, certain types of extraction sentences are quite common as I show in Lindahl (2010), where I present a corpus study of ERC in existential and presentational sentences. For the current investigation, I have not collected every written example I have noticed but I have tried to include examples of different types.

For spontaneously produced examples information about the circumstances is given directly after the example. Examples that lack a source are constructed by me.

3.1.1 Radio and television

The data from radio and television consist of 60 examples with sound files. These were recorded from radio, television, and online news media. The collection method was spontaneous listening. Sound files were obtained from online sources like the Swedish public radio, *Radio Sweden*, and Swedish national public television, *SVT*, which keep content available online. The extraction sentences and their contexts were transcribed.

When I discuss examples from this part of the collection, they are marked with a source in parentheses, giving the medium from which the example was recorded, and the date it was broadcast.

3.1.2 Conversation

The conversation part of the collection consists of 101 examples. A majority are from informal conversations, for example around the dinner or coffee table, but there are also a few examples from more formal settings, such as seminars and talks. Most of the examples are from conversations in which I was a participant, but none of the examples was produced by me. The extraction sentence was written down, and in most cases I have also written down a verbatim context sentence or a short summary of the linguistic context and the conversation topic. Examples from this part of the collection are marked with *Conversation* and information about the date of collection.

3.1.3 Examples from written language

The example collection also contains 109 examples from written language. These are from several text types, including newspapers, social media and fiction. I came across some of these examples in everyday life, but several of the examples were collected as an effort to find rare types of ERC, for example extractions involving a certain embedding predicate or type of extracted phrase. Collection methods included searching corpora in the Swedish language bank (Språkbanken) (Borin et al. 2012) and Google.

In chapter 6, I investigate whether adjuncts are ever extracted from relative clauses, and present several corpus searches. The search strings used can be found in Appendix A, and are described in more detail where the results of the searches are reported, in chapter 6, section 6.1.1.1.

For corpus examples, I give the name of the corpus in parentheses. Other examples from the written part of the collection are marked with information about source and publication date.

3.1.4 Two samples

Since the collection of ERC-sentences is heterogenous and the written data was partially collected by searching for particular types of extractions, it is not a good basis for generalizations about what is common and rare. In order to give an idea of how common certain features are in ERC-sentences, I created a sample of 100 examples from the spoken part of the collection, which I will refer to as **Sample A**. Half of the examples are from radio and television, and half from conversation. This sample is what I refer to whenever I say something about what is common and rare in ERC in general, e.g. in chapter 4.

It turns out that in all of the sentences in the example collection, the syntactic dependency that creates the long extraction from the relative clause is either T-preposing or relativization, and that T-preposing is most common (93 out of the 100 examples in Sample A). In section 5.2, I investigate the function of T-preposed phrases in ERC. As a basis for that investigation, I created a second sample, **Sample B**, with only T-preposing sentences. This sample was constructed on the basis of Sample A. In constructing the sample, I started with the 100 examples in Sample A, took out the 7 examples where the ERC was not created by T-preposing, and added 7 new examples with T-preposing instead.

The samples were created in 2015, and inclusion was based on order of collection. I use the samples in particular when I describe the distribution of certain features in the data. I sometimes refer to examples that are not included in the samples in the discussion of specific points.

The motivation for creating the samples was to be able to report some rough measure of how common certain features are in ERC-sentences, but the generalizability of results based on the samples should not be overestimated. There are several possible factors that might bias the results. For example, they are highly dependent on the types of conversations and situations I have been in during the collection period, and on what I have heard on the radio. Furthermore, what one hears and perceives might be affected by what one is listening for, or expects to hear, and certain types of ERC might be more noticeable than others. The patterns of usage should therefore only be taken as indications of what is common.[1]

[1] The data can be found at https://svn.spraakbanken.gu.se/sb-arkiv/pub/lindahl/2017/.

3.2 Additional investigations and sources

In several cases, additional data has been collected in order to answer particular questions and to complete the picture where naturally occurring examples are not available. To this end, I used a questionnaire and interviews. Data from the questionnaire are reported in chapter 4 and in chapter 6. The results of the interview are reported in section 6.2.2.

3.2.1 Questionnaire

The questionnaire was distributed electronically and consisted of 32 items. The purpose was exploratory, trying to identify factors that affect the acceptability of extraction. The questionnaire items thus include sentences both with and without extraction from relative clauses, extraction of different adjunct types, extraction from relative clauses with definite and indefinite relative complexes, and extraction out of the complements of different embedding verbs. An example item is given in (1).

(1) a. Linda: Min syster och jag är oense om när vi ska äta ikväll. Hon
Linda my sister and I disagree about when we shall eat tonight she
tycker att vi ska äta klockan nio, men jag tycker att det är alldeles
thinks that we should eat clock-DEF *nine but I think that that is* PRT
för sent.
too late
'Linda: My sister and I disagree about when to have dinner tonight. She thinks we should eat at nine, but I think that is way to late.'

b. Ida: Ja, så sent vet jag ingen som brukar äta middag.
Ida: yes that late know I no one that tends eat dinner
'Ida: Yes, I don't know about anyone who eats dinner that late.'

Each item consisted of a context sentence, here (1a), and a test sentence, here (1b). The context sentence and the test sentence were shown on the screen together, and the task for the participants was to rate the naturalness of the test sentence as a follow-up to the context sentence. Participants got to choose from three alternatives: *naturlig* 'natural', *lite konstig* 'somewhat strange', and *onaturlig* 'unnatural'. The participants were also given the possibility of adding free text comments for each test item. An underlying assumption with this questionnaire design is that ungrammatical or unacceptable test sentences would not be rated as natural.

There were four practice sentences, some of which were supposed to be

ungrammatical or unnatural in relation to the context sentence. All of the test participants saw the items in the same order. The questionnaire was answered by 16 native Swedish speakers between 24 and 79 years old, all of whom had finished high school at the time of participation. In all but one case they had also finished at least some courses at a university level. The questionnaire and an overview of the results are given in Appendix B.

The examples in the questionnaire were designed to be partially parallel in order to allow for certain comparisons, but since several different variables were explored at once, it was not possible to set up a fully controlled experimental design. For that reason, it is not always possible to connect differences between items to single variables, and I cannot report fine-grained quantitative differences between items. Instead I report only the number of participants who picked each of the given alternative ratings for each item.

3.2.2 Interview about *wh*-questions with extractions

Chapter 6 explores several issues related to the status of relative clauses as a type of weak island. Much of the previous literature engages mostly with extraction in the form of *wh*-question formation. To get comparable data from Swedish ERC, I conducted four interviews. The interview participants were graduate students of Scandinavian languages at University of Gothenburg, and all native speakers of Swedish. In the interview, the participants were asked to answer questions containing extractions given with or without context, and to describe how they went about coming up with an answer. If they did not come up with an answer for a certain question, they were given proposals for answers and were asked to judge how well they worked as answers to the question, if at all (see section 6.2.2). The interview schema can be found in Appendix C.

4. The syntax of ERC

In previous research, syntactic aspects of ERC have attracted a lot of interest and in this chapter, I discuss various issues connected to the syntax of ERC-sentences, relating examples from my collection of naturally occurring extraction sentences to proposals about the syntax of ERC and of Swedish relative clauses.

In the first part of the chapter, section 4.1, I present the view of Swedish clause structure and the left periphery of the clause which I assume in the dissertation. The aim is to introduce a simple model of main and embedded clauses, and to relate to it the preposing phenomena that are of interest here: T-preposing, left-dislocation, relativization, cleft formation, and *wh*-question formation. The overview of the clause structure that I give here will serve as background for the rest of the chapter, as well as to chapters 5 and 6.

The remainder of the chapter investigates several syntactic questions about ERC. An important contribution of the chapter is to show that the ERC-sentences studied in the thesis really involve Ā-movement out of relative clauses. This is done in two steps. First, I show evidence that there is Ā-movement out of the embedded relative clause-like constituent (section 4.2). I then argue that at least some of the relative clause-like constituents that allow such movement out of them in Swedish must be analyzed as regular restrictive relative clauses attached inside DP (section 4.3). In the remaining sections I report on the definiteness of the DP in ERC-sentences (section 4.4), discuss some outstanding issues about the structure of Swedish relative clauses (section 4.5), and examine the possible positions of the gap left by ERC (section 4.6). Section 4.7 sums up the findings of the chapter.

4.1 Swedish clause structure

This section outlines the view of the Swedish clause structure that I assume in the dissertation. I first introduce my assumptions about the structure of main and embedded clauses, and then turn to the left periphery, and to how I analyze preposing phenomena in Swedish. The discussion of the preposing phenomena is related to a recent analysis of the left periphery of main and embedded clauses by Vikner (2017) and Nyvad et al. (forthcoming).

4.1.1 Main clauses and V2

Like the other Scandinavian languages, Swedish is a verb second (V2) language, which means that the finite verb is usually in the second position in the clause. In declarative main clauses, basically any phrasal category can precede the finite verb.[1] We see some examples that illustrate V2 structure in (1).

(1) a. Anna ska nog resa till Island.
Anna will probably travel to Iceland
'Anna will probably go to Iceland.'

b. I vår ska Anna nog resa till Island.
in spring will Anna probably travel to Iceland
'Anna will probably go to Iceland this spring.'

c. Den där smörgåsen köpte jag på väg till jobbet.
the there sandwich bought I on way to work
'I bought that sandwich in my way to work.'

d. Hungrig var han inte.
hungry was he not
'He wasn't HUNGRY.'

The constituents that precede the finite verb in these examples are, respectively, a subject (1a), an adverbial phrase (1b), a direct object (1c), and a predicate complement (1d).[2] Whenever the subject is not in the preverbal position it must appear after the finite verb.

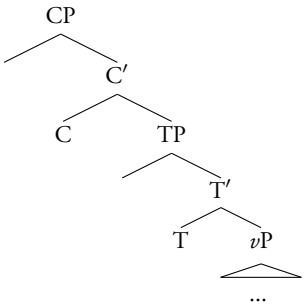
---

[1] In the Scandinavian tradition the initial position is referred to as *fundamentet* 'the foundation', following Diderichsen (1946). This positions corresponds more or less to Spec-CP.

[2] This set of examples is not supposed to give an exhaustive account ot the possibilities. Phrases of most phrasal categories and grammatical functions can appear in this position. Notable exceptions are, however, the adverbs *ju* 'indeed', and *väl* 'probably'.

Throughout the thesis, I adopt a view of Swedish clause structure which assumes that the clause consists of three domains, a verbal domain (*v*P), an inflectional domain (TP), and a complementizer domain (CP), as illustrated by (2).

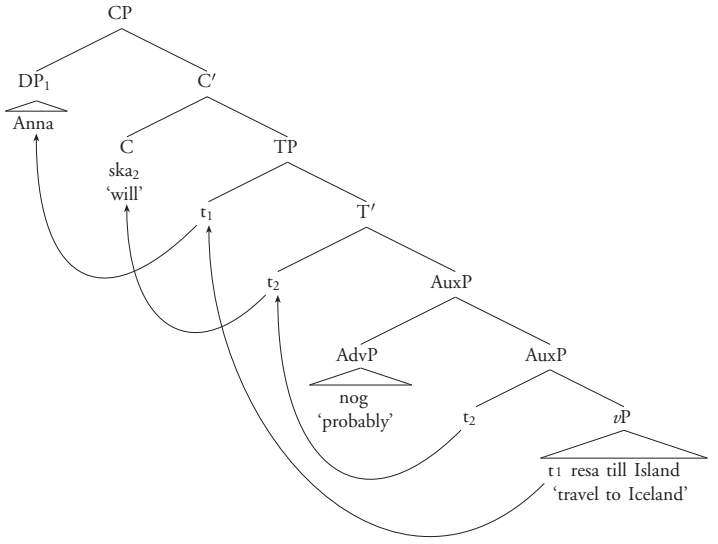
(2)



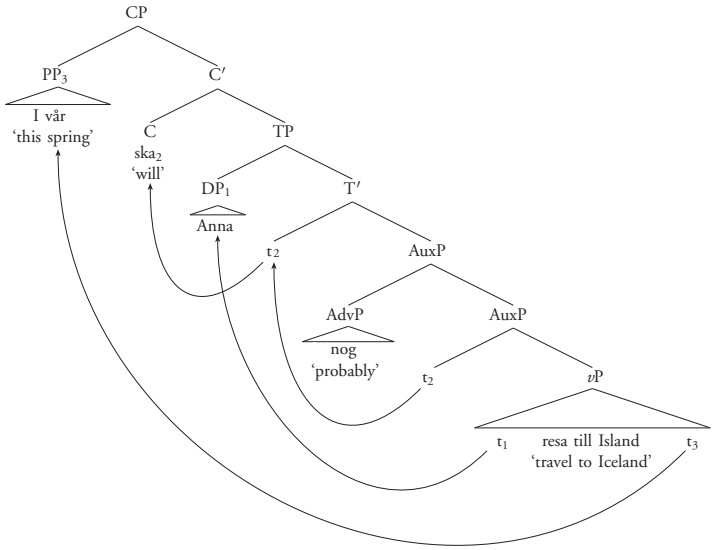
This partitioning of the clause into three domains is assumed to be universal within the Minimalist program (Chomsky 1995), and Platzack (1998) presents an extensive analysis of Swedish which adopts this view of the clause.

I furthermore adopt an analysis of V2 in Swedish developed by Platzack (1986, 1998) and others where the finite verb in main clauses appears in C, a position it reaches by way of head movement from the verbal domain. Given this view of the clause, the sentence in (1a) is represented as in (3a), and (1b) is assumed to have the structure in (3b).

(3) a.



b.



Numbered traces illustrate movement relations in these diagrams.[3] The finite verb undergoes head movement to C in both (3a) and (3b). In this particular example, the finite verb is an auxiliary[4], but if there is no auxiliary, the main verb will move to C. The subject *Anna* is first merged in Spec-*v*P. In (3a), the subject moves to Spec-CP via Spec-TP. In (3b), Spec-CP is occupied by an adverbial PP, and the subject remains in Spec-TP. This analysis, where the finite verb occupies C in both subject initial and non-subject initial clauses, is known as the symmetric V2 analysis and goes back to den Besten's (1983) analysis of V2 in German, Dutch, and Swedish.[5]

The pre-verbal position in main clauses corresponds to Spec-CP in this analysis. Even though subjects are the most common occupants of this position in declarative clauses,[6] it can be argued that Spec-CP is not really a subject position (see e.g. Svenonius 2002:215ff). It is often assumed that material in the pre-verbal position in declarative main clauses corresponds to old or

[3] I use the trace notation in order to save space, and stay uncommitted with respect to whether head movement is substitution or adjunction.

[4] I follow Platzack (2011) in assuming auxiliaries are introduced as the heads of AuxPs between TP and *v*P. Nothing in my analysis hinges on the label AuxP.

[5] There is a comprehensive discussion about whether V2 is symmetric or asymmetric. For asymmetric accounts see e.g. Travis (1991) and Mikkelsen (2015), who also provides a useful overview of the discussion up until now.

[6] Jörgensen (1976) found that the subject is in the pre-verbal position in 60–70 % of declarative main clauses in Swedish in a variety of spoken genres.

backgrounded information, the theme of the sentence in the sense of the Prague school, or a topic of some kind or another. I will return to the function of the phrase that precedes the finite verb in much more detail in chapter 5. For now I will use **T-preposing** as a cover term for movement to Spec-CP in declarative main clauses in Swedish. This way we need not commit to the exact information structural function of phrases in Spec-CP until we investigate the phenomenon in chapter 5.

There have been several proposals arguing in favor of splitting up the three domains of the clause structure in (2) into more finely layered structures. For example, it has been proposed that the inflectional domain be split up into a Tense Phrase, one or several Agreement Phrases, a Negation Phrase, and various others (e.g. Pollock 1989). I will not follow these proposals, and will assume instead that sentence adverbials like *nog* 'probably' in (3a) and (3b) are phrasal adjuncts, and that they can can either precede TP or, as in the example at hand, be adjoined between TP and AuxP, preceding the first merge position of any auxiliaries (for discussion of such an analysis of adverbials in the Scandinavian languages, see Svenonius 2002).[7] I will not discuss any additional functional projections in the inflectional domain, since they do not seem to affect the analysis of extraction from relative clauses in any obvious way. Of greater relevance to the investigation at hand is the idea that the complementizer domain is not atomic, a proposal which I will return to in section 4.1.4 and section 4.1.6.

4.1.2 Subordinate clauses

Turning to Swedish subordinate clauses, these generally do not exhibit V2, as illustrated in (4) for a selection of different subordinate clause types. Here the complementizer is directly followed by the subject, and both auxiliary verbs and main verbs appear to the right of sentential adverbs.

(4) a. Emma sa att Anna nog ska resa till Island.
Emma said that Anna probably will travel to Iceland
'Emma said that Anna probably will go to Iceland.'

b. Anna var glad över förslaget eftersom hon nog ska resa till Island.
*Anna was happy over proposal.*DEF *since she probably will travel to Iceland*
'Anna was happy about the proposal, since she will probably go to Iceland.'

[7] For an overview of what governs the ordering of adverbials in Swedish, see Andréasson (2007).

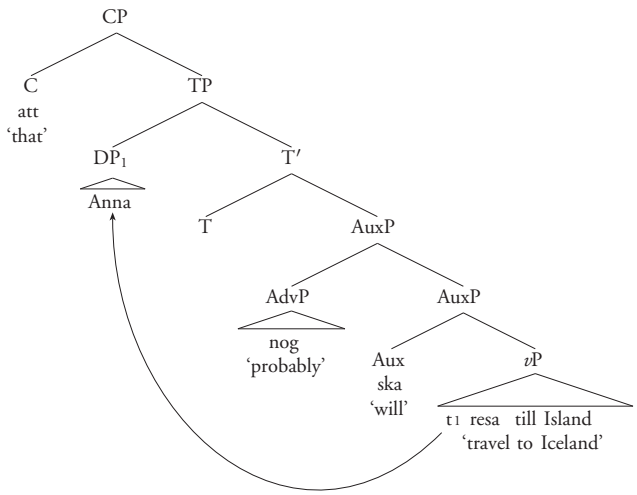
c. Jag undrar om Anna verkligen ska resa till Island.
I wonder if Anna really will travel to Iceland
'I wonder if Anna will really go to Iceland.'

d. Det där är min vän Anna, som nog ska resa till Island.
the there is my friend Anna that probably will travel to Iceland
'That is my friend Anna, who probably will go to Iceland.'

The complementizers *att* 'that', *eftersom* 'since', *om* 'if', and *som* 'that' are heads in the complementizer domain, and are thus in the same position as the one occupied by the finite verb in a main clause, i.e. C. The finite verb stays lower in the clause. Based on its relation to adverbs in the inflectional domain, we see that it stays very low, remaining inside the verbal domain, or in AuxP in the case of auxiliaries. This is the situation in all of the mainland Scandinavian languages, whereas in Icelandic, the finite verb moves into the inflectional domain preceding negation and other sentential adverbs.[8]

The subordinate clause structure in Swedish is illustrated in (5), which shows the subordinate *att*-clause in (4a).

(5)



In the next section, we turn to an exception to this structure.

[8] Modern Faroese exhibits variation, and has undergone a change from a system with verb movement into the inflectional domain towards one where verb movement to the inflectional domain in non-V2 clauses is not available to most speakers (Thráinsson 2007:58ff,Bentzen et al. 2009).

4.1.3 Embedded V2

While the general pattern for subordinate clauses is the one described in the previous section, there are well known exceptions. Specifically, subordinate clauses can exhibit V2 structure in certain environments. In particular, embedded V2 is possible in the complements of some verbs but not others. Building on the verb classification of Hooper & Thompson (1973), Andersson (1975) shows that embedded V2 is compatible with strongly assertive verbs, like *säga* 'say', *påstå* 'claim'; weakly assertive verbs *tro* 'believe', and *tycka* 'think'; and semi-factive verbs *upptäcka* 'discover'.[9] An illustration of embedded V2 with the verb *säga* 'say' is given in (6), which is from Bentzen et al. (2007b).

(6) a. Han sa att Lisa **hade inte** läst boken.
he said that Lisa had not read book-the

b. Han sa att **den här boken** hade Lisa läst.
he said that this here book-the had Lisa read

(Bentzen et al. 2007b:95)

The examples show two different cases of embedded V2, neither of which is compatible with the structure of subordinate clauses described in section 4.1.2. In (6a), the finite verb precedes negation, exhibiting the typical ordering of those elements in main clauses. In (6b), the direct object precedes the finite verb, which is then directly followed by the subject, just like in an ordinary object-initial V2 main clause.

It is interesting to note that extraction is entirely blocked from embedded V2 clauses in Swedish (e.g. Bentzen et al. 2007a), while non V2 *att*-clauses in general permits extraction, as we have already seen in chapter 2. The contrast in acceptability between (7a) and (7b) is very clear.

(7) a. *Den där boken sa han att Lisa **hade inte** läst.
*the there book.*DEF *said he that Lisa had not read*

b. Den där boken sa han att Lisa **inte hade** läst.
*the there book.*DEF *said he that Lisa not had read*
'That book, he said that Lisa hadn't read.'

Given the view of Mainland Scandinavian V2 outlined above, it is quite natural to assume that the verb has moved into the complementizer domain in the subordinate clause in examples like the ones in (6), and that the complemen-

[9] For more recent discussion of embedded V2 in the Scandinavian languages, see e.g. Bentzen et al. 2007b, Wiklund et al. 2009, and Julien 2015.

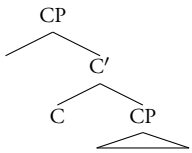
tizer domain has a more finely articulated structure, involving either several functional projections, like in the cartographic approach (e.g. Rizzi 1997), or recursion of CP. In the following section, I will turn to a particular analysis of the complementizer domain in Scandinavian languages that can be seen as a variety of the CP-recursion approach. The cartographic approach will be briefly discussed in section 4.1.6.

4.1.4 A *c*P/CP-analysis of the left periphery

Vikner (2017) and Nyvad et al. (forthcoming) propose an account of the left periphery of clauses tying together observations from research on embedded V2 and extraction in the Scandinavian languages. The model connects both of these phenomena to recursion in the C-domain.

In CP-recursion, a C-head takes a CP as complement, creating a new CP, as in (8).

(8)



CP-recursion analyses have previously been proposed for the analysis of embedded V2 by several researchers, including de Haan & Weerman (1986)[10] and Vikner (1995).

Observing that it is not possible to extract a phrase from a subordinate clause with embedded V2 word order, Vikner (2017) and Nyvad et al. (forthcoming) propose that there are two types of CPs: CP and *c*P.[11] The defining characteristic of CP is that there is movement of a finite verb to its head. There is no such verb movement in *c*P, which is headed by a functional head. CP-recursion is also of two basic kinds: recursion where *c* selects a CP, which is the kind of recursion needed to derive embedded V2 sentences like the ones in (6), and recursion where *c* selects a *c*P, which is involved in extraction from subordinate clauses.[12] Below I will describe in detail how different sentence types are analyzed in this

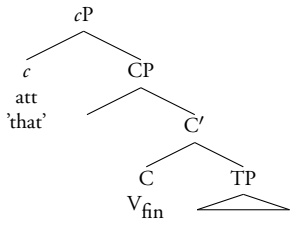
[10] In the terminology assumed at the time this was recursion of S′, which provided an empty comp-node that the finite verb moved to.

[11] The pronunciation of *c*P is "little *c*P".

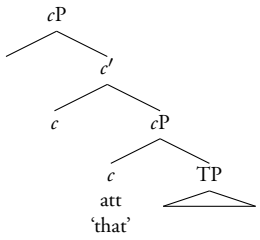
[12] I will use CP-recursion as a cover term for both of these types of recursion when it does not matter, and otherwise specify whether a big CP is involved or not.

framework. Schematically, the two types of CP-recursion are represented in (9) and (10).

(9)



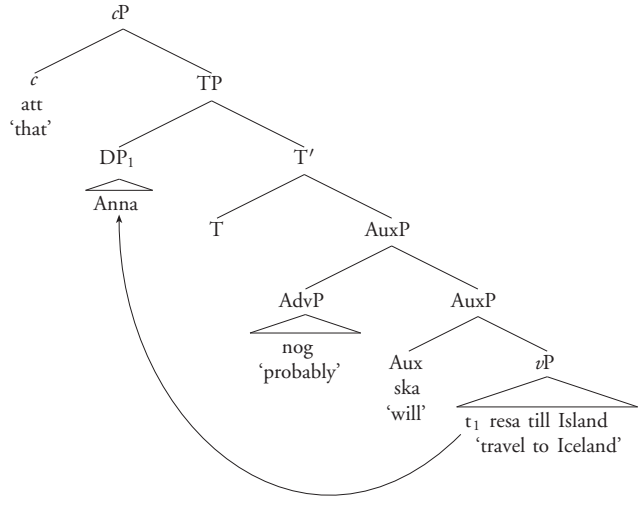
(10)



In (9), the complementizer *att* 'that' selects a big CP, where the finite verb is in C. This is the embedded V2 structure. In (10), a *c*-head selects a *c*P, i.e. a regular subordinate clause with no verb movement into its complementizer domain. As we will see below, the higher of these *c*-heads has a specifier which provides an escape hatch for extraction.

As can be seen in (9) and (10) main clause word order and embedded V2 is connected to big CP, with verb movement to C, and regular subordinate clause word order is connected to the absence of a big CP. The subordinate clause in (4a) will be headed by *c*, and have the structure in (11).

(11)

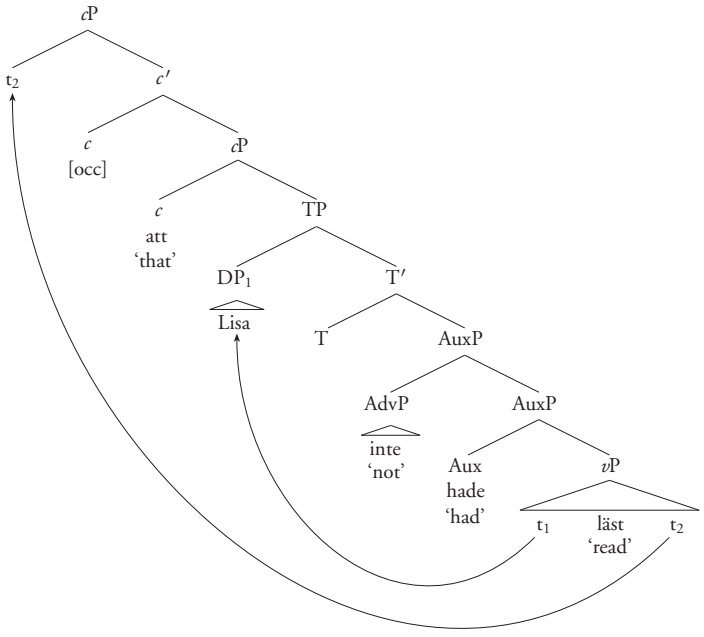


In sentences with extraction from a subordinate clause, like in (7b), movement through the complementizer domain is made possible by recursion of *c*, which provides an additional *c*-head which carries an **occurrence feature** (Chomsky

2005).[13] This feature causes movement to an extra specifier without requiring feature-matching, but prohibits phonological realization of the position created by the movement operation. This means that the phrase that is moved to the extra *c* specifier is prevented from staying there. Eventually it has to move higher in the structure.

A subordinate clause with extraction as in (7b) will have the structure in (12).

(12)



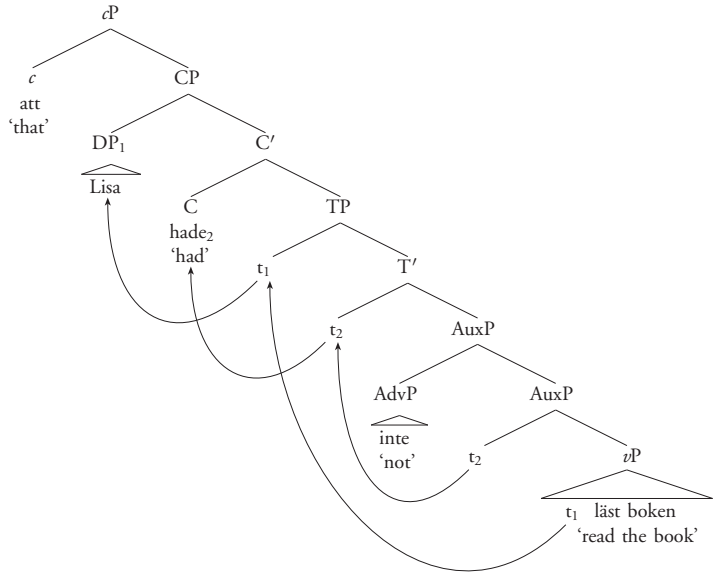
There is a trace, t_2, in the specifier of *c*[occ], left by the extracted DP *den där boken* 'that book', which eventually ends up in the Spec-CP of the highest clause.

Assuming the *c*P/CP-analysis for Swedish embedded V2, the subordinate clauses in (6a) and (6b) will have the structures in (13a) and (13b).[14]

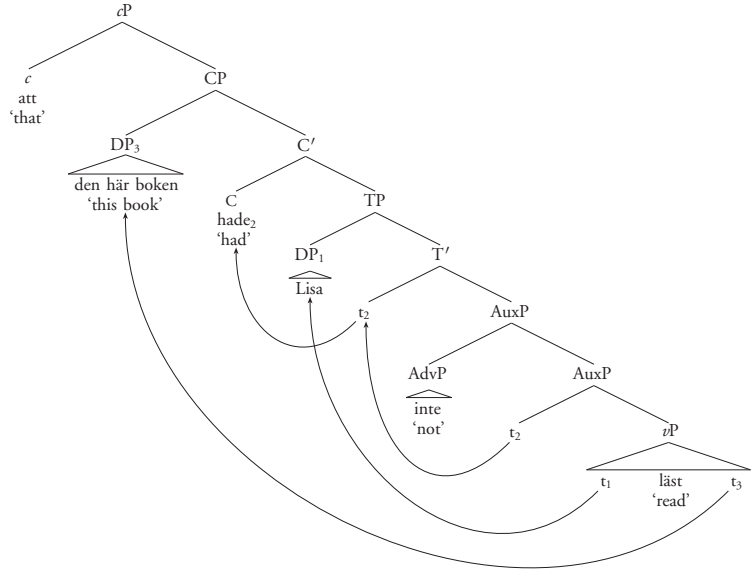
[13] The occurrence feature is an edge feature. Such features have also been called P(eripheral) features Chomsky (2000), and generalized EPP-features Chomsky (2001).

[14] I have included the negation *inte* 'not' in (13b) in order to show where it appears in such structures even though there is no negation in (6b).

(13) a.



b.



The complementizer *att* 'that' heads a *c* projection and takes a CP as its complement. This CP has the same structure as a declarative CP in a main clause. Just like in a main clause CP, the finite verb moves into the complementizer domain, to C, and there is movement of one phrase to the pre-finite position. In (13a), the subject has moved to the pre-finite position, and in (13b) the direct object has moved there, resulting in inversion between the finite verb and the subject.

As mentioned, a motivation for the distinction between *c*P and CP is the

observation that extraction and embedded V2 cannot co-occur. As we saw in (7a), extraction from an embedded V2 clause is strongly unacceptable. Vikner (2017) proposes that this follows if CP constitutes a phase on its own, and if there is no C with an occurrence-feature. Looking back to (7a) and (7b), here repeated as (14a) and (14b), we see that the difference between them is precisely that (7a), where extraction is blocked, contains a CP, while (7b), as we have previously seen, contains two *c*-heads and permits extraction.

(14) a. *[Den där boken]$_1$ sa han [$_{c\text{P}}$ att [$_{\text{CP}}$ Lisa$_2$ hade$_3$ [$_{\text{TP}}$ t$_2$ t_3 inte [$_{v\text{P}}$ t_2 läst t_1]]]]
the there book said he that Lisa had not read

b. [Den där boken]$_1$ sa han [$_{c\text{P}}$ t$_1$ [$_{c\text{P}}$ att [$_{\text{TP}}$ Lisa$_2$ inte hade [$_{v\text{P}}$ t_2 läst t_1]]]]
the there book said he that Lisa not had read

'That book, he said that Lisa hadn't read.'

As soon as the structure contains a CP, the subordinate clause will be an island since C does not have an occurrence-feature. Furthermore, the system does not permit recursion where a big C selects a CP, a restriction which corresponds to the observation that there are no clauses with two finite verbs.[15] The only kinds of CP-recursion allowed are the two types in (9) and (10).[16]

Further motivation for the CP/*c*P-distinction and the existence of a *c* with an occurrence feature comes from facts about extraction from embedded questions

[15] Sentences with VP fronting, like (i) are an exception.

(i) Han sa att han skulle läsa boken, och **läste** boken **gjorde** han.
*he said that he would read book.*DEF *and read book did he*
'He said that he would read the book, and read the book he did.'

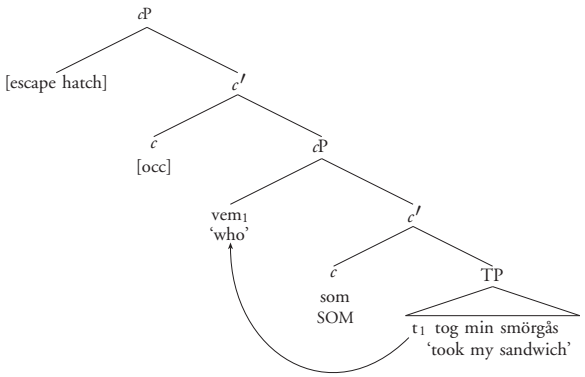
In such cases, there is still only one C-position, however. I thank Nick Kalivoda for pointing this out.

[16] As briefly mentioned in section 2.2.1.2 Bentzen et al. (2007a) Hrafnbjargarson et al. (2010) investigate extraction from embedded V2 in the mainland Scandinavian languages, as well as Icelandic and Faroese, and connect the possibility of extraction from V2 environments (or lack thereof) to whether embedded V2 is a root phenomenon in the language. This indicates that it may not be the position of the verb per se that makes CP a strong phase, but rather features connected to illocutionary force, which also happen to trigger verb movement in the mainland Scandinavian languages. Seen from this perspective, we could perhaps argue that big CP only occurs in root environments, and languages that have embedded V2 in non-root environments have another way to derive embedded V2, that does not involve big CP.

and relative clauses. Based on experimental research on extraction from several clause types, Christensen et al. (2013a) and Christensen & Nyvad (2014) argue that extraction from embedded questions and relative clauses in Danish is best analyzed as involving movement through an extra specifier in the C-domain. Vikner (2017) and Nyvad et al. (forthcoming) take this extra specifier to be just the one proposed to occur due to recursion of *c*P. Adopting the same analysis for Swedish, the embedded question in (15a) will have the structure in (15b), and the relative clause in (16a) will have the structure in (16b).

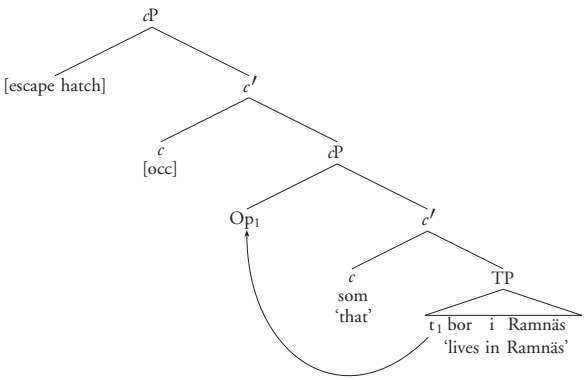
(15) a. Jag undrar vem som tog min smörgås.
I wonder who SOM *took my sandwich*
'I wonder who took my sandwich.'

b.



(16) a. Jag har en släkting som bor i Ramnäs.
I have a relative that lives in Ramnäs
'I have a relative who lives in Ramnäs.'

b.



Both embedded questions and relative clauses involve movement of a phrase

to the specifier of the lower *c*P on this account. In the embedded question in (15b), the *wh*-phrase *vem* 'who' moves there. The lower *c* head is spelled out as *som* here, which is obligatory in subordinate questions where the *wh*-phrase is the subject. In other types of questions this *som* is not obligatory, but it marginally occurs at least when the *wh*-phrase is the object or a complement of a preposition. In the Scandinavian languages, only one *wh*-phrase can be realized in the complementizer domain, as illustrated in (17).

(17) a. Jag undrar **vem** som tog med **vad**.
I wonder who SOM *took with what*
'I wonder who brought what.'

b. *Jag undrar **vad** **vem** som tog med.
I wonder what who that took with

In (17a), *vem* 'who' is in Spec-*c*P, and *vad* 'what' is left in situ. Example (17b), where both of the *wh*-phrases are in the complementizer domain, is unacceptable.

I assume, in accordance with much previous research (e.g. Platzack 2000, Stroh-Wollin 2002), that there is movement into the complementizer domain in relative clauses headed by *som* as well. In the *c*P/CP-analysis, this movement will be to spec *c*P.[17] According to Vikner (2017) and Nyvad et al. (forthcoming), an extra, optional *c*-head with an [occ]-feature provides an escape hatch for extraction in both embedded questions and relative clauses in Danish, just like in extraction from other subordinate clauses.

In what follows, I adopt the *c*P/CP-approach to the complementizer domain and explore the idea that an extra *c*-specifier is what makes ERC possible. Before turning to that investigation, however, I begin by looking at how some other preposing phenomena can be analyzed within this approach.

4.1.5 Other preposing phenomena

As described in chapter 2, topicalization (or T-preposing in the terminology I've adopted for Swedish), clefting, *wh*-questions, and relativization in both English and Swedish involve Ā-movement dependencies. In the previous section, we saw how embedded constituent questions (15b) and relativization (16b) are analyzed on the approach to the C-domain proposed by Vikner (2017) and Nyvad et al. (forthcoming) for the Scandinavian languages. At least some cases

[17] I represent the movement as movement of a relative operator here. I discuss the derivation of relative clauses in more detail in section 4.5.

of preposing in embedded V2 contexts are analyzed as movement to Spec-CP, as we saw in (13). For the sake of completeness I will also sketch the analyses of T-preposing in main clauses, left-dislocation, and main clause constituent questions here. Clefts will be discussed in section 4.3.3.

As mentioned in section 4.1.1, I use T-preposing as a cover term for movement of a phrase to the specifier of CP in main clauses, in anticipation of the more thorough analysis in chapter 5. An illustration is given in (18), where the structure is given in bracket notation.

(18) [$_{CP}$ Cykeln$_2$ [$_{C'}$ ställde$_3$ [$_{TP}$ jag$_1$ [$_{vP}$ t_1 t_3 t_2 i garaget]]]]
*bike.*DEF *put* *I* *in garage*
'The bike, I put in the garage.'

With respect to the *c*P/CP-approach, the analysis is straightforward. Main clauses are headed by the finite verb, i.e. they have a lexical head, and are big CPs.

In English, a topicalized phrase is usually contrastively stressed. In Swedish, realizing the T-preposed phrase with contrastive stress is a possibility, but it is also common for the T-preposed phrase to be non-contrastive. An example of this from a study of preposed object pronouns in mainland Scandinavian (Engdahl & Lindahl 2014) is given in (19).

(19) A: Var är cykeln?
*where is bike.*DEF
'Where is the bike?'

B: Den ställde jag i garaget.
*it put I in garage.*DEF
'I put it in the garage.'

(Engdahl & Lindahl 2014:2)

A asks about a particular bike, and Bs answer crucially does not evoke any contrast set. I assume that both contrastive and non-contrastive T-preposing targets Spec-CP.

T-preposing is distinct from **left-dislocation**, which is illustrated by (20).

(20) [Den där smörgåsen där]$_x$, jag skulle gärna äta upp den$_x$.
*the there sandwich.*DEF *there I would gladly eat up it*
'That sandwich over there, I would love to eat it.'

Here, a hanging topic precedes the element in the pre-finite position, in this

case the subject, and is co-referent with a pronoun inside the core clause.[18] In (20), this pronoun is in situ, in the verb phrase, and there is no inversion between the subject and the finite verb. While not ungrammatical in Swedish, the type of sentence in (20) is not that common (Teleman et al. 19994:446). Instead a combination of left-dislocation and pronoun fronting tends to be used, causing inversion of the subject and the finite verb, as in (21).

(21) [Den där smörgåsen där]$_x$, [den$_x$]$_1$ skulle jag gärna äta upp t_1.
the there sandwich.DEF *there it would I gladly eat up*
'That sandwich over there, I would love to eat it.'

This phenomenon is described by Andersson (1982), who calls the fronting of the pronoun **topic movement**. Since it causes inversion, it is likely to involve Spec-CP, and is then a case of T-preposing in the terminology I am assuming. The hanging topic by itself does not cause inversion as we saw in (20). I follow Platzack (2012, 2013), who argues that the hanging topic is first merged in an adjoined Ā-position, and semantically but not syntactically related to the pronoun. This gives the structure in (22).

(22) [$_{CP}$ [Den där smörgåsen där]$_x$, [$_{CP}$ [den$_x$]$_2$ [$_{C'}$ skulle$_3$ [$_{TP}$ jag$_1$ gärna
the there sandwich.DEF *there it would I gladly*
[$_{AuxP}$ t_3 [$_{vP}$ t_1 äta upp t_2]]]]]].
eat up
'That sandwich over there, I would love to eat it.'

The hanging topic and the pronoun must have the same referent.

As for direct questions, they have the finite verb in C, and are consequently CPs on the approach taken here. The structure is shown in (23).

(23) [$_{CP}$ Vem$_1$ [$_{C'}$ har$_2$ [$_{TP}$ t_1 [$_{AuxP}$ t_2 [$_{vP}$ t_1 stulit min smörgås?]]]]]
who has stolen my sandwich
'Who stole my sandwich?'

In direct constituent questions, as in (23), a *wh*-phrase Ā-moves to the specifier of C and the subject is in Spec-TP. As we saw in the previous section, only one *wh*-phrase can be fronted in embedded questions in Swedish, and the same is true in main clause questions (24).

(24) a. Vem tog med vad?
who took with what
'Who brought what?

[18] The co-reference relation is indicated by a subscript *x*.

b. *Vad vem tog med?
what who took with

If there are any other *wh*-phrases in the sentence, they must stay in situ.

The analysis I have sketched here of preposing phenomena in Swedish shows certain regularities. Main clause word order is connected to CP-structure, and embedded clause word order to the absence of a CP. For both *c*P and CP, it is the case that they can have at most one overt specifier. This is related to a restriction on the number of specifiers, which is limited to one per head. This captures both the V2 phenomenon and the fact that we can only have one overt *wh*-phrase in each clause in the complementizer domain in questions. Other specifiers are provided by the occurrence feature, and must be silent because of how the occurrence feature is defined. In cases where there is more than one phrase in the left periphery, as in left-dislocation, I have assumed this to be a position provided by adjunction.

4.1.6 The cartographic alternative

In the previous sections, I have tentatively adopted a version of a CP-recursion account of the left periphery of Swedish. An alternative to this view is the cartographic approach proposed by Rizzi (1997). Based on a detailed study of the left periphery in Italian, Rizzi argues that the C-domain should be split up into five layers, two of which (the topic layers) are recursive. This results in the structure in (25).

(25) $[_{\text{ForceP}}\ [_{\text{TopP*}}\ [_{\text{FocP}}\ [_{\text{TopP*}}\ [_{\text{FinP}}\ [_{\text{IP}}\ ]]]]]]$ (Rizzi 1997:297)

In a V2 language, all but one of the structural positions proposed would have to be obligatorily silent, as we have seen. This is not necessarily an argument against the cartographic approach, and there are proposals that adapts Rizzi's analysis to V2 languages (see e.g. Poletto 2002). However, if the V2 property can instead follow from the fact that there is only one CP (with only one specifier) in each main clause, this is more economical (for a similar discussion, see Platzack 2011:128ff). From this perspective, a CP-recursion approach to the cases where the complementizer domain is not atomic is promising, and as Vikner (2017) and Nyvad et al. (forthcoming) argue, the *c*P/CP-approach captures the Scandinavian data with fewer stipulations than the cartographic approach.

4.2 Is ERC really Ā-movement?

In chapter 2, I gave a description of Ā-dependency relations, and showed several of their properties. We saw that there are reasons to think that information about the dependencies is available in several places in a clause and that some Ā-relations involve movement, e.g. the fact that there are island violations in certain cases. In section 4.1, I outlined an analysis of Ā-movement dependencies in the local clause in Swedish, adopting an account by Vikner (2017) and Nyvad et al. (forthcoming) according to which extraction from subordinate clauses proceeds through an extra specifier in the complementizer domain, extraction from relative clauses being one case.

Like Vikner (2017), Nyvad et al. (forthcoming) and many others, I have been taking for granted the idea that the non-locally preposed phrase in ERC ended up in its position by Ā-movement. In this section, I will motivate this assumption in two ways. First, I will provide evidence that the dependency between the preposed phrase and the gap in ERC has the same characteristics as Ā-movement in the local clause and out of *att*-clauses. Second, I will consider and reject an alternative idea, namely that the relation consists not of a moved phrase and a gap, but a relation between a silent resumptive pronoun and a base generated phrase in Spec-CP (cf. the proposal by Cinque about two types of Ā-dependencies discussed in section 2.1.3). The arguments are based on observations from the previous literature and from my own investigations.

4.2.1 Evidence of Ā-movement out of RCs

In section 2.1, we saw that Ā-movement in Swedish, as well as in other languages, is characterized by the properties repeated here in (26) (e.g. Chomsky 1977, Engdahl & Ejerhed 1982, Engdahl 1986, McCloskey 1988, Haegeman 1994, Platzack 1998, 2011, Teleman et al. 1999, Pesetsky 2013).

(26) *Ā-movement*
- creates a gap
- is (apparently) unbounded
- respects island constraints
- exhibits crossover effects
- exhibits connectivity effects
- licenses parasitic gaps

With respect to the preposed phrase in ERC, a first piece of evidence, which is from spontaneous usage, comes from case connectivity. The example, given in (27), is from a novel, and the context sentence is translated to English here for the sake of simplicity.

(27) Is that so? Well, you don't look tired or hungry.

Och dig$_1$ var det ingen [som tvingade iväg t_1 till kolchoskontoret
*and you.*ACC *was there no one that forced away to the kolkhoz office*
mitt i natten och dömde t_1 till tjugofem års straffarbete].
in the middle of the night and sentenced to twenty five years penal servitude

'No one forced YOU to the kolkhoz office in the middle of the night and sentenced you to twenty five years of penal servitude.'

(Ruta Sepetys, *Strimmor av hopp*[19], novel)

The example involves an across-the-board-structure, and we find that the fronted pronoun has the form *dig* 'you.ACC'. This is expected if it originated as an object of *tvingade iväg* 'forced away' and *dömde* 'sentenced'. The corresponding example where the preposed pronoun is nominative is ungrammatical (28).

(28) *Och du var det ingen som tvingade iväg till kolchoskontoret [...]
*and you.*NOM *was there no one that forced away to the kolkhoz office*

Turning from case to binding connectivity, we find additional evidence. The examples in (29) and (30) are from the questionnaire (see section 3.2.1), and show connectivity for Condition A and Condition B. In (29), the preposed phrase contains the reflexive possessive *sina* 'their.REFLX.PL'. This example was judged as natural by eight of the questionnaire participants, as somewhat strange by seven participants, and as unnatural by one participant.

(29) men [sina$_x$ föräldrar]$_1$ känner jag inte många$_x$ [som skulle vilja ha
*but their.*REFLX.PL *parents know I not many that would want have*
med t_1].
with

'but I don't know anyone who would like to bring their parents.'

The answer pattern is very different in (30), which is the same example but with non-reflexive possessive *deras* 'their.PN' in the preposed phrase. This example was not judged as natural by any of the participants with the intended bound interpretation. Four participants judged it as somewhat strange, and twelve judged it as unnatural.

[19] Ruta Sepetys, *Strimmor av hopp*, B. Wahlström 2011:164.

(30) *men [deras$_x$ föräldrar]$_1$ känner jag inte många$_x$ [som skulle vilja ha med
*but their.*PN *parents know I not many that would want have with*
t_1].

The pair of examples shows the expected binding effects on the assumption that the preposed phrase originates inside the relative clause, in a position where the relative clause subject c-commands it. This explains how *sina* 'their.REFLX.PL' gets bound, and why *deras* 'their.PN', if co-referent with the relative clause subject, is rated as unnatural by so many participants.[20]

To complete the picture, we can also construct examples with parasitic gap licensing, connectivity with respect to Condition C, strong crossover effects, and violation of the Coordinate Structure Constraint. Example (31) shows that there is connectivity for Condition C.

(31) a. Anna$_x$ känner ingen som kan prata svenska med henne$_x$.
Anna knows no one that can speak Swedish with her
'Anna knows no one who can speak Swedish to her.'

b. [Anna$_x$]$_1$ känner hon$_{*x/y}$ ingen [som kan prata svenska med t_1].
Anna knows she no one that can speak Swedish with

An R(eferential)-expression, like *Anna* in (31) must be free. The sentence in (31a), where the proper name *Anna* binds the pronoun *henne* 'her' is perfectly fine, since there is no phrase binding *Anna*. On the other hand, (31b) is ungrammatical on the interpretation where *Anna* is co-referent with the pronoun. As discussed in section 2.1.2, the ungrammaticality can be accounted for by Condition C if the phrase *Anna* originates as the object of the preposition in the relative clause, where it is c-commanded by the pronoun, and has Ā-moved to the pre-finite position in the matrix clause.

ERC also induces strong crossover effects, which is demonstrated by (32).

(32) *[Vilken flicka]$_{x1}$ känner hon$_x$ ingen [som kan prata svenska med t_1].
which girl knows she no one that can speak Swedish with

On the reading where the *wh*-phrase *vilken flicka* 'which girl' is co-referent with the pronoun *hon* 'her', the example is not grammatical.

Since ERC apparently does not respect constrains like the Complex NP Constraint or Subjacency, we cannot use these constraints to test island ef-

[20] The example in (30) is grammatical if there is no co-reference relation between *deras* and the relative clause subject. This reading is not easily available in the context, as it would involve imagining an unmentioned antecedent for the pronoun, and the answers to the questionnaire indicate that the participants did not consider this reading.

fects. Examples like (33) show that ERC does respect some island constraints, however.

(33) *Katt$_1$ känner jag många [som har både hund och t_1].
cat know I many that have both dog and

In (33), the second conjunct of a coordinate structure has been extracted, which violates the Coordinate Structure Constraint. As illustrated by (34a), relative clauses with sentential subjects are quite ill formed. This means that the fact that extracting a phrase from a sentential subject inside a relative clause is also ungrammatical does not tell us that much.

(34) a. *Jag känner en sjuksköterska$_1$ [som [att jag tog med hatten] verkade
*I know a nurse that that I brought with hat.*DEF *seemed*
konstigt för t_1].
strange to

b. *Hatten$_2$ känner jag en sjuksköterska$_1$ [som [att jag tog med t_2]
*hat.*DEF *know I a nurse that that I brought with*
verkade konstigt för t_1]
seemed strange to

In (34a), we see a relative clause with a sentential subject, and in (34b) the version with extraction.

In addition to the other indications that the Ā-dependency in ERC is an Ā-movement dependency, it licenses a parasitic gap, which is shown in (35).

(35) [Såna där gröna bönor]$_1$ känner jag ingen [som kan äta t_1 utan att koka
such there green beans know I no one that can eat without to cook
__pg1 först].
first
'I know no one who can eat those green beans without cooking them first.'

As a comparison, we can see that no parasitic gap is licensed by left-dislocation, when no pronoun is fronted, as shown in (36).

(36) *[Såna där gröna bönor]$_x$, jag känner ingen [som kan äta dom$_x$ utan att
such there green beans I know no one that can eat that without to
koka __pg1 först].
cook first

Thus, parasitic gap licensing, like the other properties of ERC, are consistent with the relation between the preposed phrase and the gap being an Ā-movement dependency. It should be noted here that some of these diagnostics provide stronger evidence than the others. Specifically, case and binding

connectivity can also be found in left-dislocation structures, as illustrated for case connectivity by (37).

(37) a. Henne$_x$, [henne$_x$]$_1$ har jag nog inte träffat t_1 förr.
her her have I PRT *not met before*
'Her, I don't think I've met her before.'

b. ?Hon$_x$, [henne$_x$]$_1$ har jag nog inte träffat t_1 förr.
her her have I PRT *not met before*
'She, I don't think I've met her before.'

Example (37a), where the hanging topic exhibits case connectivity, is more natural sounding than (37b), with no case connectivity. For this reason, the other diagnostics are more decisive. However, the most important observation is that the tests all point in the same direction.

A final note is that all of the examples in this section involve extraction from subject relative clauses, i.e. relative clauses where the subject position has been relativized. As we will see in section 4.3.5, this is by far the most common type of relative clause to occur in extraction constructions, but Lindahl (2014) shows that the same pattern holds for extraction from non-subject relative clauses.

4.2.2 Silent pronouns?

As we saw in section 2.1.3, Cinque (1990) proposes that some syntactic dependencies which appear to involve Ā-movement should instead be analyzed as a relation between a silent pronoun and a base generated phrase which binds it, as in (38).

(38) [The article]$_x$ was too long for us to read *pro*$_x$.

A way to explain why we find ERC-sentences in Swedish would be to appeal to an Ā-binding relation of the type Cinque proposes, that is to say, that the Ā-dependency between the preposed phrase in ERC-sentences and the gap is an Ā-binding relation, and not an Ā-movement dependency. In what follows, I will call this the silent *pro*-hypothesis.

Given the overview of properties of Ā-movement in the previous section, we already have several reasons to think that the fronted phrase in ERC-sentences reaches its position by way of movement, but for the sake of completeness, it will also be useful to take a look at some types of data that bear more directly at the silent *pro*-hypothesis.

A first observation is that all of the ERC-sentences with T-preposing that we have seen so far exhibit the characteristic V2 pattern, with the finite verb in C, and the subject in Spec-TP. But as Engdahl (1997:54) argues, if such

examples involved binding of a silent pronominal in the gap site, we would have expected them to behave like left-dislocation structures, as in (39).

(39) *[Såna där gröna bönor]$_x$, jag känner ingen [som kan äta *pro*$_x$]
*such there green beans I know.*REL *no one that can eat*

However, such examples are ungrammatical.

Another argument against this type of analysis of ERC is that phrases of several categories can be preposed in ERC. One of Cinque's arguments for the silent Ā-bound *pro* analysis of certain Ā-dependencies is specifically that the gaps in the constructions he investigates are restricted to DPs, and that e.g. PP gaps do not occur in those structures. We have already seen a few examples of preposed PPs in ERC in chapter 2, among them the example from Wellander (1939) repeated here as (40).

(40) [Mot myggor]$_1$ finns det folk [som använder eukalyptusolja [$_{PP}$ t_1]].
against mosquitos exist there people that use eucalyptus oil
'There are people who use eucalyptus oil against mosquitos.' (p. 507)

A possible counter argument would be that Swedish might have a larger selection of silent *pro*-forms than other languages, and that this is why we see ERC to such a great extent. However, as Platzack (2011) shows, there is no silent PP-*pro* in Swedish. Consider the examples in (41), which are from Platzack (2011:59–60), but with my glosses.

(41) a. [$_{CP}$ *pro* funderade [$_{TP}$ jag faktiskt [$_{vP}$ aldrig [$_{vP}$ ~~jag~~ ~~funderade~~ på ~~*pro*~~]]]]
pondered I actually never on
'I never thought about that, actually.'

b. Det funderade jag faktiskt aldrig på.
that/it pondered I actually never on
'I never thought about that, actually.'

c. *Jag funderade faktiskt aldrig på.
I pondered actually never on

d. *Funderade jag faktiskt aldrig.
pondered I actually never

e. På det funderade jag faktiskt aldrig.
on that/it pondered I actually never
'I never thought about that actually.'

Examples like (41a), which illustrate a phenomenon usually called **topic drop**,

could be argued to involve a silent DP-*pro*.[21] In cases of topic drop, it is always possible to insert an overt pro-form, like in (41b). It is not possible to drop a DP that is not in Spec-CP, as is shown by (41c). Neither is it possible to drop a preposition along with the DP, as (41d) shows, even though the whole PP can be fronted (41e). If there was a silent PP-*pro*, we would have expected (41d) to be a possible sentence.

Furthermore, there is no overt PP-*pro* which can be inserted and bound in left-dislocation structures. Attempts to insert the pronoun *det* 'it' lead to strongly unacceptable sentences, which are perceived as incoherent, as shown in (42).

(42) a. *[Mot myggor]$_x$, det finns många [som använder eukalyptusolja [det$_x$].
against mosquitos there exist many that use eucalyptus oil that

b. *[Mot myggor]$_x$, [det$_x$]$_1$ finns det många [som använder eukalyptusolja
against mosquitos that exist there many that use eucalyptus oil
t_1].

Example (42) shows that a PP like *mot myggor* 'against mosquitos' cannot be the hanging topic and bind a pronominal copy in a left-dislocation structure, regardless of whether the pronominal is in situ or T-preposed. Similarly for the degree phrase *så ofta* 'that often', it can be preposed in ERC, but there is no corresponding pro-form, which is illustrated in (43).

(43) a. *[Så ofta]$_x$, jag känner ingen [som brukar tvätta bilen [det$_x$]/[då$_x$]].
*that often I know.*REL *no one that tend wash car.*DEF *that then*

b. *[Så ofta]$_x$, [det$_x$]$_1$/[då$_x$]$_1$ känner jag ingen [som brukar tvätta
*that often that then know.*REL *I no one that tend wash*
bilen t_1]
*car.*DEF

c. [Så ofta]$_1$ känner jag ingen [som brukar tvätta bilen t_1].
*that often know.*REL *I no one that tend wash car.*DEF
'I don't know anyone who washes their car that often.'

I will investigate ERC with degree phrases introduced by *så* in much greater detail in chapter 6, but for now, we can just note that they are among the phrase types that do not seem to have any natural *pro*-forms in Swedish, and that the fact that such phrases can be extracted would be hard to explain on the silent *pro*-hypothesis.

[21] See Mörnsjö (2002) for examples from spoken Swedish.

These arguments against a silent *pro* approach to ERC, together with the general characteristics of the syntactic dependency presented above, provide strong evidence that ERC is Ā-movement. In the next section, we turn to the question of the structure of the RC-like constituent in ERC-sentences.

4.3 Is the RC-like constituent a relative clause?

I turn now to the question whether the RC-like constituent in ERC is a regular restrictive relative clause. As we saw in chapter 2, this has been debated, and there are a few different proposals under consideration. The most radical of these is that the RC-like constituent is never a relative clause of any kind, but a small clause (Kush 2011, Kush et al. 2013). There are also proposals, based mainly on research on English, that suggest that relative clauses that permit extraction might not be regular relative clauses. McCawley (1981) suggests that there are "pseudo"-relative clauses in English that permit extraction, and Chung & McCloskey (1983) note that all then known examples of extraction from relative clauses in English, which were only a handful, involved subject relatives. Working within the GPSG-framework, they suggest that this is because there is a way to derive subject relative clauses that does not involve a slash-category – the GPSG-equivalent of an Ā-dependency. The proposal by Platzack (1999, 2014) that Swedish has a way of deriving subject relative clauses which does not require Ā-movement to Spec-CP is reminiscent of this proposal.

In this section, I argue for the position that there is extraction from regular restrictive relative clauses, but that there are also cases where another analysis of the RC-like constituent is plausible. The focus of the section is not the precise structure of the restrictive relative clause, a question which I return to in section 4.5. The point is to show that some of the RC-like constituents in ERC must be analyzed as restrictive relative clauses. The argument takes as a point of departure the RC-like constituents in my collection, and the organizing principle of the section is the syntactic environments where extraction from such constituents occurs.

4.3.1 Existential and presentational sentences

As noted in previous research (e.g. Engdahl 1997), a common environment for ERC is existential sentences. Two examples from my collection are given in (44) and (45).

(44) [en sån seminarieserie]$_1$ kan jag tänka mig att det finns flera [som
*a such seminar series can I think me.*REFLX *that there exist several that*
är intresserade av t_1]
are interested of
'I can imagine that there are several people who are interested in such a seminar series.'
(Conversation, Dec. 2014)

(45) [galltvål]$_1$ finns det någonting [som heter t_1]
gall soap exists there something that is called
'There is something called gall soap.'
(Conversation, June 2015)

In Swedish, existential sentences often involve the main verb *finnas* 'exist' and the expletive, formal subject, *det* 'it', as in (46a) and (46b).[22]

(46) a. Det finns kaffe i skåpet.
*there exists coffee in cupboard.*DEF
There is coffee in the cupboard.

b. Det finns inget kaffe.
there exists no coffee
'There is no coffee.'

Schematically, we can describe the structure as in (47), where the XP is optional and often gives a contextual restriction in the form of a location.[23]

(47) [Det finns DP (XP)]

---

[22] In southern Swedish varieties, the formal subject in this type of construction can also be *där* 'there', which is homophonous with a locative adverb, as in (i).

(i) Där finns gott om plats hos oss.
there exists plenty of room at our place
'There is plenty of room at our place.' (Teleman et al. 1999 4:54)

[23] Location should be understood in a metaphorical sense here, since it could involve location in space, like in (46a) but also location in time, as in (i).

(i) Det finns inga dronter numera.
there exist no dodos nowadays
'There are no dodos nowadays.'

Sundman (1980) distinguishes between "simple" existential sentences like (48a) and existentials with a relative clause (48b). The examples are hers, with my glosses.

(48) a. Det finns ett problem.
there exists a problem
'There is a problem.'

b. Det finns ett problem som vi inte kan lösa.
there exists a problem that we not can solve
'There is a problem that we can't solve.'

(Sundman 1980:59)

The ERC-sentences in (44) and (45) resemble the latter kind: in these examples there is a *som*-headed clause in which the extraction gap is located.

A very similar construction is involved in the extraction sentences in (49) and (50).

(49) ja men [det]$_1$ var det ingen [som klagade på t_1] ser du
yes but that was there no one that complained about see you
'Yes, but no one complained about that, you know.'

(Conversation, Spring 2015)

(50) [honungssenap]$_1$ är det många [som gillar t_1 till den där].
honey mustard is there many that like to the there
'There are many people who like honey mustard with that.'

(Conversation, Sept. 2014)

Here, the subject is also expletive *det*, but instead of *finnas* 'exist', the main verb is *vara* 'be'. Sundman (1980) classifies such sentences as a type of existential, and observes that they are often synonymous with the corresponding existential sentence with *finnas*. An example without extraction adapted from Sundman (1980:51) is given in (51).

(51) Det är en kråka som sitter på taket.
it/there is a crow that sits on roof DEF
'There is a crow that is sitting on the roof.'
'It is a crow that is sitting on the roof.'

As Sundman shows, this string is three ways ambiguous out of context. First, it could be interpreted as an existential sentence; second, it could be a *det*-cleft[24];

[24] As mentioned in chapter 2, Swedish has a *det*-cleft construction, which is similar to the English *it*-cleft. I will discuss the structure of such clefts in more detail in section 4.3.3.

and third, it could be interpreted as a predicative copular sentence, where *det* is not an expletive, but an anaphoric pronoun. The interpretations are roughly the ones in (52), which are translated from Sundman's Swedish ones (p. 52).

(52) a. There is a crow that is sitting on the roof.

b. The bird that is sitting on the roof is a crow.

c. (What is that? [Asked about a painting]) That is a crow sitting on the roof.

Sentences such as (51) are explored in detail by Søfteland (2013), who investigates existentials, presentational sentences and various types of clefts in spoken Norwegian. She calls the type in (51) *presenteringsutbryting* 'presentational cleft' when it has the interpretation in (52a). Like a presentational sentence, it introduces an indefinite which is typically new in the discourse. Søfteland notes that it is not always straightforward to disambiguate such strings even when they are in context, as in spontaneous discourse.

The extraction sentences in (49) and (50) are probably of the kind with the interpretation in (52a). Neither of the examples logically presupposes the content of the *som*-clause, which we would have expected if they were *det*-clefts. Furthermore it is not possible to get an anaphoric reading of the subject *det*.

I follow Sundman (1980) in assuming that this is a type of existential sentence, and will gloss the expletive in such sentences with *there* in English in order to disambiguate. When I want to talk specifically about this subtype of existentials, I will refer to them as **presentational sentences**.

The class of presentational sentences also includes examples like (53a) and (53b), which are adapted from Teleman et al. (1999 4:53–54).

(53) a. Nu flög det visst en glada över fältet igen.
now flew there PRT *a kite over field.*DEF *again*
'A kite flew over the field again, it seems.'

b. Det dök upp falska hundralappar
there dove up false hundred bills
'False 100 SEK bills showed up.'

In Swedish the set of verbs that can occur in presentational sentences is quite large. In addition to *vara* 'be', various verbs of motion, like *flyga* 'fly' in (53a), verbs of position, like *sitta* 'sit', *stå* 'stand', and verbs of appearance, like *dyka upp* 'show up' in (53b) and *komma* 'come' are used in presentational sentences. Extraction from presentational sentences with other verbs than *vara* is possible. I have no such examples in my collection, but consider (54), which is adapted from (Teleman et al. 1999 4:423).

(54) [Överblivna biljetter]$_1$ kom det en [som ville sälja t_1].
left over tickets came there one that wanted sell
'A guy who wanted to sell left over tickets came.'

In addition to the presentational sentences with *vara* and an expletive subject, my collection also contains a few examples with a referential subject instead of an expletive, in a construction that I analyze as a type of predicational copular construction with a *som*-clause, as in (55) and (56).

(55) Horses, they are lovely animals

det$_1$ är vi många [som tycker t_1]
that are we many that think

'We are many people who think that.'

(Radio Sweden, *Kropp och själ*, Aug. 30 2011)

(56) Hemingway wrote a book called *To Kill a Bull*

det$_1$ är vi inte många i den här buren [som visste om t_1]
that are we not many in the here cage.DEF *that knew about*

'We are not many people in this cage who knew about that.'

(Swedish Television, *På spåret*, Dec. 19 2015)

4.3.1.1 The attachment of the *som*-clause in existential sentences

For the present study, an important question is how the *som*-constituent in existential sentences relates to the rest of the sentence. Does it structurally belong inside the DP, or could it be analyzed as instantiation of the XP in the schematic structure of Swedish existential sentences in (47)?

Existential sentences have been the subject of intense study in modern linguistics, and there is no consensus on what structure they should be assigned, even in English. In fact, it is likely that more than one structure is necessary (McCloskey 2014:379). Consider the example in (57).

(57) There is someone sick. (Williams 1984:131)

Williams (1984) argues that *someone sick* in sentences like (57) is an NP (or a DP, in the terminology I've been assuming). This type of analysis of existential sentences is sometimes called "the NP-analysis", and can be contrasted with "the small clause analysis", argued for by among others Rezac (2006), and analyses that argue for a flatter structure (Keenan 1987), where the DP does not form a constituent with the XP at all. The analyses can be seen in (58).

(58) a. There is [$_{DP}$ someone sick].

b. There is [$_{SC}$ someone sick].

c. There is [$_{DP}$ someone][$_{AP}$ sick].

The last of these structures is parallel to the schematic structure for Swedish existential sentences proposed above in (47).

Keenan (1987) has several arguments that this is the right structure for existential sentences in English, and that the XP in existential sentences therefore is not a part of the DP, but instead is predicated of it (p. 300ff). The clearest of the arguments is maybe from sentences like (59), where the DP contains an exceptive modifier.

(59) There was no student but John in the building. (Keenan 1987:301)

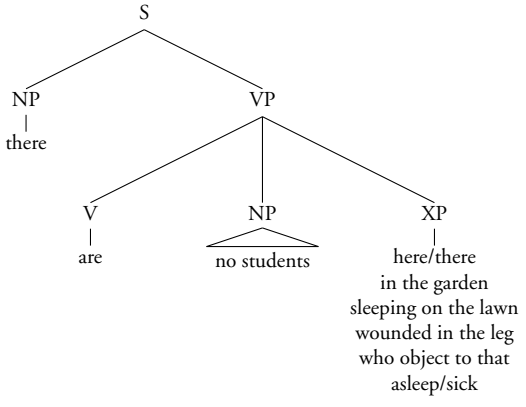
The argument is semantic. The interpretation of (59) must be the one in (60a), not the one in (60b).

(60) a. Possible interpretation: 'no student but John was in the building'

b. Impossible/unlikely interpretation: 'no student but John in the building existed'

If *in the building* was part of the DP, we would have expected the interpretation in (60b) to be possible. But this would involve either the self contradictory assertion that all students in the building except John do not have the property of existing or the assertion that there are no students except for John in the world. On the basis of this argument and others, Keenan proposes that the structure of existential sentences like (59) is the one in (61).

(61)



(cf. Keenan 1987:300)

A similar argument can be made for Swedish based on sentences like (62).

(62) Det fanns ingen student utom Johan i byggnaden.
there existed no student but Johan in the building
'There was no student but Johan in the building.'
Possible interpretation: 'no student but Johan was in the building'
Impossible/unlikely interpretation: 'no student but Johan in the building exists'

Like in English, a Swedish existential sentence with an exceptive must be interpreted as saying that Johan is the only student in the building. It is very unlikely that we want to say that none of the students in the building except Johan exists.

As shown in (61) Keenan includes RC-like constituents in existential sentences among the phrases that can be in the XP position in his structure.[25] Examples like (63) indicate that this is a possible analysis for Swedish *som*-clauses in existential sentences as well.

(63) Det finns ingen utom läraren som vet det.
*there exist no one except teacher.*DEF *that knows it*
'There is no one except the teacher who knows it.'
Possible interpretation: 'no one except the teacher knows it'
Impossible/unlikely interpretation: 'no one except the teacher who knows it exists'

Here, a possible interpretation is that no one except the teacher knows something, and it is not likely that we want to interpret this sentence as saying that no one except the teacher who knows something exists. This is a sign that the *som*-clause is not a part of the DP *ingen utom läraren*, and that it is in the XP-position in Keenan's proposed structure.

It is possible to extract a phrase from such a *som*-clause.

(64) Det_1 finns det ingen utom läraren [som vet t_1].
*it exist there no one except teacher.*DEF *that knows*
'There is no one except the teacher who knows it.'

This is perhaps not so surprising; if the *som*-constituent is not part of DP, and is not a regular restrictive relative clause, it might not be affected in the same way as other relative clauses by constraints on movement.

[25] It is not clear how the binding between the NP and the XP in such a case would come about, or if there is supposed to be any binding. Keenan does not comment further on such examples.

Keenan notes that sentences with *have* and relational nouns like *friend*, *colleague*, and *sister* constitute the same type of existential environment as do existential sentences with *be*. Consider (65).

(65) John has several friends who work in the mines. (Keenan 1987:306)

As Keenan points out this sentence is similar to the sentence *There are several friends of John who work in the mines*, and it will be true in exactly the same circumstances. Extraction sentences with *have* + relational noun + *som*-constituent are also found in my collection. An illustrative example is given in (66).

(66) det$_1$ hade vi några bekanta [som hette t_1]
that had we some acquaintances that were named
'We had some acquaintances who were named that.'
(Conversation, fall 2015)

The question now is whether all RC-like constituents that permit extraction can be analyzed as being in XP-position in Keenan structure in (61).

4.3.1.2 Extraction from *som*-clauses inside DP

As it turns out, extraction is not restricted to *som*-clauses in the XP-position in Keenan's structure. Consider for example (67).

(67) Det finns många som säger det i Göteborg.
there exist many that say that in Gothenburg
'There are many people who say so in Gothenburg.'
Possible interpretation: 'in Gothenburg, many people say so'
Possible interpretation: 'many people who say so exist in Gothenburg.'
Unlikely interpretation: 'many people who say so in Gothenburg exist'

Here, the most natural interpretation is that the *som*-clause is a relative clause which restricts the DP, whereas the PP *i Göteborg* 'in Gothenburg' is the XP in Keenan's structure, and predicates about the individuals denoted by the DP that they are in Gothenburg. Crucially, it is possible to extract a phrase from the relative clause in these cases as well (68).

(68) Det$_1$ finns det många som säger t_1 i Göteborg.
that exist there many that say in Gothenburg
'There are many people who say so in Gothenburg.'

Further evidence that it is possible to extract from a relative clause that is inside the DP comes from examples with more than one *som*-constituent. Consider

(69), where the two *som*-constituents are marked S1 and S2.[26]

(69) [en sån här]$_1$ hade min pappa en bekant [$_{S1}$ som var kommendör-
one such here had my father an acquaintance that was commander
kapten] [$_{S2}$ som körde t_1]
that drove

'My father had an acquaintance who was a commander who drove a car like this.'

(Conversation, July 2016)

The extraction gap is in S2, which could be analyzed as being in the XP position in Keenan's proposed structure. There are two *som*-constituents in (70) as well, but this time, the extraction gap is in S1.

(70) nej, [den här]$_1$ var det ingen [$_{S1}$ som hade skrivit i t_1] [$_{S2}$ som vi
no the here was there no one that had written in that we
känner]
know

'No, there is no one that has written in this that we know.'

(Conversation, June 2016)

Here, the most natural analysis is that the clause with the extraction gap, S1, is part of the DP. As in (67), the *som*-clause *som hade skrivit i* 'who had written in' restricts the denotation of the head. Since relative clauses can be stacked, it is in fact possible that both of the *som*-clauses are part of DP. Importantly, in that case no phrase is in the XP position, and we still have extraction from a *som*-clause which attaches inside DP.

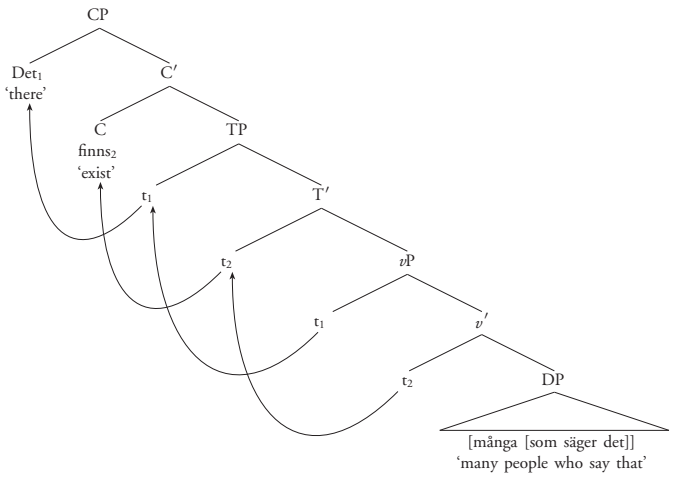
In sum, even though some *som*-constituents in existential sentences can be analyzed as XP on Keenan's approach, and hence outside DP, there are a number of examples which cannot be analyzed this way. Some *som*-clauses attach inside the DP, just like regular restrictive relative clauses. This is not surprising per se. Given that the syntax provides means to derive restrictive relative clauses, the null hypothesis should be that they could also modify DPs in existential sentences. The crucial finding is that extraction is not restricted to those *som*-clauses that could be argued to be in the XP position in Keenan's structure in (61), but is available from *som*-clauses that must be interpreted as a restrictive relative clause inside DPs as well.

The two possible analyses of the *som*-constituents in existential ERC-sentences

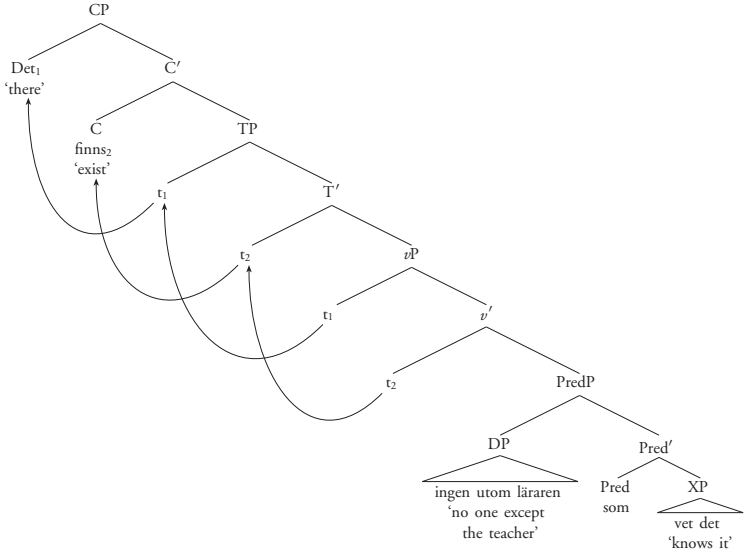
[26] This example is an illustration of the *have* + relational noun type pattern described above, and was uttered in front of a Volvo P1800 at the Volvo museum in Gothenburg.

could be represented as in (71), on the view of Scandinavian phrase and clause structure that I assume.

(71) a.



b.



In the structure in (71a), the *som*-constituent is a regular restrictive relative clause, which is in a DP-complement to *finnas* 'exist'. The internal structure of the relative complex in such relative clauses will be discussed more in section 4.5.

In (71b), I have tried to capture the insight of Keenan's structure but with a slightly more elaborate structure with binary branching instead of the ternary

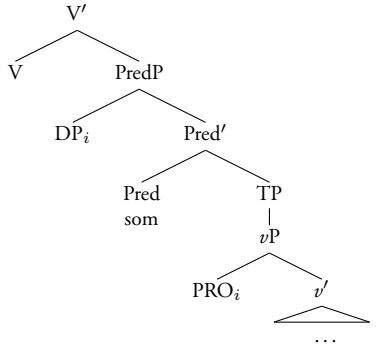
branching that Keenan proposes. On this approach, *finnas* 'exist' takes a PredP-complement, and the RC-like constituent is in the complement of Pred. I have followed Kush (2011) and Kush et al. (2013), who in turn base their analysis on Eide & Åfarli (1999), in assuming that *som* can be a predicational head. Nothing hinges on this assumption in this analysis, however, and perhaps *som* is part of XP, as *who* is in Keenan's structure. In the next section, I discuss the relationship between Keenan's structure and the structure in (71b), and the small clause structure proposed by Kush (2011) and Kush et al. (2013). I also discuss McCawley's (1981) proposal about pseudo-relatives.

4.3.1.3 Small clauses and pseudo-relatives

In relation to the observations in the previous section, it is useful to make a brief comment on how they fit with the small clause analysis of relative clauses in ERC proposed by Kush (2011) and Kush et al. (2013), and the proposal by McCawley (1981) about pseudo-relative clauses.

The structure of a small clause is obviously distinct from the DP + XP analysis which Keenan argues for. Keenan specifically argues that the DP and the XP in his analysis of existential sentences do not form a constituent, whereas on the small clause view, they form a PredP. However, when it comes to the interpretation of the XP with respect to the DP, which is what the argument in the previous section builds on, they make similar predictions, and the structure in (71b) is basically a variety of the small clause analysis, although I do not want to make claims about the "smallness" of this clause, which, I assume, only occurs in existential environments. The structure of small clauses proposed by Kush (2011) is repeated here for convenience.

(72)



In this structure, the RC-like constituent is a complement to Pred, and is predicated of the DP, which is in the specifier of Pred, just like in (71b). For this reason, the arguments from the previous section that the RC-like clause in

certain ERC-sentences is not simply predicated of the DP bears on the Small Clause Hypothesis as well.

A precursor to the idea that the RC-like constituents in ERC are not regular relative clauses is the observations about pseudo-relatives due to McCawley (1981). McCawley observes that there are some structures in English which permit extraction, and which on the surface look a lot like relative clauses. Three examples are given in (73).

(73) a. Then you look at what happens in languages that you know and languages that you have a friend who knows.

b. This is the one that Bob Wall was the only person who hadn't read.

c. Violence is something that there are many Americans who condone.

(McCawley 1981:108)

Relative clauses in English do not generally allow extraction, and McCawley furthermore points out that these RC-like clauses are different from restrictive relative clauses in other ways too. For example, while it is not generally possible to put in a parenthetical remark like *as you know* between a restrictive relative clause and its head (74a), this is fine with pseudo-relatives (74b).

(74) a. *Tom cooked a dish, as you know, that I always enjoy.

b. There are many americans, as you know, who distrust politicians.

(McCawley 1981:106)

McCawley describes the environment for pseudo-relatives as the VP-final position in existential clauses and negations of existential clauses, and he takes as a requirement that the embedding verb indicate the basis of the existential judgement (p. 107). A few verbs that are mentioned as being used in this way are *see*, *meet*, *hear of*, and *run into*.

As we see from this characterization, McCawley's notion of existential clause is somewhat broader than the one we have been assuming in this section. For example, the sentence in (75) is existential in the right sense, and the relative clause can be assumed to be a pseudo-relative, rather than a regular relative clause, according to McCawley.

(75) I've never met an American, of course, who doesn't like pizza.

(McCawley 1981:107)

On McCawley's view, we interpret this sentence as saying something like *Judging from my personal contacts, no American doesn't like pizza* (p. 107).

In the next section, I look at ERC with embedding predicates other than

finnas 'exist' and *vara* 'be' to see how the various proposals discussed in this section fare in accounting for them, and whether we can find more support for the idea that there is ERC with regular relative clauses.

4.3.2 Other main verbs

In addition to *finnas* 'exist' and *vara* 'be', other embedding predicates in my collection of ERC-sentences include verbs of knowledge or acquaintance (*veta* 'know.COG', *känna* 'know.REL'), (*träffa* 'meet'), perception (*se* 'see', *höra* 'hear'), possession (*ha* 'have'), and attitude (*gilla* 'like', *avsky* 'detest', *störa sig på* 'be annoyed by').

We can observe, as a start, that examples like (76) are of the type with *have* + relational noun described by Keenan (1987), which was mentioned in the previous section.

(76) [det]$_1$ **har** vi exempel på kollegor [som gör t_1]
that have we examples of colleagues that do
'We have examples of colleagues who do that.'

(Conversation, Aug. 2013)

Kollega 'colleague' is a relational noun, which together with *har* 'have' can combine into an existential sentence in Keenan's sense. As we saw in section 4.3.1.1, this means that the *som*-clause could be XP in Keenan's structure. The reason that (76) can be argued to have such a structure is that it can be understood to have an existence assertion reading. To see what this means, we can consider (77).

(77) a. Julie has a colleague who has been to Iceland.

b. Julie does not have a colleague who has been to Iceland.

The sentence in (77a) is naturally interpreted in the same way as *a colleague of Julie who has been to Iceland exists*, i.e. as an existential statement. The negated version in (77b) is interpreted as *no colleague of Julie's who has been to Iceland exists*. Crucially, negation of an existential sentence should result in an assertion that no entity of the type denoted by the DP exists.

Not all verbs that occur in ERC-sentences are amenable to this type of interpretation, however, which becomes clear if we consider (78), which is a perfectly acceptable extraction sentence.

(78) Sealing [your chimney] with a ceramic sealant from the inside.

Det$_1$ **har** jag inte **träffat** någon [som gjort t_1]
that have I not met someone that done

'I've never met anyone who has done that.'

(Forum thread, Jan. 2005)

To see how the ERC-sentence with *träffa* 'meet' is different from a *ha* + relational noun sentence, we can look at the two sentences in (79).

(79) a. Iris har träffat någon som har gjort det.
Iris has met someone that has done that
'Iris has met someone who has done that'.

b. Iris har inte träffat någon som har gjort det.
Iris has not met someone that has done that
'Iris hasn't met anyone who has done that.'

If we take *done that* to refer to "sealing one's chimney with ceramic sealant", we can infer from the fact that Iris has met someone who has done that that there are some people who have. However, from the negated sentence, we cannot infer that there is no one who has sealed their chimney with ceramic sealant, since it is easy to imagine a situation where people have, and where it just so happens that Iris has never met any of them. This is different from (77), where negating the sentence resulted in the assertion that no entity of the type denoted by DP existed. This means that (79a) cannot be understood on an existential assertion reading, which in turn means that we have no argument for the DP + XP structure being possible in this case.

We can conclude then, that in addition to the existence of existential clauses with ERC that should not be analyzed as involving Keenan's structure in (61), there are also ERC-sentences with other embedding predicates that do not create an existential environment as defined by Keenan, since examples with embedding predicates like *träffa* exist.

Turning instead to the Small Clause Hypothesis, it can also account for a subset of the examples with other embedding predicates. Specifically, Kush (2011) and Kush et al. (2013) identify both *känna* 'know.REL', and *höra* 'hear' as SC-selecting verbs, which would explain extraction sentences like (80) and (81).

(80) Molly (.) det_1 **känner** jag en hund i Örebro [som heter t_1]
*Molly that know.*REL *I a dog in Örebro that is called*
'Molly, I know a dog in Örebro who is called that.
(Conversation, Dec. 2015)

(81) In that case, you could call *hålla på* a progressive auxiliary,
men det_1 **har** jag aldrig **hört** nån [som har gjort t_1]
but that have I never heard someone that has done
'but I've never heard anyone do that.'
(Conversation, Jan. 2012)

Verbs like *träffa* 'meet', however, are used as the test case for a non-SC-selecting verb by Kush (2011) and Kush et al. (2013), so the fact that (78) is entirely well formed is unexpected on this approach. *Beundra* 'admire' and *störa sig på* 'be annoyed by' are equally hard to interpret as selecting a small clause, which means that (82) and (83) are also unaccounted for on the small clause view.

(82) Luckily I have an exam tomorrow, so I have something to focus on and won't have to just wait.
Det_1 **beundrar** jag folk [som klarar t_1 rent psykiskt], att bara vänta.
that admire I people that manage purely psychologically to just wait
'I admire people who can deal with that psychologically, to just wait.'
(Forum thread, Feb. 2011)

(83) Alcoholism is not a disease, however.
det_1 **stör** jag **mej** på folk [som säger t_1]
that annoy I me on people that say
'People who say that annoy me.'
(Magnus Betnér, *Folkhälsan*, stand-up show, 2010)

If we were instead to adopt the delimitation of existential sentences proposed by McCawley (1981) described in the previous section, i.e. that the embedding verb "indicates the basis of the existential judgement (p. 107)", we can explain a few more of the examples. McCawley would perhaps argue that all of the examples in (76) through (81) are covered by his pseudo-relative analysis. ERC-sentences with the embedding predicate *veta* 'know.COG', as in (84), might also be interpreted this way.

(84) Yes, all of *Rederiet*[27] is available in the open archives at Swedish television.

det$_1$ **vet** jag många [som har fastnat i t_1]
*that know.*COG *I many that have gotten stuck in*

'I know of many people who have gotten stuck on that.'

(Conversation, Nov. 2015)

The main point in this utterance is probably not to convey that the speaker knows of these people, but rather to say that they exist and to bring them into the conversation.

Nevertheless it would be very hard to interpret (82) and (83), with the embedding predicates *beundra* 'admire' and *störa sig på* 'be annoyed by', as existential sentences even with McCawley's view of them. The point of these utterances does not seem to be to say something about the existence of the set of people denoted by the DP, but rather to express that the subject holds a specific attitude about them. The existence of ERC-sentences with such embedding predicates further supports the idea that there is extraction from regular relative clauses. In the next section, we turn to extraction from the RC-like clause in clefts.

4.3.3 Clefts

Several of the ERC-sentences in my collection involve some type of cleft construction. These are included in the collection since they contain a clause which is quite similar to a relative clause. I will call this RC-like clause a **cleft clause**. Most of these examples involve extraction from *det*-clefts, as in (85) and (86).

(85) A: It's still good to keep your jacket on here. Otherwise, you'll smell of frying later.

B: det$_1$ är det [bara jackan]$_2$ som t_2 gör t_1 sen då
*that is it just jacket.*DEF *that does later then*

'It is only the jacket that will do that later then.'

(Conversation, Dec. 2015)

[27] A Swedish drama series.

(86) den allra största delen av befolkningen, bönderna, [den]$_1$ var det
*the most of all biggest part of population.*DEF, *farmers.*DEF, *that was it*
adeln$_2$ som t_2 hade domsrätt över t_1
*nobility.*DEF *that had jurisdiction over*
'It was the nobility that had the jurisdiction over the largest part of the population, the farmers.'

(Radio Sweden, *Godmorgon, världen!*, Jan. 25 2015)

Det-clefts are the Swedish counterparts of *it*-clefts, and share several properties with them; they consist of a structure with a cleft pronoun *det* which is identical to an expletive in the language, a copula, a pivot, and a cleft clause, which bears similarity to a relative clause. Furthermore, *det*-clefts, like *it*-clefts, come with logical presuppositions. We can see this in examples like (87).

(87) a. Det var i Aten vi träffades.
it was in Athens we met
'It was in Athens that we met.'

b. Det var inte i Aten vi träffades.
it was not in Athens we met
'It was not in Athens that we met.'

Both of these sentences presuppose that there is some location *x* for which the sentence *'we met at x'* is true. Furthermore, the cleft asserts that the pivot is (or is not) that *x*.

There is evidence from both case connectivity and the distribution of reflexive pronouns that the pivot in *det*-clefts originates inside the cleft clause. The two examples in (88) show that a pronominal pivot must have the form corresponding to the grammatical function filled by its gap in the cleft clause.

(88) a. Det är hon$_1$ som t_1 har ätit upp smörgåsen.
*it is she.*NOM *that has eaten up sandwich.*DEF
'She's the one who's eaten the sandwich.'

b. Det är henne$_1$ jag tar pianolektioner hos t_1
*it is her.*ACC *I take piano lessons with*
'It is her I take piano lessons with.'

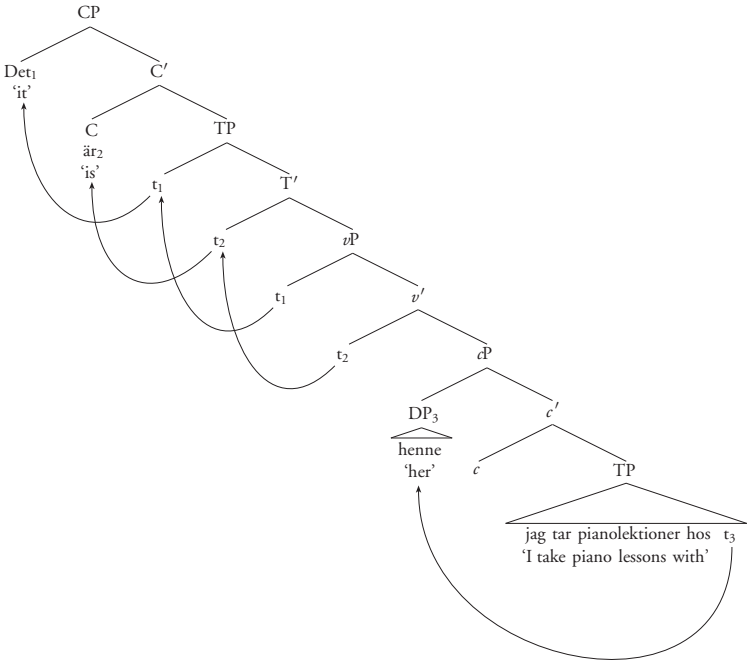
In (88a), the pivot corresponds to the subject in the cleft clause and is nominative, in (88b) it corresponds to a complement of a preposition in the cleft clause, and is accusative.

Reflexives in the pivot can be bound by the subject of the cleft clause (89).

(89) Det var bara [sin$_x$ mamma]$_1$ som hon$_x$ kunde fråga t_1 om matte.
*it was only her.*REFLX *mother that she could ask about math*
'It was only her mother that she could ask about math.'

I will assume here that the cleft clause in *det*-clefts is a *c*P complement of the copula, and that the pivot moves to Spec-*c*P, as in (90).

(90)

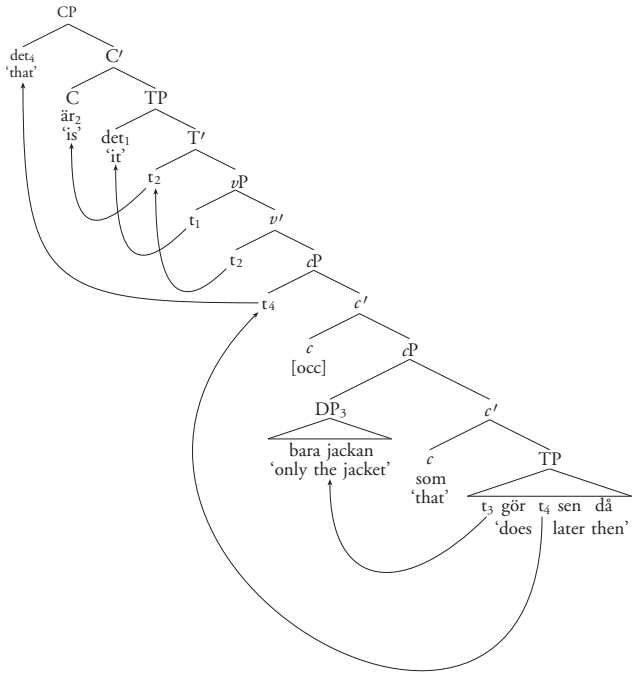


This explains the case connectivity and the fact that the subject of the cleft clause can bind a reflexive in the pivot.[28] Given the *c*P/CP-analysis, this structure also predicts that extraction should be possible, if there is a *c*[occ]-head which can select the cleft clause. In extraction sentences, *c*P-recursion will provide the escape hatch, as in other embedded clauses which allow extraction.[29] The ERC-sentence in (85) will have the structure in (91).

[28] Confer the discussion of head-raising in section 4.5.1.

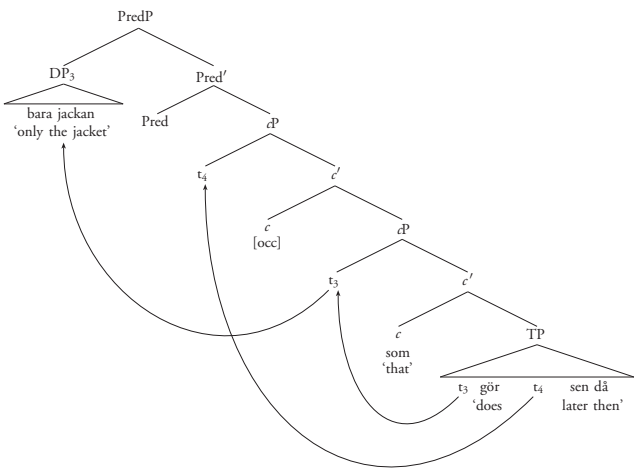
[29] An alternative is proposed by Fiedler (2014) who investigates *it*-clefts in several Germanic languages. In her structure, the cleft clause is the complement of a predicational head. The pivot is derived by movement through the complementizer domain, to the specifier of the PredP. An adaptation of this proposal within the *c*P/CP-approach to the left periphery yields the structure in (i). The PredP layer does not make a difference for my discussion.

(91)



We see here that the pivot *bara jackan* 'only the jacket' has moved to the lower Spec-*c*P, and that the extracted phrase, *det* 'that' has moved through the escape hatch provided by *c*[occ] and ended up in Spec-CP in the matrix clause.

(i)



There are several other cleft types in Swedish, and at least one other type, reverse demonstrative clefts, also permits extraction, as (92) illustrates.[30]

(92) näe det$_1$ tror jag verkligen inte att dom är dom enda som GÖR t_1
no that think I really not that they are the only that do
'No I really don't think that they are the only ones who do that.'

(Radio Sweden, *Plånboken*, 7 Nov. 2012)

Reverse demonstrative clefts differ from *det*-clefts in having a referential subject, here *dom* 'they', and the cleft clause is introduced by a demonstrative. I have a few examples of this type in my collection, but they do not seem to be very common. I will leave a more detailed analysis of the structure of reverse demonstrative clefts to future research. It is interesting to note, however, that *det*-clefts and reverse demonstrative clefts are the two cleft types that Engdahl (1997) finds in her investigation of extraction as well.

4.3.4 Overview

Table 4.1 gives an overview of how common the different types of clauses are in Sample A.[31] The example numbers in this table refer to examples discussed in this chapter.

Table 4:1. *Environments where extraction occurs*

Type	No of instances	Example
Existential (*finnas*)	13	(44)–(45)
Presentational (*vara*)	61	(49)–(50)
Predicational copular (*vara*)	4	(55)–(56)
Other main verbs	13	(83)–(84)
Det-cleft (*vara*)	8	(85)–(86)
Reverse DEM-cleft (*vara*)	1	(92)
Sum	100	

The overview shows that extraction is most common from existential sentences in the sample, especially those of the presentational variety. The two existential

[30] Other types include *wh*-clefts, reverse *wh*-clefts, and demonstrative-clefts (Engdahl 1997, Teleman et al. 1999).

[31] Given how the examples were collected, these figures should only be understood as rough frequency indications.

types make up around 75 percent of the sample. Extraction sentences with other main verbs and *det*-clefts occur, but not all that frequently, whereas predicational copular constructions and reverse demonstrative clefts are quite rare.

We saw in section 4.3.1 that *som*-clauses in some existential sentences are naturally interpreted as being predicated of the DP, but that others are most naturally interpreted as attaching inside the DP, much like restrictive relative clauses. This is problematic for analyses like the Small Clause Hypothesis which predict that extraction should be possible only from RCs or RC-like constituents that are predicated of a DP.

Another important result of this section is that there are some ERC-sentences where the embedding predicate is not compatible with either an existential assertion interpretation or a small clause analysis. McCawley's (1981) generalization that the verb should indicate the basis of the existential judgement also fails to capture all of the Swedish ERC-sentences.

4.3.5 Extraction from non-subject relatives

A question related to the structure of the RC or RC-like constituent in ERC is whether it is possible to extract from non-subject relative clauses, i.e. relative clauses where it is not the subject that has been relativized, as in (93).

(93) Jag har inte hittat någon ryggsäck$_1$ [Op$_1$ jag gillar t_1 ännu].
I have not found some backpack I like yet
'I have not found a backpack I like yet.'

In (93), the direct object of *gillar* 'like' is relativized. Extraction from such relative clauses is rare; in fact, in all of the examples we have seen so far, the extraction is from a relative clause where the subject has been relativized.

Engdahl (1997) provides the two examples of extraction from non-subject relatives in (94), however.

(94) a. Det$_1$ har jag inte haft någonting$_2$ att göra t_2 med t_1.
that I have not had anything to do with

b. Matte$_1$ var det bara pappa$_2$ (som) jag kunde fråga t_2 om t_1.
maths it was only dad that I could ask about

(Engdahl 1997:57)

The first of these examples involves extraction from an infinitival relative clause (94a), and the second involves extraction from a cleft clause where an object is the pivot (94b).

I have a few examples with extraction from finite non-subject relative clauses which are not cleft clauses. These are given in (95)–(98).

(95) It fits well, but it's not very well ventilated. I easily get a bit sweaty on my back.

Fast det_1 har jag inte hittat någon ryggsäck$_2$ [Op_2 jag inte blir t_1 av t_2].
but that have I not found some backpack I not become of

'But I haven't found any backpack that I don't get sweaty from.'

(Forum thread, June 2008)

(96) och sen så sa dom: "Aron, du får sparken.", vilket$_1$ det inte är första
and then PRT *said they Aron you get the kick which it not is first*
gången$_2$ [Op_2 jag hör nån säga t_1 t_2]
*time.*DEF *I hear someone say*
'And then they said "Aron, you're fired", which it is not the first time I hear someone say.

(Aron Flam, *Kön*, stand-up show, 2015)

(97) [just den här delen av matten]$_1$ kommer ni inte hitta nått jobb$_2$ [där$_2$
precisely the here part of the math will you not find some job where
man behöver t_1 t_2]
one needs
'You won't find any job where you need THIS part of math.'

(Conversation, June 2016)

(98) Men det här i början då, det är sånt$_1$ [Op_1 som man kanske
*but the here in beginning.*DEF *then, that is such that one maybe*
kunde få en struktur$_2$ [där$_2$ man lyfter fram t_1 t_2]].
could get a structure where one lifts forward
'What about the things in the beginning. Those are the kind of things one might create a structure such that one brings them out.'

(Conversation, Sept. 2016)

Such examples are rare. During the five years I've been collecting examples, I have only come across a handful. However, the fact that they do occur in spontaneous speech, and that they sound perfectly natural, is an additional piece of evidence that ERC is possible from full relative clause structures.

Somewhat more common is extraction from clefts where the pivot is an adverbial. An example is given in (99).

(99) [så många tårar på en och samma plats]$_1$ var det längesen$_2$ [jag såg t_1 t_2].
that many tears on one and same place was it long since I saw
'It was a long time since I saw that many tears in one and the same place.'

(Radio Sweden, P1 OS krönika dag 14, Aug. 19 2016)

The non-lexical head in *c* is not realized in (99), but in Swedish it is possible to insert *som* here, like in relative clauses. In Danish (100) and Norwegian (101) adverbial clefts, the complementizer is *at*.

(100) Det er længe siden, at en skiferie bare var en skiferie. (Da.)
it is long since that a ski break only was a ski break
'It's a long time since a ski break was just a ski break.'

(*Politiken*, Jan. 27 2013)

(101) Politiet fikk til svar at det var lenge siden at ølen ble (No.)
the police got as answer that it was long since that the beer was
tappet.
filled from the tap
'The police got the answer that it had been a long time since the beer was served.'

(*Fredrikstad Blad*, July 3 2015)

This suggests that the cleft clause in adverbial clefts may be a complement clause, but more research on adverbial clefts is needed.

Returning to (95)–(98), since they are not subject relative clauses, they show that we should not build our analysis of Swedish ERC on the idea that subject relative clauses are special (cf. Chung & McCloskey 1983, Platzack 1999). Extraction from non-subject relatives is also hard to accommodate within a small clause analysis of ERC, since small clauses are assumed to be subject-oriented, meaning that the DP in the specifier of PredP will correspond to the subject argument in a clause, and can never be interpreted as corresponding to a non-subject argument (Kush 2011, Kush et al. 2013).

In addition, these examples resist an analysis where the relative clauses are in the XP position in Keenan's structure in (61), since the relative clauses need to be interpreted as restricting their respective heads. At least one of the embedding predicates, *hitta* 'find', does not create an existential assertion environment.

Together with the observations from section 4.3.1 and section 4.3.2, the fact that there is extraction from non-subject relative clauses is evidence that the RC-like constituent is indeed a relative clause in many extraction sentences, and that it is possible to extract from uncontroversial relative clauses, i.e. relative clauses which are attached inside the DP. In section 4.5, I discuss the structure of restrictive relative clauses further, but first, we turn to the definiteness of the head of the relative clause.

4.4 Definiteness of the DP

It is observed in the previous research that the definiteness of the head of the relative clause matters for extraction possibilities, and that extraction sentences are often more acceptable if the head is indefinite and non-specific (e.g. Engdahl 1982, Andersson 1982, Engdahl 1997, Teleman et al. 1999). My investigation is in line with this observation. In Sample A, 89 of the 100 ERC-sentences have an indefinite relative complex. A large part of the examples involve quantifiers like *ingen* 'no one', *någon/nån* 'someone', *några* 'some', *något/nåt* 'something', *många* 'many', *flera* 'several', and *få* 'few', which is also in line with previous research. These quantifying expressions are often used without a noun in Swedish, and can appear in positions where a DP would be expected, like in (102), where *många* 'many' is the subject.

(102) En bok, ett konstverk, ett musikstycke blir inte sämre för att **många**
a book, a work of art a piece of music become not worse because many
tycker om det. Är det bra är det bra.
like it is it good is it good

'A book, a work of art, a piece of music does not get worse because many people like it. If it is good, it is good.'

(Bodil Malmsten, *Loggböckerna 2005–2013*[32], diaries)

In such sentences, it will usually be clear from the context what class of entities there are many or few of.

We have seen several examples of this use of the quantifiers in ERC-sentences already in the dissertation, and another is given in (103). In this section I will put the head of the relative clause in bold for perspicuity.

(103) [En sån seminarieserie]$_1$ kan jag tänka mig att det finns **flera** [som är
a such seminar series can I think me that there exist several that are
intresserade av t_1]
interested of

'I think that there are several people who would be interested in such a seminar series.'

(Conversation, Dec. 2014)

These quantifying expressions are treated as pronouns by Teleman et al. (1999). I will instead assume that they are determiners, and head the DP.[33]

[32] Bodil Malmsten, *Loggböckerna 2005–2013*, Modernista 2016:532.

[33] Alternatively, they could be treated as quantifiers heading a Quantifier Phrase. A reason to

While most of the ERC-sentences in my collection are in fact headed by a stand-alone quantifier, there are also other possibilities. A few examples with different types of indefinite relative complexes represented in Sample A are given in (104)–(107).

(104) nej inte bananer, det skalar dom inte längre, det_1 har dom **maskiner** [som
no not bananas that peel they not longer that have they machines that
gör t_1]
do
'No, not bananas. They don't peel them anymore, they have machines that do that.'

(Conversation, June 2015)

(105) det_1 finns det **inga barn** [som säger t_1]
that exist there no children that say
'There are no children who say that.'

(Conversation, Sept. 2014)

(106) [det där]$_1$ är det **många stora företag världen över** [som har gått bet
*the there is there many big companies world.*DEF *over that have failed*
på t_1]
at
'There are many big companies all over the world who have failed at that.'

(Conversation, Sept. 2014)

---

adopt this type of analysis is that the quantifiers can co-occur with more uncontroversial determiners like the definite article *de* 'the', as in (i).

(i) Ty liksom kroppen är en och har många delar och **alla de många kroppsdelarna**
*for like body.*DEF *is one and has many parts and all the many body parts*
bildar en enda kropp, så är det också med Kristus.
form one single body, so is it also with Christ
'Just like the body is one and has many parts and all the many body parts form one single body, so is it with Christ.'

(First Corinthians, 12:12, Bibel 2000)

I will overlook this more detailed split of the DP here, since it does not matter for my analysis, but see Julien (2005) for a detailed account.

(107) och [den lektionen]$_1$ skull- (.) tycker jag att det finns (.) **e- ett gäng nu**
*and that lesson.*DEF *should think I that there exist a a gang now*
aktiva politiker [som borde ta till sig av t_1] (.)
active politicians that should take to REFL *of*
'... and there are a bunch of now active politicians who should take that lesson to heart, in my opinion.'

(*Flumskolan*, podcast, April 29 2014)

In (104), the head of the relative clause is a bare plural, *maskiner* 'machines'. In (105) there is a head noun, *barn* 'children', and the DP is quantified with *inga* 'no'. Examples (106) and (107) show that there can be quite a bit of complexity in the DP in the ERC-sentences.

A special case is *alla* 'everyone/all', which occurs in three examples in Sample A. An illustration of this type is given in (108).

(108) det$_1$ är det nog inte **alla** [som blir arga eller ledsna över t_1]
that is it probably not everyone that becomes angry or sad over
'It's probably the case that not everyone will be angry or sad about that.'

(Conversation, 2011)

Notably, the matrix is negated in this example, and in the other two ERC-examples with *alla* as well. This type of sentence seems to be used to make the same types of statements as existential and presentational sentences with quantifiers, and in this particular example, the speaker was making the point that there were some people who were not going to be upset.

If the negation is removed, examples with *alla* in the head sound very strange.

(109) *det är det nog **alla** som blir arga eller ledsna över
that it is probably everyone that becomes angry or sad over

A similar effect, but in sentences without extraction, is noted by Søfteland (2013:220), who points out that *alltid* 'always' follows the same pattern, in having to be negated in the cleft sentences she investigates.

The restriction of *alla* to negated ERC-sentences is probably a reflection of the fact that they are more like *det*-clefts than existential sentences syntactically, even though they are used like existential sentences as described above. The expletive *det* in this type of sentence must be translated with *it*, rather than *there*. Furthermore, as is well known, not all types of DPs can occur in existential sentences, and the ones that can occur are those headed by **weak quantifiers**

(Milsark 1974), roughly corresponding to indefinite DPs.[34] *Alla* is a strong quantifier, which (110a) and (110b) reveal.

(110) a. *Det finns **alla** som tycker det.
there exist everyone that thinks that

b. *Det finns inte **alla** som tycker det.
there exist not everyone that thinks that

Here, negation does not make the example any better.

It is quite odd to use *alla* as a pivot in a *det*-cleft, and the effect in Swedish seems to be similar to the oddness of using *everyone* as the pivot in an English *it*-cleft out of context, as is shown in (111).

(111) *It is **everyone** that is going to school tomorrow.

This means that the fact that ERC-sentences with *alla* must be negated does not seem to have to do with the extraction per se, but follows from their cleft status.

There are also a few examples in Sample A where the DP in the construction is definite. Most of these are in cleft constructions, like in (85), (86), and (92) in section 4.3.3, and the DP is the pivot. Given the analysis of clefts that I have proposed, there is no extraction out of the definite DP in such constructions, rather, the extracted phrase moves over the pivot, which has itself raised from a position inside the cleft clause. Even though the form of the DP in the ERC-sentences thus correlates highly with the type of construction in this way, there are a few examples where the head is definite in a regular restrictive relative clause. One example is given in (112).

(112) det tycker inte JAG (.) men det_1 kanske det finns **dom** [som tycker t_1]
it think not I but it maybe there exist those that think
'I don't think so, but there may be people who do.'

(Conversation, June 2015)

Here, the head of the relative clause is the personal pronoun *dom* 'them'. This is a special construction, *det finns dom + RC* 'there are those who', which is used to say that the type of people described by the relative clause exists. The construction is described by Teleman et al. (1999 3:387), who give the example in (113).

(113) Det lär finnas de som fortfarande stöder regeringen.
*there report exist those that still support government.*DEF
'It is said that people who still support the government exist.'

[34] The restriction is often called **the definiteness restriction**.

There are a few examples like (112) in the example collection. Notably, even though the head is definite, the DP is non-specific. Apart from examples of this construction and the cleft-examples, there are only two examples with a formally definite head in my collection: (114), and (96), repeated here as (115).

(114) vi kanske ska be att få låna skräcksalen nere på konst och musik
*we maybe shall ask to get borrow horror room.*DEF *down at art and music*
där$_1$ [$_{RC1}$ *där*$_1$ man inte ser **dom**$_2$ [$_{RC2}$ Op$_2$ som t_2 sitter t_1]]
there where one not see those that sit
'Maybe we should ask if we can use the horror room down at arts and music where you don't see the students.'

(Conversation, Oct. 2014)

(115) och sen så sa dom: "Aron, du får sparken.", vilket$_1$ det inte är **första**
and then said they Aron you get the kick which it not is first
gången$_2$ [Op$_2$ jag hör nån säga t_1 t_2]
time I hear someone say
'And then they said "Aron, you're fired", which it is not the first time I hear someone say.

(Aron Flam, *Kön*, stand-up show, 2015)

In (114), the head is the personal pronoun *dom*, but it is not an instantiation of the special construction mentioned above, because it has *se* 'see' as the matrix predicate rather than *finnas* 'exist'. In (115), the head *först-a gång-en* 'first-DEF time-DEF' is definite.

In addition to confirming the general pattern that has been reported in the previous literature, my results are thus also in line with research that has suggested that there is no syntactic restriction with respect to the definiteness of the head in ERC (e.g. Andersson 1982). The reason for the large predominance of ERC-sentences with indefinite heads might be related to how ERC-sentences are used, and to the syntactic environments where ERC occurs. As we saw in 4.3, the vast majority of ERC-sentences in my collection are existential sentences. Since not all types of DPs can occur in existential sentences, as mentioned above, this means that there are only a few examples where we would expect it to be possible for the relative complex to be definite.

4.5 More about the structure of the relative complex

I have shown in previous sections that at least some ERC-sentences involve extraction from a restrictive relative clause which is attached inside the DP. An open question, however, is where inside the DP the relative clause attaches, and how relative clauses in Swedish are derived more generally. In this section, I discuss a few possibilities, namely whether Swedish relative clauses have a head-external or a head-raising structure, and whether they are adjuncts or complements. The section is divided into two parts: I first discuss head-raising in section 4.5.1, and then briefly discuss the relative clause as adjunct vs. complement in 4.5.2.[35] The focus of the discussion is, as in previous sections, whether the ERC-data commit us to any specific assumptions about the internal structure of the relative clause.

4.5.1 Head-external and head-raising relative clauses

As we have seen in chapter 2, and in this chapter, relative clauses are among the constructions which are usually argued to involved an Ā-movement dependency. It can be argued that the relative clause in Swedish ERC-sentences is created by Ā-movement as well. Consider (116).

(116) [Den där halloweenmasken]$_1$ vill jag hitta någon$_2$ [som jag kan skrämma
*the there Halloween mask.*DEF *want I find someone that I can scare*
t_2 med t_1 utan att ge _pg2 en alltför stor chock.
with without to give a too big shock
'I want to find someone who I can scare with that Halloween mask without giving them too big of a shock.'
(Lindahl 2014:12)

In this example a parasitic gap is licensed, and based on the interpretation of that gap, it is licensed by the gap in the relative clause, which can therefore be assumed to involve an Ā-movement dependency. As we see, extraction co-occurs with parasitic gap licensing here. Further evidence that the relative clause in ERC-sentences involves movement to the complementizer domain comes from examples with extraction from non-subject relative clauses like (97) and (98) in section 4.3.5, where the relative adverb *där* 'where' in each case occupies a

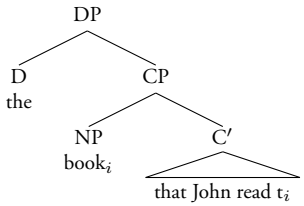
[35] The questions are partially overlapping, in that head-raising relatives are usually assumed to be complements of D or N.

specifier in the complementizer domain. I will therefore conclude that relative clauses in ERC-sentences involve Ā-movement, just like regular relative clauses (for a more thorough discussion, see Lindahl 2014).

I have followed Platzack (1999, 2000), Stroh-Wollin (2002), Vikner (2017), Nyvad et al. (forthcoming), and many others so far in representing the Ā-movement in relative clauses with a relative operator, pronoun, or adverb, and the head as originating outside of the relative clause. Recent research on relative clauses differentiates between this type of analysis, which is called a head-external analysis, and a head-raising (or promotion) analysis of relative clauses (Bhatt 2002, Hulsey & Sauerland 2006).

A simple version of a head-raising analysis is given in (117). The structure is from Hulsey & Sauerland (2006), and is not adapted to the *c*P/CP-analysis.

(117)



(Hulsey & Sauerland 2006:112)

In head-raising analyses, the head of the relative clause is assumed to originate inside the relative clause, like *book* in (117). In this particular version, the relative clause is a complement of D, and the head, which is an NP, raises to Spec-CP. Among other things, this type of analysis has been proposed to account for the fact that in English, a reflexive pronoun in the head of a relative clause can be bound by a phrase inside the relative clause, as in (118).

(118) Mary liked the picture of himself that John sent.

(Hulsey & Sauerland 2006:113)

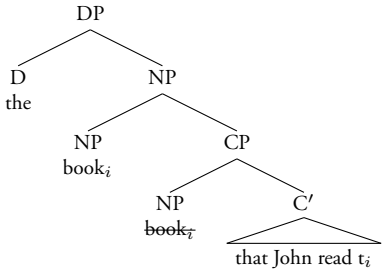
Himself is a reflexive, and interpreted as co-referent with the subject of the relative clause in this example. If the head *picture of himself* originates in the object position inside the relative clause, where it is c-commanded by the subject, we can see how the binding relation is established.[36]

A third type of analysis which is also explored in this line of research is the matching analysis, which could be seen as halfway between a head-raising and

[36] It should be mentioned that head-raising might not be necessary to account for this specific example, since it involves a picture-NP, and it has been argued that reflexives in picture-NPs are exempt from Condition A (e.g. Reinhart & Reuland 1993).

a purely head-external analysis (119).

(119)



(Hulsey & Sauerland 2006:112)

Here, the head of the relative clause is external, but there is a matching copy of it inside the relative clause, which is not pronounced. The unpronounced copy is what moves to Spec-CP.

It is often assumed by researchers who make these distinctions that several ways of deriving relative clauses can co-exist in a single language, such that both a head-raising structure and some variety of a head-external structure might be available. Head-raising is assumed, in this line of work, to be necessary for licensing reflexives and certain idiomatic interpretations of the head of the relative clause. In addition, Hulsey & Sauerland (2006) argue that extraposition blocks head-raising and is a diagnostic for the head being external.

From the perspective of ERC, these different structural possibilities are of interest in relation to a recent proposal by Sichel (to appear). Sichel argues that there is ERC in Hebrew, but only out of relative clauses derived by head-raising. In relative clauses where a head-external analysis is forced, there is no extraction.[37] A head-raising analysis is proposed for Norwegian relative clauses by Åfarli (1994). This raises the question about whether a head-raising structure is available for Swedish as well, and if so, whether we will see a similar pattern in Swedish ERC as the one Sichel finds for Hebrew. In the following sections, I consider some examples that bear on the question.

4.5.1.1 Reflexives

As we saw in the previous section, binding of a reflexive pronoun in the head of a relative clause by a subject inside the relative clause in English is taken as

[37] Sichel points out that relative clauses are weak islands in Hebrew, and that a head-raising analysis explains this, because it gives the relative clause a structure which is parallel to embedded questions in several ways. Given findings in later chapters of this dissertation, this is a very interesting observation.

evidence for a head-raising structure being available in the language. In Swedish, the evidence from the distribution of reflexive pronouns is not entirely clear. Platzack (2000:267) provides the set of examples in (120).

(120) a. *Var la du brevet från sin$_i$ lärare som Sara$_i$ fick igår?
*where put you letter.*DEF *from her.*REFL *teacher that Sara got yesterday*
'Where did you put the letter from her$_i$ teacher that Sara$_i$ got yesterday?'

b. Brevet från sin$_i$ lärare la Sara$_i$ på bordet.
*letter.*DEF *from her.*REFL *teacher put Sara on table.*DEF
'The letter from her$_i$ teacher Sara$_i$ put on the table.

c. Eva$_i$ besökte det av sina$_{i/*j}$ slott som kungen$_j$ bor i.
*Eva visited that of her.*REFL/**his.*REFL *castles which king.*DEF *lives in*
'Eva$_i$ visited the castle of hers$_i$ (*his$_j$) which the king lives in.'

d. Det av sina$_i$ slott som kungen$_i$ bor i är från 1500-talet.
*that of his castles which king.*DEF *lives in is from 16*th*-century.*DEF
'The castle of his$_i$ that the king$_i$ lives in is from the 16th century.'

In (120a), the reflexive pronoun *sin* 'her.REFLX' is not licensed in the head of the relative clause. On the other hand, the reflexive in the T-preposed phrase in (120b) is licensed, which shows that a reflexive can be bound in its trace position in Swedish. In (120c), only the reading where the reflexive is bound by Eva is possible, not the one where the relative clause-internal subject *kungen* 'the king' binds the reflexive. Lastly, (120d) reveals that it does seem to be possible for a reflexive to be co-referential with an RC-internal subject in at least some cases. Platzack offers no account of how (120d) would be derived without head-raising, but points out that if head-raising were available, we would expect (120a) to be grammatical and (120c) to be ambiguous, and uses the set of examples to argue for a head-external analysis. Another type of evidence for a head-raising structure being available is from idiom licensing, which we turn to in the next section.

4.5.1.2 Idioms

Schachter (1973), who credits an unpublished manuscript by Brame for initially bringing up this type of evidence, points out that idioms like *make headway* (121a) can provide an argument in favor of a head-raising analysis of relative clauses. He provides the set of examples in (121).

(121) a. We made headway.

b. *(The) headway was satisfactory.

c. The headway that we made was satisfactory.

(Schachter 1973:31)

Headway does not have an idiomatic meaning when it stands on its own, which is why (121b) is not acceptable, since outside of its idiomatic usage, *headway* has no meaning in everyday usage. If we take the idiomatic interpretation to be licensed by being in a local relationship with the verb *make*, a sentence like (121c) is a piece of evidence for head-raising. The idiomatic interpretation is available in this example, which indicates that there must be some local enough relationship between the licensing verb inside the relative clause and *headway* in the head. If the head originated as a complement of the verb *make*, and raised to become the head of the relative clause, we can explain this.

In his work on Norwegian relative clauses, Åfarli (1994) provides the set of examples in (122), where *ta seg vatn over hovudet* is an idiom which means 'take on too difficult or big commitments'.

(122) a. Han tok seg vatn over hovudet. (No.)
he took SELF *water over head.*DEF

b. #Vatn utviklar seg lett til alvorlige problem.
water develops SELF *easily into serious problems*

c. Vatn som ein tek seg over hovudet, utvikler seg lett til alvorlige problem.
water that one takes SELF *over head.*DEF *develops easily into serious problems*

(Åfarli 1994:86)

In Norwegian, the idiomatic interpretation of *vatn*, which otherwise means 'water', is available in (122c) and crucially unavailable in (122b), according to Åfarli.

This idiom exists in Swedish as well. However, the idiomatic interpretation of *vatten* does not seem to be available in a sentence such as (123).

(123) #Vatten som man tar sig över huvudet utvecklar sig lätt till allvarliga problem.
water that one take REFLX *over head.*DEF *develop* REFLX *often to serious problems*

I have investigated a variety of idioms, but I have found no cases where an idiomatic interpretation of a relative clause head is licensed by a verb in the relative clause. Two illustrative examples are given in (124) and (125).

(124) a. Anna la korten på bordet.
*she put cards.*DEF *on table.*DEF
'Anna laid her cards on the table.' ≈ 'Anna confessed everything/was entirely truthful.'

b. Korten förvånade alla.
*cards.*DEF *surprised everyone*
'The cards surprised everyone.' No idiomatic interpretation available

c. Korten som Anna la på bordet förvånade alla.
*cards.*DEF *that Anna laid on table.*DEF *surprised everyone*
'The cards that Anna laid on the table surprised everyone.' No idiomatic interpretation available.

(125) a. Deras hårda arbete bar frukt.
their hard work bore fruit
'Their hard work bore fruit.' ≈ 'Their hard work was worth it/gave results.'

b. Anna berömde dem för frukten.
*Anna commended them for fruit.*DEF
'Anna commended them on the fruit.' No idiomatic interpretation available.

c. #Anna berömde dem för frukten som deras hårda arbete bar.
*Anna commended them for fruit.*DEF *that their hard work bore*
'Anna commended them on the fruit that their hard work bore.' No idiomatic interpretation available.

In its idiomatic interpretation the phrase *att lägga korten på bordet* 'to lay one's cards on the table' in (124a) means 'to confess to everything' or 'to be entirely truthful'. Unlike *make headway*, the phrase has a literal interpretation, however, as does the noun *kort* 'card'. This is why (124b) is grammatical, even though it does not have an idiomatic meaning. The same is true of (124c). It is a possible sentence of Swedish, but it would have to be about some actual cards, or photographs (another meaning of *kort*).

The phrase *att bära frukt* 'to bear fruit', in (125a) has the idiomatic interpretation 'to give results' or 'to be worth it'. In (125b), however, where *frukten* 'the fruit' is not in a local relationship with the verb *bära* 'to bear', no idiomatic interpretation is licensed. Like in (124b), the example is grammatical, but Anna's praise must be about some actual fruit. The example in (124c) is odd-sounding. *Frukten* does not have an idiomatic interpretation here, and the example is odd even on its literal interpretation. The oddness is most likely due to the fact that *arbete* 'work' is not a suitable subject for the literal meaning of *bära* 'bear',

since work is an abstract concept, and *bära* in the literal sense needs a concrete subject, in particular in this example, where the preferred subject would be a fruit tree of some sort, or simply some animate subject which carries the fruit. In any case, the idiomatic interpretation of *frukten* is not licensed.

Another type of idiom is represented by (126a).

(126) a. Anna gjorde en tavla.
Anna made a painting
Idiomatic reading: 'Anna made a mistake.'

b. En tavla som Anna gjorde när hon serverade kaffet fick ödesdigra
*a painting that Anna made when she served coffee.*DEF *got fateful*
konsekvenser.
consequences
'A mistake that Anna made when she served the coffee had fateful consequences.'

Att göra en tavla, a phrase with the literal meaning 'to make a painting', can also have the idiomatic interpretation 'to make a mistake'. Interestingly, the idiomatic interpretation of *tavla* survives when it is the head of the relative clause (126b). However, this does not tell us much about head-raising, since it turns out that the idiomatic reading is not dependent on the relative clause-internal verb in this case. The idiomatic reading can be licensed without *göra* 'make', as (127) illustrates.

(127) En tavla i första halvlek fick ödesdigra konsekvenser.
a painting in first half got fateful consequences
'A mistake in the first half had fateful consequences.' (Available in a sports context)

It appears then, that in this case, the context is enough to license the idiomatic reading of *tavla* as 'mistake'. Therefore, the availability of an idiomatic reading in (126b) is not a point in favor of head-raising in Swedish.

To summarize, a verb inside the relative clause cannot be argued to license an idiomatic reading of the head noun in any of the cases I have discussed here, and consequently these cases do not provide any argument for a raising analysis. However, it is not clear what conclusions we can draw on the basis of these examples. It is possible that there is something special about the idioms I have investigated which makes the idiomatic interpretations of the head unavailable, but which is not directly linked to the availability of head-raising. A more systematic investigation would be necessary in order to find out, but is outside the scope of this study. Unless we find a suitable idiom, however, we have no

evidence from idiom licensing for head-raising, and the fact that it is hard to find such an idiom rather speaks against a head-raising structure being available for relative clauses in Swedish.

4.5.1.3 Extraposition

The last type of evidence that bears on the derivation of relative clauses that I will discuss here is data from extraposition. Hulsey & Sauerland (2006) argue that extraposition blocks head-raising, and can therefore be used as a diagnostic for a head-external or matching structure. Compare (128) and (129).

(128) a. Mary praised the headway that John made.

b. I was shocked by the advantage that she took of her mother.

(Hulsey & Sauerland 2006:114)

(129) a. *Mary praised the headway last year that John made.

b. *I was shocked by the advantage yesterday that she took of her mother.

(Hulsey & Sauerland 2006:114)

Both of the examples in (128) are grammatical and *make headway* and *take advantage* are interpreted idiomatically. In (129), an adverb has been inserted in each example in a position such that the relative clauses must be analyzed as being extraposed. In these cases, the idiomatic interpretations are blocked and the examples are ungrammatical.

Obviously the test in (129) cannot be performed in Swedish, since I have not found a suitable idiom. But if Hulsey & Sauerland's generalization about extraposition being incompatible with a head-raising structure can be transferred to other languages than English, (130) is an interesting example from the perspective of ERC.[38]

(130) Det$_1$ träffade jag många igår [som ville t_1].
that met I many yesterday that wanted
'I met many people yesterday who wanted that.'

Given Sichel's (to appear) proposal that extraction is always from a head-raising RC, the fact that sentences like (130) occur is unexpected. The adverb *igår* 'yesterday' in (130) modifies the meeting event in the matrix clause, and since

[38] For a spontaneously produced example, see (75) in chapter 5.

the relative clause appears to the right of the adverb, we can assume that it is extraposed. Yet, the extraction of *det* 'that' is entirely natural-sounding, suggesting that head-raising cannot be a precondition for extraction in Swedish.[39]

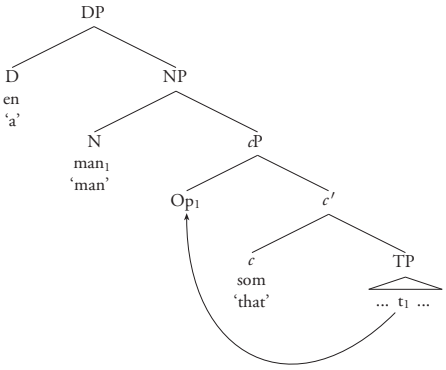
4.5.2 Complement or adjunct?

Platzack (1999, 2000) proposes that relative clauses in Swedish are complements of N, and have the structure in (131).

(131) $[_{DP}$ D ... $[_{NP}$ ... N CP]]

A motivation for this proposal is that it conforms to Kayne's (1994) antisymmetric approach to phrase structure (Platzack 2000:266). Stroh-Wollin (2002), who investigates all types of embedded clauses with *som* in Swedish, including relative clauses and embedded questions, also argues for a complement-of-N structure for restrictive relative clauses.[40] If we adapt Platzack's proposal to the *c*P/CP analysis, we will get the structure in (132) for restrictive relative clauses, and the one in (133) for non-restrictive relative clauses.

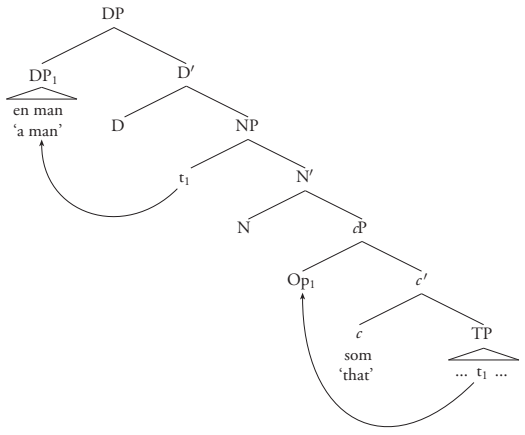
(132)



[39] I should be mentioned here that Sichel (to appear) observes that there is extraction from some extraposed relative clauses in Hebrew, and that she does not share Hulsey & Sauerland's (2006) position that head-raising and extraposition are always mutually exclusive. The argument for this is quite technical and I cannot recreate it here, but if we go down the path that Sichel (to appear) suggests, there may be a way to account for examples like (130) within her approach.

[40] For a recent proposal with extensive data about the structure of the relative complex in the Scandinavian languages, see Julien (2005).

(133)



As mentioned in the previous section, Platzack argues against head-raising in Swedish relative clauses, and assumes that an operator moves to Spec-CP in the relative clause.[41] The relative clause is a complement of N in both cases. The difference between restrictive and non-restrictive relative clauses lies in where the head is introduced. In a restrictive relative clause, the head is the N-head that takes the relative clause as a complement. In a non-restrictive relative clause, it is a full DP, and is introduced in Spec-NP, from which it moves to Spec-DP.

The complement-of-N analysis of relative clauses in the Scandinavian languages is interesting in relation to proposals about island constraints in the tradition of Huang's (1982) **Condition on Extraction Domain** (CED), where adjuncts and subjects are islands, while complements are not. If relative clauses are complements it would be less surprising that they permit extraction on such an approach. This could be taken as a point in favor of a complement-account of Swedish RCs.

However, it is not clear that the adjunct/argument distinction applies to extraction domains in Swedish, since acceptable extraction from clausal adjuncts is possible in certain cases. The example in (134) is adapted from Teleman et al. (1999 4:424).

(134) [Den här duken]$_1$ blir jag arg [om du spiller på t_1].
the here tablecloth become I angry if you spill on
'I'll get angry if you spill on this tablecloth.'

[41] As discussed in chapter 2 and in section 4.3, Platzack proposes that there is a way to derive subject relative clauses where this movement is not necessary, which would account for ERC from subject relatives, but it is not possible to extend this approach to non-subject ERC.

A recent study by Müller (to appear) confirms that extraction from at least a subset of finite adjunct clauses is acceptable to many speakers of Swedish. Müller investigats extraction from a variety of adjunct clause types in an acceptability study, and finds that extraction from both adjunct temporal clauses and adjunct purpose clauses received high acceptability ratings.

The examples provided by Teleman et al. (1999) and Müller (to appear) all involve extracting a DP, and I have not seen any reports of spontaneously produced examples with extraction of a PP, but the constructed examples in (135) are quite natural-sounding.

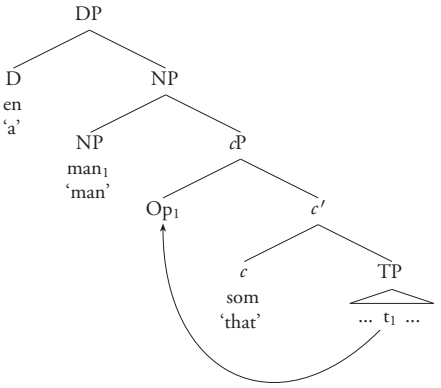
(135) a. [Från Gustav]$_1$ kan man vara glad [om man får ett livstecken om året t_1].
from Gustav can one be lucky if one gets one sign of life a year
'You can count yourself lucky if you get one sign of life a year from Gustav.'

b. [Från vem]$_1$ skulle du bli förvånad [om du fick ett Valentinkort t_1]?
from whom would you be surprised if you got a Valentine card
'Which person is such that you would be surprised if you got a Valentine's day card from that person?'

This means that the CED fails to make the right partitioning between islands and non-islands in Swedish, which in turn weakens the force of an argument for a complement account based on the CED.

A more conventional analysis is that the relative clause is an adjunct to NP, as in (136).

(136)



As an adjunct to NP, the relative clause is in a similar relation to its head as attributive adjectives are to the nouns they modify. Neither of them are selected by the noun. As far as I can tell, my investigation does not provide empirical evidence to distinguish between the complement analysis and the adjunct analysis, and since I'm not taking an antisymmetric approach to phrase

structure, I will adopt the adjunct analysis. An argument for this is that it does not require the stipulation that nouns can select relative clauses.

4.6 The position of the gap

As is evident from the ERC-sentences in this chapter, the gap in the relative clause is most often sentence-final. This is in line with the description in Teleman et al. (1999), who point out that ERC is perceived as most natural when the gap is in a relative clause in the end field of the matrix clause, and when the gap is last in the relative clause, or at least in the relative clause's end field. The end field in the schematic model of the clause that Teleman et al. use corresponds roughly to the verbal domain in the view of the clause that I assume.

Not all ERC-sentences have a sentence-final gap, however, and in this section, I briefly discuss some examples from my collection which do not. The examples are of three kinds: 1. ERC where the gap is in the end field of the RC, and the RC is itself in the end field of the matrix clause, but where there is another phrase to the right of the gap; 2. ERC where the DP with the RC is in Spec-TP, i.e. in the middle field in the model in Teleman et al.'s terminology; and 3. ERC where a subject is extracted.

The first kind is the most common. Three examples are given in (137)–(139).

(137) Hallon$_1$ är det ingen [som har t_1] här
Hallon is there no one that has here
'No one here has Hallon (as their cell phone service provider).'

(Conversation, July 2015)

(138) [De många matcherna mot Alexander Karelin]$_1$ är det många [som
*the many games.*DEF *against Alexander Karelin is there many that*
kopplar samman t_1 med Tomas Johansson].
connect together with Tomas Johansson
'There are many people who connect the many games against Alexander Karelin to Tomas Johansson.'

(Radio Sweden, article text, Oct. 19 2008)

(139) [om zinkselenid]$_1$ är det många [som forskar t_1 vid den här tiden och
about zinc selenide are there many that research at the here time and
skriver artiklar t_1], men inte om galliumnitrid
write articles but not about gallium nitride

'There are many people who do research on zinc selenide at this time, and write articles, but not about gallium nitride.'

(Radio Sweden, *Vetandets värld*, Dec. 2 2014)

In (137), the gap is followed by the locative adverb *här* 'here', in (138), it is followed by *med Tomas Johansson* 'with Tomas Johansson', which is the second argument of *koppla samman* 'connect', and in (139), it is followed by a temporal adjunct, *vid den här tiden* 'at this time'.

The second type is exceedingly rare. I have only one example of this, and it is given in (140).

(140) ... Lethe, glömskans gudinna i Grekisk mytologi men också
*Lethe forgetfulness.*DEF.POSS *godess in Greek mythology but also*
namnet på det vattendrag [vilket$_1$ under antiken sades t_1 rinna i
*name.*DEF *of the stream which under antiquity was said run in*
underjorden] och [[ur vilket]$_1$ [den [som drack t_1]] förlorade allt
Hades and out of which the one that drank lost all
minne].
memory

'Lethe, the godess of forgetfulness in Greek mythology, but also the name of the stream which during antiquity was said to run in Hades, and out of which the one who drank lost all memory.'

(*Podiet*[42], Jan 2017:47)

Here, a PP inside a relative clause modifying the subject has been relativized. The example is complicated, but sounds surprisingly acceptable, considering the description of Teleman et al. (1999), and the fact that it could be argued to involve movement out of a moved constituent, which should induce a freezing effect (Wexler & Culicover 1980).

Slightly more common is extracting the subject itself. Again, the data from my collection of naturally occurring examples conform to the pattern reported in the literature (e.g. Zaenen et al. 1981, Engdahl 1985b, 1986). Subject extraction following an overt complementizer is only possible with a resumptive pronoun, as can be seen in (141) and (142), where the resumptive pronouns are in italics.

[42] Concert magazine for the Gothenburg Symphony Orchestra.

(141) [En del av eh: dom här eh: sprängningarna och även skjutningarna]$_1$ har
a part of these here blastings and also shootings have
vi ju (.) en del utredningsuppgifter$_2$ [Op$_2$ som t_2 säger [$_{\text{CP-att}}$ att
we PRT *a part investigative information that says that*
*(*dom*$_1$) hör samman på ett eller annat sätt]]]
they belong together on one or another way

'We have some information which says that some of the bombings and shootings have something to do with each other in some way or another.'

(Radio Sweden, *Ekot*, July 26 2015)

(142) Biomedicinsk analytiker, ett yrke$_1$ [Op$_1$ som vi är många$_2$ [Op$_2$ som t_2
biomedical analyst, a profession that we are many that
lämnas frågande om [$_{\text{CP-wh}}$ vad$_3$ *(*det*$_1$) är t_3]]].
*left.*PASS *asking about what it is*

'biomedical analyst, a profession which we are many people who are left wondering what it is.'

(Radio Sweden, article text, April 16 2012)

Whatever leads to the realization of resumptive pronouns whenever *c*, or Spec-*c*P (see 142), is overt, this applies more generally in Swedish as we saw in chapter 2. If there is no overt element in the complementizer domain, subject extraction can leave a gap, as in (143).

(143) [Att bara gå rakt på jobb]$_1$ är det många$_2$ [Op$_2$ som t_2 tycker t_1 är
to just go straight on work are there many that think is
ganska tufft].
pretty rough

'There are many people who think that going straight into work is pretty rough.'

(Conversation, Aug. 2015)

In (143), the infinitival clausal subject is extracted from the subject position in a subordinate clause which is embedded in the relative clause.

4.7 Interim conclusions

In this chapter, I have outlined a view of Swedish clause structure and the left periphery of clauses that is inspired by a recent account by Vikner (2017) and Nyvad et al. (forthcoming), who propose a *c*P/CP-approach to the complementizer domain. I have also given an overview of several preposing phenomena in Swedish: T-preposing, left-dislocation, relativization, *wh*-questions, and clefts.

The *c*P/CP-analysis captures the fact that at most one specifier in the complementizer domain can be pronounced in Swedish. It also captures a distinction between V2 clauses, out of which extraction is impossible, and clauses with non-V2 word order, out of which extraction occurs. It has been proposed by Vikner (2017) and Nyvad et al. (forthcoming) that an extra specifier made available by CP-recursion provides the escape hatch needed for extraction from relative clauses.

In section 4.2, I argued that ERC involves Ā-movement to a specifier in the complementizer domain, i.e. that ERC is real extraction, and does not involve base generation and binding of a silent pronoun.

The investigation has also shown that the RC or RC-like constituent in ERC occur most frequently in existential and presentational contexts, but that there is also extraction from *det*-clefts and reverse demonstrative clefts as well as from relative clauses embedded under other verbs. These results were reported in section 4.3. It is possible that some of the *som*-clauses in ERC in existential environments are not attached inside DP, and that they could perhaps be analyzed as small clauses. But there are also examples where such an analysis is not available. This is the case both in existential environments, as the analysis in section 4.3.1 revealed, and when the *som*-constituent is embedded under another verb, which was shown in section 4.3.2. Further evidence that there is extraction from a full relative clause structure comes from ERC with non-subject relative clauses, discussed in section 4.3.5.

The definiteness of the head of the relative clauses in ERC-sentences was discussed in section 4.4, and my investigation confirms reports from the previous research: the head is almost always indefinite, and very often a quantifier like *ingen* 'no one' or *många* 'many'. ERC-sentences with definite heads are rare, but they do occur.

In section 4.5, I discussed whether a head-raising structure of the relative clause is a precondition for ERC. I considered several types of evidence that are used to argue for head-raising being available, but the results for Swedish relative clauses are inconclusive, and a head-raising account faces challenges. From the perspective of ERC, the most interesting result is that extraposition, which is taken as incompatible with head-raising by Hulsey & Sauerland (2006), does not preclude extraction. This means that Swedish might be different from languages like Hebrew, where extraction is argued to only be possible from head-raising relative clauses (Sichel to appear). However, more research is needed before we can draw any conclusions about this, since extraposition and head-raising might not actually be mutually exclusive, according to Sichel. With respect to proposals that relative clauses are complements rather than

adjuncts, I have concluded that there is nothing in my data that supports or necessitates such an analysis, and that the idea might not help us much in accounting for the possibility of ERC in mainland Scandinavian in any case, since there are other exceptions to the Condition on Extraction Domain in Swedish, as recently demonstrated by Müller (to appear).

Section 4.6 showed that the gap of extraction is usually last in the sentence, or at least in the verbal domain, but there are two exceptions. First, I have found one example with extraction from a relative clause inside a DP-subject, and second, subjects can themselves be extracted. If a subject next to an overt element in the complementizer domain is extracted, a resumptive pronoun has to occur in the gap site. This conforms to the pattern revealed in previous research about subject extractions in Swedish (Zaenen et al. 1981, Engdahl 1985b, 1986), where it is shown that Ā-bound resumptive pronouns in subject position display all of the characteristics of traces, aside from being overt.

With respect to how ERC is used and what is common, my results so far are largely in line with the description given in the *Swedish Academy grammar* (Teleman et al. 1999). There it is suggested that ERC is most acceptable if the head is indefinite and non-specific; if the matrix verb belongs to one of a few classes of verbs, e.g. of existence or presence, knowledge, observation/perception; if the relative clause is in the end field of the matrix clause; and if the gap comes last in the relative clause, or at least in its end field. However, we have also seen examples that do not conform to this description. We have seen that there are ERC-sentences with definite heads, and with matrix predicates other than the expected ones, and we have seen that the gap can be in the middle field in some cases. In each case, however, these are the less common types. Nevertheless, the examples that go against the pattern are remarkably acceptable, and they should not be seen as peripheral in the sense that we do not need to account for them.

Teleman et al. (1999) also point out that the relative clause in ERC must be restrictive. I will discuss this more in chapter 7, but I can reveal that I do not have any examples with extraction from non-restrictive relative clauses in my collection.

In the next chapter, I look at ERC in discourse and investigate ideas from the previous literature about the role of the discourse function of both the fronted phrase and the relative clause.

5. ERC in discourse

In the previous chapter, I discussed various issues about the syntax of Swedish ERC-sentences, and showed that many ERC-sentences really do involve Ā-movement from a relative clause. In this chapter, I turn to pragmatic properties of ERC-sentences, and look at how they are used in spontaneous discourse.

Previous research on ERC in the mainland Scandinavian languages has shown that the syntactic fronting involved is often motivated by discourse properties (e.g. Andersson 1982, Engdahl 1997, Lindahl 2010). For example, Engdahl (1997) notes that most examples in her collection of spontaneously occurring examples of extraction involve different types of topics, and she proposes that variation in extraction possibilities between languages might be explained by differences in discourse organization.

Erteschik-Shir (1973), Van Valin (1994), and Goldberg (2006) connect island effects to discourse properties of the constituent that extraction occurs from. The proposals concern extraction from several clause types, among them relative clauses, complements of factive verbs, complements of manner of speaking verbs, and others. As we saw in chapter 2, there are competing proposals for exactly which property matters: Erteschik-Shir (1973) argues that the embedded constituent must be dominant, Van Valin (1994) that it should be focused, and Goldberg (2006) that the extraction gap cannot be in a backgrounded part of the utterance. The accounts share the core idea that extraction is only felicitous if the extraction site is in a part of the sentence which is prominent in some respect.

There is no consensus about the role of pragmatic constraints either in constraining extraction in general or in constraining ERC in Swedish in particular. Some accounts see a role both for structural and functional constraints (e.g. Engdahl 1997, Platzack 1999), whereas others propose that functional

restrictions (perhaps in conjunction with processing constraints) can entirely replace structural restrictions (e.g. Goldberg 2006, Van Valin 1994). To my knowledge, proposals about discourse properties constraining ERC have not been tested against a corpus of spontaneously produced examples, with the exception of Engdahl (1997). With this in mind, the chapter draws on my collection of naturally occurring examples to explore what this type of data suggests for the analysis of Swedish, and the implications for theories of islands more generally.

With respect to the function of the fronted phrase in ERC, we can formulate two questions: 1. What is the discourse function of the fronted phrase, i.e. how does it relate to the utterance context? and 2. What is the information-structural function of the fronted phrase in its utterance? To address each of these questions, we need to find out whether there are stronger restrictions on fronting a phrase out of a relative clause than on fronting in general.

As for ideas about the discourse function of the embedded constituent being a way to restrict extraction possibilities, we want to find out if any of the proposed generalizations get support from the Swedish data.

In chapter 2, we saw that several syntactic mechanisms can create Ā-movement dependencies. As mentioned in chapter 3, my example collection only contains ERC-sentences with T-preposing and relativization, which means that the discussion in this chapter is restricted to these two types.

The chapter is structured as follows: I begin by giving an overview of the information-structural concepts and the concepts related to discourse functions that I use in the chapter. This is done in section 5.1. In section 5.2, I investigate how fronted phrases in T-preposing relate to the context of the utterance, and in section 5.3, I analyze their information-structural function in the clause. In section 5.4, I discuss examples of relativization out of relative clauses. Section 5.5 is dedicated to the information impact of the relative complex in ERC.

5.1 Information structure and discourse function

Information structure is intended to capture how different parts of a sentence's meaning relate to each other, and how they are to be understood to relate to the participants' information states, i.e. what they assume or take as given, at a certain point in discourse. For example, part of a sentence can be interpreted as being what the sentence "is about", and the other part can be seen as being the new information added by the sentence. Terms for this type of partitioning

include **theme-rheme**, **ground-focus** and **topic-comment** (for more detailed overviews, see Vallduví & Engdahl 1996, Gundel & Fretheim 2004).

Consider (1a) and (1b), where the capital letters indicate the nuclear accent of the sentences.

(1) a. The cat ate the HERRING.

b. The CAT ate the herring.

Both of the sentences convey the proposition 'the cat ate the herring'. But they do so in quite different ways, and they are not interchangeable in discourse. It is common to show this by juxtaposing each sentence to a question which it would be a congruent answer to, as in (2) and (3).

(2) A: What about the cat, what did she eat?
B: The cat ate the HERRING.

(3) A: What about the herring, who ate that?
B: The CAT ate the herring.

The information in B's two answers is packaged in different ways, matching A's questions. This is indicated by the intonation: the nuclear accent falls on a constituent in the part of the sentence that contains the information that A is asking for. To use the intonation in the answer in (3) to answer the question in (2) would be incongruent. The part of the sentence which corresponds to the *wh*-phrase in the question, and thus adds the information that is asked for to the current information state, is the sentence's information-structural **focus**.

Sentences expressing the proposition 'the cat ate the herring' can be used in several other contexts as well, as we see below, where the focus is marked with square brackets and a subscripted F.

(4) A: What about the cat, what did she eat?
B: The cat ate $[_F$ the HERRING].

(5) A: What about the cat, why does she look so pleased?
B: The cat $[_F$ ate the HERRING].

(6) A: What about the herring, who ate that?
B: $[_F$ The CAT] ate the herring.

(7) A: What happened?
B: $[_F$ The cat ate the HERRING].

(8) A: I wonder whether the cat ate the herring?
B: The cat $[_F$DID] eat the herring.

In (5), A is asking about the reason for the cat's contentment, and the entire verb phrase is the focus. The intonation in the answer in (5) is the same as that in (4), because the nuclear accent marks the right edge of the focus domain in English. This means that the accent will fall on *herring* in both of these answers. We see from the preceding questions, however, that their focus domains are not the same. In (7), which also shares the same intonation, the whole answer is the focus. In (8), what is emphasized is the propositional truth of the sentence. This is sometimes called verum focus.

The non-focus material in the answer sentences in each of these examples corresponds to parts of the context questions other than the *wh*-phrase. This material, which is not providing new information answering the particular question at hand, is the information-structural **ground**. Ground material is marked with a subscript G in (9)–(12).

(9) A: What about the cat, why does she look so pleased?
B: [$_G$ The cat] [$_F$ ate the HERRING].

(10) A: What happened?
B: [$_G$ After a while], [$_F$ the cat ate the HERRING].

(11) A: What about the cat, what did she eat?
B: [$_G$ The cat ate] [$_F$ the HERRING].

(12) A: What about the herring, who ate that?
B: [$_F$ The CAT] [$_G$ ate the herring].

Inside the utterance, the ground can have different functions. For instance, all or part of it can specify an entity that the rest of the sentence is saying something about, as in (9); function as a scene-setter, as in (10); or otherwise consist of known, given, or easily inferable material that the focus is adding information to, as in (11) and (12).

It will be useful to further break down ground into **topic** and other ground material. By topic, I understand roughly what the sentence is about (cf. Reinhart 1981), for example *the cat* in (9) and (11) and *the herring* in (12). The notion is further elaborated in section 5.3.1. Scene-setters, such as *after a while* in (10) and material like *ate* in (11) fall in the "other ground material" category.

The split of the ground into topic and other ground material is an utterance-level distinction; it concerns the contribution of the ground to the utterance. We can sometimes find hints in the context that help us establish which expression corresponds to the topic, as with the questions used to model the context in the examples above, which are structured so that there is a clear expectation for the answer to be about the cat or the herring mentioned in the respective contexts.

However, we will see in section 5.3 that the concept of 'utterance-level topic', or as Reinhart calls it, **sentence topic**, is largely independent of notions such as old information.

From the perspective of discourse relations, the ground can be categorized in terms of how it relates to the utterance context. To establish which part of an utterance is its focus, I made use of question-answer congruence. For topics, I will follow Erteschik-Shir (2007) who, building on Daneš (1974), distinguishes between three types of relations that can hold between a topic and a linguistic antecedent: **focus chaining**, **topic chaining**, and topics referring to part of a **hypertheme**. We see an example of the first type in (13), where the parts of the chain are in italics.

(13) There's *a girl* in the class who the teacher likes. *She* answered all the questions the teacher asked.

(Erteschik-Shir 2007:3)

The first sentence in this example is an existential sentence, a sentence of the type discussed in section 4.3.1. Since such sentences are typically used to introduce new referents, the DP *a girl* will most naturally be interpreted as part of the focus of the sentence in an out of the blue context. The topic in the second sentence is part of a focus chain, since its linguistic antecedent is a focus.

In (14), the topic is kept constant.

(14) *John* likes to read. *He* is intelligent and industrious and will go far.

(Erteschik-Shir 2007:3)

A chain is formed between the topic in the second sentence and a linguistic antecedent that is already a topic, and this is what Erteschik-Shir (2007) refers to as topic chaining.

In the third type of relation, the topic is derived from a hypertheme (15).

(15) Speaker A: Tell me about *your family*.
Speaker B: *My mother* is a teacher, *my father* works in an office, and *my sister* is a student.

(Erteschik-Shir 2007:3)

Speaker A asks about Speaker B's family, which is the hypertheme, and all of the family members become available as potential sub-themes, which are picked up in the following sentences. Such phrases are typically contrastively stressed, and function as **contrastive topics**.

The sentence's focus can also be contrastive, as in (16b), where *Anna* and *Jonathan* are contrastive topics, and *åka slalom* 'go downhill skiing' and *hälsa på*

vår faster i Stockholm 'visit our aunt in Stockholm' are **contrastive foci**. I mark contrast by capitalizing the syllable where the main stress falls.

(16) a. Vad ska dina kusiner göra på sportlovet?
what will your cousins do on sports break
'What are your cousins doing for the winter sports break?

b. ANna ska [$_F$ åka SLAlom] och JOnathan ska [$_F$ hälsa på vår FAster]
Anna will go skiing and Jonathan will visit our aunt
'Anna is going skiing and Jonathan is visiting our aunt.'

The role of contrast in ERC will be discussed in more detail in section 6.2.1.

All or part of the ground will often have a linguistic antecedent, as in (13)–(15) but it does not have to have one. To see this, we can think of examples like (17), which involves a **deictic** DP.

(17) That sandwich looks fantastic!

This sentence can be felicitously uttered out of the blue, as long as there is some sandwich present at the place of the utterance which *that sandwich* can refer to.

Erteschik-Shir (2007) also introduces a slightly different way to divide topics, which cuts across the distinction between topics with linguistic antecedents and topics which denote non-linguistic referents. Here, what is important is whether the topic is the same as in the previous utterance, or is different. Topics in topic chaining are retained from the previous sentence, and count as **continued topics**, whereas topics in focus chaining and topics derived from hyperthemes are new topics and are called **switch topics** (Erteschik-Shir 2007:10–11). We will see in section 5.2 that deictic topics can be either continued topics or switch topics.

5.2 The discourse function of T-preposing

I turn first to the discourse function of fronted phrases in T-preposing, which is the most common example type in my collection. I concentrate on the relations between the fronted phrase and the context in this section, employing a naive understanding of what a topic is, i.e. what the sentence is about, when this is necessary to discuss the phrase's relation to the context. The notion of aboutness topic is then examined in detail in 5.3.1.

In order to illustrate the relations to the discourse context that T-preposed phrases can have, I provide examples both of sentences with local T-preposing and ERC. Some of the examples with local T-preposing are from the Nordic

Dialect Corpus (NDC) (Johannessen et al. 2009), and these are marked with information about where they were recorded.

5.2.1 Fronted phrases with linguistic antecedents

Many fronted phrases in Swedish T-preposing are linked to a linguistic antecedent. As we saw in section 5.1, links between a phrase and the previous linguistic context can involve focus chaining, topic chaining, or hyperthemes.

In (18), we see **focus chaining** in a clause with local T-preposing, where a man is answering a question about what he brings on his hikes. The antecedent of the fronted phrase *det* 'that' is the underlined verb phrase in the context sentence. The antecedent is part of the focus of its utterance. The context sentence is given in English in order to save space.

(18) Some people think that you carry a lot of canned foods and other heavy things

det$_1$ gör man inte alls t_1 om man bor i stugorna
that do one not at all if one lives in cabins.def
'but you don't do that if you stay in the cabins.'

(NDC, Asby, Swe)

This type of fronting is also common in ERC. An example is given in (19), where A and B are making comments about the lyric writing of some sixties artists. The antecedent of the fronted phrase in B's utterance is the verb phrase *rimmar* rain *och* pain 'rhymes *rain* and *pain*', which is the focus of A's utterance.

(19) A: han [$_F$ rimmar *rain* och *pain*]
he *rhymes rain and pain*
'He rhymes 'rain' with 'pain'.

B: det$_1$ är det många [som har gjort t_1]
it are there many that have done
'There are many people who have done that.'

(Conversation, July 2011)

It is also possible for the fronted phrase to link back to just a part of the focus of a previous utterance, as in (20), where different phonological theories are discussed.

(20) det [$_F$ påminner om den teorin] och [den teorin]$_1$ är det väl inte alla
it reminds about that theory and that theory.def is it PRT *not everyone*
[som tror på t_1]
that believes on

'It is reminiscent of that theory, and not everyone believes in that theory.'

(Conversation, Nov. 2011)

The focus of the previous utterance is the entire verb phrase in this example as well, but the antecedent of the fronted phrase is only the DP *den teorin* 'that theory'.

Phrases involved in **topic chaining**-relations can also be fronted both in local T-preposing, as we see in (21), and in T-preposing from relative clauses, as in (22) and (23). Consider first (21).

(21) [$_F$... and her husband, who was the father of the child], he was on the front and had been involved in some type of sabotage.

och honom$_1$ hade de inte fått tag i t_1.
and him had they not got hold of

'and they had not got hold of him'

(NDC, Köla, Swe)

Honom 'him' forms a chain with *he* in the preceding sentence, which is its topic. The relation between *he* and the first expression with the same referent, *her husband*, is a focus chain.

The example in (22) is from a conversation about traveling, and the possibility of going on a roadtrip across the United States has just been mentioned.

(22) ja det är häftigt! det$_1$ vet jag en [som har gjort t_1]
yes that is really cool that know I one that has done

'Yes, that is really cool! I know someone who has done that.'

(Conversation, Aug. 2015)

In the first sentence of this example, the antecedent 'going on a roadtrip across the US' is made a topic in an instantiation of focus chaining, and in the second sentence, the one involving T-preposing from a relative clause, a topic chain is formed with the initial pronoun.

The dialogue in (23) took place at the dinner table. A suggests that the bread should be put on a plate, but B thinks that this is unnecessary, while admitting that some people might not agree.

(23) A: Shouldn't the bread be on one of these? (Shows a small plate)

B: det tycker inte JAG (.) men det$_1$ kanske det finns dom [som tycker t_1]
it think not I but it maybe there exist those that think
'I don't think so, but there may be people who do.'

(Conversation, June 2015)

Notably, no contrastive interpretation of the fronted phrase needs to be evoked in these examples for the fronting to be felicitous, which is in line with findings in Engdahl & Lindahl (2014). In an investigation of fronted non-subject pronouns in the mainland Scandinavian languages, we found that such pronouns can be non-contrastive in both topic chaining and focus chaining.

The third type of thematic progression described by Erteschik-Shir (2007) involves **hyperthemes**. This type of link between a fronted phrase in ERC and the context is also possible, as we can see in (24). In this conversation, A and B are talking about a small information film about people at the university, which B is going to be in. During the preparations for this, B got to answer some questions about herself, and A asks what B said about her previous work experience. B has been both a teacher and a UN soldier, and A thinks that she should have told them about being a UN soldier rather than about being a teacher, since teaching experience is more common.

(24) A: What did you write about previous experience?
B: Well ... that I've been a teacher.

A: men FN-soldat är bättre (.) LÄRare$_1$ är det många [som är t_1] här
but UN soldier is better teacher are there many that are here
'But UN soldier is better. There are many people who are teachers here.'

(Conversation, Spring 2015)

A's question raises the hypertheme 'previous work experience'. 'Teacher' and 'UN soldier' are two sub-themes which become available. As mentioned, the sub-theme that is chosen is contrasted against the other potential sub-themes in the set, and receives contrastive stress. The extraction sentence in (24) was uttered with contrastive stress on *lärare* 'teacher'.

A special type of linguistic antecedent is the hanging topic phrase in left-dislocation structures. As we saw in section 4.1.5, the pronoun which is bound by the hanging topic phrase in such structures is usually T-preposed. An example with local T-preposing is given in (25).

(25) Sometimes we got some small thing, but we didn't have the money to buy a lot.

båtar$_x$ [det$_x$]$_1$ gjorde jag ju själv t_1 utav bark och träd och så där
boats it made I PRT myself out of bark and trees and such

'Boats, I made that myself out of bark and trees and such.'

(NDC, St Anna, Swe)

The speaker in (25) is talking about growing up, and 'toys' is a hypertheme. The hanging topic phrase picks one potential sub-theme from this hypertheme, *boats*, and the bound pronoun is then T-preposed and serves as a continued topic, referring back to the hanging topic. There are several such examples in my collection of spontaneously produced ERC-sentences.[1] Consider (26).

(26) A: nej nu vill jag vända blad
no now want I turn page

B: [vända blad]$_x$ [det$_x$]$_1$ är det många kända svenskar [som har gjort t_1]
turn the page it are there many famous Swedes that have done

'Turn the page, there are many famous Swedes who have done that.'

(Conversation, Jan. 2015)

In this dialogue, A wants to change the topic of conversation. The idiom *vända blad* 'turn the page' means roughly the same thing as the corresponding expression in English, but it has a special import here, as at the time of utterance it had been used a few years earlier by the King of Sweden when he was addressing allegations about having had an affair.

In addition to having an antecedent in the previous linguistic context, a pronoun can in some cases precede its linguistic correlate. Such **cataphoric** pronouns can also occur in the initial position in declarative main clauses. In (27), we see an example of this, which is from an interview where the speaker is telling a story about how he was once saved from a falling branch by his helmet.

[1] That the pronoun fronting which commonly takes place in left-dislocation structures can violate the CNPC is also pointed out by Andersson (1982).

(27) dom är ju (.) [det$_x$]$_1$ vill jag säga t_1 (.)
they are PRT it want I say
[gå i skogen utan hjälm det gör jag inte]$_x$
*walk in forest.*DEF *without helmet it do I not*

'They are ... This I'll say, I don't walk in the forest without a helmet.'

(NDC, Axtorp, Swe)

T-preposing of cataphora is not common in my collection of ERC-sentences, but it does seem to be possible, and we see an instance of this in (28). The example is from a morning news show on the radio, from the introduction to a segment about the mink farming industry.

(28) We turn now to mink breeding. According to the organization Swedish Mink, the mink industry will grow between twenty and twenty-five percent this year, thanks to the higher profitability of exports of mink fur to China, among other places.

Men [det här]$_1$ är det inte precis alla [som jublar över (.) att man kan göra
but the here is it not precisely all that cheer over that one can make
pengar på t_1] (.) [hur minkarna har det på farmarna]
money on how the minks have it on the farms

'But not everyone is cheering about the fact that you can make money on this ... how the minks are doing at the farms.'

(Radio Sweden, P1 *Morgon*, July 23 2013)[2]

The phrase *det här* 'this' has been T-preposed from a position inside a *that*-clause embedded in the relative clause in this example, and is cataphorically linked to the underlined question which follows the small pause indicated by the (.)-notation.

[2] The example is interesting also from the perspective of incremental processing. After hearing the preposition *över* there is a possibility to interpret the fronted phrase as its complement. Then it would have to have *att det blir mer lönsamt att exportera mink till Kina* 'that it will become more profitable to export mink fur to China' as its antecedent. However, as the utterance continues, the intonation makes it clear that this is not the intention of the speaker. Instead, the *that*-clause that follows *över* is its complement, and the underlined phrase *hur minkarna har det på farmarna* ('how the minks are doing at the farms') has to be what the fronted phrase is cataphorically linked to.

5.2.2 Expansions

The T-preposed phrases in the previous section all have a direct relation to the linguistic context, but sometimes the relation between the fronted phrase and the context is more indirect. In (29), we see a prepositional phrase that has been extracted from a cleft clause in a *det*-cleft.

(29) There is a war on the natural rhythm of human existence. We cut all rest from our schedules and <u>burn ourselves out</u>.

[från **ett sånt haveri**]$_1$ är det bara långtråkigheten [som kan rädda oss t_1]
from a such breakdown is it only boredom-DEF *that can save us*

'Only boredom can save us from such a breakdown.'

(Radio Sweden, *Allvarligt talat*, July 31 2015)

The DP-complement of the preposition in this example, *ett sånt haveri* 'such a breakdown', has a relationship to the linguistic context: it is in a focus chain with the previously mentioned burnout. The preposition is not related to the discourse context, however, and is simply pied-piped by the T-preposed DP.

It is quite common to T-prepose an entire embedded clause in ERC. In such examples, the preposed clause often contains an expression that links back to a linguistic antecedent, and the rest of the embedded clause raises a new issue with respect to the referent of the linking expression. We see an example of this in (30), where an embedded question has been T-preposed out of a relative clause. The linking expression is in bold.

(30) The Ministry of Education and Research has awarded what corresponds to <u>14 million Swedish crowns</u>, maybe even more, to projects which Moriguchi was involved in.

[vart **pengarna** gått]$_1$ är det ingen (.) utom möjligen Moriguchi (.) [som
where money-DEF *went is there no one except maybe Moriguchi that*
vet t_1].
knows

'No one, except maybe Moriguchi, knows where that money went.'

(Radio Sweden, P1 *Vetenskapsradion*, Oct. 19 2012)

Here, a case of research fraud is discussed, and 14 million Swedish crowns which have been given to the fraudulent project are mentioned in the context. In the sentence with ERC, the T-preposed phrase raises a question about this

money: 'Where did it go?' The rest of the sentence asserts that no one knows the answer to the question.

It is also possible to front an *att*-clause out of a relative clause, as is seen in (31).

(31) I was not aware that almost all unspoiled forest was being chopped down at the same time

[att det bara skulle finnas **några ynka procent** kvar när jag själv blev
that there only would be some measly percent left when I self became
vuxen]$_1$ var det ingen [som sa t_1]
grown up was there no one that said

'No one told me that there would only be a few measly percent left when I was grown up.'

(Radio Sweden, *Tankar för dagen*, Nov. 29 2011)

In this example, the DP *några ynka procent* 'some measly percent' in the T-preposed clause links back to the unspoiled forest mentioned in the context.

A distinct type of ERC-sentence involves fronting of a quote, as in (32). This example is from a guided tour of a porphyry mill in Älvdalen, in which the guide is telling a story about Frost Anders, who founded the mill.

(32) [länsman hade ju mött **honom** på bron härborta]$_1$ är det en
rural police chief had PRT *met him on bridge-*DEF *over here is there a*
farbror från Brunnsberg [som har berättat t_1].
old man from Brunnsberg that has told

'"The rural police chief had met him on the bridge over there", an old man from Brunnsberg has told us.'

(Guided tour, Älvdalens Nya Porfyrverk, June 26 2015)

Honom in the example refers to Frost Anders, and the rest of the fronted quote says about him that he was seen by the rural police chief on the bridge. The ERC-sentence then adds information about who said this. In a study of ERC in newspaper texts (Lindahl 2010, 2011), I found that almost half of the ERC-sentences in those texts involved fronted embedded clauses.

Another type of expansion is found in (33). The example involves fronting of a degree phrase, and is from the radio show *Språket* 'The Language'. The adjective phrase *vanliga* 'common' is used in the context as a description of the words that will be the theme of the next segment of the show. In the sentence involving ERC, the phrase is expanded with the addition of *alltför* 'too'.

(33) först ska det handla om ett par små vanliga ord
first shall it be about a pair small common words

[ALLTför **vanliga**]$_1$ är det många [som tycker t_1 (.) eller tänker t_1]
too common are there many that find or think

'First, we will talk about a couple of small, common words. There are many who think that they are too common.'

(Radio Sweden, *Språket*, May 22 2012)

This example is different from the others discussed in this section, because it is not clear that it involves any potential topic. There is a link to the context – *vanliga* is given in the sense that it is in the immediate linguistics context – but the link is not between a topic referent and a linguistic antecedent. It is possible that the preposing in this example is licensed by contrast. I will return to the role of contrast in section 6.2.1, and to the question of topics in 5.3.1.

5.2.3 Deictics and demonstratives

In the examples discussed so far, the T-preposed phrases have had linguistic antecedents, or, in the case of cataphora, a linguistic correlate later in the utterance. This is not a necessity for fronted phrases in Swedish, however. Phrases with **deictic** reference can occur in the initial position in declarative clauses both in local T-preposing, as in (34), and in ERC, as in (35) and (36).

(34) Usch, [den där]$_1$ gillar jag verkligen inte t_1.
ew the there like I really not

'Ew, I really don't like that one.'

As a context for (34), we can imagine that the speaker is pointing, for example, at an unpleasant-looking statue. The example in (35) is from a short information film about things to do in your spare time in a small Swedish town.

(35) Both in winter time and summer, the energetic gliders practice their wonderful sport. And it's easy to see that the expression "air-minded" is held high. (A plane lands)

[den här perfekta punktlandningen]$_1$ finns det många segelflygare
the here perfect point landing exist there many gliders
landet runt [som beundrar t_1]
country-DEF *around that admire*

'There are many gliders all over the country who admire this perfect point landing.'

(Bo Löfberg, *Förnuftig fritid*, 1946)

The phrase *den här perfekta punktlandningen* 'this perfect point landing' refers to a sequence where a plane is landning.

The speaker uttering (36) was clearing the table in order to set it for dinner. The sentence was uttered as a bowl of sugar was picked up and removed from the table.

(36) SOCKER$_1$ är det nog ingen [som vill ha t_1]
sugar is there PRT no one that wants have
'There's probably no one who wants sugar.'

(Conversation, June 2015)

The topics in (35) and (36) are new in the sense that they were not already the topic of a previous utterance. This means that they are switch topics. Deictics can also be continued topics, as (37) shows.

(37) A: But chairs like the one you're sitting on are usually comfortable too.
B: den här är jätteskön (.) men [den här]$_1$ var det ingen [som tog
the here is very comfortable but the here was there no one that took
med sig t_1] va?
with REFLX PRT
'This one is really comfortable. But no one brought this one, right?'

(Conversation, Sept. 2015)

Here, A is a visitor in an apartment where B and several other people live. A group of people are sitting on an assortment of chairs. The merits of the different armchairs and how they were obtained are discussed.

As we have seen, many of the deictic examples involve demonstrative expressions, like *den här* 'this', *den där* 'that' or *sådana*, *sådant*, *sånna*, *sånt* 'such'. Swedish also has a demonstrative adverbial expression, *så* 'like that', which can also be T-preposed, as in (38).

(38) Enkelt: Frys ägget och såga det sedan i två delar. Så$_1$ gör jag jämt t_1.
*easy feeze egg.*DEF *and saw it then in two parts like that do I always*
'Easy: Freeze the egg and split it in two parts. That's what I always do.'

(Forum thread, Dec. 2009)

The example is from a cooking forum, and is an answer to a question about how to cut a pancake recipe with three eggs in half.

Så can also be fronted out of a relative clause, as we see in (39).

(39) “efter sitt lilla äventyr var hon ...” (.) nej *så*$_1$ är det ingen [som säger t_1
after her little adventure she was *no like that is there no one that says*
i mitt material]
in my material
‘ “After her little adventure, she was ...” No, there is no one who says things like that in my data.’
(Conversation, May 6, 2011)

This example is from the question period after a conference presentation, and *så* is used to link back to the previous utterance which exemplifies a way of speaking, or type of thing to say, which is not represented in the data that the presentation was based on. In (38) and (39), there is a linguistic antecedent, which is not uncommon with demonstratives, even though they have the potential to be used deictically.

5.2.4 Other types

The most common patterns with respect to relations between the discourse and the T-preposed phrase in ERC have been described above, but they do not exhaust the possibilities. In this section, I describe some some other possible discourse functions of T-preposing.

First, there are some examples where the T-preposed phrase has a linguistic antecedent, but not in the immediate context. This is the case in (40).

(40) ... men som sagt [något omedelbart hot för eh: kärnvapenKRIG]$_1$ är det
but as said some immediate threat for eh nuclear war are there
få [som tror på t_1] här ...
few that believe in here
‘ ... as mentioned, there are few people here who believe there is any immediate threat of a nuclear war ... ’
(Radio Sweden, P4 *Extra*, Apr. 9 2013)

In this example, the T-preposed phrase is an indefinite noun phrase *något omedelbart hot för kärnvapenkrig* ‘any immediate threat of a nuclear war’. The example is from a segment of a news show, where the host of the show is interviewing a correspondent in Asia. The threat of a nuclear war is not mentioned in the context immediately preceding the example, but a few turns back in the interview, threats about nuclear activity from North Korea and the possibility of a nuclear war are mentioned several times.

Second, in examples with bridge anaphora, world knowledge is necessary to establish the relation between the T-preposed phrase and the discourse. Erteschik-Shir (2007) gives the example in (41), where knowledge about the

world is used to derive the topic *the composer* from the previously mentioned *concert* (41).[3]

(41) John heard a beautiful concert. *The composer* directed it.

(Erteschik-Shir 2007:18)

Similar examples are found in Swedish ERC. Consider (42).

(42) A: I think everyone is at home consuming all the internet because of the wind.
B: How come it works for me, then?

A: Hallon$_1$ är det ingen [som har t_1] här
Hallon is there no one that has here

'No one here has Hallon (as their cell phone service provider).'

(Conversation, July 2015)

The context of this example is that A, B and some other people are on an island. Most of the people, including A, cannot access the internet on their cell phones. The T-preposed phrase *Hallon* is the name of a cell phone service provider, and the topic is derived from the mention of the internet. Specifically, A assumes that the choice of service provider could affect the internet access and he knows that B's service provider is Hallon, which he furthermore assumes is rare.

In addition to the types of relations between the fronted phrase and the context that we have seen so far, it is also possible for the fronted phrase to provide the information that has been asked for in a context question. I have one such example in my collection, which is given in (43).

(43) Radio show host: Wait, what was your idea?

Caller: [att det skulle NAPPA väl i (.) skymningen eller nåt sånt
that it would bite well in the dusk or something such
där till exempel]$_1$ var det nån [som hade en teori kring t_1]
there for example was there someone that had a theory around

'Someone suggested that it might mean that it would bite well in the dusk or something like that for example.'

(Radio Sweden, *Språket*, Sept. 9 2015)

The fronted phrase here is the information focus of the utterance, and I will

[3] This example may also lend itself to an analysis where *concert* is a hypertheme, introducing a set of concert-related concepts as sub-themes, but deriving this set is not as straightforward as in the case of the family–family members that we saw in (15) in section 5.1.

come back to such examples in section 5.3.2, where I discuss the information-structural role of the T-preposed in its clause.

5.2.5 Summing up

To give an idea of how common the different types of discourse functions for the fronted phrase are, I have annotated the 100 examples in Sample B. As described in chapter 3, this is a sample from the larger collection which consists of 100 ERC-sentences with T-preposing, all from spoken language. The results are summarized in table 5.1. The example numbers in this table refers to examples of the respective type discussed in this chapter.

Table 5:1. *Discourse functions of T-preposed phrases*

Discourse relation	No of instances	Example
Focus chaining	45	(19)–(20)
Topic chaining	15	(22)–(23)
of which left-dislocation	*8*	(26)
Hypertheme	8	(24)
Cataphor	1	(27)
Expansion	13	(29)–(33)
Deictic	5	(35)–(36)
Other	5	(40)–(43)
Unclear	8	(44)
Sum	100	

I have only categorized as deictic the examples where the T-preposed phrase denotes something in the non-linguistic context, i.e. in the spatial context where the conversation takes place. Examples with T-preposed demonstratives which have a linguistic antecedent have been categorized based on the relation to this antecedent.

The category "unclear" in table 5.1 comprises examples where it is not possible to establish the discourse relation with any certainty. In one of these examples, the phrase that must be analyzed as T-preposed, given the gap inside the relative clause, is also part of the focus of a previous sentence, i.e. a case of *apokoinou* (cf. Norén 2007). The example, (44), is from a radio segment about mushrooms. The phrase *lite problem med magen och lite sånt* 'some stomach problems and such' functions both as the object of the verb *få* 'get' in the previous sentence, and as the T-preposed phrase in the extraction sentence.

(44) A: Are they poisonous?

B: ja dom är så pass att du kan få [lite problem med magen och
yes they are enough so that you can get a little problem with stomach and
lite sånt]$_1$ har jag hört någon [som har fått t_1] [som har av misstag
a little such have I heard someone that has gotten that has accidentally
ätit det]
eaten it

'Yes, they are poisonous enough that you can get a bit of indigestion, I've heard people say who have eaten it by mistake.'

In the other unclear cases, either the context is missing and the ERC-sentence in and of itself does not provide enough clues to establish the discourse function of the fronted phrase, or there is a context, but it is unclear which phrase is the antecedent. All in all, the investigation has shown that the T-preposed phrase in ERC can have the same discourse functions as the fronted phrase in local T-preposing.

5.3 The information-structural functions of T-preposing

I turn now to the information-structural relations that hold within the utterance in ERC-sentences. The aim of this section is to explore which information-structural functions the fronted phrase in T-preposing can have. In connection to the overall aim of the thesis, an important goal is to find out whether there are information-structural functions that a T-preposed phrase can have in local T-preposing that it cannot have in T-preposing with ERC.

In section 5.2, I call several of the T-preposed phrases topics. That topics should be common in ERC is not surprising. Engdahl (1997) reports that most extractions from relative clauses in her investigation involve some type of topic, but that extracting a phrase that is the information focus of the sentence from a relative clause is also possible. In local T-preposing, topics are also common, but here, the initial position is also often filled by a scene-setting adverbial (see e.g. Andréasson 2007). With respect to constraints on extraction, then, we can ask whether T-preposing in ERC always involves fronting topics or foci, or whether we ever find phrases that function as scene-setting adverbials being T-preposed from relative clauses.

I will begin by examining the notion of topic. As shown by Vallduví & Engdahl (1996) and Gundel & Fretheim (2004), there is considerable disagreement about most information-structural concepts, and topics are no exception. They

have been defined, for example, as old, given, familiar, what the sentence is about, the pivot of truth value assessments, the referent of the phrase that comes first in the sentence, or as exhibiting some combination of these properties. I will explore the view proposed by Reinhart (1981), that a topic is what an utterance is about.

5.3.1 Aboutness topics

In her article about sentence topics, Reinhart (1981) argues that topic is best conceived of as "the expression whose referent the sentence is about" (Reinhart 1981:57). A proposition, as we have seen, can be packaged in several ways, and which referent is interpreted as being what the sentence is about depends on this packaging, and on the previous discourse. For this reason, aboutness cannot be a semantic notion, but has to be pragmatic, Reinhart argues (p. 58). Following Stalnaker (1978), she conceives of a discourse as the conversation participants' "joint-procedure of constructing a context set". Stalnaker (1978) defines the context set of a given discourse at a given point as the set of possible worlds compatible with the propositions that the discourse participants pragmatically presuppose at that point: slightly simplified, the propositions that they assume to be true, or at least act as if they assume to be true for the purpose of the conversation.

For Reinhart, the context set is not an unstructured set of worlds or propositions, but has internal organization. As a metaphor for this structuring, she uses a library catalogue metaphor, where each proposition corresponds to a book entry. The entries in a library catalogue can be organized and accessed in several ways. For example we can look at authors' last names or book titles in alphabetical order. But this is not very useful if we want to find out which books on a certain subject are available. To do this, we would rather access the subject catalogue.

Reinhart's idea is that in constructing a context set, speakers are likely to organize the propositions in a way that is useful for evaluating upcoming propositions, and that makes the information easier to remember. It would be reasonable, from such a perspective, to organize the propositions in a way that resembles the subject catalogue in a library, and for speakers to classify propositions that are admitted into the context set into subsets, stored under defining entries. In Reinhart's model, topics are precisely such entries (p. 79–80).

Reinhart uses a rewrite test to identify topics. To see if a certain phrase can serve as the topic of a sentence, one places the sentence in an *about*-sentence: *He said about* {*DP*} *that* {*YP*} (Reinhart 1981:65). The resulting sentence will

only be a natural paraphrase if the DP following *about* can be its topic.

Using this test to evaluate the cat-and-herring-sentences from section 5.1, we see that the ground in (9), *the cat*, seems to be the topic of that sentence. The example is repeated here as (45a), and the paraphrase in (45b) is perfectly natural. In fact, *the cat* seems to be the topic in (11), here repeated as (46a), as well, which supports the split of the ground into topic and other ground material that was proposed in section 5.1.

(45) a. A: What about the cat, why does she look so pleased?
B: [$_{G}$ [$_{TOP}$ The cat]] [$_{F}$ ate the HERRING].

b. B said about {the cat} that {she ate the herring}.

(46) a. A: What about the cat, what did she eat?
B: [$_{G}$ [$_{TOP}$ The cat] ate] [$_{F}$ the HERRING].

b. B said about {the cat} that {she ate the herring}.

In (45) and (46), the topics are initial, but the test supports the idea that topics do not have to be realized initially: in (47), the test indicates that *the herring* is a natural topic in B's answer in (47a).

(47) a. A: What about the herring, who ate that?
B: [$_{F}$ The CAT] [$_{G}$ ate [$_{TOP}$ the herring]].

b. B said about {the herring} that {the cat ate it}.

c. #B said about {the cat} that {it ate the herring}.

Placing the initial phrase *the cat*, which is the sentence's information focus, in an *about*-sentence in this context yields an infelicitous result (47c).

Many of the fronted phrases in my ERC collection are aboutness topics according to Reinhart's test. Take for example (20), here repeated as (48a).

(48) a. det påminner om den teorin och [den teorin]$_{1}$ är det väl inte
it reminds about that theory-DEF *and that theory*-DEF *is it* PRT *not*
alla [som tror på t_{1}]
everyone that believes on
'It is reminiscent of that theory, and not everyone believes in that theory.'

b. Det påminner om den teorin, och Anna sa om {den teorin} att {det
it reminds about that theory, and Anna said about that theory that it
inte är alla som tror på den}.
not is everyone that believes in it.
'It is reminiscent of that theory, and Anna said about that theory that not everyone believes in it.'

In (48b), the example has been placed in an *about*-sentence, and the result is felicitous.

The fronted phrase in T-preposing ERC is often the pronoun *det* 'it', which is often used to refer to events and properties in addition to neuter referents, and is usually a natural topic. A representative example is (19), repeated in (49a).

(49) a. A: han rimmar rain och pain
he rhymes rain and pain

B: det$_1$ är det många [som har gjort t_1].
it are there many that have done

b. A sa att sångaren rimmar rain och pain, och B sa om {det} att
A said that singer-DEF *rhymes rain with pain and B said about that that*
{det är många som har gjort det}.
there are many who have done that
'A said that the singer rhymes 'rain' with 'pain', and B said about that that many people have done that.'

Again, the *about*-sentence is natural-sounding (49b).

Reinhart (1981) restricts her discussion to DP-topics, but she suggests that her analysis can be extended to other topic expressions as well. In doing so, we will have to amend the application of the *about*-test. An amendment of the test is also necessary when a DP-topic pied-pipes a preposition. As we have seen in 5.2.2, a topic in ERC can pied-pipe a preposition, as in (29), repeated here as (50a).

(50) a. There is a war on the natural rhythm of human existence. We cut all rest from our schedules and burn ourselves out.
[från ett sånt haveri]$_1$ är det bara långtråkigheten [som kan rädda oss t_1]
from a such breakdown is it only boredom-DEF *that can save us*
'Only boredom can save us from such a breakdown.'

b. Han sa om {ett sånt haveri} att {det bara är långtråkigheten som
he said about a such breakdown that it only is boredom-DEF *that*
kan rädda oss från det}.
can save us from it
'He said about such a breakdown that the only thing that can save us from it is boredom.'

The *about*-test is only suitable for DPs, for syntactic reasons that have nothing to do with the topichood of the constituent, and for this reason, the preposition which is pied-piped in (50a) must stay in situ in the *about*-sentence in (50b).

With this adjustment, the test can be applied, and confirms that the DP can be seen as an aboutness topic, where the sentence is asserting something about a type of breakdown, namely being burnt out. I will discuss the kind of topic which refers to a type more in 5.3.1.2.

Another way to adjust the test is to see if it is possible to rewrite the potential topic expression as a DP with the same denotation, and see if an *about*-sentence about this DP works as a paraphrase in the context. In (51), I have rephrased (39), which involves a fronted adverbial. A paraphrase with *så* in the *about*-sentence would not work, which is expected, but the version in (51b) is a reasonable paraphrase and sounds acceptable, although perhaps not perfect.

(51) a. "efter sitt lilla äventyr var hon ..."
after her little adventure she was

nej så$_1$ är det ingen [som säger t_1] i mitt material.
no like that is there no one that says in my material

'There is no one who says things like that in my data.'

b. $^{OK/?}$Hon sa om {det uttryckssättet} att {det inte var någon som sa så i hennes material}.
*she said about that way of expressing oneself.*DEF *that there not were someone who said so in her material*

'She said about that way of expressing oneself that no one in her data expressed themselves in that way.'

In some cases, modification of the test is needed even though the T-preposed phrase is a DP. One example of this is when a proper name is used to denote the name itself. Unmodified in *about*-sentences, names will quite naturally be taken to denote an individual rather than that individual's name. Consider (52).

(52) a. It has to be the right list because ...

Idun$_1$ har vi inte många [som heter t_1]
Idun have we not many that are named

'we don't have many people named Idun.'

b. $^{?}$Han sa om {Idun} att {vi inte har många som heter så}.
he said about Idun that we don't have many who are named that

c. Han sa om {namnet Idun} att {vi inte har många som heter så}.
*he said about name.*DEF *Idun that we not have many who are called that*

'He said about the name Idun that we don't have many people named that.'

(Conversation, Spring 2015)

The speaker in this example is trying to find a particular class list among other lists. Finding the name *Idun* on the list tells him that he has found the right one, since there is only one student named Idun in the department. *Idun* in the extraction sentence is clearly meant to be interpreted as 'the name Idun'. The *about*-sentence in (52b), however, is odd, because intuitively it sounds as if it is about a person. To make the sentence a good paraphrase, we need to make it explicit that it is about the name, as in (52c).

Similarly the *about*-sentence based on (24), repeated here as (53), sounds slightly better if we rephrase the predicate nominal *lärare* 'teacher' as *yrket lärare* 'the profession teacher'.

(53) a. A: What did you write about previous experience?
B: Well ... that I've been a teacher.

A: men FN-soldat är bättre (.) LÄRare$_1$ är det många [som är t_1] här
A: but UN soldier is better teacher are there many that are here

'But UN soldier is better. There are many people who are teachers here.'

b. ?A sa om {lärare} att {det är många som är det här}.
A said about teacher that there are many that are it here

c. A sa om {yrket lärare} att {det är många som har det här}.
*A said about profession.*DEF *teacher that there are many that have it here*
'A said about the teaching profession that many people have it here.'

(Conversation, Spring 2015)

Since the expression is *ha ett yrke* 'have a profession', we need to change the verb in the YP-clause.

As we have seen so far several of the T-preposed phrases in my ERC-sentence collection can be identified as aboutness topics using Reinhard's rewriting test, if we allow ourselves to be a bit flexible in applying it. In some cases minor modifications are needed. In the next section, we will turn to the information-structural function of the T-preposed phrase in left-dislocation structures, which I have already mentioned briefly in section 5.2.1.

5.3.1.1 Left-dislocation

In left-dislocation, the hanging topic will naturally be interpreted as what the sentence is about (54a).[4]

[4] The bound pronominal copy in the left-dislocation sentence occurs in situ in the *about*-sentence here, but it could also be preposed, as in (i).

(54) a. [Den där smörgåsen där]$_x$, [den$_x$]$_1$ skulle jag gärna äta upp t_1.
the there sandwich-DEF *there that would I gladly eat up*

'That sandwich over there, I would love to eat it.'

b. Jag sa om {den där smörgåsen där} att {jag gärna skulle vilja äta
I said about the there sandwich-DEF *there that I gladly would want eat*
upp den}.
up it

'I said about that sandwich over there that I would love to eat it.'

The phrase that is extracted from the relative clause in ERC with left-dislocation is the pronoun, as mentioned in previous sections. In (55a), we see an example of this. To be able to try the *about*-test, I have rephrased the *att*-clause, which is the hanging topic here, as a DP in (55b).

(55) a. [att fjällämlar som är trängda både kan skälla och gå till
that Norway lemmings who are in a tight spot both can bark and go to
attack]$_x$, [det$_x$]$_1$ är det många [som har erfarenhet av t_1]
attack that is there many that have experience of

'Many people have experienced that Norway lemmings that are in a tight spot can both bark and attack.'

b. Anna sa om {det att fjällämlar som är trängda både kan
Anna said about that that Norway lemmings who are in a tight spot both can
skälla och gå till attack} att {det är många som har erfarenhet av det}.
bark and go to attack that there are many people who have experience of it

'Anna said about the fact that Norway lemmings that are in a tight spot can bark and go to attack that there are many people who have experienced that.'

(Radio Sweden, *Naturmorgon*, June 4 2011)

The *about*-sentence is natural-sounding, which indicates that the *att*-clause, which is the hanging topic in the ERC-sentence, is an aboutness topic. This makes it quite natural to analyze the preposed pronoun in the ERC-sentence as a continued topic, forming a topic-topic chain with the hanging topic. As in (54b), the pronoun occurs in situ in the *about*-sentence here.

(i) Jag sa om {den där smörgåsen där} att {den skulle jag gärna vilja äta upp}.
I said about the there sandwich-DEF *there that it would I gladly want eat up*

'I said about the sandwich over there that I would love to eat it.'

5.3.1.2 Non-specific indefinites

A particularly interesting phenomenon is T-preposing of non-specific indefinite DPs out of relative clauses. As Reinhart (1981) discusses, indefinite DPs can serve as topics if they have a specific interpretation, or what she calls "pragmatic reference", by which she means that the DP can be interpreted as denoting a set. In such cases the sentence can be interpreted as asserting something about this set, or about its members. An example is given in (56), where *all grownups* is a hanging topic.

(56) Parents don't understand. But *all grownups*, they do it to kids, whether they're your own or not.

(Reinhart 1981:65)

The fact that *they* can take *all grownups* as an antecedent here reflects that the phrase has pragmatic reference, even though it is quantified.

In my example collection, I have a few examples where indefinite DPs which do not have such a specific interpretation are T-preposed from a relative clause. Consider the example in (57), where A and B are discussing whether there are dangerous animals in California. Mountain lions have been mentioned earlier in the conversation.

(57) A: I think there are some that are a bit more dangerous than the ones in Sweden at least.

B: ja [ett LOdjur]$_1$ har jag inte hört talas om nån [som blivit
B: yes a lynx have I not heard spoken about someone that has been
uppäten av t_1]
eaten by

'Yes, I've never heard of anyone who was eaten by a lynx.'

(Conversation, Summer 2013)

In this sentence, a specific interpretation is not possible. B is clearly not making a statement about some particular lynx, rather he is saying that he has never heard of anyone who has been eaten by any lynx. If we embed the ERC-sentence in an *about*-sentence, as in (58), we get a natural-sounding result, but the *about*-sentence is not a good paraphrase.

(58) Han sa om {ett lodjur} att {han aldrig hade hört talas om nån
he said about a lynx that he never had heard spoken about someone
som blivit uppäten av det}.
who had been eaten by it.

'He said about a lynx that he had never heard about anyone who was eaten by it.'

The *about*-test forces a specific interpretation: We can only take the *about*-sentence in (58) to be about a specific lynx. This effect of the test is also noted by Reinhart (1981), who points out that generic singular indefinite DPs, like *a shark*, fail the *about*-test, while plural generic indefinites, like *sharks* do not. She provides the examples in (59) as an illustration.

(59) a. She said about sharks that they won't attack unless they are very hungry.

b. She said about a shark that he won't attack unless he is very hungry.

c. A shark won't attack unless he is very hungry.

(Reinhart 1981:88, footnote 6)

Example (59b) must be interpreted as a statement about a specific shark. Nevertheless, it is possible to interpret (59c) as making a generic statement about sharks, and singular generic indefinites thus seem to be possible aboutness topics.

It is not entirely straightforward to interpret example (57) as a generic statement about lynxes, but the assertion that the speaker hasn't heard of anyone who has been eaten by a lynx might be taken to imply, in the context of the utterance, that lynxes are not very dangerous. The example is interesting since it shows that a specific interpretation is not necessary for ERC, and that a non-specific interpretation can survive in T-preposing from a relative clause. Contrast seems important here. Example (57) was produced with contrastive stress on *lodjur*, contrasting lynxes against mountain lions, which had been mentioned in the conversation, and which can be seen as sub-themes of the hypertheme 'wild animals'.

The lynx case bears some resemblance to left-dislocation with **type anaphora** (Borthen 2003). Consider the Norwegian example in (60), which is adapted from Borthen (2003:41).

(60) Ola har (en) fin bil. Det har Kari også. (No.)
*Ola has (a) nice car.*MASC *that.*NEUT *has Kari too*

'Ola has a nice car. So does Kari.'

Det 'that.NEUT' in this example is neuter, even though the antecedent is masculine. Borthen calls this a type anaphor, because the anaphoric relation involved is not at the token level, but at the type level; i.e. in (60), both Ola and Kari have a nice car, but not the same nice car. Looking again at (25), repeated here as (61), we see that it too involves such a type anaphor.

(61) båtar$_x$ [det$_x$]$_1$ gjorde jag ju själv t_1 utav bark och träd och så där
boats it made I PRT myself out of bark and trees and such
'boats, I made that myself out of bark and trees and such'

The T-preposed pronoun *det* is a type anaphor, which refers to the type *båtar* 'boats'. The hanging topic is indefinite and non-specific, like the T-preposed phrase in the lynx sentence, and the phrase is a sub-theme to the hypertheme 'toys' given in the context, as we saw earlier.

5.3.2 Answers to questions – foci

Fronted phrases in local T-preposing can sometimes be the answer to a question in the linguistic context, i.e. the focus of the utterance as I have defined it. An example of this is given in (62), which is from a radio segment about the terror attacks in Paris 2015, where a reporter describes an interview. The reporter has asked the person being interviewed how she felt, and the answer was that she was not afraid.

(62) ... och när jag frågade hur hon kände det idag sa hon
and when I asked how she felt it today said she

"nä [$_F$ RÄDD]$_1$ är jag inte t_1"
no afraid am I not

'... and when I asked how she felt today she said "No, I'm not afraid."'

Another example, with extraction from an *att*-clause is given in (63). This example is from a discussion about falsification of documents at a government agency, the Central Ethical Review Board.

(63) vad tycker du om det här förfarandet?
what think you about the here procedure

[$_F$ PINsamt]$_1$ (.) tycker jag [att det är t_1]
embarrassing think I that it is

'What do you think about this procedure?' 'I think it's embarrassing.'

(Radio Sweden, P1 *Morgon*, Aug. 31 2012)

Each of these examples involves an adjective being T-preposed. This may be the most common situation, but other types of phrases can serve as foci in T-preposing as well. The dialogue in (64) took place in a grocery store at the checkout counter. The cashier is holding up a bag of buns, which can be bought by the piece.

(64) Cashier: hur många var det här?
how many was the here

Customer: [$_F$ TIO STYcken]$_1$ tror jag [jag tog t_1].
ten pieces think I I took

Cashier: 'How many of these are there?' Customer: 'I think I took ten.'

(Conversation, Dec. 2015)

Tio stycken 'ten pieces' in the customer's reply provides the answer to the question from the cashier.

My collection of ERC-sentences contains only one clear examples in which a phrase that is the answer to a question is extracted from a relative clause, as briefly mentioned in section 5.2.4. The example, repeated in (65), is from a radio show about language, where people can call in and ask questions. Here the caller has asked about an expression, *kura mörtning*, which the radio show host thinks might mean 'to sit in stillness in the dusk with the lights off'. It is clear from the discussion that the caller has a different idea about what it may mean, but she is reluctant to tell the host. In the part of the conversation that we see below, the host is asking about the caller's idea.

(65) Radio show host: Wait, what was your idea?

Caller: [$_F$ att det skulle NAPPA väl i (.) skymningen eller nåt sånt
that it would bite well in the dusk or something such
där till exempel]$_1$ var det nån [som hade en teori kring t_1]
there for example was there someone that had a theory around

'Someone suggested that it might mean that it would bite well in the dusk or something like that for example.'

(Radio Sweden, *Språket*, Sept. 9 2015)

The fronted phrase, which is the answer to the question, is a *that*-clause.

Although fronting information foci is not common in ERC, this example is entirely natural-sounding, and similar examples have been reported in previous literature. Engdahl (1997) provides the example in (66).

(66) a. Vad tycker de flesta om att dricka?
what do most people like to drink?

b. KAFFE känner jag många som tycker om.
coffee I know many that like

(Engdahl 1997:71)

In (66b), *kaffe* 'coffee' provides the information which is asked for in (66a), and is the utterance's information focus.

5.3.3 Scene-setting adverbials

As we have seen, Spec-CP in a declarative main clause in Swedish can be filled by an adverbial, as in (67).

(67) Vad hände sen?
what happened then

[G Efter ett tag] åt katten en fisk.
*after a while ate cat.*DEF *a fish*

'After a while, the cat ate a fish.'

These adverbials often serve the function of setting up a scene for the rest of the sentence, and are sometimes called **scene-setting** adverbials (Lambrecht 1994, Andréasson 2007).

Several types of adverbials can function as such scene-setters, including time adverbials, like *plötsligt* 'suddenly' and *efter en stund* 'after a while'. The example in (68) is from a news article about a fire, and (69) is a constructed example based on (68).

(68) Plötsligt skrek nån "jävlar", "jävlar".
suddenly yelled someone devils devils
'Suddenly, someone yelled "damn", "damn".

(*Göteborgs-Posten*, Nov. 25 2007)

(69) Efter en stund skrek nån "jävlar", "jävlar".
after a while yelled someone devils devils
'After a while, someone yelled "damn", "damn".'

My collection of extraction sentences contains no clear example of a T-preposed adverbial that modifies an RC-internal predicate and has a scene-setting function. If we try to construct sentences based on (68) with T-preposing from an *att*-clause and from an RC, we see that *plötsligt* can only modify the verb in the higher clause, much like in English.

(70) Plötsligt tror jag [att nån skrek "jävlar", "jävlar"].
suddenly think I that someone yelled devils devils
'Suddenly, I'm thinking that someone yelled "damn", "damn".'
Impossible reading: 'I think that someone suddenly yelled "damn", "damn".'

(71) Plötsligt kom det nån [som skrek "jävlar", "jävlar"].
suddenly came there someone that yelled devils devils
'Suddenly, someone who yelled "damn", "damn" came.'
Impossible reading: 'Someone who suddenly yelled "damn", "damn" came.'

Interestingly, however, a contrast between *att*-clauses (72) and relative clauses (73) can be seen with the adverbial *efter en stund*.

(72) Efter en stund tror jag [att nån skrek "jävlar", "jävlar"].
after a while think I that someone yelled "damn", "damn".
'I think that someone yelled "damn", "damn" after a while.'

(73) Efter en stund kom det nån [som skrek "jävlar", "jävlar"].
after a while came there someone that yelled devils devils
'After a while, someone who yelled "damn", "damn" came.'
Impossible reading: 'a person x came and after a while x yelled "damn", "damn"'

Example (72) shows that *efter en stund* 'after a while' can modify the yelling event expressed by the verb in the complement *att*-clause, but in (73), we cannot interpret it as modifying the yelling event in the relative clause.

Jensen (2002) makes a similar observation. Based on her study of extraction in spoken Danish, she argues that time adverbials in the initial position in main clauses cannot be interpreted as modifying a predicate inside a relative clause. Her corpus does, however, contain examples where such a time adverbial is interpreted inside a complement clause. The example in (74) is from Jensen, with my glosses.

(74) fra ni og fyrre tror jeg nok de boede der (Da.)
from nine and forty think I PRT *they live there*
'I think they lived there from forty-nine.'

(Jensen 2002:114)

Here the time adverbial *fra ni of fyrre* 'from forty nine' refers to the time from which someone began living somewhere, and not to when the speaker's belief about this began. Jensen attributes this possibility to the fact that the embedded constituent *de boede der* 'they lived there' is a complement, and to the fact that the matrix is not contributing to the propositional content, but rather expresses epistemic modality.

Both Jensen's and my observations are in line with the statement of Teleman et al. (1999) that the extracted phrase in ERC usually does not refer to the situational background where the proposition of the utterance holds (p. 424–431).

I explore the question of whether unselected adjuncts like these time adverbials can ever be extracted more thoroughly in chapter 6. The point here is merely to show that there seems to be more rigorous restrictions on T-preposing from relative clauses than on T-preposing in local clauses and from *att*-clauses.

5.3.4 Summing up

This section has shown that T-preposing in ERC is like local T-preposing in that both topics and foci are fronted. In ERC, T-preposed foci are unusual, but possible. These findings follow the pattern observed in the previous literature about extraction in the mainland Scandinavian languages (e.g. Engdahl 1997, Jensen 2002). Like in local T-preposing, the topic-expression can pied-pipe a preposition.

Most T-preposed phrases fit Reinhart's definition of aboutness topics, especially if the *about*-test is employed with some flexibility. A case where the *about*-test does not work is when the T-preposed phrase is a singular non-specific indefinite. In such cases, contrast seems to be necessary.

T-preposing in ERC is more restricted than local T-preposing and T-preposing from *att*-clauses in that certain scene-setting adverbials cannot be interpreted as modifying a predicate inside the relative clause.

5.4 Relativization

Most of the ERC-sentences in my collection involve T-preposing a phrase out of a relative clause, but there are a few exceptions (7/100 in Sample A). These examples instead involve relativization of a position inside another relative clause, as in (75), which is from a discussion about a department library.

(75) vi har till exempel flera hyllmeter namnforskning$_1$ (.) [$_{RC1}$ Op$_1$ som
we have to example several running meters onomastics which
vi inte har en enda$_2$ på vår institution [$_{RC2}$ Op$_2$ som t_2 håller på med t_1]].
we not have one single on our department that holds on with
'For example, we have several running meters of books on onomastics, which no one in the department works on.'
(Conversation, Oct. 2014)

In (75), both of the relative clauses, here marked RC1 and RC2, are introduced by *som* 'that'. The speakers in (76) and (77) use other relativization strategies to form RC1, involving the relative adverb *där* 'where' and a relative pronoun *vilket* 'which' respectively.

(76) vi kanske ska be att få låna skräcksalen nere på konst och musik
*we maybe shall ask to get borrow horror room.*DEF *down at art and music*
där$_2$ [$_{RC1}$ *där*$_1$ man inte ser dom$_2$ [$_{RC2}$ Op$_2$ som t_2 sitter t_1]]
there where one not see those that sit
'Maybe we should ask if we can use the horror room down at arts and music where you don't see the students.'
(Conversation, Oct. 2014)

(77) ... och sen så sa dom "Aron du får sparken" (.) [$_{RC1}$ vilket$_1$ det inte är
and then PRT *said they Aron you get kick*-DEF *which it not is*
första gången$_2$ [$_{RC2}$ Op$_2$ jag hör nån säga t_1 t_2]]
*first time*DEF *I hear someone say*
'... and then they said "Aron, you are fired", and that's not the first time I've heard someone say that.'
(Aron Flam, *Kön*, stand-up show, 2015)

In some accounts of restrictions on extraction (e.g. Van Valin & LaPolla 1997, Goldberg 2006), it is important to find parallels between relativization and other A′-dependencies in terms of discourse function, explaining why the pragmatic constraints that these accounts propose would affect all of these operations. Van Valin & LaPolla (1997), for example, assume that in relativization, the head of the RC is in a topical relation to the RC, and that the RC is "about" the head (p. 626–627). They propose that topicalization and relativization are both subject to a pragmatic condition on aboutness (with specific reference to Reinhart's concept of pragmatic aboutness):

(78) *Pragmatic aboutness restriction on topicalzation and relativization*
The sentence fragment following the topical element in the pre-core slot or a restrictive relative clause must be pragmatically interpretable as being about the pre-core slot element or the head noun. (Van Valin & LaPolla 1997:627)

Goldberg (2006) builds her account on the idea that all long-distance dependencies involve discourse-prominent positions, and that the head of a relative clause, even though it does not in general function as a topic or focus in the main clause, is prominent in relation to the relative clause, and that therefore relativization also involves a discourse prominent position. There are also analyses of relative clauses on which the relative pronoun is argued to function as a topic of the relative clause (Dalrymple 2001).

I will discuss the account proposed by Goldberg (2006) in more depth in section 5.5. Here, I will just point out that it is not obvious that one can unify topic-comment structures and relativization in terms of discourse function. While a sentence like *Anna känner en kvinna som har varit i Paris* 'Anna knows a woman who has been to Paris' might be seen as in some sense adding information about a certain woman, the relationship between the restrictive relative clause and its head is not identical to an aboutness relation, where a comment is added about the referent of a DP. Consider (79).

(79) Anna kan räkna upp namnen på alla människor som har bestigit
*Anna can count up names.*DEF *of all human.*PL *who have climbed*
Kebnekaise.
mount Kebnekaise
'Anna can name all the people who have climbed mount Kebnekaise.'

The fact that (79) does not entail the sentence *Alla människor har bestigit Kebnekaise* 'all humans have climbed Kebnekaise' indicates that the relative clause is not predicated of the denotation of the DP *alla människor* 'all humans'. Rather, the relative clause modifies only the noun (or NP) *människor* 'humans'. The most straightforward way to capture this is to let the noun and the relative clause form a constituent to the exclusion of the determiner. Furthermore, restrictive relative clauses do not just add information about their heads, since they serve to restrict the head's denotation. For that reason, if we see *människor* in *människor som har bestigit Kebnekaise* 'humans who have climbed mount Kebnekaise' as a kind, the relative clause is not adding information "about" this kind generally, but lets us pick out a subset of the kind.

The examples in (75)–(77) require forming two relativization dependencies, one that creates the most deeply embedded relative clause, RC2, and one that relativizes a position inside this relative clause, creating the larger relative clause, RC1. The first of these, RC2, has to be of the restrictive type, as is pointed out by Engdahl (1997). In (75)–(77), the second relativizing dependency, RC1 is non-restrictive in each case. However, this is not the only possibility. We can construct an example inspired by (75) where both RC2 and RC1 are restrictive.

(80) Vi makulerade alla tidskrifter om ämnen [$_{RC1}$ Op$_1$ som vi inte har
we cancelled all journals about subjects that we not have
någon på institutionen [$_{RC2}$ Op$_2$ som t$_2$ håller på med t_1]].
*someone on department.*DEF *that holds on with*
'We cancelled all journals about subjects that no one in the department works on.'

While it is clearly important to find some property which unifies operations that can or cannot displace a constituent from certain domains, relativization cannot be equated with a topic-comment structure at the discourse level. Rather, what unifies T-preposing and relativization is their syntax. T-preposing is Ā-movement to Spec-CP in a main clause, relativization is Ā-movement to spec-*c*P in a relative clause.

5.5 The information impact of the relative complex

I turn now to proposals which relate the felicity of ERC to the pragmatic function of the relative clauses. There are a few different ideas about what defines domains that allow extraction. Here, I will discuss the Dominance Condition, proposed by Erteschik-Shir (1973, 1982, 2007), and the proposal by Goldberg (2006, 2013) that domains that permit extraction are focal or non-backgrounded.[5]

5.5.1 The Dominance Condition

Erteschik-Shir (1973, 1982, 2007) observes that Danish permits ERC in certain cases (81a), but not others (81b).

(81) a. Det_1 kender jeg mange [der har gjort t_1] (Da.)
that know I many who have done
'I know many people who have done that.'
(Erteschik-Shir 1973:63)

b. *Det_1 har jeg peget paa mange [der har gjort t_1] (Da.)
that have I pointed at many who have done
(Erteschik-Shir 1973:64)

The example in (81a), with the matrix verb *kende* 'know.REL' is acceptable, whereas the one in (81b), where the matrix predicate is *pege på* 'point to', is unacceptable. She argues that there is no structural difference between the relative clauses which permit extraction, and those that do not, and concludes that there must instead be some pragmatic or semantic condition on extraction. **Dominance**, as defined in (82), is proposed to be the relevant notion.

(82) A clause or phrase is semantically dominant if it is not presupposed and does not have contextual reference. (Erteschik-Shir 1973:22)

Dominant constituents are not presupposed or given in the context. From Erteschik-Shir's discussion, it seems as if presupposition should not be taken as

[5] I will not discuss the proposal by Van Valin (1994) and Van Valin & LaPolla (1997) that the constituent that a phrase is extracted from has to be a potential focus domain. The reason for this is that the proposal rests on the idea that the potential focus domain is defined by the syntactic structure in a way which rules out extraction from relative clauses.

logical presupposition, but rather as given information. Extraction possibilities are constrained by the Dominance Condition on extraction (83).

(83) Extraction can occur only out of clauses or phrases which can be considered dominant in some context. (Erteschik-Shir 1973:27)

In a later article, Erteschik-Shir & Lappin (1979) propose a different definition of dominance, which concerns the speaker's intentions. This definition is given in (84).

(84) A constituent c of a sentence S is dominant in S if and only if the speaker intends to direct the attention of his hearers to the intension of c by uttering S. (Erteschik-Shir & Lappin 1979:43)

Both definitions of dominance are supposed to pick out constituents whose content is a potential topic for future discourse, and in both accounts, dominance is operationalized with a test designed to capture this idea. The test is called "the lie test", because it diagnoses whether it is possible to contradict a certain utterance or part of utterance by saying that it is a lie. This tests which part or parts of a sentence are potential topics for future discourse, by actually making them the topic of the contradicting utterance. We see an example of how the test is used in (85) and (86).

(85) Speaker A: I know many people who like that.
Speaker B: That's a lie – you don't.
That's a lie – nobody likes that.

(86) Speaker A: Peter knows the girl who likes the boy.
Speaker B: That's a lie – he doesn't.
*That's a lie – she doesn't.

In (85), Speaker B can choose to comment on either the whole sentence, contradicting the assertion that Peter knows such people, or just on the content of the relative complex, as is done in the answer *That's a lie, nobody likes that*. In (86), on the other hand, Speaker B cannot as naturally deny the content of the relative complex. At any rate, the answer *That's a lie, she doesn't* seems much less congruent. According to the logic of the test, this shows that the relative complex can be considered dominant in (85), but not in (86), and predicts that extraction should only be possible from the relative clause in (85).[6]

[6] This is not the right result for English, since you still cannot extract from the relative clause in (85), but the distinction is the one that Erteschik-Shir (1973) finds in Danish and which the test is designed to capture.

The dominance approach can account for a large part of the examples in my collection. To show this, we can apply the lie test to the types of sentences that were identified in section 4.3 as environments where extraction is used spontaneously. As we see in (85) and (86) above, the test is performed using sentences without extraction, and then makes a prediction about whether extraction should be possible or not. When I apply the test below I use adaptations of ERC-sentences which we have seen are acceptable. If the Dominance Condition is correct, the lie test should show that the relative complex is dominant.

First, we consider the relative complexes or DP + *som*-clauses in existential sentences. These can be dominant according to the test, as illustrated by (87) and (88) (example (45) and (50) in chapter 4).

(87) Speaker A:

Det finns något som heter galltvål.
there exists something that is called galltvål
'There is something called *galltvål*.'

Speaker B:

a. Det stämmer inte, det finns det inte.
'That's not right, there isn't.'

b. Det stämmer inte, inget heter så.
'That's not right, nothing is called that.'

(88) Speaker A:

Det är många som gillar honungssenap till den där.
there are many that like honey mustard to the there
'There are many people who like honey mustard with that.'

Speaker B:

a. Det stämmer inte, det är det inte.
'That's not right, there isn't.'

b. Det stämmer inte, ingen gillar det.
'That's not right, no one likes that.'

As we see, Erteschik-Shir's test identifies the relative complex as potentially dominant in both (87) and (88), since in each case the b-answer is possible. This is the right outcome, since extraction is acceptable.

Similarly, the test correctly predicts that extraction from a predicative copular construction is permissible in Swedish (example (56) in chapter 4).

(89) Speaker A:

Vi är inte många som vet det.
we are not many that know that
'There aren't many of us that know that.'

Speaker B:

a. Det stämmer inte, ni är jättemånga
'That's not right, there are a lot of you (that know that).'

b. Det stämmer inte, jättemånga vet det.
'That's not right, a lot of people know.'

Again, Speaker B can comment either on the content of the whole sentence, or on the assertion that not many people know. This shows that the relative clause can be dominant, which is in accordance with the fact that we find extraction from such sentences.

The relative complex in examples with verbs like *känna* 'know.REL', *se* 'see', and *höra* 'hear' can also be shown to be dominant. This is illustrated for *känna* by (90) (example (80) in chapter 4).

(90) Speaker A:

Jag känner en hund som heter Molly
I know a dog that is called Molly
'I know a dog called Molly.'

Speaker B:

a. Det stämmer inte, det gör du inte.
'That's not right, you don't.'

b. Det stämmer inte, inga hundar heter Molly.
'That's not right, no dogs are called Molly.'

Both (90a) and (90b) are congruent answers here.

Erteschik-Shir's approach make the correct prediction for *det*-clefts as well. We saw in chapter 4 that it is possible to extract a phrase from the cleft-clause in a *det*-cleft, and in (91) we see that the lie test identifies the cleft clause as potentially dominant (example (86) in chapter 4).

(91) Speaker A:

Det var adeln som hade domsrätt över bönderna.
it was nobility-DEF *that had jurisdiction over farmers*-DEF
'It was the nobility that had the jurisdiction over the farmers.'

SpeakerB:

a. Det stämmer inte, det var det inte.
'That's not right, it wasn't.'

b. Det stämmer inte, kungen hade det.
'That's not right, the king did.'

c. Det stämmer inte, ingen hade det.
'That's not right, no one did.'

As (91a) and (91b) show, it is possible for Speaker B to either contradict the matrix, or to claim that someone else had the jurisdiction over the farmers, e.g. the king. It also seems quite congruent for Speaker B to say that no one had jurisdiction over the farmers at all, as in (91c).

But there are also examples in my collection which cannot be explained by dominance as easily, because the result of the lie test is less clear cut. Consider (92) (example (82) in chapter 4).

(92) Speaker A:

Jag beundrar folk som klarar det rent psykiskt.
I admire people that manage it purely psychologically
'I admire people who can deal with that psychologically.'

SpeakerB:

a. Det stämmer inte, det gör du inte.
'That's not right, you don't.'

b. #Det stämmer inte, folk klarar inte det.
'That's not right, people can't.'

Here it is acceptable for Speaker B to contradict Speaker A's assertion that she admires certain people (92a). But the second answer, (92b), is quite odd. This suggests that the relative complex cannot be dominant in (92), and it is surprising that we find extraction from such relative complexes. Löwenadler (2015) makes the same observation. He also provides extraction sentences like (93).

(93) Såna låtar är jag verkligen imponerad av de som kan skriva.
such songs am I really impressed by those who can write
'I'm really impressed by those who can write such songs.'

(Löwenadler 2015:43)

Here the embedding predicate is more complex. If we apply the lie test to (93) it becomes very clear that the relative complex in this sentence is not dominant (94).

(94) Speaker A:

Jag är verkligen imponerad av de som kan skriva såna låtar.
I am really impressed by those that can write such songs
'I'm really impressed by those who can write such songs.'

SpeakerB:

a. Det stämmer inte, det är du inte.
'That's not right, you don't.'

b. #Det stämmer inte, de kan inte det.
'That's not right, people can't.'

The answer in (94b), where Speaker B denies that the people in question can write songs of a certain type, is not felicitous at all. This means that extraction should not be possible, but as we saw in (93), it is.

Non-restrictive relative clauses pose another challenge for the dominance-based account of extraction. As we saw in chapter 2, non-restrictive relative clauses do not permit extraction at all. However, the lie test seems to predict that they should. Consider (95).

(95) Speaker A:

Jag träffade Ida, som förresten också har en sån bil.
I know Ida that by the way also has a such car
'I met Ida, who, by the way, also has a car like that.'

Speaker B:

a. Det stämmer inte, det gjorde du inte.
'That's not right, you didn't.'

b. Det stämmer inte, det har hon inte.
'That's not right, she hasn't.'

Surprisingly, it turns out that (95b) is a quite felicitous answer. This is unexpected, since extraction from non-restrictive relative clauses is strongly unacceptable, which we see if we try extracting *en sån bil* 'such a car' from the relative clause in (95).

(96) *En sån bil träffade jag Ida, [som förresten också har t_1].
a such car know I Ida that by the way also has

To summarize, the Dominance Condition accounts for the acceptability of several types of extraction, namely ERC with existentials, predicational copular constructions, *det*-clefts, and ERC with *känna* 'know.REL' but it makes the wrong prediction about ERC with embedding predicates like *beundra* 'admire'

and *vara imponerad av* 'be impressed by', which express the speaker's attitude towards the people or things denoted by the relative complex. The lie test also fails to capture the fact that non-restrictive relative clauses are islands for extraction.

5.5.2 Goldberg: Backgrounded constructions are islands

Goldberg (2006) proposes a somewhat similar hypothesis for what constrains extraction. Her proposal is within the framework of Construction Grammar, where Ā-dependencies are thought of in terms of different constructions combining with each other. These constructions are assumed to have particular information-structural properties. Pragmatic clashes arise when these properties are not consistent with one another, and these clashes are responsible for the unacceptability of sentences with island violations (p. 132). More specifically, according to Goldberg, phrases that are extracted are placed in discourse-prominent positions, and this puts constraints on where their gaps can be. Goldberg argues that backgrounded parts of sentences cannot contain gaps. She argues that the generalization in (97) predicts several island effects.

(97) Backgrounded constructions are islands (BCI). (Goldberg 2006)

Backgroundedness is conceived of in negative terms: any part of a sentence is backgrounded if it corresponds neither to part of the potential focus domain, nor to the primary topic of the sentence. The potential focus domain is identified as the part of a sentence which is "interpretable as being asserted" (Goldberg 2006:130). Conversely, it is assumed that "elements that are part of presupposed clauses are backgrounded" (Goldberg 2006:130).

Instead of the lie test, Goldberg employs a negation test, which picks out what she calls the focus domain (98).

(98) Test for being within the focal domain: propositions expressed within the potential focus domain can be understood to be negated by sentential negation. (Goldberg 2006:130)

To see how this test works, we can look at a non-restrictive relative clause, which is one of the cases Goldberg discusses (99).

(99) a. I saw John, who I told you about last week.
→ I told you about John last week.

b. I did't see John, who I told you about last week.
→ I told you about John last week.

(Goldberg 2006:145)

It follows from the sentence in (99a) that the speaker told the listener about John last week. This also follows from the negation of (99a), i.e. (99b). This means that the proposition *I told you about John last week* is not in the focal domain of the sentence, since it cannot be negated by negating the matrix predicate. This makes the correct prediction about non-restrictive relative clauses, namely that they are islands.

Applying the negation test to the ERC-sentence types, we find that the test correctly identifies both the existential sentences in (100) and (101), and the predicative copular construction (102), as sentences that should permit extraction.

(100) a. Det finns något som heter galltvål.
there exists something that is called galltvål
→ Something is called *galltvål*

b. Det finns inget som heter galltvål.
there exists nothing that is called galltvål
↛ Something is called *galltvål*

(101) a. Det är många som gillar honungssenap till den där.
there are many that like honey mustard to the there
→ Many people like honey mustard to that.

b. Det är inte många som gillar honungssenap till den där.
there are not many that like honey mustard to the there
↛ Many people like honey mustard to that.

(102) a. Vi är många som vet det.
we are many that know that
→ Many people know that.

b. Vi är inte många som vet det.
we are not many that know that
↛ Many people know that.

In each of these cases, it is possible to negate the propositional content expressed by the RC or RC-like constituent by negating the matrix. However, Goldberg's test incorrectly predicts that *det*-clefts should be islands since they trigger logical presuppositions (103), which cannot be cancelled by negating the matrix.

(103) a. Det var adeln som hade domsrätt över bönderna.
it was nobility-DEF *that had jurisdiction over farmers*-DEF
→ Someone had jurisdiction over the farmers.

b. Det var inte adeln som hade domsrätt över bönderna
it was not nobility-DEF *that had jurisdiction over farmers*-DEF
→ Someone had jurisdiction over the farmers.

As we saw in chapter 4, section 4.3.3, extraction from cleft clauses is fine. The example which corresponds to the tested sentence in (103) is given in (104).

(104) den allra största delen av befolkningen, bönderna, den$_1$ var det
the most of all biggest part of population.def, farmers-DEF, *that was it*
adeln$_2$ [som t_2 hade domsrätt över t_1]
nobility-DEF *that had jurisdiction over*
'It was the nobility that had the jurisdiction over the largest part of the population, the farmers.'

(Radio Sweden, *Godmorgon, världen!*, Jan. 25 2015)

Since cleft clauses are backgrounded according to the negation test, sentences like (103) are predicted to be unacceptable. Examples (92) and (93), which were problematic for Erteschik-Shir's account, turn out to be unexplained on the BCI account as well, which (105) and (106) show.

(105) a. Jag beundrar folk som klarar det rent psykiskt.
I admire people that manage it purely psychologically
'I admire people who can deal with that psychologically.'
→ People can deal with that psychologically.

b. Jag beundrar inte folk som klarar det rent psykiskt.
I admire not people that manage it purely psychologically
'I don't admire people who can deal with that psychologically.'
→ People can deal with that psychologically.

(106) a. Jag är verkligen imponerad av de som kan skriva såna låtar.
I am really impressed by those that can write such songs
'I'm really impressed by those who can write such songs.'
→ Someone can write such songs.

b. Jag är verkligen inte imponerad av de som kan skriva såna låtar.
I am really not impressed by those that can write such songs
'I'm really not impressed by those who can write such songs.'
→ Someone can write such songs.

It follows from both (105a) and (105b) that people can deal with whatever *det* 'that' refers to here , and similarly it follows from both (106a) and (106b) that there are people who can write a certain type of song. This means that the BCI-account predicts that the relative complexes in these sentences should be islands, but as we have seen, they are not. Like the Dominance account, the BCI-account makes the right predictions for the common types of ERC-sentences, i.e. the existential sentences, but it fails when it comes to ERC

with embedding predicated like *beundra* 'admire' and *vara imponerad av* 'be impressed by'. In addition, it makes the wrong prediction about *det*-clefts, since the negation test identifies the cleft clause as backgrounded, even though extraction is possible.

5.5.3 Summing up

Summarizing the results of the section, we have seen that both the Dominance Condition and the BCI generalization make correct predictions about a part of the ERC-sentence types that are found in Swedish. The types that they capture are among the most common ones, namely existential and presentational sentences.

However, there are some problematic cases for each of the accounts. If the Dominance Condition really were operative in language in general, we should not find examples of extraction in which the relative complex fails the dominance test. But this is exactly what we find in examples like (93). The situation is similar for the BCI generalization. The negation test indicates that the relative complex in sentences like (93) is backgrounded, which means that the BCI-account predicts that extraction should be blocked. In addition, if a generalization like the BCI were at work, we would expect extraction from *det*-clefts to be impossible as well, contrary to fact.

On a more general level, these proposals face difficulties when it comes to accounting for variation between languages. Goldberg (2006) mentions that languages are on a continuum, and that extraction can be more acceptable in one language than in another. For this to be an explanation, we need some factor that predicts where on the scale a given language will be. Why should presupposition create an unaccessible domain in English but not in Swedish? Or formulated slightly differently, why should extracting from a backgrounded part of the sentence create less of a pragmatic clash in some languages, or the speakers of some languages be more tolerant of pragmatic infelicity? In general, an account that places all explanatory weight on the information-structural impact of the embedded constituent seems ill-suited to describe the variation that exists between languages.[7]

[7] See Löwenadler (2015) for a similar observation.

5.6 Summary

In this chapter, I have investigated various facts about the pragmatics and information structure of ERC-sentences, and provided illustrations of how they are used in spontaneous discourse. I began by investigating the discourse function of T-preposed phrases, and showed that T-preposing in ERC is much like local T-preposing in this regard. I then turned to the function of the T-preposed phrase in its utterance. Here, the investigation showed that the T-preposed phrase in ERC is often an aboutness-topic, but that in some cases, it is the information focus of the sentence. This is in line with previous research (e.g. Engdahl 1997, Jensen 2002). The most interesting finding in this part of the chapter is that there seems to be some restriction on ERC that makes it impossible to interpret a T-preposed scene-setting adverbial as modifying an event expressed by a verb inside a relative clause.

In addition to T-preposing ERC-sentences, there are examples which involve relativization out of relative clauses in my collection. The relation between the head of the relative clause, and the relative clause, does not need to be an aboutness-relation, contrary to some claims in the previous literature (Van Valin & LaPolla 1997). The similarity between T-preposing ERC and ERC created by relativization out of a relative clause is syntactic; both create Ā-movement dependencies.

With respect to the information impact of the relative complex, I investigated previous proposals by Erteschik-Shir (1973, 1982) and Goldberg (2006, 2013). As it turns out, both proposals capture large parts of the ERC-data, but neither of them can capture all of it.

I will discuss what the findings of this chapter imply about the role of discourse, pragmatics, and information structure in an analysis of ERC in chapter 7. Before doing this, however, I explore the idea that relative clauses constitute some type of weak island in chapter, 6, where I also examine extraction of T-preposed adjuncts, like *efter en stund* 'after a while', and extraction from relative clauses in *wh*-questions.

6. Very weak islands

We have seen in the previous chapters that there is licit Ā-movement of several types of phrases from relative clauses in Swedish, and that this can be captured by a recently proposed account by Vikner (2017) and Nyvad et al. (forthcoming), where an extra specifier in the complementizer domain of the relative clause provides an escape hatch. Even so, there are certain constraints on such extractions; for example there are Comp-trace effects, and extraction has to be from a restrictive relative clause. In addition, there are signs that even extraction of non-subjects from restrictive relative clauses is constrained. Specifically, we saw in chapter 5 that it is impossible or at least very hard to interpret certain T-preposed adverbial expressions, like *efter en stund* 'after a while' as modifying a predicate inside a relative clause. In this chapter, I examine this phenomenon in more detail, revisiting a few distinctions that have commonly been made in research on locality to see whether they can shed light on Swedish ERC.

As introduced in chapter 2, a distinction is often made between strong and weak islands. Broadly speaking, strong islands are those domains from which no extraction can occur, and weak islands those which permit some but not all types of phrases to extract (Szabolcsi 2006). Given this admittedly informal characterization and the findings in the previous chapters, we may ask whether Swedish relative clauses constitute some type of weak island.

Since ERC differs from the typical case of extraction from weak islands in a few important ways, this connection might seem non-obvious at first glance. First, the class of domains that are identified as weak islands in the previous literature does not include relative clauses, which are strong islands in many languages. Second, previous research on weak islands has mostly been focused on extraction of *wh*-phrases, and as we have seen, ERC in spontaneous usage involves mostly T-preposing and relativization.

Even so, there are interesting ways in which Swedish relative clauses are similar to weak islands. For example, as we have seen in previous chapters, they permit PP-gaps, which on some accounts in itself is enough to place them in the class of weak islands, as compared to strong islands which at most permit DP-gaps (cf. Cinque 1990, Szabolcsi 2006). Furthermore, even though there are no instances of extraction of *wh*-phrases in my collection of spontaneously produced extraction sentences, acceptable examples can be constructed and, as we will see, such extractions exhibit some effects familiar from the literature on weak islands.

The main aim of the chapter is to investigate additional data representing types of extractions which do not occur in my collection of ERC-sentences, but which have been important for the development of theories of Ā-dependencies. On a more general level, the aim is to situate ERC in the broader landscape of types of Ā-dependencies that have been described in the previous literature. In section 6.1, I investigate the possibility of T-preposing adjuncts from Swedish relative clauses, reporting on data from corpus searches and a questionnaire. In section 6.2, I turn to approaches to weak islands that describe restrictions on extraction in terms of pragmatics and/or semantics, and present data about question formation into relative clauses.

6.1 Extraction of adjuncts

Previous research has identified a general pattern with respect to extraction where, typically, extraction of an adjunct is harder than extraction of an argument, and extracting a phrase *from* an adjunct is harder than extracting a phrase from an argument (e.g. Huang 1982, Chomsky 1986, Cinque 1990). In this section I take a closer look at the possibility of extracting adjuncts from relative clauses, in order to see if this general pattern is instantiated in Swedish ERC. Before turning to the empirical question, I introduce some background about how the pattern has been analyzed in the syntactic literature, and how it has been refined by the introduction of further distinctions.

As a more specific example of the argument/adjunct distinction in extraction, we can consider (1a) and (1b), where it is illustrated with extraction from embedded questions. The examples are from Haegeman (1994).

(1) a. *Why$_i$ do [$_{IP}$ you wonder [$_{CP}$ whom$_j$ [$_{IP}$ John will invite t$_j$ t$_i$]]]?
b. ??Whom$_j$ do [$_{IP}$ you wonder [$_{CP}$ why$_i$ [$_{IP}$ John will invite t$_j$ t$_i$]]]?

(Haegeman 1994:518)

In (1a) the adjunct *why* has been extracted from the embedded question, which results in an entirely unacceptable sentence. In (1b), we see extraction of the argument *whom*, which is marginally possible. At any rate, the resulting sentence is more acceptable than (1a).

One way of capturing contrasts like that between (1a) and (1b) is in terms of the **Empty Category Principle (ECP)**[1] (Huang 1982, Lasnik & Saito 1984, Chomsky 1986). The ECP states that traces must be properly governed. A trace is properly governed if a head theta-marks it, or if it has a sufficiently local antecedent. Crucially, no barriers can intervene between a trace and its antecedent in antecedent-government.

From the ECP it follows that adjunct and argument traces will pattern in different ways with respect to extraction from embedded clauses. Since arguments are theta-marked by the head that selects them, argument traces are properly governed.[2] Adjunct traces, on the other hand, must be antecedent governed, since they are not theta-governed. This means that they must be governed by a coindexed XP from which they are not separated by a barrier;

---

[1] There are several formulations of the ECP, the one in (i) is from Haegeman (1994):

(i) **ECP**
Traces must be properly governed.
A properly governs B if and only if A theta-governs B or A antecedent-governs B.
A theta-governs B if and only if A governs B and A theta-marks B.
A antecedent-governs B iff A governs B and A is coindexed with B.

(Haegeman 1994:478)

Government is defined as in (ii).

(ii) **Government**
A governs B if and only if
(i) A is a governor;
(ii) A m-commands B;
(iii) no barrier intervenes between A and B;
(iv) minimality is respected.
where governors are: (a) heads,
(b) coindexed XPs.

(Haegeman 1994:478)

[2] The ECP is designed to capture a complement-non-complement distinction, but I am leaving subjects out of the discussion here. Like complements, subjects are theta-marked, but they are not governed by the head that theta-marks them, which means that they are not theta-governed, and thus must be antecedent-governed.

the theory predicts that adjuncts will be affected by interveners in ways that arguments are not.

To make this more concrete, consider again (1a) and (1b), here repeated as (2a) and (2b). In these examples, traces are γ-marked following Lasnik & Saito (1984), where [+γ] is assigned to a properly governed trace, and [−γ] is assigned to a trace which is not properly governed.

(2) a. *Why$_i$ do [$_{IP}$ you wonder [$_{CP}$ whom$_j$ [$_{IP}$ John will invite t_j [+γ] t_i [−γ]]]]?

b. ??Whom$_j$ do [$_{IP}$ you wonder [$_{CP}$ why$_i$ [$_{IP}$ John will invite t_j [+γ] t_i [+γ]]]]?

(Haegeman 1994:518–519)

Why is (2a) worse than (2b) on this view? The argument traces are properly governed in both of the examples, because they are theta-marked. The difference has to do with the adjunct trace. In (2b), the adjunct trace t_i is antecedent-governed by *why*, since there is no barrier intervening between them.[3] In (2a), on the other hand, *why* cannot properly govern its trace since CP is a barrier.

The ECP accounts for differences in extractability between adjuncts like *why* and arguments like *whom*. But it turns out that a few qualifications are needed in the description of the pattern, because the distinction between extractable and non-extractable phrases does not entirely overlap with the distinction between arguments and adjuncts. First, being an argument is not a sufficient condition for extractability. Second, not all adjunct extraction is equally bad (Ross 1984, Engdahl 1986, Comorovski 1989, Cinque 1990, Frampton 1990, Rizzi 1990, Szabolcsi 2006). We turn first to the observation that not all arguments can be extracted. As illustrated by sentences involving verbs like *behave* (3), being selected by a lexical head is not enough.

(3) a. How$_1$ did he behave t_1?

b. *How$_1$ did he ask whether to behave t_1?

It could be argued that *behave* selects the manner phrase in these examples, but if this is the case, why is (3b) not acceptable? An even clearer illustration is provided by verbs like *weigh*. As pointed out by Rizzi (1990), *weigh* is ambiguous and can be agentive, selecting a regular DP object (4a), or stative, selecting a measure phrase (4b).

[3] The reason that (2b) is nevertheless not fully grammatical is, on this view, due to the subjacency violation which is induced by *whom* crossing two bounding nodes (IP).

(4) a. John weighed apples.

b. John weighed 200 lbs.

(Rizzi 1990:78)

Questioning the complement of either of these is possible. That is, (5) is ambiguous and could be answered either with "*Apples*" or with "*200 lbs*".

(5) A: What$_1$ did John weigh t_1? (Rizzi 1990:78)

a. B: Apples.

b. B: 200 lbs.

If questioning involves creating a dependency into an embedded question, however, one of these readings disappears. The question in (6) is marginal according to Rizzi, but it is still possible to see that it can be answered only with a regular object, like "*Apples*", not with a measure phrase, showing that the reading of *weigh* which allows questioning of its complement must be the agentive one.

(6) A: ?What$_1$ did John wonder how to weigh t_1? (Rizzi 1990:78)[4]

a. B: Apples.

b. B: #200 lbs.

A further indication that measure phrases are different from regular objects is that they resist passivization.

(7) a. The apples were weighed by John.

b. *200 lbs were weighed by John.

[4] It should be noted here that the corresponding declarative clause is quite odd: John wondered how to weigh 200 lbs. There may be situations where similar sentences are not entirely impossible, though. For example, we can think of a boxer worrying about ending up in the right weight class.

(i) The boxer wondered how to weigh 200 lbs before weighing in on Wednesday.

Furthermore, we can reproduce the contrast with a different embedded question:

(ii) ?What did John wonder who weighed?

a. Apples

b. #200 lbs

While the passivization of (4a) is fine, as (7a) reveals, passivizing (4b) results in an odd sentence (7b).

Another relevant contrast, involving *weigh* and amount phrases, like *how many*, can be observed in (8). The examples are from Frampton (1990:2), who attributes them to Cinque.

(8) a. *[How many pounds]$_1$ do you wonder whether he weighed t_1?
b. ?[How many fish]$_1$ do you wonder whether he weighed t_1?

Whereas (8a) is not a well formed question, (8b) is marginally grammatical on a reading where what is asked about are specific fish.

These observations about selected manner phrases, like *how* in (3), and amount phrases, or **amount-quantified questions**, as they are sometimes called (e.g. Kroch 1989), led previous researchers to the idea that an argument has to be in some sense "referential" to be extracted. "Referentiality" in this context has been defined in several different ways. Rizzi (1990) aims to make the notion specific by tying it to thematic roles. Assuming that selected phrases are theta-marked, this approach is grounded in the idea that there are two types of theta-roles: referential and non-referential. The distinction is drawn between selected phrases that refer to "participants in the event described by the verb" (p. 86) and those that rather qualify the event, like manner and measure phrases.

The observed contrasts can then be accounted for in terms of referential indexing, given the principle in (9).

(9) A referential index must be licensed by a referential theta-role.

(Rizzi 1990:86)

On Rizzi's approach, bearing a referential index is a precondition for participating in binding relations, which means that only phrases which have referential theta-roles may bind variables. Binding can hold at an arbitrary distance, and so if a phrase can participate in binding relations, it can be part of *wh*-dependencies into embedded questions. Phrases which cannot carry a referential index – manner- and amount-phrases and the like – cannot enter into binding relations, and must be associated with their traces by government, which is local. The explanation for why (3b) and (8a) are not well formed, then, is that since these phrases bear non-referential theta-roles, they cannot license a referential index, and will have to have a local antecedent. We end up with a system where such phrases pattern with adjuncts with respect to extraction.

Like Rizzi, Cinque (1990) connects ability to participate in long-wh-movement to being able to bind a variable, but Cinque proposes that in addition to

having a referential theta-role, a necessary condition for being a binder is that the phrase refer to a specific member (or specific members) of a pre-established set in the mind of the speaker. The concept of a pre-established set is closely related to that of **D(iscourse)-linking** (Pesetsky 1987), and Cinque argues that D-linking could be subsumed under his notion of "referentiality" (Cinque 1990:8). Cinque's analysis builds on observations by Comorovski (1989), who argues that D-linking is the characteristic that distinguishes *wh*-phrases that can be extracted from embedded questions in Romanian from those that cannot. I will return to this line of explanation in section 6.2.

In addition to some arguments patterning with adjuncts in long-extractions, adjuncts tend to differ in how hard they are to extract. In the case of *wh*-questions, *how* (manner) and *why* (reason) are very hard or impossible to extract, whereas *when* (time) and *where* (location) are slightly better. Rizzi (1990) argues that this is because the event specification might license temporal and locative indices, which would allow such adjuncts to take part in operator-variable chains. Szabolcsi (2006) gives a good illustration of the pattern, shown in (10).

(10) a. *How did John ask <which problem to phrase __>?
'What is the manner such that John asked which problem to phrase in that manner?'

b. *Why did John ask < whether to fire him __ >?
'What is the reason such that John asked whether it is a good reason for firing him?'

c. ??When did John ask < whether to fire him __ >?
'What is the time such that John asked whether it is a good time for firing him?'

d. ?Where did John ask < whether to put/read the book __ >?
'What is the location such that John asked whether to put this book there/-whether to read this book there?'

(Szabolcsi 2006:494–495)

In summary, the data presented in this section so far shows the distinction between phrases that can and cannot extract from embedded questions does not reduce to the argument-adjunct distinction. Arguments are in general easier to extract than adjuncts, but some arguments (measure/amounts and manners) resist extraction, and some adjuncts (those that denote times and places) are easier to extract than others.

With respect to argument-adjunct asymmetries in long extraction in the Scandinavian languages, one of the studies by Christensen et al. which was mentioned in chapter 2 is particularly relevant here. In two acceptability judgement

experiments about *wh*-questions in Danish, Christensen et al. (2013a) found no evidence of an argument-adjunct asymmetry in extraction from embedded questions. Their first experiment involved extraction of bare *wh*-phrases, as shown in (11).

(11) a. Hvad ved hun godt hvor man kan leje? (Da.)
What knows she well where one can rent?
'What does she know where you can rent?'

b. Hvor ved hun godt hvad man kan leje? (Da.)
Where knows she well what one can rent?
'Where does she know what you can rent?'

(Christensen et al. 2013a:55)

The *wh*-phrase in the adjunct condition in this experiment was always *hvor* 'where'.[5] In a second experiment more complex *wh*-phrases like *hvilken båd* 'which boat' and *hvor billigt* 'how cheap' were tested, again with no evidence for an adjunct-argument asymmetry. The results are in line with the observation about place (and time) adjuncts patterning like arguments with respect to extraction possibilities (e.g. Rizzi 1990), but given what we have seen so far, it is surprising that there is no difference between more complex arguments, like *hvilken båd* 'which boat' and measure-adjuncts like *hvor billigt* 'how cheap'.

The results of Christensen et al. (2013a) indicate that what is important for extraction out of embedded questions in Danish, like in English and other languages, is not a straightforward argument/adjunct distinction. However, there are certain adjuncts that do resist extraction, and we should not conclude from the study that all phrases are equally extractable from embedded questions in Scandinavian languages. Consider Swedish examples like (12), from Engdahl (1985a).

---

[5] It should be mentioned that there is a potential confound in this experiment. It is not clear how to control for the fact that (i) is a possible parse of (11b).

(i) Hvor$_1$ ved hun godt [$_{\text{CP-wh}}$ hvad$_2$ man kan leje t_2] t_1?
where knows she well what one can rent
'Where does she know what you can rent?'

Here, *hvor* 'where' originates as an adjunct at the matrix level. If this is how the question is parsed, it would be read as asking about where someone is when she knows the answer to the embedded question. This is probably not the most expected interpretation, but it is not excluded in principle in this experimental setting, where no context is available to the participant.

(12) Varför undrar du vem som skrev boken?
why do you wonder who SOM *wrote the book*

a. För att jag gärna vill ha tag på fler böcker av honom/henne.
because I would like to get hold of more books by him/her
(*varför* in matrix clause: OK)

b. #För att han/hon hoppades tjäna pengar på den.
because he/she hoped to make money on it
(*varför* in embedded clause: *)

(Engdahl 1985a:1–2)

In Swedish, like in other languages that allow some extraction from embedded questions, it is still not possible to extract the counterpart of *why*. This is illustrated here by the fact that while the question in (12) can be answered as in (12a), the answer in (12b) is not felicitous. For the answer in (12b) to make sense, the question would have to be interpreted as asking about the reason for writing the book, which would involve extraction from the embedded question. This interpretation is not available.

Why-questions were not included in the study by Christensen et al. (2013a), but Danish is no exception to the pattern illustrated in (12); *why*-questions cannot be interpreted as involving extraction from an embedded question in Danish either, as (13) reveals.

(13) Hvorfor vil du gerne vide hvem der har skrevet den her bog? (Da.)
why want you PRT *know who that has written this here book*
'Why do you want to know who wrote this book.'

a. Fordi jeg gerne vil læse flere bøger af ham eller hende.
because I PRT *want read more books by him or her*
'Because I would like to read more books by him or her.'
(*varför* in matrix clause: OK)

b. #Fordi han/hun håbede at tjene penge på det.
because he/she hoped to make money on it
'Because he/she hoped to make money on it.'
(*varför* in embedded clause: *)

(Anne Mette Nyvad p.c.)

Like in the corresponding example in Swedish, the b-answer, which gives the reason for writing the book, is not a felicitous answer to the question in (13). This means that even tough Christensen et al. (2013a) found no

adjunct/argument asymmetry in their study, not all phrases are equally easy to extract.

6.1.1 ERC and adjuncts

I now turn to the question of how Swedish ERC fits into the background I have outlined in the previous section. A first observation is that the overwhelming majority of examples in my collection involve extraction of phrases that are complements either of a verb or a preposition. It is not entirely clear how to interpret this fact. One possibility is that extracting an adjunct from a relative clause is impossible, or that such extractions are marginal. This would explain why I did not find any T-preposed phrases which function as scene-setters in ERC when I investigated this in chapter 5, since scene-setting adverbials, like *efter en stund* 'after a while', are adjuncts.

However, (Engdahl 1997:57) provides the example in (14), in which a manner adjunct is extracted from a relative clause.

(14) Så$_1$ känner jag ingen som kan måla t_1.
*like that know.*REL *I no one that can paint*
'I don't know anybody who can paint like that.'

The fact that such examples sound completely natural shows that adjunct extraction from relative clauses is possible.

Another possibility is that the small number of adjunct extractions is a result of the way the examples were collected. Most of the collection is made up of sentences that I heard or read and immediately recognized as involving a gap. It is quite likely that this collection method skews the results, since a gap might be more noticeable if it is selected or is syntactically obligatory. It could also be the case that adjunct extraction is just much rarer than argument extraction and that it does not occur often enough for my collection method to capture many examples, even though it is in principle possible.

Indeed, my example collection provides additional evidence that adjunct extraction from relative clauses sometimes occurs, but with adjuncts that do not serve a scene-setting function. There are some cases where an RC-external adjunct must be interpreted as modifying a predicate inside the relative clause, and not the matrix predicate. Consider (15) and (16).

(15) Arbetsförmedlingen har inte gjort något, men [genom Per]$_1$ är det
employment service has not done something but through Per are there
många [som fått jobb t_1].
many that got job
'The employment service has not done anything, but many people have gotten jobs through Per.'
(Radio Sweden, article text, June 3 2013)

(16) [Så mycket]$_1$ finns det ingen [som bryr sig t_1].
that much exist there no one that cares REFLX
'There is no one who cares that much.'
(Forum thread, Jan. 2013)

In (15), the context reveals that the manner-phrase *genom Per* 'through Per' must modify *få jobb* 'get a job', since it is contrasted with getting help from the employment service. In addition, *genom Per* 'through Per' cannot be an adjunct at the main clause level here, since this would lead to the strange interpretation *vara genom Per* 'exist through Per'. Similarly, the degree phrase *så mycket* 'that much' in (16) must be modifying *bry sig* 'care', since existence is presumably not a gradable property. Like in Engdahl's (1997) example (14), the T-preposed adjunct in (15) specifies a manner. The adjunct in (16) denotes an amount.

To learn more about extraction of adjuncts from relative clauses in spontaneous usage, I conducted a corpus study, which I report on in the next section.

6.1.1.1 Corpus searches

In order to find out more about possible interpretations of adjuncts in potential ERC-sentences with a variety of matrix predicates, I conducted a small corpus study using the search interface Korp provided by Språkbanken (the Swedish Language Bank) at the University of Gothenburg (Borin et al. 2012), and searched the corpora Blog mix 2012 and 2013, which contain blog text (approximately 8,5 million sentences/134 million words), and the corpora GP 2010–2013, which are made up by newspaper text from *Göteborgs-Posten*, a daily newspaper published in Gothenburg (approximately 5,3 million sentences/71 million words).

A group of matrix verbs was selected based on my collection and on the previous literature: *veta* 'know.COG', *känna* 'know.REL', *behöva* 'need', *träffa* 'meet', *finnas* 'exist', *avsky* 'detest', *beundra* 'admire', *hitta* 'find', *höra* 'hear'.

For each verb in this group, I constructed search strings based on the schema in (17).

(17) [] [(word = verb_present | word = verb_preterite)] [pos = “PN”] []{1,4} [word = “som”]

Search strings based on this schema will return hits which contain a verb, either in present or preterite tense, followed by a pronoun, followed by 1–4 unspecified words, followed by *som*. What comes before the verb, or after *som*, is not specified. In addition, I required there to be at least one word before the finite verb in order to avoid finding polar questions. The unspecified word can be sentence-initial, but it does not have to be.

The particular search string used for the verb *veta* ‘know.COG’ is given in (18).

(18) [] [(word = ”vet” | word = ”visste”)] [pos = ”PN”] [pos != ”HP” & pos != ”HS” & pos != ”HA” & pos != ”HD”]{1,4} [word = ”som”]

The search string here includes a ban on interrogative pronouns (HP), interrogative possessives (HS), interrogative adverbs (HA) and interrogative determiners (HD) in the 1–4 unspecified positions, to avoid getting examples where *veta* has an embedded question complement, which proved to be common in test searches. Such modifications were needed for some of the other verbs as well. The specific search strings used for each embedding verb can be found in Appendix A.

The order of the returned hits was randomized, and I read through them and noted all examples where a clause-initial adjunct had to be interpreted as modifying the event expressed by a predicate inside a relative clause. For each search, I read through the 50 first returned strings. Several of the searches returned less than 50 strings, however.

The most obvious result of the searches is that it is not common to begin a sentence with a T-preposed adjunct that is interpreted as modifying exclusively the predicate inside a relative clause. Many of the examples begin with a T-preposed adjunct, but in the vast majority of the cases, it is not meaningful to talk about extraction from the relative clause. Consider the examples in (19).

(19) a. – Det här är värst i storstäderna, för där finns det alltid fler
*the here is worst in big cities.*DEF *because there exist there always more*
ungar som står på kö.
kids that stand on line
‘The situation is worst in the big cities, because there are always more kids in line there.’ (GP2013)

b. På 1 maj hörde vi plötsligt mässingsinstrument som spelade marschmusik
on 1 May heard we suddenly brass instruments that played march music
utanför hennes köksfönster.
outside her kitchen window
'On May first, we suddenly heard brass instruments playing march music outside her kitchen window.' (Blog mix 2012)

c. Men nu vet jag en gård som kommer jubla.
*but now know.*COG *I a farm that will cheer*
'But now I know of a farm that will cheer.' (GP2013)

In (19a) and (19b), the events and states in the matrix and the embedded clauses are simultaneous. In (19c), it is possible to distinguish the time of the matrix event and the event in the embedded clause which occurs in the future relative to the time of the matrix clause event, and the adjunct specifies the time of the matrix event. Jensen (2002:114) provides two similar constructed examples from Danish (20). I have adapted the example-formatting and added glosses.

(20) a. den aften var der en kritiker der så forestillingen (Da.)
*that evening was there a critic who saw performance.*DEF
'There was a critic who saw the show that evening.'

b. i går talte jeg med hende der har repareret vores vaskemaskin. (Da.)
yesterday talked I with her who has repaired our washing machine
'Yesterday I talked to the woman who repaired our washing machine.'

The example in (20a) is similar to (19a) and (19b) in that the matrix event and the event in the embedded clause are simultaneous. In (20b) we can distinguish the events, and Jensen reports that *i går* 'yesterday' must be interpreted as modifying the matrix event.[6] She concludes that time adverbials cannot be extracted from embedded clauses which are not arguments of the verb in the matrix clause, a group where she includes relative clauses. However, as I will show below, there are in fact some clear ERC-cases.

While most of my searches did not return any clear examples of extraction, searches with the verb *veta* 'know.COG' is an exception. Here, I found two examples that involve adjuncts which specify the time of the event expressed in the relative clause, as in (21) and (22).

[6] In this particular example there is a tense clash between the temporal adverb and the perfect tense in the embedded clause, which means that the example is not as clear an illustration as it could have been.

(21) Barnen har ju äntligen börjat sova nån halvtimme längre (mkt för
*kids.*DEF *have* PRT *finally began sleep some half hour longer a lot for*
småbarnsföräldrar)
parents of young children

så [imorgon]$_1$ vet jag tre trötta barn [som kommer vara ganska
*so tomorrow know.*COG *I three tired kids that will be pretty*
sega t_1].
tired

"The children have finally began sleeping half an hour longer (a lot for parents of young children) so I know of three tired kids who will be pretty tired tomorrow.'

(Blog mix 2013)

(22) [När Findus kommer in]$_1$ så vet jag en liten grabb [som blir
when Findus comes in PRT *know.*COG *I a little guy that becomes*
glad t_1].
happy
'I know of a little guy who will be very happy when Findus comes inside.'

(Blog mix 2013)

In each of these examples, there is an initial time expression that must be interpreted as modifying only the predicate inside the embedded clause, based on the interpretation of the rest of the sentence. The example in (21) is from a blog post where a parent is writing about getting her children back on a semester schedule after the Christmas break. The children have been getting used to sleeping in during the holidays, and for that reason, the parent knows that the three children will be *sega* 'tired' the next day. The adjunct *imorgon* 'tomorrow' is interpreted as modifying *vara sega* 'be tired' here, and not the matrix verb. If the adjunct *imorgon* were modifying the matrix *veta* 'know.COG', the blog writer would be making a statement about what she will know of on the following day, which is not very plausible in this context.

The second example (22) involves a cat, Findus, and a toddler who gets very happy when the cat comes inside. As in (21), the adjunct has to be interpreted as modifying the predicate inside the relative clause, and not the matrix *veta*, for similar reasons. Both of the examples sound entirely natural to my ear. It seems then, that even though it is very rare in spontaneous use, at least adjuncts that express time can show this behavior when the embedding verb is *veta* 'know.COG'.

The finding from chapter 5 which motivated the investigation of extraction of adjuncts from relative clauses was that there are certain adverbial expressions

which can be T-preposed in the local clause and out of *att*-clauses, but which cannot be interpreted as modifying a predicate inside the relative clause in contexts which otherwise permit ERC. The relevant examples are repeated in (23).

(23) a. Efter en stund skrek nån ”jävlar”, ”jävlar”.
after a while yelled someone damn damn
‘After a while, someone yelled “damn”, “damn”.’

b. [Efter en stund]$_1$ tror jag [$_{CP}$ att någon skrek ”jävlar”, ”jävlar” t_1].
after a while think I that someone yelled “damn”, “damn”.
‘I think that someone yelled “damn”, “damn” after a while.’

c. *[Efter en stund]$_1$ kom det nån [$_{RC}$ som skrek ”jävlar”, ”jävlar” t_1].
after a while came there someone that yelled damn damn
Impossible reading: ‘a person x came and after a while x yelled “damn”, “damn”’

The adverbial *efter en stund* ‘after a while’ in (23b) is most naturally interpreted as modifying the yelling-event, i.e. the event denoted by the predicate inside the *att*-clause. But in (23c), where the predicate *skrika* ‘yell’ is inside a relative clause, *efter en stund* cannot be interpreted as modifying the yelling-event; it can only modify the coming-event in the matrix clause.

It should be noted that tense matters for the adjunct interpretation here. If the tense in the matrix clause in (23b) is changed to the preterite, the interpretation where *efter en stund* modifies the matrix event would be preferred. The point of (23b), however, is that when that interpretation is not available or is unlikely due to incompatibility between the adjunct and the matrix predicate event structure, the adjunct can be interpreted as modifying the event in the embedded clause. This makes *att*-clauses different from relative clauses, where a similar situation results in an ungrammatical sentence, as is illustrated in (24).

(24) *[Efter en stund]$_1$ känner jag en [$_{RC}$ som skrek ”jävlar”, ”jävlar” t_1].
*after a while know.*REL *I one that yelled damn damn*

If the reason that *efter en stund* cannot be interpreted as modifying the yelling-event in (23c) were simply that there is a preference for such adverbials to modify the closest predicate, we would expect (24) to be grammatical. Here the matrix verb is stative, which makes the matrix interpretation of *efter en stund* unavailable. Even so, the adjunct cannot be interpreted as modifying the yelling-event in the relative clause. Instead the example is unacceptable on any reading.

Efter en stund is not the only adjunct that follows the pattern in (23) and (24). Two others that exhibit the same effect are *till slut* 'at last' and *snart* 'soon'. Consider the examples in (25), which are from the corpus searches.

(25) a. Till slut hittade jag några flickor i folkdräkt som sålde kanderad
at last found I some girls in traditional costume that sold candied
mandel.
almonds
'At last, I found some girls in traditional costumes who sold candied almonds.'
(Blog mix 2013)

b. Snart träffade hon en man som hon fick barn med.
soon met she a man who she received child with
'Soon thereafter she met a man whom she had a child with.' (GP 2010)

In both of these cases, the adjunct has to be interpreted as modifying the matrix events, *träffa* 'meet' and *hitta* 'find'. If we construct examples parallel to those in (23a) and (23c), we find a similar pattern with respect to possibilities to interpret the adjuncts as with *efter en stund* 'after a while', as shown for *till slut* 'at last' in (26).

(26) a. Till slut skrek nån på domaren.
*at last yelled someone at referee.*DEF
'At last, someone yelled at the referee.'

b. [Till slut]$_1$ tror jag [$_{CP}$ att nån skrek på domaren t_1].
*at last think I that someone yelled at referee.*DEF
'I think that someone yelled at the referee at last.'

c. *[Till slut]$_1$ kom det nån [$_{RC}$ som skrek på domaren t_1].
*at last came there someone that yelled at referee.*DEF
Intended reading: 'a person x came and x yelled at the referee at last"

d. *[Till slut]$_1$ känner jag en [$_{RC}$ som skrek på domaren t_1].
*at last know.*REL *I one that yelled at referee.*DEF
Intended reading: 'I know a person x and x yelled at the referee at last'

I have adjusted these examples slightly in order to make them sound more natural. Interestingly, as with *efter en stund* 'after a while', it is possible to interpret *till slut* 'at last' as modifying only the predicate in an *att*-clause embedded under *tro* 'think', but not the predicate inside a relative clause.[7] The

[7] If the adjunct is contrastively stressed, other possibilities of interpretation become available, as

pattern is the same for *snart* 'soon'.

Like *efter en stund*, *till slut* and *snart* are time adverbials. Other time-denoting expressions, for example adverbs of frequency, behave in a similar way. Examples with T-preposed adverbs of frequency do occur in the corpora that I have checked. In the examples that I have found they modify the entire sentence, not just the relative clause (27), quite like the other adverbial expressions I have discussed here. In (27) we see an example with *ofta* 'often'.

(27) Ofta finns det skidtester som är gratis [...]
often exist there ski tests that are free
'There are often ski tests which are free.'
Impossible interpretation: 'There are ski tests which are often free.'

(Blog mix 2012)

This example is about the possibility of trying out new skis at a ski resort. Crucially, the example is about the frequency with which there are free testing opportunities, not about any particular ski tests which are often free, but sometimes paid for, which we would expect if *ofta* could modify the predicate inside the relative clause.

A key question is what sets the adjuncts in (21) and (22), *imorgon* 'tomorrow' and *när Findus kommer in* 'when Findus comes inside' apart from adjuncts like *efter en stund*, *till slut*, *snart*, and *ofta*, which are also time expressions. One difference is that they refer to specific points in time, either defined deictically, as with *imorgon*, or by way of description by the adjunct itself, as *när Findus kommer in*. The expressions *efter en stund*, *snart* and *till slut* are more complicated. While they do relate to points in time, to get to the time they refer to, we need to look at the preceding linguistic context and the sequence of events, and the full denotation involves relating points in time to each other. Adverbs of frequency, like *ofta*, are also more complicated than simple points in time. In order to give the semantics of *ofta* we would need to say something about some interval, and several points on it.

Summing up, we have learned from the corpus searches that T-preposing an adjunct that belongs inside a relative clause is very rare. Even so, certain adjuncts denoting times can nevertheless be interpreted as modifying the event expressed by a predicate in a relative clause, in particular if they are deictic and/or refer to a point in time. But what about other types?

Given the findings of Cinque (1990) and Rizzi (1990) regarding extraction

---

we will see in section 6.2.1.

from syntactic domains like embedded questions, it is not surprising that the adjunct ERC examples we find in the corpora are precisely time-denoting adjuncts. Perhaps the event specifications in these cases license a temporal index and allow the adjuncts to take part in operator-variable chains, and this is what makes the extractions possible. If this were the case, we might expect that extracting adjuncts denoting amounts, manners and reasons should be impossible. However the three previously discussed examples (14), (15), and (16) indicate that this is not the correct conclusion, since they involve adjuncts that denote manners (*så* 'like that' and *genom Per* 'through Per') and an amount (*så mycket* 'that much'), and yet are very natural sounding. This question requires further investigation, as does the question regarding which embedding verbs are compatible with T-preposed adjuncts in ERC, since only the search string with *veta* 'know.COG' yielded any results. Are there acceptable examples with other embedding predicates than *vara* 'be', *finnas* 'exist' and, *veta* 'know.COG'?

6.1.1.2 Questionnaire

In order to find out more about ERC for a variety of adjunct types, and with matrix predicates other than *vara* 'be', *finnas* 'exist', and *veta* 'know.COG', I constructed a questionnaire. As described in section 3.2.1, the questionnaire was exploratory, and included several types of ERC-sentences, in order to identify different factors that affect acceptability of extraction. The items included sentences with and without extraction, extraction of different adjunct types, extraction from relative clauses with definite and indefinite relative complexes, and sentences with a variety of matrix predicates. The main purpose was not to compare the effects of different factors, but to see whether certain types of extraction are possible at all and whether they sound natural to Swedish speakers.

The questionnaire items are partially parallel to allow for certain comparisons, but since several different variables are explored at once, I could not set up a full experimental design. This means that it is not always possible to link differences between items to single variables. Instead of reporting fine grained quantitative differences between items, I therefore report only the number of participants who picked each of the given alternative ratings for each item.

The questionnaire was distributed electronically and consisted of 32 items. An example item is given in (28).

(28) a. Linda: Min syster och jag är oense om när vi ska äta ikväll. Hon
Linda: my sister and I disagree about when we shall eat tonight she
tycker att vi ska äta klockan nio, men jag tycker att det är alldeles
*thinks that we should eat clock.*DEF *nine but I think that that is* PRT
för sent.
too late
'Linda: My sister and I disagree about when to have dinner tonight. She thinks we should eat at nine, but I think that is way too late.'

b. Ida: Ja, så sent vet jag ingen som brukar äta middag.
*Ida: yes that late know.*COG *I no one that tend eat dinner*
'Ida: Yes, I don't know of anyone who eats dinner that late.'

Each item consisted of a context sentence, here (28a), and a test sentence, here (28b). The context sentence and the test sentence were shown on the screen together, and the task for the participants was to rate the naturalness of the test sentence as a follow-up to the context sentence. Participants got to chose from three alternatives: "natural", "somewhat strange", and "unnatural". The participants were also given the possibility of adding free text comments for each test item. The questionnaire design builds on the assumption that ungrammatical or unacceptable test sentences would not be rated as natural.

There were four practice sentences, two of which were supposed to be ungrammatical or unnatural in relation to the context sentence. All of the test participants saw the items in the same order. The questionnaire was answered by 16 native Swedish speakers between 24 and 79 years old, all of whom had finished high school. In all but one case they had also finished at least some courses at a university level. The questionnaire and an overview of the results are given in Appendix B.

A first observation is that there are contexts where extraction of an adjunct is rated as being just as natural as a version of the sentence with the adjunct in situ. This is illustrated by the two test sentences in (29), which were presented as continuations to the same context sentence.[8] The numbers in parentheses represent the number of participants who chose the answers "natural", "somewhat strange", and "unnatural". For the sake of clarity, I have included brackets around the relative clause and indicated the trace position of the extracted phrase in the examples in this section, but the items were presented to the questionnaire participants without such markings.

[8] The two test sentences shown together in (29) were separated by other test sentences in the questionnaire (see Appendix B).

(29) Linda: Min syster och jag är oense om när vi ska äta ikväll. Hon tycker
Linda: my sister and I disagree about when we shall eat tonight she thinks
att vi ska äta klockan nio, men jag tycker att det är alldeles för sent.
*that we should eat clock.*DEF *nine but I think that that is* PRT *too late*
'Linda: My sister and I disagree about when to have dinner tonight. She thinks we should eat at nine, but I think that is way to late.'

a. Ida: Ja, [så sent]$_1$ vet jag ingen [som brukar äta middag t_1].
*Ida: yes that late know.*COG *I no one that tend eat dinner*
'Ida: Yes, I don't know of anyone who eats dinner that late.'
(Questionnaire 11, 5, 0)

b. Ida: Ja, jag vet ingen [som brukar äta middag så sent].
*Ida: yes I know.*COG *no one that tend eat dinner that late*
'Ida: Yes, I don't know of anyone who eats dinner that late.'
(Questionnaire 12, 2, 1)

Both (29a), where the adjunct *så sent* 'that late' has been extracted, and (29b) without extraction, are rated as natural by most participants, and the number of people who rated the two continuations as natural are comparable.

In other cases, there is a small difference between the versions of the sentence with and without extraction with respect to ratings. Consider (30).

(30) Anna: Mina barnbarn sjöng så fint att alla i kyrkan grät.
*Anna: my grand children sang so nicely that everyone in church.*DEF *cried*

a. Sven: [Så fint]$_1$ känner jag ingen [som kan sjunga t_1]!
*Sven: that nicely know.*REL *I no one that can sing*
'I don't know anyone who can sing that well!'
(Questionnaire 9,6,1)

b. Sven: Jag känner ingen [som kan sjunga så fint]!
*Sven: I know.*REL *no one that can sing that nicely*
'I don't know anyone who can sing that well.'
(Questionnaire 14,1,1)

In this example a large proportion of the participants (9 out of 16) rates the extraction sentence in the continuation as natural, but the proportion is smaller than for the sentence in which the adjunct is in situ (14 out of 16). It is worth noting that only one participant found each of these two items unnatural, and that this was the same participant in both cases. The rest of the participants who did not find (30a) natural all rated it as somewhat strange.

A similar test item is shown in (31). Here the embedding verb is *träffa* 'meet' instead of *känna* 'know', and we see a similar pattern, but the extracted phrase

is an argument rather than an adjunct.

(31) Olle: Jag var på en marknad igår där de hade en mjölkdrickar-
Olle: I was at a market yesterday where they had a milk drinking
tävling. Några av tävlingsdeltagarna drack mer än fem liter mjölk!
*contest some of competitors.*DEF *drank more than five liters milk*

a. Ida: [Så mycket mjölk]$_1$ har jag aldrig träffat någon [som kan dricka t_1
Ida: that much milk have I never met someone that can drink
på en gång]!
at once
'I have never met anyone who can drink that much milk at once!'
(Questionnaire 9,5,2)

b. Ida: Jag har aldrig träffat någon [som kan dricka så mycket mjölk
Ida: I have never met someone that can drink that much milk
på en gång]!
at once
'I've never met anyone who can drink that much milk at once!'
(Questionnaire 15,1,0)

The fronted phrase in (31a) is an amount-phrase, and the example was included because of the observation from the previous literature about amount-phrases in *wh*-questions (Kroch 1989, Cinque 1990, Rizzi 1990). As with (30), the proportion of participants who rated the version of (31) with extraction as natural is smaller than the proportion who rated it as natural with the phrase in situ. Importantly however, a majority of the participants reported that the version with extraction is a natural continuation, which means that we can not conclude that extraction of amount phrases from relative clauses is ungrammatical in general. It is probably important that the phrase is T-preposed and not a fronted *wh*-phrase. We take a closer look at ERC in *wh*-questions in 6.2.2.

With respect to the extractability of different adjunct types, the questionnaire answers provide additional evidence that time adjuncts are not the only type of adjunct that can be extracted. The ratings for (30) demonstrate that many of the participants find T-preposing of adjunct manner phrases natural. Another test item that shows this is given in (32).

(32) IPCC säger att det är 95 procent säkert att det är vi människor som har
IPCC says that it is 95 percent sure that it is we humans that have
orsakat merparten av klimatförändringarna sen 1950.
*caused most of climate changees.*DEF *since 1950*

[Tydligare än så]$_1$ går det inte att hitta någon forskare [som vill uttala
clearer than that goes it not to find some researcher that want express
sig t_1].
REFLX

'It is not possible to find a researcher who wants to express themself any more clearly on the matter.'

(Questionnaire 11, 4,1)

The continuation here involves T-preposing of the manner phrase *tydligare än så* 'clearer than that', which can only be interpreted as an adjunct of *uttala sig* 'express oneself' in the relative clause. A manner phrase, *så fint* 'that nicely', is also T-preposed in the continuation in (30a). This phrase is most naturally interpreted as modifying *sjunga* 'sing' in the relative clause.

These two T-preposed adjuncts, as well as *så sent* 'that late' in (29a), and the amount phrase *så mycket mjölk* 'that much milk' in (31a) all include the adverb *så* 'that'. *Så* is demonstrative, as described in chapter 5, and finds a linguistic antecedent in the context in each of these examples, but can also be used deictically. In (29a), the antecedent is the phrase *klockan nio* 'nine o'clock' which denotes the controversial dinner time, and in (31a) *så mycket mjölk* 'that much milk' refers back to the amount five liters, which is mentioned in the context sentence.

What is denoted in (30a) and (32) is actually a specific degree of a manner introduced in the preceding sentence. In (30a) a degree of niceness of singing is denoted (nice to the degree that all of the people in the church cried), and in (32) a degree of clarity is referred to. To get at this degree is slightly more complex than in (30a), because the reader has to figure out that speaking with 95 percent certainty is being clear to a certain degree.

The continuation in (33) shows extraction of a reason adjunct.

(33) Hundägare anger flera fördelar med att ha hund. Hundar kan hålla en sällskap, och att ha hund leder till en mer aktiv livsstil för många. De kan också tränas att hämta tidningen på morgonen,
'Dog owners state several benefits of having a dog. Dogs can keep you company, and having a dog leads to a more active lifestyle for many people. They can also be trained to fetch the newspaper in the morning,'

och [av precis den anledningen]$_1$ vet jag många [som skaffat
*and of precisely that reason.*DEF *know.*COG *I many that got*
hund t_1]
dog
'and I know of many people who have gotten a dog for precisely that reason.'
(Questionnaire 10,6,0)

In (33) the context is set up so that *den anledningen* 'that reason' in the T-preposed adjunct has a linguistic antecedent in the preceding sentence, and functions as an aboutness topic in a focus chaining relation (cf. section 5.1). The reason for getting a dog which is referred to is that dogs can be trained to bring the newspaper in the morning, and we can think of the sentence as saying something about that particular reason. In fact, we can interpret all of the fronted phrases with *så* as aboutness topics as well, if we see each of the ERC-sentences with *så* as saying something about a specific degree of niceness, a specific time to eat dinner, a set amount of milk, and so on.

The stimulus materials for the questionnaire were designed to avoid matrix predicates which allow for existential interpretations of their complement DP, in the sense specified by Keenan (1987) (see section 4.3.1.1). The answers show that a large part of the participants find ERC of adjuncts natural with embedding predicates like *känna* 'know.REL' and *gå att hitta* 'be possible to find', in addition to with *veta* 'know.COG', which we had already seen occurs in spontaneously produced ERC-sentences with T-preposed time adjuncts (see section 6.1.1.1). None of these predicates allow for the existential interpretations of DP. Extraction of amount-phrases with the embedding predicate *träffa* 'meet' is also judged as natural by a majority of the participants, as (31a) showed.

Due to the exploratory nature of the questionnaire, it is impossible to draw conclusions about the role of the embedding predicate, but it does seems to matter. Consider (34).

(34) Ines: Min mamma åker till jobbet klockan sex varje morgon.
*Ines: my mother goes to work.*DEF *clock.*DEF *six each morning*

Erik: [Så tidigt]$_1$ beundrar jag verkligen folk [som orkar gå upp t_1]!
Erik: that early admire I really people that manage go up

'Erik: I really admire people who manage to get up that early!'
(Questionnaire 3,11, 2)

Like in (29a), a time adverbial has been extracted in this example, but instead of *veta* 'know.COG', the embedding predicate is *beundra* 'admire' (34). Whereas (29a) was rated as a natural continuation in its context by a majority of the

participants, (34) is perceived as a 'somewhat strange' continuation to the context sentence by most participants, and several participants wrote in the free text comments that they would prefer the sentence with the adverbial in situ.

As we saw there is some individual variation in what the participants accept. Interestingly, there is more variation in the sentences with extraction than in the sentences without extraction. Different extractions also elicit different answer patterns, such that some extractions are accepted by almost everyone, whereas others are rated 'natural' by fewer participants. From this study it is not possible to tell whether adjunct type or embedding verb has the strongest effect. This would require a larger study.

It is worth mentioning that it might matter for the results that the questionnaire was presented in written form. Since the participants read the test items, it is impossible to know which prosody they imagined, and whether the fronted phrase was contrasted, for example. Even though they were provided with a sentence which was supposed to indicate the discourse function of the fronted phrase, and the information structure of the sentence, it is likely that some of the variation between speakers can be explained by the prosody each speaker imagined.

6.1.1.3 Summing up

In light of the results of the corpus searches in section 6.1.1.1 and of the questionnaire in section 6.1.1.2 we can conclude that extraction of adjuncts is not common in Swedish ERC, but that it is nevertheless possible. Given the right matrix predicate, certain adjuncts can be extracted with a natural-sounding result. The embedding context does not have to be existential in the sense of Keenan (1987). We also saw that there does not seem to be a specific pattern such that only adjuncts that denote times and places are possible to extract. Manner and reason adjuncts can be extracted as well, and so can amount-phrases.

More generally, the results indicate that what matters for being extracted from a relative clause is not syntactic function, even with the qualification regarding referential and non referential theta-roles and referential indices that are available to a subset of the adjuncts, i.e those that specify time and place. Rather, extractability seems to be connected to the semantic type of the extracted phrase, and whether it is linked to the context. Extractable adjuncts are deictic, or denote a point in time or a specific degree on a scale. They often have a linguistic antecedent, or are contrasted. Seen from the perspective of information structure, it is possible to think of the fronted adjuncts as aboutness topics. In the following section, I will review a few pragmatic and semantic approaches to weak islands, and relate them to these results.

6.2 D-linking, individuals, and semantic restrictions

In this section, I relate Swedish ERC to accounts of weak islands that connect extractability to **D(iscourse)-linking** (Pesetsky 1987) or to the semantic type of the gap left by the preposed phrase, and to possible interactions with scopal elements (Comorovski 1989, Frampton 1990, Szabolcsi & Zwarts 1993, Szabolcsi 2006). In doing so I also aim to complete the picture somewhat by discussing examples with *wh*-questioning out of relative clauses.

As mentioned in section 6.1.1, several proposals about extraction from weak islands involve the notion of D-linking (Comorovski 1989, Cinque 1990). The term is due to Pesetsky (1987). He observes that *wh*-phrases like *which book* on the one hand, and *who* and *what* on the other, exhibit differing behavior with respect to syntactic conditions on question formation, such as **subjacency** and **superiority** (Chomsky 1973). He relates the differences to Discourse linking (D-linking). *Which*-phrases are linked to the discourse in a way that *who* and *what* are usually not. Consider (35).

(35) Which cup would you like?

In a question like (35), the *which*-phrase ranges over a set of cups. According to Pesetsky, both the speaker and hearer must have this set in mind in order for the question to sound natural. The set could have been brought into the minds of the conversation participants by the preceding linguistic context, or be otherwise contextually salient. For instance, (35) could be uttered if someone has offered a guest something to drink, and the guest gets to chose from a number of cups in the kitchen. A phrase which is linked to a contextually defined and salient set in this way is D-linked. If a *which*-question is asked when there is no such set in the mind of the hearer, the question will sound odd (Pesetsky 1987:108).[9] This shows that *which*-phrases are inherently D-linked. *Who* and *what*, on the other hand *can* be D-linked, but they can also be used in open questions where there is no need for a contextually defined set.

Several researchers note that question formation out of embedded questions requires that the *wh*-phrase be D-linked. For example, Comorovski (1989) argues that Swedish, Italian, Bulgarian, and Romanian provide evidence for this. She gives the examples in (36) showing the contrast between extraction of

[9] Of course it might be possible for the hearer to accommodate or somehow find a set in the context after the question is asked, in which case the question should be felicitous.

D-linked and non-D-linked *wh*-phrases in Romanian. The examples, which are from Comorovski (1989:82), are slightly reformatted.

(36) a. Despre care$_i$ [$_S$ stii [cine$_j$ [$_S$ e$_j$ i - a povestit e$_i$]]]? (Rom.)
about which you-know who to-him has told
'Which one do you know who told him about?'

b. *Despre ce$_i$ [$_S$ stii [cine$_j$ [$_S$ e$_j$ i - a povestit e$_i$]]]? (Rom.)
about what you-know who to-him has told

The D-linked question in (36a) is felicitous, but its non-D-linked counterpart in (36b) is not. Comorovski (1989:81) quotes the two examples below from Maling (1978) and Engdahl (1980a), which illustrate a parallel case in Swedish.

(37) a. *Vad visste ingen vem som skrev?
'What does no one know who wrote?'
(Maling 1978:84)

b. Sven undrar vilken bok alla studenter minns vilken författare som skrev.
'Sven wonders which book all students remember which author wrote.'
(Engdahl 1980:65)

The question with *vad* 'what' in (37a), is reported to be ungrammatical by Maling, whereas the embedded question with *vilken bok* 'which book' in (37b) is acceptable. As Comorovski points out, one difference between the two examples is that the *wh*-phrase in (37b) is D-linked.

According to Comorovski, the reason for contrasts such as the ones in (36) and (37) is that there is a general felicity condition on questions:

(38) A question is felicitous only if the hearer can check the truth of its presupposition(s). (Comorovski 1989:90)

As an example, constituent questions presuppose existential sentences. Consider (39a), which presupposes (39b).

(39) a. Who ate my sandwich?

b. Someone ate my sandwich.

The felicity condition in (38) states that in order for a question to be felicitous, it must be possible to check whether its presuppositions are actually true. It is straightforward to do this in a question like (39a), which contains only one *wh*-phrase. In this particular case, we just need to know whether the sandwich in question has been eaten or not. The presuppositions of questions which

involve extraction from an embedded question are harder to check. Comorovski (1989:92) gives the example in (40a), which has the presupposition in (40b).[10]

(40) a. *Ce$_i$ *pro* stii cine$_j$ e$_j$ a distrus e$_i$? (Rom.)
what you-know who has destroyed

b. (∀x)(thing(x)) → (∃y)(person(y) ∧ destroy (y,x))

Comorovski's proposal is the following: if the *wh*-phrase in sentences like (40a) is not D-linked, it is not possible to check the presupposition, since the set over which the phrase quantifies is unknown. Therefore the question is infelicitous. In sentences like (36a), checking the corresponding presupposition that it introduces is made possible by the D-linking. This is so because if there is a specific set of referents that the hearer has in mind, it is possible to check for each referent if the presupposed proposition holds of it.

As we saw in section 6.1.1, Cinque, like Comorovski, argues that D-linking is a requirement on extraction from weak islands, and he connects this to "referentiality", which he argues subsumes D-linking. Let's look again at the example in (8), here repeated as (41):

(41) a. *[How many pounds]$_1$ do you wonder whether he weighed t_1?

b. ?[How many fish]$_1$ do you wonder whether he weighed t_1?

From the perspective of Cinque, (41a) is unacceptable because for it to be grammatical, we would have to be asking about some specific pounds whether someone weighed them, which is a quite odd thing to be asking. As mentioned previously, the marginally grammatical (41b) can presumably only have precisely such a reading. It is a question about specific fish, not about an amount. A similar discussion about extraction of amount phrases can be found in Kroch (1989). Kroch frames the discussion in terms of existential presuppositions. Consider (42).

(42) a. How much did the book cost? (Kroch 1989:2)

b. There is an amount of money such that the book cost that amount. (Kroch 1989:7)

The proposition in (42b) is a presupposition of the grammatical and fully acceptable amount-quantified question in (42a). In (43), we see a question which involves an amount phrase being extracted out of an embedded question

[10] For a detailed argument showing that (40b) is a presupposition of the question in (40a), see Comorovski (1989:86–90).

(43a), along with its presupposition in (43b). The question is unacceptable, and, as we have seen, similar questions have been treated as ungrammatical (Cinque 1990, Frampton 1990, Rizzi 1990).

(43) a. *How much money was John wondering whether to pay?

b. There was an amount of money such that John was wondering whether to pay it.

(Kroch 1989:8)

According to Kroch (1989), this example is not ungrammatical, but is perceived as unacceptable because it is unusable in almost all situations. On his account, questions like (43a) are unacceptable because their presuppositions are implausible. In support of his analysis, he points out that if a context is set up where there is a situationally specified amount, amount-quantified questions become more acceptable. Consider (44).

(44) a. How many points are the judges arguing about whether to deduct?

b. There is a number of points about which the judges were arguing whether to deduct that number.

(Kroch 1989:8)

Kroch (1989) proposes that (44a) can be made acceptable under certain circumstances, namely if we imagine a context, e.g. a sports tournament in a sport where the judges deduct specified amounts of points for rule violations. In such a situation it would be possible for the judges to be debating whether to deduct a certain amount of points, and for a spectator to ask another spectator the question in (44a).

Kroch (1989) also provides the example in (45), which he takes as an argument against encoding complement/adjunct extraction asymmetries in the syntax.

(45) That quickly, I wonder whether anybody can run.

(Kroch 1989:15, footnote 4)

In this example, a manner adjunct has been topicalized. The example clearly bears some similarity to the ERC-sentences with T-preposed adjuncts with *så* 'that' or 'like that' that we have seen.

Cinque (1990) and Kroch (1989) largely aim to capture the same type of data, but their accounts differ in how they explain what goes wrong in the unacceptable examples. For Cinque, the reason that the non-D-linked version of extraction from embedded questions is not acceptable is a matter of syntax. Since such phrases cannot bind variables at a distance in the framework that Cinque develops, the trace needs to be licensed by an intermediate trace via

antecedent-government, which there is no room for in the embedded question's C-domain. For Kroch, the unacceptability of such questions is a matter of failed presuppositions, and the unacceptability is not encoded in the syntax at all. Kroch's account is thus similar to Comorovski's. Essentially, both Comorovski and Kroch attribute the restriction to the pragmatics/semantics, arguing that there is no need to posit a syntactic constraint on question formation from embedded questions.

In part building on the work we have reviewed here, Szabolcsi & Zwarts (1993) and Szabolcsi (2006) propose a scopal account of weak islands, which makes reference to the algebraic structure of the denotation domain of the extracted phrase. On this view, weak island effects arise when a scopal element intervenes between an extracted phrase and the extraction gap. Scopal elements, like *wh*-phrases, negation, and so on, are associated with certain operations. If a *wh*-phrase scopes over another scopal element, the operations associated with the intervening scopal element will be performed in the denotation domain of the *wh*-phrase. *Wh*-phrases that can escape all weak islands range over individuals, which can be collected into Boolean algebras. Phrases that do not denote in the right kind of domain can be extracted only if their domain is turned into an unordered set by D-linking. Consider (46).

(46) [How much milk$_i$ [$_D$ did [$_\beta$n't] you drink [$_\alpha$ _$_i$]]]? (Szabolcsi 2006:519)

In (46), the *wh*-phrase *how much milk* scopes over negation. Given the syntactic configuration created by *wh*-movement, we can talk about an extraction domain D, a scopal element, here negation, β, and a gap, α. To calculate the denotation of the whole sentence, we must calculate the denotation of D. In order to do this, we start from the bottom. Szabolcsi (2006) assumes that the gap of *how much milk* denotes an amount, for example one glass, or three liters. In (46) we then get *you drink* [$_\alpha$ _$_i$]. Given that *how much milk* denotes an amount, *you drink* [$_\alpha$ _$_i$] should denote the amount of milk you drank, for instance one glass. So far, everything is alright. The problem occurs when we get to the scopal element. To calculate the denotation of *n't you drink* [$_\alpha$ _$_i$], we should take the complement of the amount of milk that you drank. This is not possible since the complement of *how much milk* isn't defined, according to Szabolcsi (2006). This is because the *wh*-phrase in this case ranges over an amount, and Boolean operations like negation can only be performed on unordered sets. On this view then, we cannot extract a phrase like *how much milk* here, because the scopal intervener would force us to perform an operation which the denotation of the gap is not defined for.

If instead we let the *wh*-phrase range over a set of D-linked amounts, for ex-

ample three specific, individual glasses of milk, things would be different. Then we could take the complement of the glass of milk that had been consumed, in this case two glasses. Kroch's (1989) example in (44) within the context of the sports tournament is one situation where a *wh*-phrase ranges over such a set of D-linked of amounts.

Szabolcsi (2006:497) notes that D-linking is a way to make almost all phrases immune to weak islands, because it turns non-individuated domains into discrete individuals. This makes it quite hard to distinguish the predictions of the pragmatic D-linking accounts of Comorovski (1989) and Kroch (1989), on the one hand, and Szabolcsi & Zwarts (1993) and Szabolcsi (2006) semantic account, on the other.

6.2.1 Taking stock

At this point it will be useful to return to the types of adjuncts that we have seen in this chapter, and reflect on what sets apart the extractable adjuncts from those that resist extraction. An overview of the types of adjuncts we have seen in the chapter so far is given in table 6.1.

Table 6:1. *Adjunct types*

Adjunct	Extractable	Denotation	Example
imorgon 'tomorrow' , *när Findus kommer in* 'when Findus comes inside'	yes	point in time	(21), (22)
så mycket 'that much', *så sent* 'that late', *så fint* 'that nicely', *tydligare än så* 'clearer than that'	yes	specific degree on a scale	(16), (29a), (30a), (32)
av precis den anledningen 'for precisely that reason'	yes	specific reason	(33)
genom Per 'through Per'	yes	member of contrast set	(15)
efter en stund 'after a while', *snart* 'soon thereafter', *till slut* 'at last', *ofta* 'often'	no	mixed	(23c), (25a), (25b), (27)

The extractable adjuncts denote a point in time, a degree on a scale, a specific reason, or a member of a contrast set. The fronted phrase can also pied-pipe

a PP, as in (33), where *av precis den anledningen* 'for precisely that reason' is extracted.

As already mentioned, there are several of the extractable adjuncts that make it difficult to account for the distinction between extractable and non-extractable phrases in terms of theta-roles. The question is whether the extraction possibilities can be better understood on the pragmatic and semantic approaches to weak islands that I have reviewed in this section.

From the perspective of the algebraic approach, the question is whether the phrases can be interpreted as denoting individuals, i.e. as denoting entities that can be collected into unordered sets. With respect to points in time and degrees on a scale, this might seem counterintuitive at first, since they both seem to be naturally perceived as ordered with respect to other points in time, and to other degrees on the same scale. However, the T-preposed adjuncts seem to refer to specific individual points or degrees in each case.

Turning first to the extracted time adjuncts, these refer to just one specific point in time in the cases where they are extracted. In the examples that I have found, they are either deictic, like *imorgon* 'tomorrow', or else refer to a specific point which is defined by the adjunct itself by description, like *när Findus kommer in* 'when Findus comes inside'.

T-preposed degree phrases all involve the demonstrative adverb *så* 'that', which is most likely important for their extractability. *Så* picks out a certain degree that is available in the context, either linguistic or non-linguistic. Such cases can be seen as instances of D-linking. Similarly, the phrase *av precis den anledningen* 'for precisely that reason' clearly refers to a specific reason that is available in the immediate context, and is thus also D-linked.

What about phrases that invoke a contrast, like (15), where the manner phrase *genom Per* 'through Per' is extracted? In section 5.2.1 and section 5.2.2, I mentioned that T-preposed phrases in ERC sometimes evoke a contrast, and are contrastively stressed. Contrastive stress is used when a phrase is picked out from a comparison set and emphasized (e.g. Rooth 1992). An example is given in (47).

(47) a. Den nya Star Wars-filmen har premiär ikväll, ska du gå?
the new Star Wars movie has premier tonight shall you go
'The new Star Wars movie premiers tonight, are you going?'

b. Nej, jag tänkte se den iMORgon.
no I think see it tomorrow
'No, I'm going to see it toMORRow.'

Imorgon 'tomorrow' in this example contrasts against *ikväll* 'tonight', which is

given in the context question. The comparison set is thus given in the linguistic context. If there is no comparison set available in the linguistic context, the contrastive stress will evoke the idea of such a set, as is shown in (48). The example is from Andréasson (2007), with my glosses.

(48) a. – Vad irriterad du ser ut. Är det något särskilt?
what annoyed you look PRT *is there something particular*
'You look annoyed. What's the problem?'

b. – Ja, MIN hund får inte komma med på julfesten.
*yes my dog can not come with on Christmas party.*DEF
'Yes, MY dog can't come to the Christmas party.'

(Andréasson 2007:87)

Here, the possessive *min* 'my' in *MIN hund* 'MY dog' receives contrastive stress. There is no comparison set in the context, but the prosody indicates that *MIN hund* is contrasting with something; it evokes the idea that someone else's dog is allowed to come to the Christmas party.

If we take contrast to create an unordered set of alternatives, examples involving contrast also behave as expected from the perspective of Szabolcsi & Zwarts's (1993) and Szabolcsi's (2006) approach, since this is precisely the property they point to as important. On this theory, all of the extracted adjuncts can be interpreted (in the context) as ranging over domains that have the right algebraic structure for escaping weak islands.

In section 5.2.2, where I discuss the information structure of ERC, I suggested that contrast can license T-preposing when the fronted phrase is not a good topic and there is no linguistic antecedent or deictic interpretation available. In fact, there are examples that indicate that this is the case. One such example is given in (49).

(49) En möjlig förklaring är att det skulle ha att göra med det tyska
a possible explanation is that it would have to do with the German
ordet *wickeln*, som betyder 'linda in',
*word.*DEF wickeln *which means wrap*

men [med SÄKERhet]$_1$ är det ingen [som vet t_1] ...
*but with certainty is there no one that knows.*COG

'But no one knows with CERtainty.'

(Language Council of Sweden, web page)

This example is from the web page of the Language Council of Sweden, and is part of a speculation about the etymology of the word *vickning* 'late night

snack/dinner at a party', which is of unclear origin. A manner adjunct, *med säkerhet* 'with certainty', has been T-preposed. If the example is read with no contrast on the fronted phrase, the adverb would be interpreted as modifying the whole clause, and the utterance would mean 'But it is certainly the case that no one knows', i.e. it would be a statement about the likelihood of someone knowing the origin of the word. This interpretation, if it is even possible, is clearly not warranted by the context, however, since the context makes clear that what is at issue is the certainty of the knowledge.

If contrast has the function of creating the right kind of environment for T-preposing out of a relative clause, we predict that other bad examples should improve if contrastive stress is added to the fronted phrase. This prediction seems to be borne out, as can be seen in (50).

(50) a. *Plötsligt$_1$/^{ok}PLÖTSligt$_1$ vet jag ingen [som har insjuknat t_1].
*suddenly/ suddenly know.*COG *I no one that has gotten sick*
'I don't know of anyone who has gotten sick SUDDenly.'

b. *Ofta$_1$/^{ok}OFta$_1$ känner jag ingen [som brukar tvätta bilen t_1].
*often/ often know.*REL *I no one that tends to wash car.*DEF
'I don't know anyone who washes their car OFTen.'

Plötsligt 'suddenly' and *ofta* 'often' cannot generally be extracted, but if they receive contrastive stress, they become extractable. In fact, if a context is created where *efter en stund* 'after a while' can be interpreted as contrastive, even this sort of example sounds better (51).

(51) a. Var det någon som smög ut redan när filmen började?
*was there someone who snuck out already when movie.*DEF *started*
'Did someone sneak out already when the movie started playing?

b. ?Nej, men [efter en liten STUND]$_1$ känner jag en tjej [som smög ut t_1].
*no but after a little while know.*REL *I a girl that snuck out*
'No, but I know a girl who snuck out after a little WHILE.'

In the context, *efter en liten stund* 'after a little while' contrasts with *redan när filmen började* 'already when the movie started'.

The adverb *så* is also interesting in this regard. If it is indeed the case that *så* is involved in making examples like (16), (29a), (30a), and (32) in section 6.1.1 acceptable, we predict that it too should be able to transform otherwise unacceptable examples into acceptable ones.

(52) a. [Så plötsligt]$_1$ vet jag ingen [som har insjuknat t_1].
*that suddenly know.*COG *I no one that has gotten sick*
'I don't know of anyone who has gotten sick that suddenly.'

b. [Så ofta]$_1$ känner jag ingen [som brukar tvätta bilen t_1].
*that often know.*REL *I no one that tends to wash car.*DEF
'I don't know anyone who washes their car that often.'

This seems to be the case. While unstressed *plötsligt* 'suddenly' or *often* cannot on their own be T-preposed (50a), (50b), *så plötsligt* and *så ofta* are possible to T-prepose (52a), (52b).

What we have seen in this section is that the notion of unordered sets of various kinds can help in distinguishing between extractable and non-extractable phrases in ERC. This way of looking at extractability is inspired by the approach to weak islands taken by Szabolcsi & Zwarts (1993) and Szabolcsi (2006), but we can also understand the generalization in terms of aboutness topics. As mentioned previously in the chapter, many of the T-preposed adjuncts can be seen as aboutness topics. It might be the case that being an individual in the relevant sense is a property which makes a phrase a good aboutness topic, i.e. something which it is reasonable to say something about.

So far, I have mainly discussed examples with T-preposing in this chapter. In the next section, I expand the investigation by looking at *wh*-questions.

6.2.2 *Wh*-questions

Given our present understanding, it would be useful to know more about question formation into relative clauses. When are such examples acceptable, and are there differences in acceptability in line with the weak island pattern identified in the previous literature?

As previously mentioned, there are no examples of question-formation out of relative clauses in my collection of spontaneously produced ERC-sentences. In order to find out what types of *wh*-phrases can be extracted from relative clauses, if any, I conducted interviews with four graduate students in Scandinavian languages at the University of Gothenburg, who are native speakers of Swedish. In the interview, I read questions aloud, and the interview participants were asked to provide an answer to the question. Some of the questions were asked "out of the blue", and for some of them I provided a context before reading the question. The questions and contexts can be found in Appendix C. After the participants had provided an answer to a question, I asked them to describe how they went about coming up with the answer, and in what situation they

could imagine that such a question could be used. If they did not come up with an answer for a certain question, I suggested answers and asked them to judge how well they worked as answers to the question, if at all.

This methodology was chosen in an effort to provide a naturalistic task, where it would become clear whether the questions were interpretable or not. It also allowed me to find out whether and how certain contexts made the questions easier to interpret and answer.

If Swedish relative clauses are like other weak islands, we expect a certain pattern to emerge. *Wh*-questions which correspond to examples with T-preposing of reason-, manner-, and amount-phrases should be much harder to interpret than examples which already denote in the right type of domain and/or are D-linked. Such examples should be ameliorated if the participant is provided with a context where they can be interpreted as being D-linked. Out of context, inherently D-linked *wh*-phrases should be easiest to interpret. Extraction over negation is expected to be impossible on non-D-linked readings, since negation creates a negative island. As we will see in this section, these expectations turn out to be largely met.

As we saw in section 6.1, *varför* 'why' cannot be extracted from embedded questions in Swedish (see (12)). The case is parallel for extraction from relative clauses. Consider (53).

(53) Varför känner du många [som har skrivit böcker]?
*why know.*REL *you many that have written books*
'Why do you know many people who have written books?'

a. För att jag är med i Författarförbundets styrelse.
*because I am with in Writer union.*DEF*'s board*
'Because I'm on the board of The Swedish Writer's Union.'

b. #För att chockera sin samtid.
*in order to shock their.*REFLX.SG *contemporaries*
'In order to shock their contemporaries.'

If extraction of *varför* from a position inside the relative clause were possible, then the question would be ambiguous. The two possible answers in (53a) and (53b) reflect these different readings, where (53a) corresponds to a short movement of *varför* and (53b) corresponds to an interpretation where *varför* is extracted from a position inside the relative clause. The interview participants all provided answers of the type in (53a), and judged answers of the type in (53b) impossible or very strange. As a comparison, extraction of *varför* from an *att*-clause is acceptable. The example in (54) is from Engdahl (1985b), but

the glosses are mine. Again, there are two proposals for how to answer the question, (54a), which entails short movement of *varför*, and (54b), which entails extraction from the *att*-clause. In this case both types of answers were accepted by all of the interview participants.

(54) Varför tror du att Strindberg skrev *Fröken Julie*?
why think you that Strindberg wrote Fröken Julie
'Why do you think that Strindberg wrote *Miss Julie*?'

a. Därför att ordvalet och stilen liknar Strindbergs övriga verk.
*because that word choice.*DEF *and style.*DEF *resembles Strindberg's other works*
'Because the choice of words and the style resembles Strindberg's other works.'

b. För att chockera sin samtid.
*in order to shock his.*REFLX *contemporaries*
'In order to shock his contemporaries.'

(Engdahl 1985a:4)

In this case, the participants all either proposed answers of the type in (54b), or accepted them when asked about them, suggesting that long extraction of *varför* is fine.

Of all the *wh*-phrases included in the interview, *varför* is the one where resistance to extraction from a relative clause is the strongest. In most other cases, it was possible for the participants to come up with some context where the question could be answered, but with *varför* this was not the case. It is interesting to note in relation to this that *varför* sticks out in another way as well. Brandtler (submitted) reports on a large corpus study of Swedish constituent questions with cleft constructions, i.e. questions like the ones in (55).

(55) a. Vem$_1$ var det [$_{c\text{P}}$ t_1 som t_1 kom]?
who was it that came
'Who was it that came?'

b. När$_1$ var det [$_{c\text{P}}$ t_1 vi skulle äta middag t_1]?
when was it we should eat dinner
'When was it we were going to have dinner?'

In both of these questions, the *wh*-phrase is the pivot of a *det*-cleft and has been fronted to form the question. Brandtler (submitted) investigates which types of *wh*-phrases occur in such clefted questions. It turns out that all types of *wh*-phrases are used in this construction to some extent, except *varför*. Almost all clefted questions with *varför* have another phrase in the pivot position, like in (56), which is from Brandtler (submitted), but with my annotation.

(56) Varför$_2$ är det alltid [$_{cP}$ skitförsvar$_1$ som t_1 sänker svenska lag] t_2,
why is it always shit-defense that sinks Swedish teams
egentligen?
really
'Why is it that Swedish teams always lose because of a crappy defense?'

Here, *skitförsvar* 'crappy defense' is the pivot. The sentence has two presuppositions. In virtue of being a cleft, it logically presupposes that there is something which always causes Swedish teams to lose. The pivot specifies that this something is a crappy defense. The resulting proposition 'Swedish teams always lose because of a crappy defense' is the presupposition of the entire question, which asks for the reason for this state of affairs.

In his study of 6000 clefted questions, Brandtler finds only three examples where *varför* is the pivot. One of his examples (Brandtler submitted) is given in (57), again with my annotation.

(57) haha juste varför$_1$ var det [$_{cP}$ t_1 han slutade nu igen t_1], dom drev med
haha exactly why was it he quit now again they drove with
han i serien väl?
*he in series.*DEF PRT
'Haha, exactly! What was the reason he quit, again? They made fun of him in the series, right?'

Brandtler points out that such questions are very rare, and that the three examples he found are all used in a 'reminding' function, where the speaker signals that he or she could already be expected to know the answer.

Brandtler argues that the *wh*-phrase in clefted questions is an **identificational focus** in the sense of Kiss (1998). On this view, an identificational focus exhaustively identifies a subset of a set of situationally or contextually given elements (p. 245). It is tempting to interpret the fact that *varför* is so infrequent in clefted questions as a sign that it resists denoting elements in such a contextually or situationally given set, which seems very similar to saying that it resists D-linking. This observation also fits into the picture about *wh*-extraction from weak islands, if D-linking is a requirement.[11]

Turning to other adjunct questions, the interview participants were asked to provide answers to the questions in (58a)–(58c).

[11] This is probably not the entire story about *varför*, since *därför* 'therefore', which refers to a contextually salient reason, also seems to resist extraction from relative clauses in T-preposing, see example 26 in the questionnaire, Appendix B.

(58) a. Var$_1$ är det många [som brukar jobba övertid t_1]?
where are there many that tend work overtime
'Which place is such that there are many people who tend to work overtime there?'

b. ?När$_1$ vet du många [som brukar vara trötta t_1]?
*when know.*COG *you many that tend be tired*
'Which time is such that you know of many people who tend to be tired then?'

c. ??Hur$_1$ vet du många [som fick sitt första jobb t_1]?
*how know.*COG *you many that got their.*REFLX.SG *first job*
'Which way is such that you know of many people who got their first job that way?'

I have tried to provide translations which capture the meaning of the questions as well as possible, but they are not very natural-sounding, since questions like these are generally not acceptable in English.

The participants were able to answer (58a)–(58b) with relative ease even without context. Example (58b) was slightly harder to answer than (58a), but this could be an effect of the embedding verb. In fact (58a) allows for an existential interpretation, whereas (58b) does not.[12] It was pointed out by two of the interview participants that it is possible to interpret the (58b) in two ways, i.e. either as a question about when you know something, or as a question about when people tend to be tired. Both interpretations were accepted.

The manner question in (58c) offered more of a challenge. Three of the interview participants said that they could not answer it if this were supposed to be a question about the way people got their first job, which means that the question is quite marginal out of context (but see section 6.3.1).

The pattern is very reminiscent of the one we saw in (10) in section 6.1, with respect to *wh*-question formation out of embedded questions in English. The pattern there is that *how*- and *why*-questions are strongly ungrammatical if they involve extraction from an embedded question, whereas *when*- and *where*-questions are slightly better. The fact that the questions in (9) and (58c) are so hard to interpret means that *wh*-questioning out of relative clauses resembles adjunct extraction in other languages, and more clearly so than T-preposing out of relative clauses.

The interviews moreover provide evidence that D-linking facilitates interpre-

[12] On the existential interpretation, the structure would be *Var*$_1$ *är det* [$_{\text{DP}}$*många som brukar jobba övertid*] [$_{\text{XP}}$ t_1], according to Keenan (1987).

tation. First, without context, the inherently D-linked questions in (59) were the ones which the participants had the easiest time in answering.

(59) a. [Vilka språk]$_1$ är det många [som talar t_1 i Sverige]?
which languages are there many that speak in Sweden
'Which languages are such that there are many people who speak them in Sweden?'

b. [Vilka språk]$_1$ känner du nån [som talar t_1]?
*which languages know.*REL *you someone that speaks*
'Which languages are such that you know someone who speaks them?'

The argument question in (60), which is not inherently D-linked, was perceived as harder to answer, but was nevertheless accepted as an interpretable question by three of the participants, who also provided answers to it.

(60) ?Vad$_1$ såg du någon [som förgiftade t_1]?
what saw you someone that poisoned
'What is such that you saw someone who poisoned it?'

As mentioned above, the participants were also asked how they went about finding an answer to the question, and under what circumstances such a question could be used. The answers here made it clear that the strategy used in order to find an answer to the questions was to come up with a suitable context, and to think about why such a question would be asked. For instance, in answering (58b), the question about when many people are tired, the participants mentioned that it could be used if the conversation topic was tiredness, and if one thinks about times when people might be tired. One participant said that perhaps the question could be used in a discussion about what time to schedule a night class.

The situations that were described often involved a set of a suitable type based on the *wh*-phrase, and the answer involved picking something from this set. Out of context, it is naturally easier to find such a situation for the inherently D-linked questions. If the *wh*-phrase is *which languages*, it is straightforward to say that the question can be used in a conversation where different languages are being discussed. For questions that are not inherently D-linked, and where no context is given, finding a situation with a set to pick from is harder. The strategies for answering such questions in responses to the questionnaire still involved trying to come up with such a situation, however. Often this involved imagining a situation where the *wh*-phrase would be D-linked, like in the example with the night classes above, when different possible times to schedule the class are up for discussion. For (60), one participant imagined a witness

interrogation in a crime investigation, where it is known that the witness saw someone poison something. This seems to correspond to Kroch's (1989) idea about presuppositions. The witness interrogation as described by the interview participant provides us with a situation where the presupposition 'there is some x such that you saw someone poison x' holds.

Another context that was helpful in answering the questions in the interview was a context where the question was interpreted as an echo question. One of the participants came back to this interpretation for almost every question. This is expected given the previous literature, where it is pointed out that long *wh*-movement is much freer in echo questions than in regular *wh*-interrogatives (see Comorovski 1989, Kroch 1989, Szabolcsi & Zwarts 1993).

With respect to amount-quantified questions, the clearest result is that negation has a very strong effect. Whereas the participants could answer the question in (61a) with a number, they found the question in (61b) extremely hard or impossible to answer, reporting that they could not tell what was asked for, and what the question was supposed to mean.

(61) a. ?Hur många kilo känner du många som väger?
how many kilo know.REL *you many that weigh*
'How many kilos do you know many people who weigh?'

b. *Hur många kilo känner du ingen som väger?
how many kilo know.REL *you no one that weighs*
'How many kilos do you know no one who weighs?'

In (61a), it does not seem possible that the kilos that are referred to are specific ones. Rather the participants interpreted the question as an information question looking for a common, or average, weight among their acquaintances. The relative ease with which this question was answered is perhaps surprising given the general pattern, since no context was given. However, the reason might be that the presupposition 'there is some weight such that you know many people who weigh that weight' is plausible given how the question was interpreted.

D-linking seems to facilitate interpretation of amount-quantified questions with negation somewhat. Presented with a scenario where two people have had a dinner party and are discussing how much of the wine is left over the next day, participants could interpret the question in (62) and proposed answers such as *tre flaskor* 'three bottles'.

(62) ??/*[Hur mycket vin]$_1$ var det ingen [som drack t_1]?
how much wine was there no one that drank
'How much of the wine did no one drink?'

Even so, the example seems to be marginal and the interview participants all reported that it is a very strange question, and that it demands a very specific context.

In sum, extraction of interrogative *wh*-phrases from relative clauses exhibits several of the traits which are usually found in extraction from weak islands such as embedded questions: it shows the pattern with respect to different types of adjuncts and amount questions observed in earlier studies of extraction from embedded questions, and D-linking and contrastive interpretations facilitate extraction. An echo question reading will also ameliorate the example, and negation seems to negatively affect amount-quantified questions, as expected.

6.2.3 Accounting for the pattern

The summary in the previous section reveals that there are some clear parallels between ERC and extraction from domains which have been analyzed as weak islands, specifically when it comes to what types of phrases can be extracted. The analyses I discussed, Comorovski (1989), Kroch (1989), Szabolcsi & Zwarts (1993), and Szabolcsi (2006), were developed for extraction from domains like embedded questions and negative islands, and they mostly concern extraction of *wh*-phrases. The question is now why relative clauses and other weak islands should exhibit similar patterns, or slightly differently put, why analyses of weak islands should be relevant for the analysis of T-preposing and *wh*-questions in ERC.

Thinking in terms of the proposal of Kroch, we might wonder whether the strangeness of certain of the *wh*-questions we have seen can be explained in terms of odd presuppositions, i.e. whether they can be said to have presuppositions that make them unusable in most situations. Considering (63a) and (64a), which have the presuppositions in (63b) and (64b), this seems quite a reasonable assumption.

(63) a. ??Hur$_1$ vet du många [som fick sitt första jobb t_1]?
*how know.*COG *you many that got their.*REFLX.SG *first job*
'Which way is such that you know of many people who got their first job that way?'

b. There is a way such that you know many people who got their first job that way.

(64) a. ?Hur många kilo känner du många som väger?
how many kilo know.REL *you many that weigh*
'How many kilos do you know many people who weigh?'

b. There is an amount of kilos such that you know many people who weigh that amount of kilos.

I think it is fair to say that the presuppositions in (63b) and (64b) are odd, and that these questions are useful only in quite specific circumstances. The interviews even showed that in order to answer the question in (64a), the participants had to interpret the question such that it rather meant something like *for what weight, you know many people who have that weight*, which would have the presupposition 'there is some weight such that you know many people who weigh that weight.' This type of question would be possible to answer with e.g. '65 kilos'.

It is not obvious how this approach could be extended to say anything about the T-preposing cases, however, since the odd presuppositions which would make the questions unusable arise from *wh*-questions. We are then left wondering why such an "odd" presupposition should arise in T-preposing from relative clauses, in order to explain the restrictions there, and why such presuppositions do not arise in T-preposing from *att*-clauses.

On Szabolcsi & Zwarts's (1993) and Szabolcsi's (2006) approach, we can understand the T-preposing cases as well as the *wh*-question cases. As we have seen, on this approach, weak island effects occur because in calculating the denotation of a sentence, a Boolean operation needs to be applied in a domain where it cannot be performed due to the semantic type of the gap. Weak islands are thus those domains for which some Boolean operation is needed to calculate the denotation.

The reason that relative clauses behave as weak islands in this regard follows from the fact that they are intersective modifiers. This means that in order to calculate the denotation of the relative complex, we need to intersect the denotation of the relative clause and the denotation of the head. Since intersection is a Boolean operation, it is expected that we should see weak island effects.

Furthermore, the relative complexes in ERC-sentences usually contain quantifiers, like *många* 'many' and *någon* 'someone', which are also scopal elements. Negation, which is also present in many ERC-sentences, induces a weak island effect on its own.

6.3 Some remaining issues

As we saw in the previous section, the restrictions on ERC that we have seen in this chapter make sense from the perspective of semantics and pragmatics, against the background of the proposals by Comorovski (1989), Kroch (1989), Szabolcsi & Zwarts (1993), and Szabolcsi (2006). In this section, I discuss two issues that I have not been able to address so far but that deserve further comment. The first issue has to do with a difference between Swedish relative clauses and embedded questions in English and other languages with respect to possible interpretations of the fronted phrase in *wh*-questions. I discuss this in section 6.3.1, and then shift the focus to the role of the embedding predicate in extraction of adjuncts from relative clauses in section 6.3.2.

6.3.1 Functional readings

Some analyses in the research on long *wh*-extraction from embedded questions connect extractability specifically to the ability of a phrase to bind a variable of type *e*, i.e. an individual variable (Frampton 1990, Cresti 1995). Cresti (1995) proposes an analysis of extraction from *wh*-islands that takes this approach. On her account, phrases can move out of weak islands via an escape hatch, adjoining to Spec-CP. There is a semantic filter on this escape hatch which disallows all traces that cannot be interpreted as variables of type *e*. Higher order variables are not allowed. Formally, the restriction takes the form of the LF-filter in (65).

(65) *$[_{CP}\ [_{\delta}\ X]\ [_{CP}\ \ldots\ ]]$ where X is not of type *e*

On Cresti's account all phrases that can escape weak islands move in a successive cyclic fashion through the escape hatch, and this is why we see the types of effects described in the previous sections of this chapter.

Cresti (1995) connects her account of weak islands to **functional readings** of questions (cf. Engdahl 1986). Consider (66).

(66) I know which book you think no student read. (Cresti 1995:89)

The sentence is ambiguous: it has one reading where there is a unique book that no student read, and one where the set of books is determined based on the set of students. The second reading, which is the functional reading, would say something similar to "I know that you think no student$_x$ read her$_x$ mother's book".

Cresti found that functional readings are sensitive to weak islands. Consider (67).

(67) I know which book you wonder whether no/any student read.

(Cresti 1995:89)

Cresti's consultants did not take (67) to be ambiguous, unlike (66). The reading that is available is something like "I know that you wonder whether no/any student read *War and Peace*". A reading like "I know that you wonder whether no/any student$_x$ read her$_x$ mother's book" is not present.[13] The unavailability of functional readings is predicted by accounts like Cresti's, on the assumption that such readings require a variable of a higher type than the permitted type *e*.[14] Some later accounts of weak islands have taken the absence of functional readings as an explanandum (Szabolcsi 2006 and references therein).

If Swedish relative clauses are some type of weak island, we might expect them to block functional readings, in addition to displaying the other weak island effects discussed in previous sections. The interviews suggest that Swedish relative clauses do not block functional readings, however.

The first piece of evidence for this is that one of the participants in fact volunteered an answer that involves a functional reading in the first part of the interview, before any such examples had been presented. Specifically, the interview participant answered the question in (68a) with the answer in (68b).

(68) a. Hur vet du många [som fick sitt första jobb]?
*how know.*COG *you many that got their.*REFLX.SG *first job*
'Which way is such that you know of many people who got their first job that way?'

b. Via sin förälder.
*through their.*REFLX.SG *parent*
'Through their parent.'

The question in (68) was among the ones that the three other interview participants had the hardest time answering. Nevertheless, this participant had no trouble with it, and the answer reveals that the participant has a functional interpretation of the question in mind, since the denotation of *sin* 'their.REFLX.SG' will depend on the set of people. In a later section of the interview, all participants were asked whether such answers were possible, in questions with

[13] Cresti reports that there is one semanticist for whom this reading is available.

[14] In her account of functional readings in Swedish, Engdahl (1986) uses a Skolem function, of type <e,e>, from individuals to individuals.

different types of *wh*-phrases, and with several different embedding predicates. Consider (69) and (70).

(69) [Vilka åsikter]$_1$ finns det många studenter [som tar avstånd från t_1]?
which opinions exist there many students that take distance from
'Which opinions are there many students who condemn?'

a. Högerradikala åsikter.
right radical opinions
'Extreme right-wing views.'

b. Dom som deras föräldrar står för.
those which their parents stand for
'The ones that their parents hold.'

(70) [Vilken bok]$_1$ var det ingen student [som ville läsa t_1]?
which book was there no student that wanted read
'Which book was there no one who wanted to read?'

a. Krig och fred.
war and peace
'War and Peace.'

b. Sin mammas bok.
their.REFLX.SG *mother's book*
'Their mother's book.'

In each of these examples, the answer in (b) involves a functional reading, where the set of opinions and the set of books are dependent on the set of students. The interview participants were asked about the possibility of both (a)-type and (b)-type examples, and all of them accepted both types of answers. As (70) shows, negation does not seem to affect the availability of the reading.

Functional readings are also possible with embedding predicates like *känna* 'know.REL'; the answer in (71b) was accepted by three of the four interview participants, even when they reported that it was easier to come up with a specific movie.

(71) [Vilken film]$_1$ känner du inte nån [som kan glömma t_1]?
which film know.REL *you not someone that can forget*
'Which film don't you know anyone who can forget?'

a. Titanic.

b. Den första skräckfilmen hen såg.
the first horror movie.DEF *they*.SG *saw*
'The first horror movie they saw.'

In light of the interviews, we can conclude that relative clauses in Swedish do not block functional readings of questions, at least not for all speakers.[15] This means that an LF-filter like the one Cresti proposes for extraction from embedded questions in English cannot be right for extraction from Swedish relative clauses, since traces of higher order are clearly possible.

The observation that relative clauses and embedded questions do not block functional readings in long extraction might seem problematic, if we are to analyze relative clauses as weak islands. An interesting question for further research is how to achieve an analysis which captures the things Swedish relative clauses have in common with weak islands, i.e. the D-linking requirement of *wh*-questions, and the restriction on the semantic type of extracted adjuncts, without ruling out functional readings of questions. An interesting starting point for such an investigation is a recent proposal by Ruys (2015).

Ruys (2015) proposes a unified account for various constraints on reconstruction, among them reconstruction of *wh*-phrases into *wh*-questions, and functional readings of questions. The proposal is of interest from the perspective of the approach to ERC that I have taken in this thesis because it connects the availability of reconstruction to the availability of specifiers in the C-domain. Specifically, Ruys proposes that reconstruction requires determining whether the moved XP and its trace are identical. This is determined locally; it can only be accomplished if the XP and the trace are in the same phase. If they are, the trace can be of any semantic type that the moving phrase can have. If they are not, the trace will be assigned a default semantic type, which is type *e* in Ruys' analysis.[16]

---

[15] Functional readings are also available in extraction from embedded questions in Swedish (see Engdahl 1986).

(i) Vilken bok undrade du om nån student hade läst?
which book wonder you if some student had read?
'Which book did you wonder whether any student had read?'

a. Krig och fred.
war and peace
'War and Peace.'

b. Den som hennes lärare rekommenderade.
that which her teacher recommended
'The one that her teacher recommended.'

[16] As Ruys points out, something more needs to be said about the Phase Impenetrability

Like much previous research Ruys assumes that the C-domain in embedded questions in English can only host one specifier. For that reason, a phrase that is extracted cannot stop there, and its trace lower down will default to type *e*. The availability of an additional local landing site in the C-domain, like in Vikner's (2017) and Nyvad et al.'s (forthcoming) proposal, would explain why functional readings are available in Swedish ERC. Since recursion of *c*P provides an extra landing site in the C-domain, the trace will not have to default to type *e*. If this is the case, the availability of functional readings provides additional evidence that movement out of relative clauses in Swedish proceeds in a successive-cyclic fashion. On such an approach the important locus of variation between English and Swedish with respect to weak island behavior is that the C-domain of weak islands in English can host only one specifier, whereas Swedish permits recursion of little *c*P in relative clauses and embedded questions, providing the extra landing sites needed both for grammatical extraction and for functional readings of questions. I return to the issue of additional specifiers in 7.2.1.3.

6.3.2 The embedding predicate

An observation that I have not yet discussed in much detail is that the embedding predicate seems to matter for how natural adjunct extraction is, and for how easy a certain question is to interpret in ERC. This is in line with findings from previous research on extraction of non-arguments from embedded questions (Engdahl 1985b), and even if I cannot say anything definitive, my investigations allow for some further remarks about the situation with respect to ERC.

As we saw in section 6.1.1.2, there was some variation in how natural a certain extraction was perceived to be by the questionnaire participants. It was not clear from the questionnaire what the role of the embedding predicate was in relation to the role of the adjunct type with regard to this variation, however. The embedding predicates that were used in the questionnaire were *känna* 'know.REL', *veta* 'know.COG', *gå att hitta* 'be possible to find', *träffa* 'meet', and *beundra* 'admire'. The clearest result was that the example with *beundra* as the embedding predicate got much worse ratings than the other examples.

In the interviews, there was also some variation. Some participants were more liberal than others, and accepted extraction with all of the embedding

Condition (PIC) and the types and number of specifiers that can occur in the C-domain. Moving a *wh*-phrase out of the embedded question without an intermediate landing site would violate the PIC (Ruys 2015:463–464, footnote 10).

predicates, whereas some participants only liked extractions with some of them. The embedding predicates used in the interview items were *vara* 'be', *finnas* 'exist', *veta* 'know.COG', *känna* 'know.REL', and *träffa* 'meet'. The variation was not random. Participants generally had an easier time answering the question, and finding a good context, when the verb was *vara*, *finnas*, or *veta* than when the verb was *känna* or *träffa*. If a participant accepted a certain type of extraction with *känna* and *träffa*, they would also accept it with *vara*, *finnas*, and *veta*. Two of the interview participants almost never fully accepted extraction with embedding predicates other than *vara* and *finnas*, remarking that the questions were strange sounding, and that they didn't sound like something they would say.

Taken together with the observations from the questionnaire, we might tentatively propose an implicational hierarchy for adjunct ERC like that in (72), suggesting that if a participant accepts extraction of an adjunct over an embedding predicate higher in the hierarchy, they will also accept extraction over predicates lower in the hierarchy.

(72) vara, finnas < veta < känna, gå att hitta, träffa < beundra
*be exist know.*COG *know.*REL *be possible to find meet admire*

My investigation does not explain why there should be variation in acceptability, or why the implicational hierarchy should look like this. Perhaps there is some substantial difference in the grammars of speakers, such that some speakers actually only accept adjunct extraction in existential contexts, or when a small clause analysis is available, as proposed by Kush et al. (2013). My interviews lend some support to this idea. It might also be the case, however, that the variation is simply an effect of frequency. Christensen & Nyvad (2014) show in an acceptability study that there is such an effect on extraction of arguments from relative clauses in Danish.

Both of these hypotheses would need to be further explored before any conclusion could be made, and in particular, the interview study should be seen as a pilot investigation, considering that only four people were interviewed. A factor that is hard to control for is the ambiguity that some of the interview participants noticed with verbs like *veta* and *känna*. The task to answer the question is quite naturally harder if there are two possible ways to parse the sentence.

My investigations suggest that verb type in ERC-sentences is not lexically restricted. This conclusion also holds for extraction of arguments from relative clauses, as shown by the variety of embedding predicates in naturally occurring ERC-sentences (see chapter 4).

6.4 Concluding remarks

The starting point of this chapter was an observation from chapter 5 about adverbial expressions like *efter en stund* 'after a while' which indicated that there may be stronger restrictions on ERC than on preposing from *att*-clauses and on local preposing. I then considered how asymmetries in (roughly) argument and adjunct extraction have been treated in previous research on extraction from embedded questions in languages like English and Italian. Several accounts encode the distinction between extractable and unextractable phrases in the syntax, in terms of theta-marking (e.g. Cinque 1990, Rizzi 1990). The investigations in this chapter show that this approach cannot easily be adapted to Swedish ERC, since we find extraction of certain adjuncts as well as arguments, and I suggested that a semantic or pragmatic approach is more likely to work.

In particular, my investigation showed that in T-preposing out of relative clauses, there is extraction of time and place adjuncts, as well as extraction of manner and reason adjuncts, and amount phrases. This is a different pattern than what might have been expected from the previous literature on *wh*-questioning out of embedded questions. Even though adjuncts can be extracted from relative clauses, they are subject to certain restrictions. In T-preposing, the extracted phrases are typically deictic or contrastive, or refer to a specific point on a scale. A phrase with such a denotation can pied-pipe a preposition or larger phrase. Expressions like *efter en stund* 'after a while', *till slut* 'at last', and *snart* 'soon', which can only be T-preposed in local fronting and in extraction from *att*-clauses, have more complicated denotations.

I have found no spontaneously produced *wh*-question ERC-sentences, but my interview study indicates that such questions are interpretable for most speakers under certain circumstances. Here we see evidence for the weak island pattern which is reported in previous research (e.g. Szabolcsi 2006). Without context, a question involving time (*när* 'when') or place (*var* 'where') is easier to interpret than a question involving manner (*hur* 'how'). The reason-adjunct *varför* 'why', cannot be extracted, nor can it be the pivot in a clefted question. D-linking facilitates interpretation, and at least improves the interpretability of amount-quantified questions somewhat.

Although Swedish relative clauses exhibit these weak island-like effects, there are also differences between ERC and extraction from domains such as embedded questions in languages like English. In particular, Swedish relative clauses do not block functional readings of questions. The C-domain in relative clauses in Swedish is thus more transparent than the C-domain in embedded questions in languages like English.

In the next chapter, I summarize the findings of the dissertation and discuss how the syntactic analysis I adopted in chapter 4 interacts with semantics/pragmatics in the account of ERC. I also discuss what my findings mean for the analysis of ERC and island constraints more broadly.

7. Main findings and discussion

The aim of the work presented here has been to investigate extraction from relative clauses (ERC) in Swedish, and in doing so to contribute to our knowledge about extraction phenomena in natural languages, and how they are represented in our mental grammars. The undertaking has been guided by four questions. I have investigated what characterizes the dependency between the preposed phrase and the empty position in ERC-sentences; what the structure of the relative clauses in ERC-sentences is; which role discourse, information structure, and pragmatic constraints play in ERC; and how ERC is constrained, in particular whether there are constraints on the types of phrases that can be extracted from a relative clause.

I have used a combination of data sources in addressing these questions. Spontaneously produced examples make up the foundation of the investigation, and acceptability judgements of constructed ERC-sentences have been used to complete the picture and test hypotheses. A secondary goal has been to give access to the Swedish extraction data to linguists who do not speak a Scandinavian language.

In this chapter, I summarize and discuss the findings of the investigation: in section 7.1, I review the main findings; in section 7.2, I discuss their consequences for the analysis of ERC; and in section 7.3, I point out some directions for future research.

7.1 Main findings

The investigations reported on here have led to several findings. An important first conclusion, which serves as a background to the rest of the inquiry and

motivates it, is that many ERC-sentences really do involve extraction from relative clauses. There are two parts to the argument behind this conclusion. The first part of the argument concerns the dependency between the preposed phrase and the empty position inside the relative clause in ERC. This dependency is an Ā-movement dependency, which is shown in section 4.2.1. The basis for this claim is that the dependency shows connectivity effects, licenses parasitic gaps, exhibits crossover effects, and obeys certain island constraints. For example it obeys the Coordinate Structure Constraint. All of these are properties that are characteristic of Ā-movement in many languages (for overviews about Ā-movement in English see Chomsky 1977, McCloskey 1988, Haegeman 1994, Pesetsky 2013), and hold of Swedish Ā-movement dependencies as well, as was shown in chapter 2. The type of data that was used to show that the extraction dependencies in ERC-sentences share these properties is exemplified by (1a)–(1c).

(1) a. [Såna där gröna bönor]$_1$ känner jag ingen [som kan äta t_1 utan att
*such there green beans know.*REL *I no one that can eat without to*
koka __$_{\text{pg 1}}$ först].
cook first

'I know no one who can eat those green beans without cooking them first.'

b. *[Vilken flicka]$_{x1}$ känner hon$_x$ ingen [som kan prata svenska med t_1].
*which girl knows.*REL *she no one that can speak Swedish with*

c. *Katt$_1$ känner jag många [som har både hund och t_1].
*cat know.*REL *I many that have both dog and*

In (1a), the phrase *såna där gröna bönor* 'those green beans' has been extracted from the relative clause, and the extraction licenses a parasitic gap. Example (1b) is unacceptable when the *wh*-phrase which has been extracted from the relative clause is co-referential with the pronoun *hon* 'she', which means that the extraction induces a strong crossover effect. Example (1c) involves extraction of the phrase *katt* 'cat' from a coordinate structure inside a relative clause. The example is unacceptable, i.e. the extraction induces an island effect.

Furthermore, an account in the spirit of Cinque (1990) involving base generation of the preposed phrase in the left periphery and binding of a silent pronominal in the gap site is not feasible, as was argued in section 4.2.2. Here, the argument is based on facts from inversion of the subject and the finite verb, and observations about the possibility of extracting phrases which cannot be the antecedents of any pro-forms otherwise available in Swedish. Looking again at (1a), we see that the finite verb *känner* 'know.REL' precedes the subject *jag* 'I' in Spec-TP. This is characteristic of structures which involve Ā-movement to

Spec-CP in declarative main clauses. If such examples involved binding of a silent pronominal in the gap site, we would have expected them to behave like left-dislocation structures, and an example like (2) would be predicted to be a possible Swedish sentence.

(2) *[Såna där gröna bönor]$_x$, jag känner ingen [som kan äta *pro*$_x$]
*such there green beans I know.*REL *no one that can eat*

However, (2) is unacceptable. Taken together with data like that in (1a)–(1c), this indicates that the ERC-sentences under investigation should be analyzed as involving an Ā-movement dependency, i.e. that they really are extractions.

The second part of the argument that ERC-sentences truly involve extraction from relative clauses consists of showing that the RC-like constituent in such sentences can be a full-blown relative clause, and that it is not always a small clause, as has been proposed by Kush (2011) and Kush et al. (2013). This part of the argument was laid out in section 4.3. The central finding here is that the RC-like constituent in many ERC-sentences should be analyzed as a relative clause. An example of the type of data that is relevant to showing this is given in (3).[1]

(3) Det$_1$ har jag inte träffat någon [som gjort t_1]
that have I not met someone that done
'I haven't met anyone who has done that.'

In (3), there is extraction from a relative clause embedded under *träffa* 'meet', a verb which does not select a small clause.

The investigation furthermore showed that the relative clauses in ERC-sentences are like regular restrictive relative clauses; there does not seem to be anything that distinguishes them structurally from other restrictive relative clauses in Swedish. Extraction is possible from relative clauses which are attached inside the DP and restrict the denotation of the head of the relative clause. In addition, there are acceptable examples of extraction from non-subject relative clauses, as illustrated by (4).

(4) [just den här delen av matten]$_1$ kommer ni inte hitta nåt jobb$_2$ [där$_2$
precisely the here part of the math will you not find some job where
man behöver t_1 t_2]
one needs
'You won't find any job where you need THIS part of math.'

[1] In this chapter, I repeat many examples from earlier chapters. See the original discussion for information about the source of the examples.

Here we have extraction from a relative clause in which the relative operator is an adverb.

In addition to the ERC-sentences with regular relative clauses just mentioned, I have found that there are certain ERC-sentences with existential predicates that can be analyzed as involving extraction from an XP which is predicated of a DP. The XP in such a structure is similar to a small clause in terms of interpretation (cf. Keenan 1987). Evidence for this comes from examples like (5), with an exceptive modifier in the DP.

(5) Det$_1$ finns det ingen utom läraren [som vet t_1].
*it exist there no one except teacher.*DEF *that knows*
'There is no one except the teacher who knows it.'

This was discussed in section 4.3.1. It is not always clear which analysis is the right one for a particular example. Sentences can be ambiguous with respect to the interpretation, such that there are strings which could be interpreted either as involving extraction from a restrictive relative clause, or from an XP predicated of the DP.

Examining a few different possible structures of relative clauses which have been proposed in the previous literature (e.g. Platzack 2000, Hulsey & Sauerland 2006), I have found that there is no clear evidence for a head-raising structure for relative clauses in Swedish, although the evidence against a head-raising structure is also inconclusive. These arguments were reviewed in section 4.5. Evidence from extraction from extraposed relative clauses, as seen in (6), suggests that head-raising is not a precondition for ERC in Swedish in any case, at least on the assumption that head-raising and extraposition are mutually exclusive (Hulsey & Sauerland 2006; for an opposing view, see Sichel to appear).

(6) Det$_1$ träffade jag många igår [som ville t_1].
that met I many yesterday that wanted
'I met many people yesterday who wanted that.'

This means that ERC in Swedish is potentially different from ERC in languages like Hebrew, where ERC is argued to be possible only from relative clauses derived by head-raising (cf. Sichel to appear).

Examining ERC in spontaneous usage, we saw that the syntactic operation creating the Ā-movement dependency out of the relative clause involves either movement to Spec-CP in declarative main clauses, which I have called T-preposing, or relativization. T-preposing is most common; in a collection of 100 ERC-sentences, 93 involved T-preposing and 7 relativization. All of the

examples in this chapter so far are of the T-preposing-kind. Examples with relativization are given in (7a) and (7b).

(7) a. vi har till exempel flera hyllmeter namnforskning$_1$ (.) [$_{RC_1}$ Op$_1$
we have to example several running meters onomastics
som vi inte har en enda$_2$ på vår institution [$_{RC_2}$ t_1 Op$_2$ som t_2 håller
which we not have one single on our department that holds
på med t_1]].
on with
'For example, we have several running meters of books on onomastics, which no one in the department works on.'

b. Vi makulerade alla tidskrifter om ämnen [$_{RC_1}$ Op$_1$ som vi inte har
we cancelled all journals about subjects that we not have
någon$_2$ på institutionen [$_{RC_2}$ t_1 Op$_2$ som t_2 håller på med t_1]].
*someone on department.*DEF *that holds on with*
'We cancelled all journals about subjects that no one in the department works on.'

In each of these examples, there are two relativization dependencies indicated by Op$_1$ ... t_1 and Op$_2$ ... t_2, forming two relative clauses, RC$_1$ and RC$_2$. Example (7a) is from my collection of naturally occurring examples, and in it, RC$_1$ is non-restrictive. The slightly adjusted example (7b) shows that RC$_1$ can also be restrictive.

The most deeply embedded relative clauses, marked RC$_2$ in these examples, are restrictive in both (7a) and (7b). The relative clauses in T-preposing ERC-sentences are also always restrictive in my collection of spontaneously produced examples, which is consistent with what has been reported in previous research (Engdahl 1997, Platzack 1999). Extraction from non-restrictive relative clauses is not acceptable.

We saw in section 5.2 that the fronted phrase in T-preposing can be related to the discourse context in a variety of ways. In general, we can say that the preposing in ERC functions a lot like preposing in local clauses, with respect to discourse. A specific observation is that T-preposed phrases in ERC do not need to invoke a contrastive interpretation, rather they often have a cohesive function in discourse. The same observation has previously been made by Engdahl & Lindahl (2014) for local fronting. Non-contrastive T-preposing is quite common in my collection of ERC-sentences, and the T-preposed phrase is often a pronoun, as in (8).

(8) A: Shouldn't the bread be on one of these? (Shows a small plate)
B: det tycker inte JAG (.) men det$_1$ kanske det finns dom [som tycker t_1]
it think not I but it maybe there exist those that think
'I don't think so, but there may be people who do.'

The fronted pronoun in the second sentence in B's utterance is non-contrastive, and functions as a continuous topic in a topic chain.

In its utterance, a T-preposed phrase often functions as an aboutness topic, but it can also be the information focus. This was shown in section 5.3, and is in line with findings in previous research (Engdahl 1997). In section 5.3.3, we saw that certain adverbials which often function as scene-setters in local T-preposing, like *efter en stund* 'after a while', cannot be extracted from relative clauses. This suggests that ERC is more restricted than local preposing and extraction from complement *att*-clauses. The relevant contrasts are shown in (9).

(9) a. [Efter en stund]$_1$ skrek nån ”jävlar”, ”jävlar” t_1.
after a while yelled someone damn damn
'After a while, someone yelled "damn", "damn".'

b. [Efter en stund]$_1$ tror jag [$_{CP}$ att nån skrek ”jävlar”, ”jävlar” t_1].
after a while think I that someone yelled damn damn
'I think that someone yelled "damn", "damn" after a while.'

c. *[Efter en stund]$_1$ känner jag en [$_{RC}$ som skrek ”jävlar”, ”jävlar” t_1].
*after a while know.*REL *I one that yelled damn damn*

Example (9a) shows *efter en stund* 'after a while' in Spec-CP in a clause with local T-preposing, and (9b) shows extraction of *efter en stund* from an *att*-clause. Notably it is most natural to interpret the adverbial as modifying the event in the embedded clause in (9b). Yet example (9c) shows that such a "low" reading of the adverbial is not available in a relative clause.

With respect to the information impact of the relative complex, which has been proposed to determine extraction possibilities (Erteschik-Shir 1973, 1982, Goldberg 2006), I have examined two ideas: that the relative complex has to be dominant for extraction to be possible (Erteschik-Shir 1973, 1982), and that it has to be non-backgrounded (Goldberg 2006). While both of these proposals capture substantial and mostly overlapping parts of the ERC-data, neither proposal can account for the full range of extraction sentences. A type of example which is problematic for both of the proposals is given in (10).

(10) Det$_1$ beundrar jag folk [som klarar t_1 rent psykiskt], att bara vänta.
that admire I people that manage purely psychologically to just wait
'I admire people who can deal with that psychologically, to just wait.'

I showed in section 5.5 that the relative complex in examples like (10) is not dominant according to the lie test which Erteschik-Shir (1973) employs to

identify legitimate extraction domains, and that it is identified as backgrounded by Goldberg's (2006) test for backgroundedness.

Prompted by the observation about the possible interpretations of adverbials like *efter en stund* 'after a while' in examples like the ones in (9), I investigated whether adjuncts can be extracted at all from relative clauses. The investigation was reported in chapter 6, where I also explored the idea that Swedish relative clauses may be weak islands. The investigation showed that adjunct extraction is possible, but only if the adjunct is of the right type. Extractable adjuncts tend to be deictic, refer to a specific point in time or space, or in the case of measures and amounts, to some specific degree on a scale. Another way to make an adjunct extractable is to contrast it, thereby creating a comparison set. Example (11) shows an example with a fronted deictic adjunct.

(11) så [imorgon]$_1$ vet jag tre trötta barn [som kommer vara ganska sega t_1].
*so tomorrow know.*COG *I three tired kids that will be pretty tired*
'so I know three tired kids who will be pretty tired tomorrow.'

Imorgon 'tomorrow' in (11) modifies the event in the relative clause, not the matrix event.

Since there are no questions in my collection of spontaneously produced ERC-sentences, I conducted a small interview study. The study, which is reported in section 6.2.2, showed that it is possible to form interpretable constituent questions with ERC. The interviews revealed a pattern familiar from the literature on weak islands such as embedded questions in English. Questions involving time (*när* 'when') or place (*var* 'where') are easier to interpret than questions involving manner (*hur* 'how'), especially out of context. Measure phrases are also harder to extract, even when they are complements and not adjuncts. The reason-adjunct *varför* 'why', cannot be extracted at all, as illustrated by (12).

(12) Varför känner du många som har skrivit böcker?
*why know.*REL *you many that have written books*
'Why do you know many people who have written books?'

a. För att jag är med i Författarförbundets styrelse.
*because I am with in Writers union.*DEF *' board*
'Because I'm on the board of The Swedish Writer's Union.'

b. #För att chockera sin samtid.
*in order to shock their.*REFLX.SG *contemporaries*
'In order to shock their contemporaries.'

Providing a context where there is a salient set that the *wh*-phrase ranges over, i.e. D-linking the phrase in the sense of Pesetsky (1987), makes questions more acceptable and easier to interpret. For example, the question in (13) is easiest to interpret if you imagine a conversation where a set of times is salient.

(13) ?När$_1$ vet du många [som brukar vara trötta t_1]?
*when know.*COG *you many that tend be tired*
'Which time is such that you know of many people who tend to be tired then?'

The facts from the investigation about adjunct extraction and constituent questions, taken together with the more general observation that relative clauses in Swedish permit extraction of some phrases but not all, indicate that they constitute weak islands in Swedish (Cinque 1990, Szabolcsi 2006). However, in previous research on weak islands in languages like English, such domains are shown to block functional interpretations of questions (Engdahl 1986, Cresti 1995). This has been taken as a fact about weak islands which should follow from the same account as the other weak island effects, e.g. the D-linking effects mentioned above. Swedish relative clauses do not block functional interpretations of questions, as shown in section 6.3.1. An illustrative example is given in (14), where the spontaneously produced answer reveals that the interview participant interpreted that *how*-question as functionally dependent on the subject of the relative clause.

(14) a. Hur$_1$ vet du många [som fick sitt första jobb t_1]?
*how know.*COG *you many that got their.*REFLX.SG *first job*
'Which way is such that you know of many people who got their first job that way?'

b. Via sin förälder
*through their.*REFLX.SG *parent*
'Through their parent'

This suggests that functional interpretations and extraction from weak islands are not necessarily restricted in the same way. As Engdahl (1986) has previously shown, questions retain functional readings in long extraction from Swedish embedded questions as well.

In the following sections, I discuss what these results imply for the analysis of ERC, and for a theory of islands more broadly.

7.2 Consequences for the analysis of ERC

We now turn to the question of what kind of analysis the data we have seen in the dissertation point towards. In this section, I discuss the consequences that my findings have for the analysis of ERC and how ERC fits into the broader theory of islands. In particular, which facts should receive a syntactic explanation, and which facts should be accounted for by a pragmatic and/or semantic explanation? I first discuss issues specific to the syntactic analysis of ERC, and what such an analysis should capture in section 7.2.1, and then consider the role of pragmatics, information structure, and semantics in section 7.2.2. In section 7.2.3, I turn to the implications my findings have for a theory of islands.

7.2.1 Syntax

The observation which has caused researchers to be interested in ERC in the first place is that is does not seem to respect constraints like Subjacency, the Phase Impenetrability Condition, or the Condition on Extraction Domain, constraints which are otherwise thought to be universal. Why would it be that Ā-movement in Swedish does not respect such constraints?

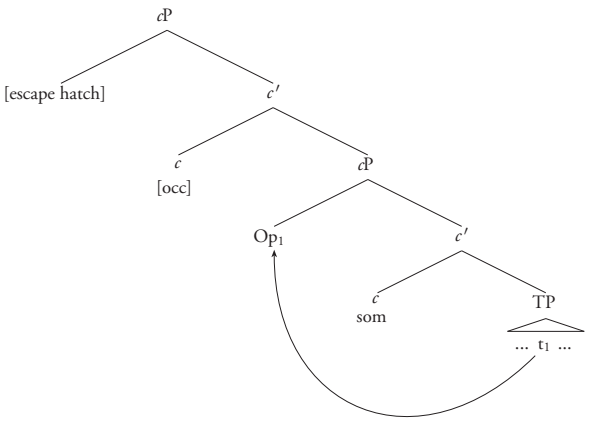
As we saw in chapter 2, previous answers to this question have sometimes involved the proposed structure of the relative clause in ERC-sentences. Platzack (1999) proposes that Swedish has a way to derive subject relative clauses without creating an Ā-dependency. Kush (2011) and Kush et al. (2013) suggest that the *som*-constituent in ERC-sentences is not a real relative clause. On each of these accounts, the constituent is a complement, either to N, as in Platzack's proposal, or to the embedding (small clause-selecting) verb, as in the proposal by Kush (2011) and Kush et al. (2013). Another possibility is that the relative clause is indeed a regular relative clause, but that apparent extraction from RCs is not true extraction—that is to say, the dependency between the fronted phrase and the empty position inside the relative clause is not a movement dependency, involving instead binding, as proposed by Cinque (1990).

In this dissertation, I have argued that neither of these proposals is a feasible way forward. Given the findings in chapter 4, another type of analysis which has recently been proposed by Vikner (2017) and Nyvad et al. (forthcoming) is more promising. This approach, which I described in more detail in section 4.1.4, makes a distinction between *c*P and CP. CP occurs in V2 clauses, where the finite verb moves to C. The account makes a connection between verb movement to C and islandhood; if there is a CP in the clause, extraction is not

possible. It is worth pointing out that I have not come across any extractions from sentences with V2 order in my investigation, neither relative clauses, nor *that*-clauses.

Vikner (2017) and Nyvad et al. (forthcoming) furthermore connect extraction to there being extra specifiers available in the complementizer domain in a number of embedded clauses in the mainland Scandinavian languages, among them relative clauses and embedded questions. The extra specifiers are made available by a certain type of recursion in the complementizer domain. In the case of relative clauses, the structure that is responsible for the availability of grammatical extraction from relative clauses is shown in (15).

(15)



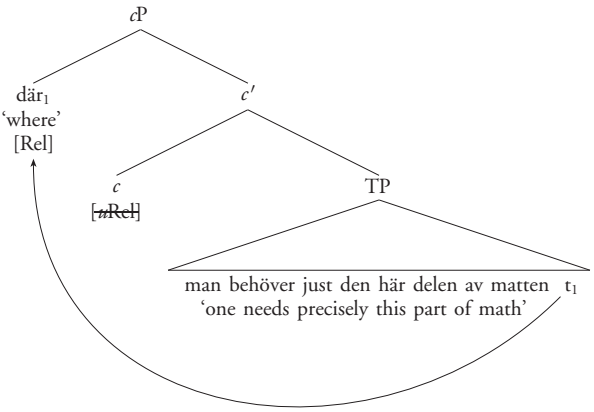
The tree in (15) shows a fairly standard analysis of relative clauses, with respect to the lower *c*P, where a relative operator has moved into the complementizer domain, to Spec-*c*P. Extraction proceeds through the higher Spec-*c*P, an edge position, made available by the **occurrence feature** on the higher *c*-head in the structure. Vikner (2017) and Nyvad et al. (forthcoming) describe the occurrence feature as an edge feature which triggers internal merge to a specifier of the head which carries it, without requiring feature matching (see also Chomsky 2005).

This structure can accommodate the types of relative clauses we have seen in the ERC-sentences in the dissertation. To take a specific example, consider the derivation of example (4) in this chapter, which is repeated here as (16).

(16) [just den HÄR delen av matten]$_1$ kommer ni inte hitta nått jobb [där$_2$
precisely the here part of the math will you not find some job where
man behöver t_1 t_2]
one needs
'You won't find any job where you need THIS part of math.'

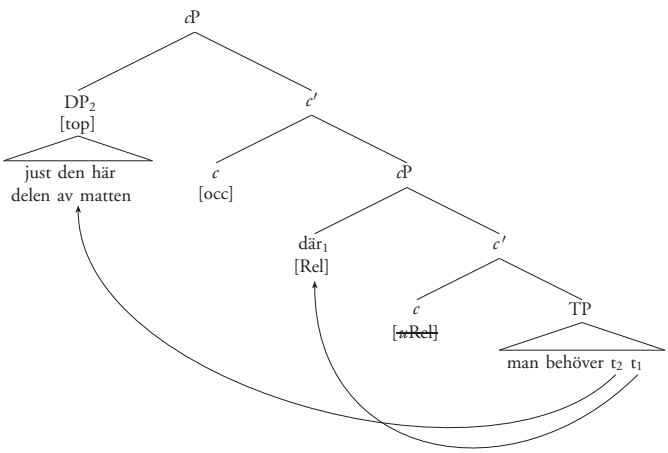
The sentence involves T-preposing of a DP from a non-subject relative clause introduced by the relative adverb *där* 'where'. The DP is a contrastive aboutness topic. For the sake of simplicity, I will assume that movement to the higher specifier is driven by a topic feature. The tree in (17) shows the point in the derivation where the lower *c* head has been merged, and has attracted the relative adverb to its specifier.

(17)



The movement is driven by the feature [*u*Rel] on *c*. In the next step, *c*[occ] is merged. As this feature provides a specifier for the head which carries it without requiring feature matching, the phrase *just den här delen av matten* 'precisely this part of math' can move to its specifier without itself carrying an [occ]-feature, as illustrated in (18).

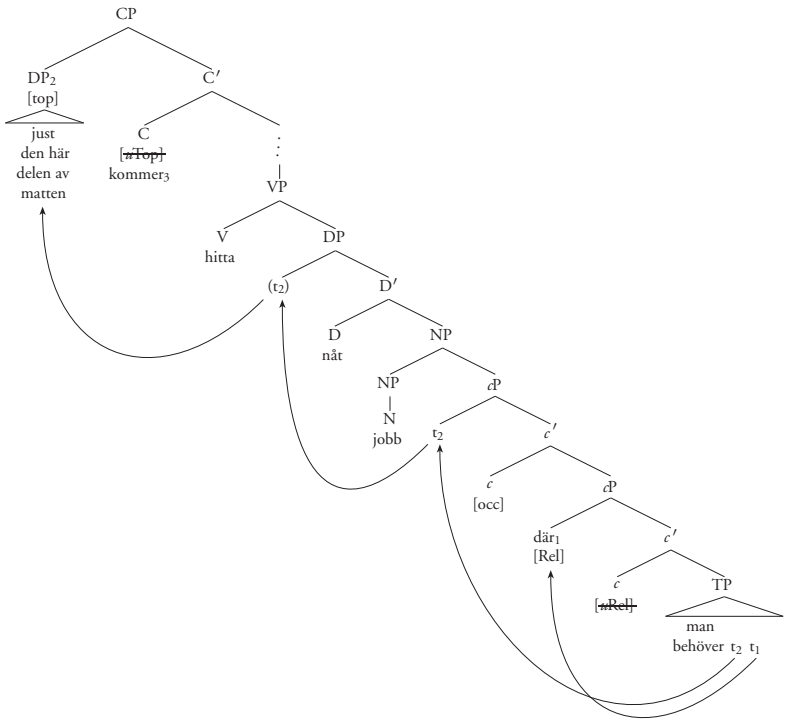
(18)



Since the phrase is now on the edge of *c*P, it is available for subsequent move-

ment out of the DP, and eventually ends up in the specifier of CP in the matrix clause, as shown in (19).

(19)



The movement to Spec-CP is driven by a feature on C in this analysis. According to Vikner (2017), this feature is not inherent in C, however, but C acquires it through the movement of the finite verb.

I have included a trace within parenthesis in the specifier of DP, as well as the trace at the edge of *c*P showing the position where the T-preposed phrase stops on its way out of the relative complex. Whether the T-preposed phrase has to stop at the edge of DP depends on whether DP is assumed to be a phase or not. I will return to the question of whether such an assumption is made necessary by facts about ERC in section 7.2.1.1.[2]

The extra specifier analysis provides a means to allow movement to proceed successive-cyclically through the complementizer domain, even when there is already a relative operator in spec-*c*P in the relative clause. A descriptive general-

[2] I gloss over the question about what happens at the *v*P edge here, even though some or all *v*Ps are phases (see Chomsky 2000, Legate 2003), since the *v*P domain seems to be irrelevant for ERC.

ization that holds in Swedish is that at most one specifier in the complementizer domain in embedded clauses can be overt. As mentioned in section 4.1.5, we never see multiple overt filled specifiers in the case of multiple questions, like in (20), or a phrase stranding in an intermediate specifier on its way out of the clause, like in (21). If any phrase other than the one that moves to the lowest spec-*c*P is stranded, the example will be unacceptable.

(20) a. Jag undrar vem som tog med vad.
I wonder who SOM *took with what*
'I wonder who brought what.'

b. *Jag undrar vad vem som tog med.
I wonder what who that took with

(21) a. [Den där boken]$_1$ undrar jag vem$_2$ som t_2 tog med t_1.
*the there book.*DEF *wonder I who that took with*
'I wonder who brought that book.'

b. *Jag undrar [den där boken]$_1$ vem$_2$ som t_2 tog med t_1.
*I wonder the there book.*DEF *who that took with*

ERC is the same in this respect: we do not find sentences like (22), which is an adaptation of (16), but with the phrase that is T-preposed in (16) stranded in the intermediate specifier.

(22) *Ni kommer inte hitta nåt jobb$_2$ [[just den här delen av matten]$_1$
you will not find some job precisely the here part of the math
där$_2$ man behöver t_1 t_2]
where one needs

In the account proposed by Vikner (2017) and Nyvad et al. (forthcoming), this generalization is captured by properties of the occurrence feature. The feature prevents the phrase that occurs in its specifier from being spelled out there (Nyvad et al. forthcoming).

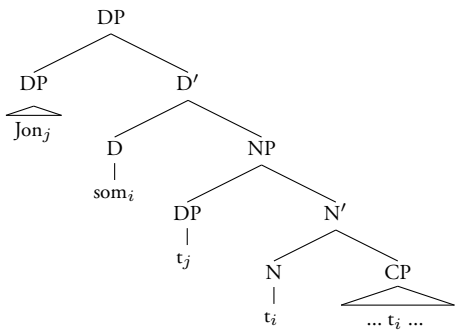
Given the extra specifier account, it is no longer surprising that we find ERC sentences in Swedish. The [occ]-feature provides us with a way to circumvent constraints like Subjacency or Relativized minimality. In the following sections, I discuss three other issues in relation to this account: extraction from non-restrictive relative clauses, multiple extractions, and extractions of adjuncts and arguments.

7.2.1.1 Non-restrictive relative clauses

An observation from the previous literature about ERC is that only restrictive relative clauses permit extraction (e.g. Engdahl 1997, Teleman et al. 1999, Platzack 1999, 2000). Extraction from a non-restrictive relative clause is not possible. As mentioned above, the results of my investigation are in line with this. There are no examples of ERC in my collection in which the movement is out of a non-restrictive relative clause, and constructed examples are invariably unacceptable. The analysis I have developed so far does not capture this generalization, so this deserves some comment.

There are at least two ways of thinking about this restriction. First, it might have to do with the position of the head of the relative clause, as suggested by Platzack (1999, 2000). Platzack (2000:274) proposes that non-restrictive relative clauses have the structure in (23).

(23)



The head of the relative clause starts out in Spec-NP in this structure, and moves to Spec-DP. If DP is a phase, as is sometimes assumed, the head blocks extraction through the edge of DP in this position (Platzack 2000:275). This would explain why non-restrictive relative clause complexes are islands even though extraction from relative clauses is not impossible per se in Swedish. Crucially, this account relies on two assumptions: that DP is a phase, as already mentioned, and that there is no little *d* with an [occ] feature with a function parallel to that of *c*[occ]. As we will see in section 7.2.1.3, this line of explanation faces a challenge in accounting for multiple extractions.

Another possible line of explanation has to do with the parenthetical nature of non-restrictive relative clauses. A phrase inside a parenthetical is generally not extractable, irrespective of whether the parenthetical is a relative clause or not. Consider the set of examples in (24).

(24) a. Jag har haft arachnofobi i hela mitt liv, alltså fobi för spindlar.
I have had arachnophobia in all my life that is phobia for spiders
'I have had arachnophobia all my life, phobia of spiders, that is.'

b. *[Vilket slags djur]$_1$ var det du hade haft arachnofobi i hela ditt liv,
which kind animal was it you had had arachnophobia in all your life
[alltså fobi för t_1]?
that is phobia for

c. *Spindlar$_1$ har haft jag arachnofobi i hela mitt liv, [alltså fobi för t_1].
Spiders have I had arachnophobia in all my life that is phobia for

d. [Vilket slags djur] var det du hade haft fobi för t_1 i hela ditt liv, sa du?
which kind animal was it you had had phobia for in all your life said you
'Which kind of animal did you say you've had a phobia of all your life?'

e. Spindlar$_1$ har jag haft fobi för t_1 i hela mitt liv.
spiders have I had fobia for in all my life
'I have had a phobia of spiders all my life.'

Example (24a) contains a parenthetical, *fobi för spindlar alltså* 'phobia of spiders, that is'. Examples (24b) shows that it is not possible to form a question based on a phrase inside the parenthetical, and (24c) that it is not possible to T-prepose a phrase from within one. In (24d) and (24e), we see that it is possible to prepose the complement of *för* 'for' in this construction, however.

Goldberg (2006) uses examples with parentheticals to argue for the BCI-generalization (Backgrounded Constructions are Islands), suggesting that the problem with such examples is that the parenthetical part of a sentence is not the main point of the utterance, and that they are presupposed. This makes them backgrounded, and phrases inside them inaccessible for extraction (Goldberg 2006:136). We saw in chapter 5 that extraction from presupposed or backgrounded clauses is not a problem *per se*, so this could not be an explanation for the ban on extraction from non-restrictive relative clauses (or parentheticals, for that matter). The more general observation that parenthetical structures are islands in and of themselves is relevant, however, since it indicates that it would be futile to try to find a syntactic account that rules out extraction specifically from non-restrictive relative clauses.

7.2.1.2 Extraction from adjuncts

Another question that deserves a brief comment is the structural position of relative clauses in Swedish. In this thesis, I have adopted a fairly conventional analysis, where the relative clause is an adjunct to NP, contra Platzack (2000) and Stroh-Wollin (2002), who analyze relative clauses as complements. Such analyses are interesting in relation to proposals about island constraints like the Condition on Extraction Domain (Huang 1982), which take adjuncts

and subjects to be islands, whereas complements are non-islands. If Swedish relative clauses are adjuncts, ERC is in violation of the CED, but if they are complements, it is expected that ERC should be possible. At first glance, this seems like an argument in favor of a complement analysis.

However, as I concuded in chapter 4, this argument is not very strong, because it is not clear that the adjunct/argument distinction applies to extraction domains in Swedish. There are cases of acceptable extraction even from clausal adjuncts. The example in (25) is adapted from Teleman et al. (1999 4:424).

(25) [Den här duken]$_1$ blir jag arg [om du spiller på t_1].
the here tablecloth become I angry if you spill on
'I'll get angry if you spill on this tablecloth.'

A recent study by Müller (to appear) confirms that many speakers of Swedish accept extraction from at least a subset of finite adjunct clauses. This means that the CED fails to make the right partitioning between islands and non-islands in Swedish, which in turn weakens the argument for Swedish relative clauses being complements.

7.2.1.3 Multiple extractions

A question I have not addressed so far is whether there is a limit on the number of intermediate landing sites that are available in the complementizer domain. Vikner (2017) and Nyvad et al. (forthcoming) discuss only cases where one phrase is extracted from a subordinate clause, and for this reason, only one *c*[occ] is ever needed to derive their examples. Engdahl (1980b), however, discusses this question in detail, on the basis of extractions from embedded questions and relative clauses. For relative clauses she gives the constructed examples in (26) and (27), where I have adapted the glossing and notation slightly (Engdahl 1980b:100).

(26) Jag har flera studenter [Op$_1$ som det inte finns någon [Op$_2$ som t_2
I have several students that there not exist someone that
vågar prata med t_1 om politiska frågor]].
dares talk to about political issues.
'I have several students that there is no one who dares to talk to about political issues.'

(27) [Sådana känsliga politiska frågor]$_1$ har jag flera studenter [Op$_2$ som det
such touchy political issues have I several students that there
inte finns någon [Op$_3$ som t_3 vågar prata med t_2 om t_1]].
not exist someone that dares talk to about
'I have several students who there is no one who dares talk to about such touchy political issues.'

The sentence in (26) is similar to some examples that we have seen before, and involves two relative clauses, one inside the other. In (27), a phrase has been T-preposed out of the most deeply embedded of these relative clauses. What is at stake here is how many landing sites the complementizer domain in the most deeply embedded relative clause can provide. In (26), there must be two positions, one for creating the first relative clause, and one escape hatch for the movement involved in creating the second relative clause. In (27), a third position is needed for the T-preposed phrase. Insofar as examples like (27) are grammatical in Swedish, an account that relies on intermediate specifier positions as escape hatches for successive-cyclic movement must accommodate at least three specifiers here. Example (28) shows the intermediate landing sites in the relevant complementizer domains.

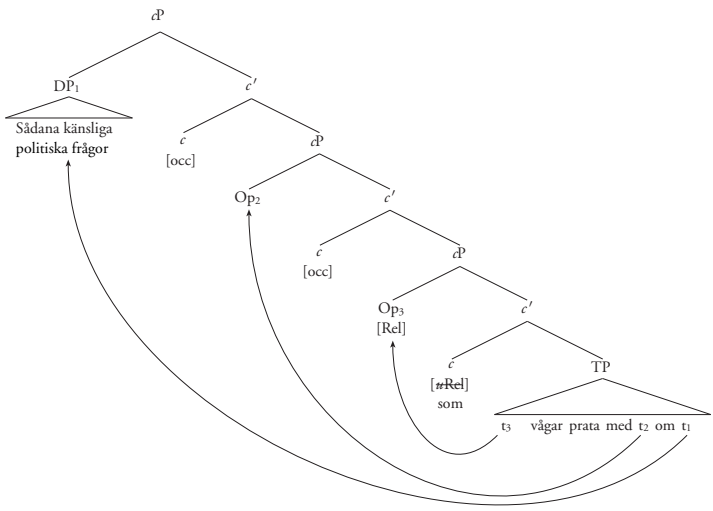
(28) [Sådana känsliga politiska frågor]$_1$ har jag flera studenter [$_{c\mathrm{P}}$ t_1 [$_{c\mathrm{P}}$ Op$_2$
such touchy political issues have I several students
som det inte finns någon [$_{c\mathrm{P}}$ t_1 [$_{c\mathrm{P}}$ t_2 [$_{c\mathrm{P}}$ Op$_3$ som t_3 vågar prata med t_2
that there not exist someone that dares talk to
om t_1]]]]].
about
'I have several students who there is no one who dares talk to about such touchy political issues.'

If movement is successive-cyclic through *c*P, we need two *c*[occ]-heads in the structure to provide the landing sites that would allow the phrases to pass out of the lowest relative clause.

On the basis of this type of example, Engdahl (1980b) argues that it is not meaningful to restrict the number of specifiers in the complementizer domain to one or two, in both embedded questions and in relative clauses. In fact, given complexity and constraints on working memory, it is probably neither necessary nor meaningful to place an upper limit on the number of specifiers in the complementizer domain at all.

Note, however, that such examples raise questions about the approach I have adopted in the dissertation, since they seem to force us to a rather specific view of phases. Consider the tree in (29), which is a representation of the lowest relative clause in (28).

(29)



In order for both of the phrases in the specifiers of the *c*[occ]-heads in (29) to be accessible to operations later in the derivation, they both have to be on the edge of the phase. This necessitates the view that the lowest *c* is the phase head. However, this means that the phase as a whole includes several heads on the left of the phase head. As far as I can tell, the only alternative to this would be that *c*[occ] is sometimes a phase head and sometimes not (see Bošković (2014) for a recent proposal on which a projection's phasal status is determined contextually).

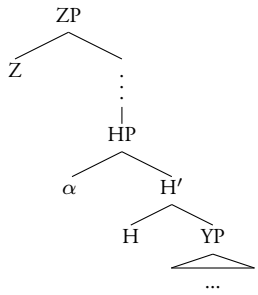
Another possible way forward would be to adopt another version of the Phase Impenetrability Condition, PIC 2, which is weaker than the version we saw in section 2.1.1, repeated here as (30).

(30) **Phase Impenetrability Condition 1**
The domain of H is not accessible to operations outside HP; only H and its *edge* are accessible to such operations. (Chomsky 2001:13)[3]

Consider the structure in (31), where Z and H are phase heads, and the definition of PIC 2 in (32).

[3] The stronger version of the PIC was first introduced in Chomsky (2000), and the weaker PIC 2 in Chomsky (2001), but I quote both definitions from Chomsky (2001) here for the sake of parallelism.

(31)

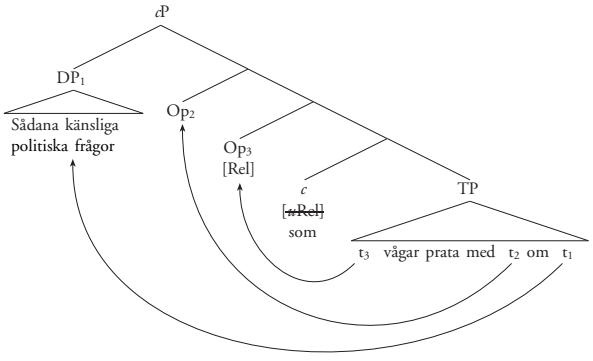


(32) **Phase Impenetrability Condition 2**
The domain of H is not accessible to operations at ZP; only H and its edge are accessible to such operations. (Chomsky 2001:13–14)

The difference between the two versions of the PIC has to do with how long the complement of the phase head is accessible. In the first version, the complement of H becomes inaccessible as soon as HP is complete. In the second version, all of HP is accessible until the higher phase head Z is merged.[4] If the complement of the *c*[*u*Rel] is not spelled out until the next phase head is merged, and *c*[occ] is not a phase head, then all of the specifiers of *c*[occ] will be available in the derivation of the next phase, since the next phase head is either D (or *d*), or *v*. This seems to work, mechanically. However, we are effectively saying that the *c*Ps above the lowest *c*P, i.e. the core relative clause, are not part of the same phase, but part of the higher DP (or *v*P)-phase.

A more standard analysis given these multiple extractions would be that there is only one head in the complementizer domain, but that it tolerates an unlimited number of specifiers, like in (33).

(33)



[4] See Citko (2014:31–41) for further discussion about different definitions of the PIC.

This type of approach should be compatible with the CP/*c*P distinction proposed by Vikner (2017) and Nyvad et al. (forthcoming), and we could still appeal to the distinction to account for the fact that there is no extraction from embedded V2 clauses.

An argument for the *c*P-recursion approach to the complementizer domain is that *c*P-recursion is necessary for the analysis of Scandinavian clauses anyway, since it provides the head positions necessary for complementizer stacking. This means that this approach together with the *c*P/CP-distinction can capture a wide range of phenomena at once. Whether multiple specifiers of a single *c*-head or *c*P-recursion is preferable in a broader perspective will have to be decided on the basis of a wider range of data than just the ERC-sentences. In any case, Engdahl (1980b) seems to be right in pointing out that there is no obvious grammatical cut-off point between (26) and (27). For my analysis, which adopts the recursion approach, this means that we should not put a limit on the number of *c*[occ] that can occur. If a more standard analysis with one *c* with multiple specifiers is adopted instead, there should be no fixed number of specifiers this head tolerates.

Another thing to note about examples with multiple extractions is that they render the filled Spec-DP account of the ban on extraction from non-restrictive relative clauses unexplanatory. If movement has to proceed through the edge of DP, examples like (27) show that DP will have to provide at least two specifiers. This type of example thus further supports the view that extraction from non-restrictive relative clauses is unacceptable due to their parenthetical nature, and not because a sole specifier in the DP-phase is blocked. A more general point can also be made in relation to the last observation: any account that tries to explain Swedish ERC by appealing to a sole specifier which is occupied in other languages being unoccupied in Swedish is challenged by the multiple extraction examples.

7.2.1.4 Extraction of adjuncts and arguments

As we saw above, one of the main findings of the dissertation is that ERC exhibits effects which are reminiscent of effects found in extraction from weak islands, a pattern which was examined in detail in chapter 6, and which has to do with asymmetries in (roughly) adjunct and argument extraction. In ERC, the pattern is only relevant to cases of extraction from restrictive relative clauses; as discussed above in section 7.2.1.1, non-restrictive relative clauses appear to be strong islands, but probably not for syntactic reasons, as we saw above.

In the following sections, I discuss the pattern that the examination in chapter 6 revealed, together with the findings from chapter 4 and chapter 5,

bringing together observations from the investigation as a whole. Table 7.1 sums up which types of phrases can and cannot be extracted from adjunct and argument positions in restrictive relative clauses.

Table 7:1. *Extractable and non-extractable phrases in ERC*

Extractable	Not extractable	Example
T-preposing		
aboutness topics, foci, contrastive phrases		Ch. 5 (48), (65), Ch. 5 (33), (57)
adjuncts		
- specific		Ch. 6 (22)
- deictic		Ch. 6 (21)
- contrastive		Ch. 5 (15), (49)
	other adjuncts	Ch. 6 (23c), (25a), (25b), (27)
Relativization		
Relative Operator in non-restrictive and restrictive RCs		Ch. 5 (75), (80)
***Wh*-questions**		
D-linked *wh*-phrases		Ch. 6 (59)
	non-D-linked *wh*-phrases (in particular *varför* 'why')	Ch. 6 (9)

In T-preposing, extractable phrases function as aboutness topics or foci, which was shown in chapter 5. Phrases that are neither aboutness topics nor foci can still be T-preposed if they are contrastive. Looking specifically at T-preposed adjuncts, we found that they can be extracted if they are specific, deictic, or have a contrastive interpretation. Other adjuncts seem to be impossible to associate with the event expressed by a predicate inside a relative clause.

As shown in example (7) in this chapter, I assume that long movement of a relative operator takes place in ERC with relativization. The relative clause that is formed can be either restrictive or non-restrictive.

In *wh*-question formation, we see a pattern very similar to that which

has been described before for extraction from embedded questions in English, Italian, Romanian, and other languages. As described in section 6.2.2, questions involving time (*när* 'when') or place (*var* 'where') are easier to interpret than questions involving manner (*hur* 'how'), especially out of context. Measure phrases are also harder to extract. *Varför* 'why' cannot be extracted at all. I argued that we can understand these facts about *wh*-questions in terms of D-linking, which seems to make *wh*-questions in ERC easier to interpret quite generally.

How can we explain these properties of Swedish ERC? Previous research has encoded weak island effects in the syntax (Cinque 1990, Frampton 1990, Rizzi 1990), or given them a pragmatic or semantic account (Comorovski 1989, Kroch 1989, Szabolcsi & Zwarts 1993, Szabolcsi 2006). I will discuss the former type of account in this section and argue that even though these accounts provide important insights, they do not entirely explain the restrictions on Swedish ERC. Worse still, they also predict that the ERC examples should be less acceptable than they in fact are. Pragmatic and semantic accounts are then discussed in section 7.2.2.

Observing that not all extractions from embedded questions are equally unacceptable, Cinque (1990), Frampton (1990), and Rizzi (1990) connect extractability to "referentiality" and theta-roles. Simplifying slightly, a phrase which does not carry a referential theta-role will have to move in highly local steps, because its trace will have to be licensed by antecedent government. Phrases which carry a referential theta-role are alleviated of this requirement, because they can enter into binding relations. Long-distance argument questions out of *wh*-islands are still marginal in e.g. English, and on these accounts, this is because they still violate other locality conditions, like Subjacency and Relativized minimality.

In Swedish ERC, we have seen that it is not only possible to extract adjuncts which could be argued to carry referential theta-roles, like *imorgon* 'tomorrow', as in (34a); it is also possible to extract phrases such as manner adjuncts, as in (34b), and measure phrases, as in (34c). It would be hard to argue that such phrases carry referential theta-roles without rendering the theory vacuous.

(34) a. så [imorgon]$_1$ vet jag tre trötta barn [som kommer vara ganska
*so tomorrow know.*COG *I three tired kids that will be pretty*
sega t_1].
tired
'so I know three tired kids who will be pretty tired tomorrow.'

b. Arbetsförmedlingen har inte gjort något, men [genom Per]$_1$ är det
employment service has not done something but through Per are there
många [som fått jobb t_1].
many who got job

'The employment service hasn't done anything, but many people have gotten jobs through Per.'

c. Ida: [Så mycket mjölk]$_1$ har jag aldrig träffat någon [som kan dricka t_1
Ida: that much milk have I never met someone that can drink
på en gång]!
at once

'I have never met anyone who can drink that much milk at once!'

The fact that such examples occur in Swedish suggests that using theta-marking to encode the distinctions between phrases that can and phrases that cannot be extracted will not work easily for Swedish ERC. It also seems that even if proposals which refer to "referentiality" point to something interesting, it is not clear that the distinction between phrases that are "referential" and phrases that are "non-referential" is a syntactic one. Rather, it seems to have a semantic basis, and it is also affected by contextual factors.

Furthermore, Swedish ERC is not the same as long extraction from the prototypical weak islands, i.e. embedded questions in English and Italian, in that ERC-sentences seem to be fully acceptable in many cases. As we have seen, T-preposing and relativization occur in spontaneously produced discourse, without any indication that something unacceptable has been said or written.[5] I take this as support for an approach like the one taken by Vikner (2017) and Nyvad et al. (forthcoming), where extraction from relative clauses proceeds through additional specifiers in the complementizer domain. This type of

[5] In an appendix, Rizzi (1990) points out that examples like the one in (i) are quite well formed in Italian.

(i) Per questa ragione, non immagino chi potrebbe essere licenziato
'For this reason, I don't imagine who could be fired.' (Rizzi 1997:104)

The example involves preposing of a reason adverbial PP out of an embedded question, and the example is thus similar to the T-preposed reason and manner adjuncts that we have seen being T-preposed in ERC. This type of example is surprising given Rizzi's account, since the reason adjunct does not carry a referential theta-role, and there is an intervening Ā-specifier, which should mean that the adjunct cannot move in local steps. Rizzi offers no fully worked out solution to this problem, but suggests that it might be possible to account for it by redefining Relativized minimality.

approach puts ERC on a par with extraction from embedded questions and *att*-clauses in the syntax, since extraction from all of these clause types is assumed to proceed through intermediate landing sites provided by a *c*[occ]. In effect, this is a way to circumvent conditions like Subjacency and Relativized minimality by providing a specifier that will offer a position for a phrase even though there is an intervening Ā-specifier, like the relative operator in relative clauses and the *wh*-phrase in embedded questions.[6]

An extra Ā-specifier accounts for the full acceptability of grammatical extractions from relative clauses. But if extraction from restrictive relative clauses is free in the syntax, then the weak island-like effects in ERC must come from somewhere else. In the next section, I first briefly revisit pragmatic and semantic factors that could explain these effects which were discussed in chapter 6, and then discuss the role of pragmatics, information structure, and semantics in the account of ERC more generally.

7.2.2 Pragmatics and semantics

As I have argued in the dissertation, the previous syntactic accounts of weak island effects are not very helpful for the case of Swedish ERC, since they build on theta-roles. Looking at table 7.1, it is also clear that any other syntactic generalization is elusive, other than the fact that the preposed phrases all appear in spec-CP (or *c*P). Rather, what determines whether argument/adjunct extraction will be possible or not seems to be pragmatics, semantics, and contextual factors which make certain interpretations of a phrase possible.

As discussed in chapter 6, there are several proposals which argue that the argument/adjunct pattern in extraction from embedded questions (and other weak island) need not be encoded in the syntax. For example, Comorovski (1989) argues that the D-linking requirement on *wh*-extraction from embedded questions comes about due to a felicity condition on questions: it has to be possible to check whether their presuppositions are true. For Kroch (1989), what makes certain *wh*-extractions from embedded questions unacceptable is that they would be unusable in most circumstances, because they have existential presuppositions which are almost never true. If we can come up with a situation where the presupposition is true, however, the question becomes

[6] Note that this approach can even handle examples like (7b), with restrictive relativization out of a restrictive relative clause because of the [occ]-feature. The acceptability of such examples would otherwise be very hard to explain on recent conceptions of Relativized minimality (as described by e.g. Boeckx 2012), since the two relative operators must be assumed to carry the same types of features.

usable. This general approach fits well with the results from the interview about *wh*-questions reported in 6.2.2. A certain question often sounded strange or unanswerable to the interview participants, but when they came up with a context where the *wh*-phrase was D-linked, or were provided with such a context, the question sounded a lot better. Consider (35), which is given here together with its existential presupposition.

(35) a. ??Hur$_1$ vet du många [som fick sitt första jobb t_1]?
*how know.*COG *you many that got their.*REFLX.SG *first job*
'Which way is such that you know of many people who got their first job that way?'

b. There is a way such that you know many people who got their first job that way.

Out of context, this question is hard to interpret and answer. If we imagine a situation in which we are talking about ways of getting a job, however, the question becomes easier to answer. In the reasoning of Kroch (1989), the presupposition in (35b) is "implausible", and the question is only usable if we can find a situation where such a presupposition holds.

Another way to look at the restrictions on ERC, which also allows us to say something about the restriction on T-preposing, is Szabolcsi's (2006) approach. From this perspective, where the problem with extraction from weak islands is discussed in terms of scopal interactions, weak island effects arise when a scopal element intervenes between an extracted phrase and the gap (see section 6.2). On this view, it is not surprising that precisely deictic and contrastive phrases, as well as phrases that denote specific times or degrees, are extractable, since they can be shown to denote individuals in Szabolcsi's sense. Non-extractable phrases like *efter en stund* 'after a while' have more complicated denotations; they cannot be interpreted as denoting, for example, some specific "while". I have also suggested, that we could think about this restriction on T-preposing as a restriction on what would make a good topic. The D-linking requirement on questions also follows from this perspective, given that D-linking creates the type of unordered domain which Szabolcsi assumes that extractable *wh*-phrases must range over.

7.2.2.1 The role of pragmatics, information structure, and semantics in ERC

After having discussed the findings from the whole dissertation, we are now in a position to discuss the broader question of the role of pragmatics, information structure, and semantics in ERC. Previous research indicates that such factors

are important in the account of ERC and in extraction in general (Erteschik-Shir 1973, Allwood 1982, Andersson 1982, Engdahl 1997, Jensen 2002, Goldberg 2006), but there is no consensus as to what role they play. The results of my study are largely in line with many of the findings in the literature, but since I have had access to spontaneously produced examples in context, I have also been able to test, and reject, some hypotheses.

Beginning with observations in line with previous research, we saw that, from a discourse-functional perspective, the preposing in ERC is a lot like local preposing, in that the preposed phrases in both types of preposing have the same relations to the context of the utterance.

There are also some very clear patterns when it comes to information structure in the spontaneously produced examples in my collection. These patterns have been observed and well described in the previous literature, and my study aligns with the previous research. Prototypically, the fronted phrase is an aboutness topic, and the rest of the sentence functions as a comment. ERC-sentences share these properties with other extraction sentences as well (see Jensen 2002). However, there are also extraction sentences which do not fall into this pattern, e.g. ERC-sentences with relativization out of a relative clause, and sentences where the fronted phrase is the information focus.

My study has also shown that we need to reevaluate the role of the information impact of the relative complex in constraining extraction. Some previous proposals argue that it is the information impact of an embedded constituent that decides whether it is an island for extraction (Erteschik-Shir 1973, 1982, Goldberg 2006). In brief, they predict that extraction is only possible when the relative complex is dominant or non-backgrounded. Even though the vast majority of ERC-sentences in my collection are of the expected type, there are perfectly acceptable and natural sounding examples which do not follow the pattern (see e.g. (10) in this chapter). Discourse-functional proposals have thus identified a clear pattern of usage, but such generalizations cannot account for the full range of data, and should not be thought of as absolute constraints on extraction.

The weak island effects described in the previous section show that there is nevertheless an important role for pragmatics/semantics in explaining certain constraints on ERC. However, there are also constraints that do not have any obvious pragmatic/semantic explanation, and which seem to be purely syntactic, such as the Comp-trace effects which we saw in chapter 2, section 2.2.1.1 and chapter 4, section 4.6. These facts about subject extraction are hard to understand from a discourse-functional perspective, and even more so since the structurally very similar variety of Swedish *finlandssvenska*

(spoken in Finland) does not have the same Comp-trace effects.

Another challenge for a discourse-functional account of ERC is to capture the differences between languages like English and Swedish, but it is perhaps possible (see the proposal by Engdahl 1997). I have suggested that the variation has a partially syntactic explanation, but the acquisition problem might involve use of discourse strategies. Even though I have suggested a pragmatic/semantic explanation for certain extraction effects, there are also syntactic restrictions which a purely discourse-functional account cannot handle. If anything, my study has shown that there is a need for both syntactic and pragmatic/semantic constraints in the account of ERC.

7.2.3 Implications for a theory of islands

One of the overarching motivations for the present investigation has been to make a contribution to our knowledge of extraction phenomena in natural languages. In this section, I briefly discuss how Swedish ERC fits into the larger study of island phenomena and constraints on movement. My intention has not been to present an entirely new theory of islands, an undertaking that should not be based on an in-depth study of extraction out of just one type of clause in one language. However, by discussing my observations on the case of Swedish ERC, I hope to have contributed insights pointing toward how such a theory should look, and what variation such a theory has to be able to accommodate.

The most important result in this respect is the fact that there is fully acceptable Ā-movement from relative clauses in Swedish, and that creating the relative clause also involves creating an Ā-dependency. From the perspective of syntactic variation, the existence of Swedish ERC indicates that we need to find a way to parametrize the islandhood of relative clauses, and that theories of islands need to provide a means to circumvent constraints like Subjacency and the PIC.

I have linked the acceptability of ERC to the availability of extra specifiers in the complementizer domain. In the specific approach I have pursued here, these specifiers are provided by a head *c*[occ] which can appear at the top of a relative clause. Such a *c*[occ] head is presumably not available in languages like English. If we do not want to take the *c*P/CP approach to the left periphery, and want to maintain a more standard view of the phase edge, we can phrase the difference as follows: the relative C-head in English only tolerates one specifier, whereas the Swedish one tolerates multiple specifiers, so long as only one is overt. Of course in the broader scheme of things, the really interesting question here is what it is that allows Swedish speakers to acquire this type of grammar, while

English speakers do not.

This is a challenge, since Ā-movement in Swedish is in general quite similar to Ā-movement in other languages. It exhibits many familiar characteristics, like parasitic gap licensing, crossover phenomena, and connectivity effects, and is sensitive to certain islands. Furthermore, ERC is constrained in the same way as other types of long extraction in Swedish, as we saw in the overview of previous research on extraction in chapter 2. The type of Ā-dependency that we see out of relative clauses in ERC-sentences cannot violate the Coordinate Structure Constraint or the Sentential Subject Constraint, and it exhibits Comp-trace effects (at least in most varieties of Swedish). Extraction of a subject adjacent to a filled C or Spec-CP cannot leave a gap; the trace has to be spelled out as a resumptive pronoun. These observations are not new, but for the broader discussion of syntactic movement, they are important.

Swedish also provides a challenge for theories of islands that appeal to the distinction between complements and non-complements as extraction domains (Huang 1982). We have no indication that relative clauses are complements in Swedish other than the fact that they permit extraction, and in addition, as we have seen, other finite clausal adjuncts also permit extraction.

As we saw in the previous section, the case of ERC in Swedish is furthermore problematic for purely functionalist/pragmatic theories of island constraints, especially given the cross-linguistic variation in the acceptability of such extractions.

7.3 Future research

Relative clauses are generally assumed to be universally strong islands. In this dissertation, I have reached a different conclusion about relative clauses in Swedish. I have shown that in this language, relative clauses are a type of weak island,[7] and I have argued that certain restrictions on ERC are pragmatic/semantic in nature rather than syntactic. I believe that an analysis of these restrictions inspired by previous research on weak islands, in particular by proposals by Comorovski (1989), Kroch (1989), Szabolcsi & Zwarts (1993), and Szabolcsi (2006), would be most promising. In pursuing this line of inquiry, a natural next step would be to compare extraction from different clause types, looking

[7] This conclusion receives some support from recent experimental research on extraction in Swedish (Tutunjian et al. under review). Tutunjian et al. find that ERC-sentences make up a separate category, in between non islands (that clauses) and strong islands (non-restrictive and subject).

closely at the interaction of the types of fronted phrases with the type of extraction domain. Further comparisons with embedded questions seem particularly motivated. Combining such an investigation with a study of extraction from *att*-clauses with different embedding predicates would also be a natural next step.

Engdahl (1985a) discusses the possibility of interpreting non-argument questions in long extraction from embedded questions and *att*-clauses, and argues that the embedding predicate constrains the available interpretations. The previous literature on weak islands has also recognized this, and identified complements of certain classes of verbs, e.g. factive verbs, as domains that constitute weak islands (Szabolcsi 2006; see also Cattell 1978). If the weak island-like effects that we see in ERC indeed reflect difficulties in checking presuppositions or in finding a situation where a presupposition would be true, then looking more closely at different classes of embedding verbs might prove informative.

This type of study could inform both our understanding of ERC in mainland Scandinavian, and of weak islands in general. Particularly interesting in this case would be to look more closely at functional interpretations of questions. Since such interpretations are not blocked in Swedish weak islands, we might be able to tease apart parts of the account which we might otherwise have thought to require the same explanation.

From the perspective of the general theory of islands and universal grammar, this study leaves us with an intriguing puzzle. I have suggested that Swedish ERC should be accommodated in a general theory of islands by assuming an additional functional head in the complementizer domain, following Vikner (2017) and Nyvad et al. (forthcoming). On their account, this functional head triggers recursion, but the head could also be slightly more standard, and just be a variety of a relative C-head which tolerates multiple specifiers. In either case, the puzzle is the same. How do Swedish speakers acquire the grammars that we have for ERC, and how do speakers of languages like English arrive at their more rigorous restriction?

The distribution of the different types of ERC-sentences is of particular interest here. Common types, like the existential and presentational types, occur in everyday conversation, and should presumably be sufficient for speakers of Swedish to acquire ERC. More atypical ERC-sentences, with an unusual embedding predicate, unusual information structure, or extraction from a non-subject relative clause, seem to be exceedingly rare. Yet such sentences are perfectly acceptable when they fulfill the general restrictions we have seen: that the extracted phrase be of the right type, that the relative clause be restrictive,

that we do not extract a subject without spelling out a resumptive pronoun, and so on.

The sparsity of the non-canonical types of ERC is potentially a problem for accounts like the one put forth by Löwenadler (2015), where it is assumed that the acceptability of an ERC-sentence is decided by its compatibility with language-specific conventionalized constructional schemas. To take a specific example, Löwenadler (2015) assumes that ERC-sentences with embedding predicates like *beundra* 'admire' and *avsky* 'detest' are instances of a "Cause-related complex NP extraction construction" in Swedish. But given the findings in this dissertation, it is unclear how such a schema could ever be conventionalized or acquired at all. I actively collected ERC-sentences for five years, and came across less than a handful of this type. In my view, a more plausible approach is to assume that what is acquired is an abstract syntactic generalization along the lines of Vikner (2017) and Nyvad et al. (forthcoming).

However, to find out more about the acquisition problem, we should, of course, obtain additional data on the frequency of ERC. The data collection method in this study makes it impossible to say anything very specific about the frequency of certain types of extraction, even if we can say what is common and what is not common, broadly speaking. A future study could look into the frequency of the different types of extractions in corpora of different kinds. A prediction based on the impressionistic pattern I can see in the types of data that I have is that less formal genres should offer more variation. The present study can contribute ideas about which types of extractions are good candidates to look at more closely. When we know more about what can be found in different kinds of linguistic input, we can get a better idea about how to approach the acquisition problem.

Corpora of child-directed speech are the most important source of information regarding acquisition. A possible way forward is to examine such corpora for Danish, Norwegian, and Swedish, and compare them to similar corpora for English, Icelandic, Faroese, German, and Dutch, to see whether there is anything in the input which could lead to the differences in acquisition. One noticeable fact is that many ERC-sentences involve fronting of non-contrastive pronouns, and it would be interesting to investigate whether non-contrastive pronouns are fronted in child-directed speech in the different languages, and if so, then how often.

ERC-sentences with existential predicates are also potentially of interest here. Future research should find out whether extraction from the type of ambiguous existential sentences discussed in section 4.3.1 is found in child-directed speech in Swedish, Norwegian, and Danish, but absent in child-directed speech in languages like English.

Sammanfattning (Summary)

I svenska, liksom i danska och norska, går det att bilda satsfläta med relativsats, som i (1).

(1) a. men ingen av dom är ju varmblodiga (.) det$_1$ finns det inga insekter [som är t_1]
b. ja [ett lodjur]$_1$ har jag inte hört talas om nån [som blivit uppäten av t_1]
c. där$_1$ hade du ju en svartvit flugsnappare [som brukade bo t_1]

Leden *det*, *ett lodjur* och *där* i (1) har sin grammatiska funktion i relativsatsen i respektive exempel, men realiseras inte där, utan är **spetsställda**. I positionen där ledens grammatiska funktion bestäms finns en lucka, här markerad med *t*, och ett sifferindex markerar att det finns en relation mellan det spetsställda ledet och luckan.

Satsflätor som i (1) har beskrivits i den skandinaviska grammatiktraditionen av bland andra Mikkelsen (1894), Cederschiöld (1897), Wellander (1939), Faarlund et al. (1997) och Teleman et al. (1999). De har också väckt uppmärksamhet i internationell språkvetenskaplig forskning, framförallt inom den teoretiska lingvistiken och inom psykolingvistiken (t.ex. Erteschik-Shir 1973, Allwood 1976, Engdahl 1997, Kush 2011, Müller 2015, Vikner 2017, Nyvad et al. forthcoming, Tutunjian et al. under review). Satsfläta med relativsats är nämligen typologiskt ovanligt, och i de flesta språk där det undersökts är sådana satsflätor oacceptabla. För engelskans del observerades detta tidigt av Chomsky (1964) och Ross (1967), som diskuterade exempel som de i (2).

(2) a. *The man who I read a statement that was about is sick. (Ross 1967:119)
b. *Who does Phineas know a girl who is working with. (Ross 1967:124)

Att satser som (2a) och (2b) är omöjliga att bilda i många språk har föreslagits bero på syntaktiska restriktioner (t.ex. Chomsky 1964, Ross 1967), pragmatiska

restriktioner (t.ex. Erteschik-Shir 1973, Goldberg 2006) eller på hur människan processar språk (t.ex. Deane 1991, Hofmeister & Sag 2010). Ofta antas restriktionerna vara universella. Eftersom de fastlandsskandinaviska språken fungerar annorlunda än engelska och många andra språk när det gäller satsflätor kan vi, genom att studera spontana satsflätor i svenskan, få mer kunskap både om vilken variation som förekommer och om vad denna variation innebär för teorier om den mentala grammatiken.

Även i de fastlandsskandinaviska språken, där satsfläta med relativsats accepteras av många talare, finns det begränsningar. Exempelvis kan man inte bilda satsfläta med icke-restriktiva relativsatser, vilket visas i (3b).

(3) a. Biljard$_1$ fanns där många [som spelade t_1].

b. *Biljard$_1$ fanns där väldigt många människor [som alla spelade t_1].

(Teleman et al. 1999 4:423)

Satsflätan i (3a), med en restriktiv relativsats, är välformad, medan den i (3b), med en icke-restriktiv relativsats är oacceptabel.

Tidigare forskning pekar på att informationsstruktur och pragmatik spelar en stor roll för möjligheterna att bilda satsfläta (t.ex Erteschik-Shir 1973, Allwood 1976, Andersson 1982, Engdahl 1997, Goldberg 2006). Erteschik-Shir (1973) observerar exempelvis att relativkomplexets informationsstatus spelar roll för satsflätans acceptabilitet, och Engdahl (1997) visar med hjälp av exempel från talspråk och skrift att spetsställningen i satsfläta med relativsats ofta är diskursmotiverad.

Större delen av den tidigare forskningen har utgått från konstruerade exempel utan kontext. Ur det perspektivet är de fastlandsskandinaviska språken intressanta eftersom satsfläta med relativsats inte bara accepteras av många svensktalande, utan också förekommer i spontant språkbruk.

I avhandlingen undersöker jag satsfläta med relativsats från ett allmänlingvistiskt perspektiv, och föreslår en analys av sådana satsflätor på basis av systematiserade spontandata och acceptabilitetsomdömen. Undersökningen har vägletts av fyra frågeställningar:

1. Vilken sorts relation är det mellan det spetsställda ledet och luckan i satsfläta med relativsats?
2. Vilken struktur har relativsatserna i sådana satsflätor?
3. Vilken roll spelar diskurs och pragmatik för bildandet av satsflätor?
4. Vilka begränsningar finns på satsfläta med relativsats?

Ett övergripande syfte har varit att bidra till kunskapen om de fastlandsskandinaviska språkens satsstruktur, och till kunskapen om hur den mänskliga språkförmågan är organiserad. Ett delsyfte är att tillgängliggöra relevanta svenska data för den internationella forskargemenskapen.

Avhandlingens huvudmaterial är en samling om 270 spontant producerade satsflätor från talad och skriven svenska som samlats in mellan åren 2011 och 2016. Talspråksmaterialet (163 satsflätor) kommer från vardagssamtal, radio och tv, och skriftspråksmaterialet (107 satsflätor) från skönlitteratur, tidningstext och sociala medier. Talspråksexemplen är nedtecknade efter spontant lyssnande, dvs. att jag nedtecknat exempel jag observerat i min vardag. Vad gäller skriftspråksexemplen är vissa av dessa insamlade via korpussökningar, men liksom med talspråksexemplen har jag också antecknat exempel jag stött på i vardagen. Data har också samlats in genom en enkät, där 16 modersmålstalare fick bedöma satsflätor i kontext, samt genom en mindre intervjustudie med fyra deltagare.

I kapitel 4 diskuteras ett flertal frågeställningar relaterade till satsflätornas syntaktiska struktur. Bland annat undersöks relationen mellan det spetsställda ledet och luckan i relativsatsen, relativkomplexets definithet, relativsatsens struktur och luckans position, samt vilka matrisverb relativkomplexet kan vara inbäddat under. Undersökningen visar att satsflätning oftast förekommer i existentialsatser och presenteringssatser, men det finns också andra typer, såsom satsflätor där relativkomplexet är inbäddat under verb som *ha*, *känna*, *se*, *beundra* och *störa sig på*. Spetsställning förekommer också ur den relativsatsliknande konstituenten i några olika typer av utbrytningssatser.

I kapitel 5 presenteras en undersökning av hur satsflätorna används i kontext, med fokus på det spetsställda ledets funktion och relativkomplexets informationsstatus. Spetsställning i satsfläta med relativsats jämförs med spetsställning i den lokala satsen. Undersökningen visar att det spetsställda ledet i satsfläta med relativsats relaterar till kontexten på samma sätt som det spetsställda ledet i lokal spetsställning. När det gäller det spetsställda ledes funktion i yttrandet utgör det oftast aboutness-topik. Det kan också vara satsens rema, men undersökningen visar att det inte kan fungera som scenbyggare, vilket är en vanlig funktion för spetsställda led i lokal spetsställning. Relativkomplexet i satsfläta med relativsats är ofta diskursprominent, men undersökningen visar att det också finns fullt acceptabla undantag från generaliseringar om relativkomplexets informationsstatus, exempelvis satsflätor med matrisverb som *beundra* och *störa sig på*.

En studie av spetsställda argument och adverbial som presenteras i kapitel 6 visar att det finns begränsningar på vilka led som kan spetsställas ur relativsatser,

och att begränsningarna är starkare än på satsfläta med *att*-sats. Ett tydligt exempel finns i (4).

(4) Varför känner du många som har skrivit böcker?

a. För att jag är med i Författarförbundets styrelse.

b. #För att chockera sin samtid.

Yttrandet i (4a) är ett naturligt svar på frågan, men yttrandet i (4b) är ett märkligt svar, vilket tyder på att *varför* inte kan tolkas såsom hörande till relativsatsen. I studien undersöks förutom konstituentfrågor ett flertal olika typer av spetsställda adverbial, och studien visar att begränsningarna har att göra med det spetsställda ledets semantiska/pragmatiska egenskaper.

I avhandlingen diskuteras satsflätorna i relation till olika förslag från den allmänlingvistiska forskningen, med utgångspunkt i en analys som nyligen föreslagits av Vikner (2017) och Nyvad et al. (forthcoming) inom det Minimalistiska programmet. Utifrån min undersökning argumenterar jag för att satsfläta med relativsats är möjligt i de fastlandsskandinaviska språken, men inte i engelska och liknande språk, på grund av en skillnad i relativsatsernas C-domän. Jag visar också att restriktionerna som vi finner i svenskan verkar vara av olika karaktär; det finns både syntaktiska och pragmatiska/semantiska begränsningar.

Bibliography

Åfarli, Tor A. 1994. A promotion analysis of restrictive relative clauses. *The Linguistic Review* 11:2, pp. 81–100.

Allwood, Jens 1976. The complex NP constraint as a non-universal rule and some semantic factors influencing the acceptability of Swedish sentences which violate the CNPC. In: Stillings, Justine (ed.), *University of Massachusetts Occasional Publications in Lingustics* II. Amherst, MA: University of Massachusetts, pp. 1–20.

Allwood, Jens 1982. The Complex NP Constraint in Swedish. In: Engdahl & Ejerhed (1982), pp. 15–32.

Andersson, Lars-Gunnar 1975. *Form and Function of Subordinate Clauses*. (Gothenburg Monographs in Linguistics 1.) Gothenburg: Department of Linguistics, University of Gothenburg.

Andersson, Lars-Gunnar 1982. What is Swedish an Exception to? In: Engdahl & Ejerhed (1982), pp. 33–45.

Andréasson, Maia 2007. *Satsadverbial, ledföljd och informationsdynamik i svenskan* [*Sentence adverbials, Constituent order, and Information dynamics in Swedish*]. (Göteborgsstudier i nordisk språkvetenskap 7.) Gothenburg: Department of Swedish, University of Gothenburg.

Bentzen, Kristine, Piotr Garbacz, Caroline Heycock & Gunnar Hrafn Hrafnbjargarson 2009. On variation in Faroese verb placement. *Nordlyd* 36:2, pp. 78–102.

Bentzen, Kristine, Gunnar Hrafn Hrafnbjargarson, Þorbjörg Hróarsdóttir & Anna-Lena Wiklund 2007a. Extracting from V2. *Working Papers in Scandinavian Syntax* 79, pp. 119–128.

Bentzen, Kristine, Gunnar Hrafn Hrafnbjargarson, Þorbjörg Hróarsdóttir & Anna-Lena Wiklund 2007b. The Tromsø guide to the Force behind V2. *Working Papers in Scandinavian Syntax* 79, pp. 93–118.

den Besten, Hans 1983. On the Interaction of Root Transformations and Lexical Deletive Rules. In: Abraham, Werner (ed.), *On the Formal Syntax of the Westgermania. Papers from the "3rd Groningen Grammar Talks (3e Groninger Grammatikgespräche)", Groningen, January 1981*. (Linguistik Aktuell 3.) Amsterdam: John Benjamins, pp. 47–131.

Bhatt, Rajesh 2002. The Raising Analysis of Relative Clauses: Evidence from Adjectival Modification. *Natural Language Semantics* 10:1, pp. 43–90.

Boeckx, Cedric 2012. *Syntactic Islands*. (Key topics in syntax.) Cambridge: Cambridge University Press.

Borin, Lars, Markus Forsberg & Johan Roxendal 2012. Korp – the corpus infrastructure of Språkbanken. In: *Proceedings of LREC 2012*. Istanbul: ELRA, pp. 474–478.

Borthen, Kaja 2003. *Norwegian Bare Singulars*. Ph.D. thesis, Norwegian University of Science and Technology, Trondheim.

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* 45:1, pp. 27–89.

Brandtler, Johan submitted. The question of form in the forming of questions. The meaning and use of clefted *wh*-interrogatives in Swedish. Ms. 2017 Ghent University.

Bresnan, Joan 1977. Variables in the Theory of Transformations. In: Culicover et al. (1977), pp. 157–196.

Cattell, Ray 1978. On the Source of Interrogative Adverbs. *Language* 54:1, pp. 61–77.

Cederschiöld, Gustaf 1897. *Om svenskan som skriftspråk* [*Swedish as a written language*]. (Populärt vetenskapliga föreläsningar vid Göteborgs Högskola 4.) Gothenburg: Wettergren & Kerber.

Chomsky, Noam 1964. *Current Issues in Linguistic Theory*. (Janua Linguarum. Series Minor 38.) The Hague: Mouton.

Chomsky, Noam 1973. Conditions on Transformations. In: Andersson, Stephen R. & Paul Kiparsky (eds), *A Festschrift for Morris Halle*. New York: Holt, Rinehart and Winston, pp. 232–286.

Chomsky, Noam 1977. On WH-Movement. In: Culicover et al. (1977), pp. 71–132.

Chomsky, Noam 1981. *Lectures on Government and Binding*. (Studies in Generative Grammar 9.) Dordrecht: Foris.

Chomsky, Noam 1986. *Barriers*. (Linguistic Inquiry Monographs 13.) Cambridge, MA: MIT Press.

Chomsky, Noam 1993. A Minimalist Program for Linguistic Theory. In: Hale, Kenneth & Samuel Jay Keyser (eds), *The View from Building 20. Essays in Linguistics in Honor of Sylvain Bromberger*. (Current Studies in Linguistics 24.) Cambridge, MA: MIT Press, pp. 1–52.

Chomsky, Noam 1995. *The Minimalist Program*. (Current Studies in Linguistics 28.) Cambridge, MA: MIT Press.

Chomsky, Noam 2000. Minimalist Inquiries: The Framework. In: Martin, Roger, David Michaels & Juan Uriagereka (eds), *Step by Step. Essays on Minimalist Syntax in Honour of Howard Lasnik*. Cambridge, MA: MIT Press, pp. 89–115.

Chomsky, Noam 2001. Derivation by Phase. In: Kenstowicz, Michael (ed.), *Ken Hale. A Life in Language*. Cambridge, MA: MIT Press, pp. 1–52.

Chomsky, Noam 2005. Three Factors in Language Design. *Linguistic Inquiry* 36:1, pp. 1–22.

Chomsky, Noam & Howard Lasnik 1977. Filters and Control. *Linguistic Inquiry* 8:3, pp. 425–504.

Christensen, Ken Ramshøj, Johannes Kizach & Anne Mette Nyvad 2013a. Escape from the Island: Grammaticality and (Reduced) Acceptability of *wh*-island Violations in Danish. *Journal of Psycholinguistic Research* 42:1, pp. 51–70.

Christensen, Ken Ramshøj, Johannes Kizach & Anne Mette Nyvad 2013b. The processing of syntactic Islands – An fMRI study. *Journal of Neurolinguistics* 26:2, pp. 239–251.

Christensen, Ken Ramshøj & Anne Mette Nyvad 2014. On the nature of escapable relative islands. *Nordic Journal of Linguistics* 37:1, pp. 29–45.

Christensen, Kirsti Koch 1982. On Multiple Filler-Gap Constructions in Norwegian. In: Engdahl & Ejerhed (1982), pp. 77–98.

Chung, Sandra 1982. Unbounded Dependencies in Chamorro Grammar. *Linguistic Inquiry* 13:1, pp. 39–77.

Chung, Sandra 1994. Wh-Agreement and "Referentiality" in Chamorro. *Linguistic Inquiry* 25:1, pp. 1–44.

Chung, Sandra & James McCloskey 1983. On the Interpretation of Certain Island Facts in GPSG. *Linguistic Inquiry* 14:4, pp. 704–713.

Cinque, Guglielmo 1990. *Types of Ā-dependencies*. (Linguistic Inquiry Monographs 17.) Cambridge, MA: MIT Press.

Citko, Barbara 2014. *Phase Theory. An Introduction*. (Research Surveys in Linguistics.) Cambridge: Cambridge University Press.

Comorovski, Ileana 1989. Discourse Linking and the WH-Island Constraint. In: *Proceedings of the 19th Annual Meeting Northeastern Linguistic Society*, vol. 19. pp. 78–96.

Cresti, Diana 1995. Extraction and Reconstruction. *Natural Language Semantics* 3:1, pp. 79–122.

Croft, William 2001. *Radical Construction Grammar. Syntactic Theory in Typological Perspective*. Oxford: Oxford University Press.

Culicover, Peter W., Thomas Wasow & Adrian Akmajian (eds) 1977. *Formal Syntax*. New York: Academic Press.

Dalrymple, Mary 2001. *Syntax and Semantics 34. Lexical Functional Grammar*. New York: Academic Press.

Daneš, František 1974. Functional sentence perspective and the organization of the text. In: Daneš, František (ed.), *Papers on Functional Sentence Perspective*. (Janua Linguarum. Series Minor 147.) Prague: Academia, pp. 106–128.

Deane, Paul 1991. Limits to attention: A cognitive theory of island phenomena. *Cognitive Linguistics* 2:1, pp. 1–63.

Diderichsen, Paul 1946. *Elementær Dansk Grammatik* [*Elementary Danish Grammar*]. Copenhagen: Gyldendal.

Eide, Kristin M. & Tor A. Åfarli 1999. The Syntactic Disguises of the Predication Operator. *Studia Linguistica* 53:2, pp. 155–181.

Engdahl, Elisabet 1980a. *The Syntax and Semantics of Questions in Swedish*. Ph.D. thesis, University of Massachusetts Amherst, Amherst, MA.

Engdahl, Elisabet 1980b. WH-Constructions in Swedish and the Relevance of Subjacency. In: *Proceedings of the 10th Annual Meeting of the North East Linguistic Society*. (Cahiers Linguistiques d'Ottawa 9.) Ottawa: University of Ottawa, pp. 98–108.

Engdahl, Elisabet 1982. Restrictions on Unbounded Dependencies in Swedish. In: Engdahl & Ejerhed (1982), pp. 151–174.

Engdahl, Elisabet 1983. Parasitic Gaps. *Linguistics and Philosophy* 6:1, pp. 5–34.

Engdahl, Elisabet 1985a. Non-argument questions. Talk given at the Comparative Germanic workshop in Reykjavik, June 1985.

Engdahl, Elisabet 1985b. Parasitic gaps, resumptive pronouns, and subject extractions. *Linguistics* 23:1, pp. 3–44.

Engdahl, Elisabet 1986. *Constituent Questions. The Syntax and Semantics of Questions with Special Reference to Swedish*. (Studies in Linguistics and Philosophy 27.) Dordrecht: Reidel.

Engdahl, Elisabet 1997. Relative Clause Extractions in Context. *Working Papers in Scandinavian Syntax* 60, pp. 51–79.

Engdahl, Elisabet & Eva Ejerhed (eds) 1982. *Readings on Unbounded Dependencies in Scandinavian Languages*. (Umeå Studies in the Humanities 43.) Stockholm: Almqvist & Wiksell International.

Engdahl, Elisabet & Filippa Lindahl 2014. Preposed object pronouns in mainland Scandinavian. *Working Papers in Scandinavian Syntax* 92, pp. 1–32.

Erteschik-Shir, Nomi 1973. *On the Nature of Island Constraints*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Erteschik-Shir, Nomi 1982. Extractability in Danish and the Pragmatic Principle of Dominance. In: Engdahl & Ejerhed (1982), pp. 175–191.

Erteschik-Shir, Nomi 2007. *Information Structure. The Syntax-Discourse Interface*. (Oxford Surveys in Syntax and Morphology 3.) Oxford: Oxford University Press.

Erteschik-Shir, Nomi & Shalom Lappin 1979. Dominance and the Functional Explanation of Island Phenomena. *Theoretical Linguistics* 6:1–3, pp. 41–86.

Faarlund, Jan Terje, Svein Lie & Kjell Ivar Vannebo 1997. *Norsk referansegrammatikk* [*Norwegian reference grammar*]. Oslo: Universitetsforlaget.

Fiedler, Judith 2014. *Germanic It-Clefts: Structural Variation and Semantic Uniformity*. Ph.D. thesis, University of California, Santa Cruz, Santa Cruz, CA.

Fodor, Janet Dean 1978. Parsing Strategies and Constraints on Transformations. *Linguistic Inquiry* 9:3, pp. 427–473.

Frampton, John 1990. The Fine Structure of Wh-Movement and the Proper Formulation of the ECP. Ms. Northeastern University.

Gazdar, Gerald, Ewan Klein, Geoffrey K. Pullum & Ivan A. Sag 1985. *Generalized Phrase Structure Grammar*. Oxford: Blackwell.

Goldberg, Adele E. 2006. *Constructions at Work. The Nature of Generalization in Language*. New York: Oxford University Press.

Goldberg, Adele E. 2013. Backgrounded constituents cannot be "extracted". In: Sprouse & Hornstein (2013), pp. 221–238.

Gundel, Jeanette K. & Thorstein Fretheim 2004. Topic and Focus. In: Horn, Laurence R. & Gregory Ward (eds), *The Handbook of Pragmatics*. (Blackwell Handbooks in Linguistics 16.) Malden, MA: Blackwell, pp. 175–196.

de Haan, Germen & Fred Weerman 1986. Finiteness and Verb Fronting in Frisian. In: Haider & Printzhorn (1986), pp. 77–110.

Haegeman, Liliane 1994. *Introduction to Government and Binding Theory*. 2nd edn. (Blackwell Textbooks in Linguistics 1.) Oxford: Blackwell.

Haider, Hubert & Martin Printzhorn (eds) 1986. *Verb Second Phenomena in Germanic Languages*. (Publications in Language Sciences 21.) Dordrecht: Foris.

Hofmeister, Philip & Ivan A. Sag 2010. Cognitive constraints and island effects. *Language* 86:2, pp. 366–415.

Hooper, Joan B. & Sandra A. Thompson 1973. On the Applicability of Root Transformations. *Linguistic Inquiry* 4:4, pp. 465–497.

Hrafnbjargarson, Gunnar Hrafn, Kristine Bentzen & Anna-Lena Wiklund 2010. Observations on extraction from V2 clauses in Scandinavian. *Nordic Journal of Linguistics* 33:3, pp. 299–309.

Huang, C.-T. James 1982. *Logical Relations in Chinese and the Theory of Grammar*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Hulsey, Sarah & Uli Sauerland 2006. Sorting out Relative Clauses. *Natural Language Semantics* 14:2, pp. 111–137.

Jensen, Anne 2002. Sætningsknuder i dansk [Knot sentences in Danish]. *NyS - Nydanske Studier & Almen kommunikationsteori* 29, pp. 105–124.

Johannessen, Janne Bondi, Joel Priestley, Kristin Hagen, Tor A. Åfarli & Øystein A. Vangsnes 2009. The Nordic Dialect Corpus – an advanced research tool. In: Jokinen, Kristiina & Eckhard Bick (eds), *Proceedings of the 17th Nordic Conference of Computational Linguistics NODALIDA 2009*. (NEALT Proceedings Series 4.) pp. 73–80.

Jörgensen, Nils 1976. *Meningsbyggnaden i talad svenska* [*Sentence structure in spoken Swedish*]. (Lundastudier i nordisk språkvetenskap C 7.) Lund: Studentlitteratur.

Julien, Marit 2005. *Nominal Phrases from a Scandinavian Perspective*. (Linguistik Aktuell 87.) John Benjamins.

Julien, Marit 2015. The force of V2 revisited. *Journal of Comparative Germanic Linguistics* 18:2, pp. 139–181.

Kayne, Richard S. 1994. *The Antisymmetry of Syntax*. (Linguistic Inquiry Monographs 25.) Cambridge, MA: MIT Press.

Keenan, Edward L. 1987. A Semantic Definiton of "Indefinite NP". In: Reuland & ter Meulen (1987), pp. 286–317.

Kiss, Katalin É 1998. Identificational focus versus information focus. *Language* 74:2, pp. 245–273.

Kluender, Robert & Marta Kutas 1993. Subjacency as a Processing Phenomenon. *Language and Cognitive Processes* 8:4, pp. 573–633.

Kroch, Anthony 1989. Amount quantification, referentiality, and long wh-movement. Ms. University of Pennsylvania.

Kush, Dave 2011. On the (In?)Escapability of Islands in Scandinavian. Ms. University of Maryland.

Kush, Dave, Akira Omaki & Norbert Hornstein 2013. Microvariation in islands? In: Sprouse & Hornstein (2013), pp. 239–264.

Lambrecht, Knud 1994. *Information structure and sentence form. Topic, focus and the mental representations of discourse referents*. (Cambridge Studies in Linguistics 71.) Cambridge: Cambridge University Press.

Lasnik, Howard & Robert Fiengo 1974. Complement Object Deletion. *Linguistic Inquiry* 5:4, pp. 535–571.

Lasnik, Howard & Mamoru Saito 1984. On the Nature of Proper Government. *Linguistic Inquiry* 15:2, pp. 235–289.

Legate, Julie Anne 2003. Some Interface Properties of the Phase. *Linguistic Inquiry* 34:3, pp. 506–516.

Lindahl, Filippa 2010. Spetsställda led och rematiska relativer. En korpusstudie av satsfläta med presenteringsomskrivning/existentialsats [A corpus study of extraction from presentational clefts and existential constructions]. MA Thesis, Department of Swedish, University of Gothenburg.

Lindahl, Filippa 2011. Spetsställda bisatser i satsfläta med relativsats. [Preposed subordinate clauses in relative clause extraction]. *Språk och stil* 21, pp. 199–204.

Lindahl, Filippa 2014. Relative clauses are not always strong islands. *Working Papers in Scandinavian Syntax* 93, pp. 1–25.

Lindstedt, Torvald 1926. Om satssammanflätning i svenskan [About long extractions in Swedish]. *Nysvenska studier* 6, pp. 1–20.

Ljunggren, Carl August 1917. *Stilarter och språkriktighetsregler. Ett försök till hjälpreda vid uppsatsskrivningen i modersmålet* [*Style and rules for proper language use*]. Lund: Gleerupska Universitetsbokhandeln.

Löwenadler, John 2015. Relative clause extraction: Pragmatic dominance, processing complexity and the nature of crosslinguistic variation. *Nordic Journal of Linguistics* 38:1, pp. 37–65.

Maling, Joan M. 1978. An Asymmetry with Respect to Wh-Islands. *Linguistic Inquiry* 9:1, pp. 75–89.

McCawley, James D. 1981. The syntax and semantics of English relative clauses. *Lingua* 53:2–3, pp. 99–149.

McCloskey, James 1979. *Transformational Syntax and Model Theoretic Semantics: A Case Study in Modern Irish*. (Synthese Language Library 9.) Reidel: Dordrecht.

McCloskey, James 1988. Syntactic theory. In: Newmeyer, Frederick J. (ed.), *Linguistic Theory: Foundations*. (Linguistics: The Cambridge Survey I.) Cambridge: Cambridge University Press, pp. 18–59.

McCloskey, James 2001. The Morphosyntax of WH-extraction in Irish. *Journal of Linguistics* 37:1, pp. 67–100.

McCloskey, James 2014. Irish Existentials in Context. *Syntax* 17:4, pp. 343–384.

Mikkelsen, Kristian Mathias 1894. *Dansk Sproglære med sproghistoriske Tillæg. Haandbog for Lærere og Viderekomne* [*Danish Grammar*]. Copenhagen: Lehmann & Stage.

Mikkelsen, Line 2015. VP anaphora and verb-second order in Danish. *Journal of Linguistics* 51:3, pp. 595–643.

Milsark, Gary 1974. *Existential Sentences in English*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Mörnsjö, Maria 2002. *V1 Declaratives in Spoken Swedish. Syntax, Information Structure, and Prosodic Pattern*. (Lundastudier i nordisk språkvetenskap A 59.) Lund: Department of Scandinavian Languages, Lund University.

Müller, Christiane 2015. Against the Small Clause Hypothesis: Evidence from Swedish relative clause extractions. *Nordic Journal of Linguistics* 38:1, pp. 67–92.

Müller, Christiane to appear. Extraction from Adjunct Islands in Swedish. To appear in *Norsk Lingvistisk Tidskrift* 35.

Norén, Niklas 2007. *Apokoinou in Swedish talk-in-interaction. A family of methods for grammatical construction and the resolving of local communicative projects*. (Linköping studies in arts and science 405.) Linköping: Department of Culture and Communication, Linköping University.

Nyvad, Anne Mette, Ken Ramshøj Christensen & Sten Vikner forthcoming. CP-Recursion in Danish: A *c*P/CP Analysis. Forthcoming in *The Linguistic Review*.

Pesetsky, David 1987. Wh-in-Situ: Movement and Unselective Binding. In: Reuland & ter Meulen (1987), pp. 98–129.

Pesetsky, David 2013. Phrasal Movement and Its Discontents: Diseases and Diagnoses. In: Cheng, Lisa Lai-Shen & Norbert Corver (eds), *Diagnosing Syntax*. Oxford: Oxford University Press, pp. 123–157.

Phillips, Colin 2006. The real-time status of island phenomena. *Language* 82:4, pp. 795–823.

Platzack, Christer 1986. The Position of the Finite Verb in Swedish. In: Haider & Printzhorn (1986), pp. 27–47.

Platzack, Christer 1998. *Svenskans inre grammatik – det minimalistiska programmet. En introduktion till modern generativ grammatik* [*The internal grammar of Swedish – The minimalist program*]. Lund: Studentlitteratur.

Platzack, Christer 1999. Satsfläta med relativsats [Extraction from relative clauses]. In: Haskå, Inger & Carin Sandqvist (eds), *Alla tiders språk. En vänskrift till Gertrud Pettersson november 1999* [*A festschrift for Gertrud Pettersson*]. (Lundastudier i nordisk språkvetenskap A 55.) Lund: Department of Scandinavian Languages, Lund University, pp. 189–199.

Platzack, Christer 2000. A Complement-of-N^0 Account of Restrictive and Non-Restrictive Relatives. The case of Swedish. In: Alexiadou, Artemis, Paul Law, André Meinunger & Chris Wilder (eds), *The Syntax of Relative Clauses*. (Linguistik Aktuell 32.) Philadelphia, PA: Benjamins, pp. 265–308.

Platzack, Christer 2011. *Den fantastiska grammatiken. En minimalistisk beskrivning av svenskan* [*The fantastic grammar – A minimalist description of Swedish*]. Stockholm: Norstedts.

Platzack, Christer 2012. Split Topicalization in Swedish and doubling with *det* 'it.NEUTER'. In: Brandtler, Johan, David Håkansson, Stefan Huber & Eva Klingvall (eds), *Discourse and Grammar. A Festschrift in Honor of Valéria Molnár*. Lund: Centre for Languages and Literature, Lund University, pp. 411–431.

Platzack, Christer 2013. Spurious topic drop in Swedish. In: Lohndal, Terje (ed.), *In Search of Universal Grammar. From Old Norse to Zoque*. (Linguistik Aktuell 202.) Amsterdam: John Benjamins, pp. 27–50.

Platzack, Christer 2014. A Feature Driven Account of Mainland Scandinavian Extraction from Relative Clauses. Talk given at the Grammar seminar at the Centre for Languages and literature, Lund University, April 3 2014.

Poletto, Cecilia 2002. The left-periphery of V2-Rhaetoromance dialects: a new view on V2 and V3. In: Barbiers, Sjef, Leonie Cornips & Susanne van der Kleij (eds), *Syntactic Microvariation*. (Meertens Institute Electronic Publications in Linguistics 2.) Amsterdam: Meertens Institute, pp. 214–242.

Pollard, Carl Jesse & Ivan A. Sag 1994. *Head-Driven Phrase Structure Grammar*. (Studies in Contemporary Linguistics 3.) University of Chicago Press.

Pollock, Jean-Yves 1989. Verb Movement, Universal Grammar, and the Structure of IP. *Linguistic Inquiry* 20:3, pp. 365–424.

Postal, Paul 1971. *Cross-Over Phenomena*. New York: Holt, Rinehart and Winston.

Reinhart, Tanya 1981. Pragmatics and linguistics: An analysis of sentence topics. *Philosophica* 27:1, pp. 53–94.

Reinhart, Tanya & Eric Reuland 1993. Reflexivity. *Linguistic Inquiry* 24:4, pp. 657–720.

Reuland, Eric J. & Alice G. B. ter Meulen (eds) 1987. *The Representation of (In)definiteness*. (Current Studies in Linguistics 14.) Cambridge, MA: MIT Press.

Rezac, Milan 2006. The Interaction of Th/Ex and Locative Inversion. *Linguistic Inquiry* 37:4, pp. 685–697.

Richards, Norvin 1998. The Principle of Minimal Compliance. *Linguistic Inquiry* 29:4, pp. 599–629.

Rizzi, Luigi 1990. *Relativized Minimality*. (Linguistic Inquiry Monographs 16.) Cambridge, MA: The MIT Press.

Rizzi, Luigi 1997. The Fine Structure of the Left Periphery. In: Haegeman, Liliane (ed.), *Elements of Grammar. Handbook in Generative Syntax*. (Kluwer International Handbooks of Linguistics 1.) Dordrecht: Kluwer Academic, pp. 281–337.

Rooth, Mats 1992. A Theory of Focus Interpretation. *Natural Language Semantics* 1:1, pp. 75–116.

Ross, John R. 1967. *Constraints on Variables in Syntax*. Ph.D. thesis, Massachusetts Institute of Technology, Cambridge, MA.

Ross, John R. 1984. Inner Islands. In: Brugman, Claudia & Monica Macaulay (eds), *Proceedings of the Tenth Annual Meeting of the Berkeley Linguistics Society*. Berkeley, CA: Berkeley Linguistics Society, pp. 258–265.

Ruys, E. G. 2015. A Minimalist Condition on Semantic Reconstruction. *Linguistic Inquiry* 46:3, pp. 453–488.

Schachter, Paul 1973. Focus and Relativization. *Language* 49:1, pp. 19–46.

Sichel, Ivy to appear. Anatomy of a Counterexample: Extraction from relative clauses. To appear in *Linguistic Inquiry*.

Søfteland, Åshild 2013. *Utbrytningskonstruksjonen i norsk spontantale* [*The cleft construction in spoken Norwegian*]. Oslo: Department of Linguistics and Scandinavian Studies, University of Oslo.

Sprouse, Jon 2007. Continuous Acceptability, Categorical Grammaticality, and Experimental Syntax. *Biolinguistics* 1, pp. 123–134.

Sprouse, Jon & Norbert Hornstein (eds) 2013. *Experimental Syntax and Island Effects*. Cambridge: Cambridge University Press.

Sprouse, Jon, Matt Wagers & Colin Phillips 2012. A test of the relation between working-memory capacity and syntactic island effects. *Language* 88:1, pp. 82–123.

Stalnaker, Robert 1978. Assertion. In: Cole, Peter (ed.), *Syntax and Semantics 9. Pragmatics*. New York: Academic Press, pp. 315–332.

Stroh-Wollin, Ulla 2002. *Som-satser med och utan* som [*Som-clauses with and without* som]. (Skrifter utgivna av institutionen för nordiska språk vid Uppsala universitet 58.) Uppsala: Department of Scandinavian Languages, Uppsala University.

Sundman, Marketta 1980. *Existentialkonstruktionen i svenskan* [*The existential construction in Swedish*]. (Meddelanden från stiftelsens för Åbo akademi forskningsinstitut 57.) Åbo: The Research Institute of the Åbo Akademi Foundation.

Svenonius, Peter 2002. Subject Positions and the Placement of Adverbials. In: Svenonius, Peter (ed.), *Subjects, Expletives, and the EPP*. (Oxford Studies in Comparative Syntax.) Oxford: Oxford University Press, pp. 201–242.

Szabolcsi, Anna 2006. Strong vs. Weak Islands. In: Everaert, Martin & Henk van Riemsdijk (eds), *The Blackwell Companion to Syntax IV*. (Blackwell Handbooks in Linguistics 19.) Malden, MA: Blackwell, pp. 479–531.

Szabolcsi, Anna & Frans Zwarts 1993. Weak Islands and an Algebraic Semantics for Scope Taking. *Natural Language Semantics* 1:3, pp. 235–284.

Teleman, Ulf 1991. Om normföreställningars födelse [On the birth of the conceptions of norms]. In: Malmgren, Sven-Göran & Bo Ralph (eds), *Studier i svensk språkhistoria 2* [*Studies in the history of Swedish*]. (Nordistica Gothoburgensia 14.) Gothenburg: Acta Universitatis Gothoburgensis, pp. 214–229.

Teleman, Ulf, Staffan Hellberg & Erik Andersson 1999. *Svenska Akademiens grammatik* [*Swedish Academy grammar*]. Stockholm: Norstedts.

Thráinsson, Höskuldur 2007. *The Syntax of Icelandic*. (Cambridge Syntax Guides.) Cambridge: Cambridge University Press.

Travis, Lisa deMena 1991. Parameters of Phrase Structure and Verb-Second Phenomena. In: Freidin, Robert (ed.), *Principles and Parameters in Comparative Grammar*. (Current Studies in Linguistics 20.) Cambridge, MA: The MIT Press, pp. 339–364.

Tutunjian, Damon, Fredrik Heinat, Eva Klingvall & Anna-Lena Wiklund 2015. An eye-tracking study of Swedish filler-gap dependencies: Processing relative clause extractions. Poster presented at Architectures and Mechanisms for Language Processing, University of Malta, Malta.

Tutunjian, Damon, Fredrik Heinat, Eva Klingvall & Anna-Lena Wiklund under review. Processing relative clause extractions in Swedish. Ms. under review. Lund University and Linnaeus University.

Vallduví, Enric & Elisabet Engdahl 1996. The linguistic realization of information packaging. *Linguistics* 34:3, pp. 459–519.

Van Valin, Robert D. 1994. Extraction Restrictions, Competing Theories and the Argument from the Poverty of the Stimulus. In: Lima, Susan D., Roberta L. Corrigan & Gregory K. Iverson (eds), *The Reality of Linguistic Rules*. (Studies in Language Companion Series 26.) Philadelphia: Benjamins, pp. 243–259.

Van Valin, Robert D. 1996. Toward a functionalist account of so-called extraction constraints. In: Devrient, Betty, Louis Goossens & Johan van der Auwera (eds), *Complex Structures. A Functionalist perspective*, *Functional Grammar Series*, vol. 17. Berlin/New York: De Gruyter, pp. 29–60.

Van Valin, Robert D. & Randy J. LaPolla 1997. *Syntax: Structure, meaning and function*. (Cambridge Textbooks in Linguistics.) Cambridge: Cambridge University Press.

Verhagen, Arie 2009. The conception of constructions as complex signs: Emergence of structure and reduction to usage. *Constructions and Frames* 1:1, pp. 119–152.

Vikner, Sten 1995. *Verb Movement and Expletive Subjects in the Germanic Languages*. (Oxford Studies in Comparative Syntax.) New York: Oxford University Press.

Vikner, Sten 2017. CP-recursion and the derivation of verb second in Germanic main and embedded clauses. In: Freitag, Constantin, Oliver Bott & Fabian Schlotterbeck (eds), *Two perspectives on V2. The invited talks of the DGfS 2016 workshop "V2 in grammar and processing: Its causes and its consequences"*. Konstanz: Universität Konstanz, pp. 1–26.

Wellander, Erik 1939. *Riktig svenska: En handledning i svenska språkets vård* [*Real Swedish*]. Stockholm: Norstedt & Söner.

Wexler, Kenneth & Peter W. Culicover 1980. *Formal Principles of Language Acquisition*. Cambridge, MA: MIT Press.

Wiklund, Anna-Lena, Kristine Bentzen, Gunnar Hrafn Hrafnbjargarson & Þorbjörg Hróarsdóttir 2009. On the distribution and illocution of V2 in Scandinavian *that*-clauses. *Lingua* 119:12, pp. 1914–1938.

Williams, Edwin 1984. *There*-insertion. *Linguistic Inquiry* 15:1, pp. 131–153.

Zaenen, Annie, Elisabet Engdahl & Joan M. Maling 1981. Resumptive Pronouns Can Be Syntactically Bound. *Linguistic Inquiry* 12:4, pp. 679–682.

Appendix A
Search strings

Avsky (‘detest’)

[] [(word = "avskyr" | word = "avskydde")] [pos = "PN"] []{1,4} [word = "som"]

Beundra (‘admire’)

[] [(word = "beundrar" | word = "beundrade")] [pos = "PN"] []{1,4} [word = "som"]

Behöva (‘need’)

[ ] [(word = "behöver" | word = "behövde")] [pos = "PN"] []{1,4} [word = "som"]

Finnas (‘exist’)

[] [(word = "finns" | word = "fanns")] [word = "det"] []{1,3} [word = "som"]

This search string has *det* ‘it’ following the verb instead of an unspecified pronoun, since expletive *det* is the only possible subject in an existential construction in the standard dialect of written Swedish. Since this search generates many hits, the number of words representing the head of the relative clause was reduced to 1–3.

Hitta (‘find’)

[][(word = "hittar" | word = "hittade")] [pos = "PN"] []{1,4} [word = "som"]

An additional search string was used in order to find any examples where *hitta* ‘find’ was in perfect tense.

[] [(word = "har" | word = "hade")] [pos = "PN"] [pos = "AB"]{0,1} [word = "hittat"] []{1,4} [word = "som"]

Höra (‘hear’)

[] [(word = "hör" | word = "hörde")] [pos = "PN"] []{1,4} [word = "som"]

Känna (‘know.REL’)

[] [(word = "känner" | word = "kände")] [word ! = "oss" & word ! = "dig" & word ! = "mig" & word ! = "sig" & pos = "PN"] [word ! = "oss" & word ! = "dig" & word ! = "mig" & word ! = "sig"] {1,4} [word = "som"]

To avoid reflexive uses of *känna*, I excluded the pronoun forms *mig*, *sig*, *dig*, and *oss* from the position following the verb.

Träffa (‘meet’)

[] [(word = "träffar" | word = "träffade")] [pos = "PN"] []{1,4} [word = "som"]

An additional search string was used in order to find examples where *träffa* ‘meet’ was in the perfect tense.

[] [(word = "har" | word = "hade")] [pos = "PN"] [pos = "AB"]{0,1} [word = "träffat"] []{1,4} [word = "som"]

Se (‘see’)

[] [(word = "ser" | word = "såg")] [pos = "PN"] []{1,4} [word = "som"]

Veta (‘know.COG’)

[ ] [(word = ”vet” | word = ”visste”)] [pos = ”PN”] [pos != ”HP” & pos != ”HS” & pos != ”HA” & pos != ”HD”]{1,4} [word = ”som”]

To exclude examples where *veta* has an embedded question complement, interrogative pronouns (HP), interrogative possessives (HS), interrogative adverbs (HA) and interrogative determiners (HD) were excluded in the 1–4 unspecified positions.

Appendix B
Questionnaire

The questionnaire items are given below in the order that they were presented on the screen to the participants. A short rendition in English is given for each test sentence. In the leftmost column, a reference consisting of chapter and example number is provided for the examples which are discussed in the dissertation. In a few cases, the total number of answers is less that 16. This is because the test item was skipped by some participant.

Ex	Test item	Natural	Somewhat strange	Unnatural
	Practice items			
	Erik frågade mig hur mycket klockan var **men det visste jag inte**. *but I didn't know*	11	4	1
	Jag var hos min bror igår. Han bor nära mig, **så jag ofta träffar honom**. *so I often meet him*	0	4	12
	Anna: Finns det någon cykelverkstad här i närheten? **Sven: Det finns varje affär som lagar cyklar**. *there is every store that repairs bikes*	0	4	12
	Lisa tror att Anna gillar katter, **men jag vet det inte**. *but I don't know it*	2	6	8

Ex	Test item	Natural	Somewhat strange	Unnatural
	Questionnaire items			
	1. IPCC säger att det är 95 procent säkert att det är vi människor som har orsakat merparten av klimatförändringarna sen 1950. **Tydligare än så går det inte att hitta forskaren som vill uttala sig**. *clearar than that it is not possible to find the researcher that wants to express themselves*	1	2	13
	2. Erik: Vet du vad? Två av mina kompisar ska åka till Abisko för att gifta sig. Det är visst vanligt nuförtiden. **Elin: Jasså, av just den anledningen känner jag ingen som har åkt till Abisko.** *oh yeah? for precisely that reason I don't know anyone that has gone to Abisko*	7	7	2
	3. Alla på mitt jobb klagar på att kaffemaskinen inte fungerar, **men jag vet ingen som har försökt göra ett dugg åt det**. *but I don't know of anyone that has tried do anything about it*	9	6	0
4.30	4. Jag vet folk som tar med sig en kompis på första dejten, **men deras föräldrar känner jag inte många som skulle vilja ha med**. *but their.*PN *parents I don't know many people that would want to have there*	0	4	12
6.30a	5. Anna: Mina barnbarn sjöng så fint att alla i kyrkan grät. **Sven: Så fint känner jag ingen som kan sjunga.** *that nicely I don't know anyone that can sing*	1	6	9
	6. Strindberg var en kontroversiell författare. Han ville chockera sin samtid med provokativa teman. **Av den anledningen tror jag att han skrev Fröken Julie.** *for that reason I think that he wrote Miss Julie*	10	6	0
	7. Bengt: Vem kom sist till festen? Lisa kom inte förrän klockan nio. **Erik: Så sent vet jag också några som kom.** *that late I too know of some people that came*	8	7	1

Ex	Test item	Natural	Somewhat strange	Unnatural
	8. Olle: Jag var på en marknad igår där de hade en mjölkdrickartävling. Några av tävlingsdeltagarna drack mer än fem liter mjölk! **Ida: Så mycket mjölk har jag aldrig träffat den som kan dricka på en gång!** *that much milk I have never met the one that can drink at once*	0	4	12
6.29a	9. Linda: Min syster och jag är oense om när vi ska äta ikväll. Hon tycker att vi ska äta klockan nio, men jag tycker att det är alldeles för sent. **Ida: Ja, så sent vet jag ingen som brukar äta middag.** *yes, that late I don't know of anyone that eats dinner*	11	5	0
	10. Felicia: Kan du skicka ordboken? **Julia: Nu igen! Va i helvete vet du fortfarande inte hur det stavas?** *now again, what the hell do you still not know how to spell?*	5	8	3
	11. Jag har letat efter forskning om varför det finns fler högerhänta än vänsterhänta. En professor i Umeå menar att det kan ha att göra med utvecklingen under fosterstadiet, **men med säkerhet har jag inte hittat professorn som vet**. *but with certainty I have not found the professor that knows*	0	4	12
6.29b	12. Linda: Min syster och jag är oense om när vi ska äta ikväll. Hon tycker att vi ska äta klockan nio, men jag tycker att det är alldeles för sent. **Ida: Ja, jag vet ingen som brukar äta middag så sent.** *yes, I don't know of anyone that eats dinner that late*	12	2	1
	13. Bengt: Vem kom sist till festen? Lisa kom inte förrän klockan nio. **Erik: Så sent var det ingen annan som kom.** *that late there was no one else that came*	13	1	1

Ex	Test item	Natural	Somewhat strange	Unnatural
	14. Hundägare anger flera fördelar med att ha hund. Hundar kan hålla en sällskap, och att ha hund leder till en mer aktiv livsstil för många. De kan också tränas att hämta tidningen på morgonen, **men av den anledningen vet jag ingen som har skaffat hund**. *and for that reason I don't know of anyone that has gotten a dog*	7	8	1
4.29	15. Jag vet folk som tar med sig en kompis på första dejten, **men sina föräldrar känner jag inte många som skulle vilja ha med**. *but their.*REFLX *parents I don't know many people that would want have there*	8	7	1
6.32	16. IPCC säger att det är 95 procent säkert att det är vi människor som har orsakat merparten av klimatförändringarna sen 1950. **Tydligare än så går det inte att hitta någon forskare som vill uttala sig.** *clearar than that it is not possible to find any researcher that wants to express themselves*	11	4	1
	17. Strindberg var en kontroversiell författare. Han ville chockera sin samtid med provokativa teman. **Det är därför jag tror att han skrev Fröken Julie.** *it is therefore I think that he wrote Miss Julie*	8	6	2
6.34	18. Ines: Min mamma åker till jobbet klockan sex varje morgon. **Erik: Så tidigt beundrar jag verkligen folk som orkar gå upp!** *that early I really admire people that manage to get up*	3	11	2
	19. Nuförtiden kommer många sent till fester. Igår kom inte Lisa förrän klockan elva, **men varför var det ingen som visste**. *why there was no one that knew*	10	6	0
	20. Linda: Min syster och jag är oense om när vi ska äta ikväll. Hon tycker att vi ska äta klockan nio, men jag tycker att det är alldeles för sent. **Ida: Ja, så sent vet jag familjen som brukar äta middag.** *yes that late I know of the family that eats dinner*	0	5	11

Ex	Test item	Natural	Somewhat strange	Unnatural
	21. Strindberg var en kontroversiell författare. Han ville chockera sin samtid med provokativa teman. **Därför tror jag att han skrev Fröken Julie.** *therefore I think that he wrote Miss Julie*	12	4	0
6.30b	22. Anna: Mina barnbarn sjöng så fint att alla i kyrkan grät. **Sven: Jag känner ingen som kan sjunga så fint!** *I don't know anyone that can sing that nicely*	14	1	1
6.31b	23. Olle: Jag var på en marknad igår där de hade en mjölkdrickartävling. Några av tävlingsdeltagarna drack mer än fem liter mjölk! **Ida: Jag har aldrig träffat någon som kan dricka så mycket mjölk på en gång!** *I have never met anyone that can drink that much milk at once*	15	1	0
	24. Jag har letat efter forskning om varför det finns fler högerhänta än vänsterhänta. En professor i Umeå menar att det kan ha att göra med utvecklingen under fosterstadiet, **men på vilket sätt är det ingen som vet.** *but in which way there is no one that knows*	13	2	1
	25. Hundägare anger flera fördelar med att ha hund. Hundar kan hålla en sällskap, och att ha hund leder till en mer aktiv livsstil för många. De kan också tränas att hämta tidningen på morgonen, **och därför vet jag många som skaffat hund.** *and therefore I know of many people that have gotten a dog*	4	9	3
	26. Erik: Vet du vad? Två av mina kompisar ska åka till Abisko för att gifta sig. Det är visst vanligt nuförtiden. **Elin: Jasså, därför känner jag ingen som har åkt till Abisko.** *oh yeah? therefore I don't know anyone that has gone to Abisko*	1	6	9
	27. Anna: Mina barnbarn sjöng så fint att alla i kyrkan grät. **Sven: Så fint känner jag flickan som kan sjunga!** *that nicely I don't know the girl that can sing*	0	7	9

Ex	Test item	Natural	Somewhat strange	Unnatural
	28. Felicia: Kan du skicka ordboken? **Julia: Nu igen! Va i helvete ska du kolla upp den här gången?** *now again! what the hell are you going to look up this time?*	16	0	0
6.31a	29. Olle: Jag var på en marknad igår där de hade en mjölkdrickartävling. Några av tävlingsdeltagarna drack mer än fem liter mjölk! **Ida: Så mycket mjölk har jag aldrig träffat någon som kan dricka på en gång!** *that much milk I have never met anyone that can drink at one time*	9	5	2
	30. Jag har letat efter forskning om varför det finns fler högerhänta än vänsterhänta. En professor i Umeå menar att det kan ha att göra med utvecklingen under fosterstadiet, **men med säkerhet har jag inte hittat någon som vet.** *but with certainty I have not found anyone that knows*	5	7	4
6.33	31. Hundägare anger flera fördelar med att ha hund. Hundar kan hålla en sällskap, och att ha hund leder till en mer aktiv livsstil för många. De kan också tränas att hämta tidningen på morgonen, **och av precis den anledningen vet jag många som skaffat hund**. *and for precisely that reason I know of many people that have gotten a dog*	10	6	0
	32. Ines: Min mamma åker till jobbet klockan sex varje morgon. **Erik: Så tidigt beundrar jag verkligen henne som orkar gå upp!** *that early I really admire her that manages to get up*	0	6	10
	33. Jag vet folk som tar med sig en kompis på första dejten, **men jag känner inte många som skulle vilja ha med sina föräldrar**. *but I don't know many people that would want to have their.*REFLX *parents there*	14	1	0

Appendix C: Interview

I presented the following questions orally to the informants. The examples that are discussed in the dissertation are shown in their annotated version.

Section A: Answering questions out of context

Instruction:

Den här undersökningen handlar om frågor, och jag är intresserad av om de frågor som jag kommer att ställa går att svara på, och i så fall hur.

This investigation is about questions, and I am interested in whether the questions I will ask are answerable, and if so, how they can be answered.

(1) [Vilka språk]$_1$ är det många [som talar i t_1 Sverige]?
which languages are there many that speak in Sweden
'Which languages are such that there are many people who speak them in Sweden?'

(2) [Vilka språk]$_1$ känner du nån [som talar t_1]?
*which languages know.*REL *you someone that speaks*
'Which languages are such that you know someone who speaks them?'

(3) Var$_1$ är det många [som brukar jobba övertid t_1]?
where are there many that tend work overtime
'Which place is such that there are many people who tend to work overtime there?'

(4) När$_1$ vet du många [som brukar vara trötta t_1]?
*when know.*COG *you many that tend be tired*
'Which time is such that you know of many people who tend to be tired then?'

(5) Hur$_1$ vet du många [som fick sitt första jobb t_1]?
*how know.*COG *you many that got their.*REFLX.SG *first job*
'Which way is such that you know of many people who got their first job that way?'

(6) [Hur mycket]$_1$ finns det ingen [som bryr sig t_1]?
how much is there no one that cares
'How much is there no one that cares?'

(7) [Hur fint]$_1$ känner du många [som kan sjunga t_1]?
*how nicely know.*REL *you many that can sing*
'How nicely do you know many people that can sing?'

Section B: Questions and suggested answers

Nu kommer det några fler frågor. Jag är fortfarande intresserad av att veta om och hur man kan svara på frågorna. I den här delen av intervjun kommer jag att ge förslag på hur man skulle kunna svara på frågorna, och är intresserad av om du tycker att det låter som ett möjligt svar.

Here are some more questions. I am still interested in finding out how the questions can be answered. In this part of the interview, I will provide suggestions for how to answer the questions, and I am interested in whether you think they sound like possible answers.

(8) Varför tror du att Strindberg skrev *Fröken Julie*? (Engdahl 1985a:4)
why think you that Strindberg wrote Fröken Julie
'Why do you think that Strindberg wrote *Miss Julie*?'

a. Därför att ordvalet och stilen liknar Strindbergs övriga verk.
*because that choice of words.*DEF *and style.*DEF *resembles Strindberg's other works*
'Because the choice of words and the style resembles Strindberg's other works.'

b. För att chockera sin samtid.
*in order to shock his.*REFLX *contemporaries*
'In order to shock his contemporaries.'

(9) Varför känner du många [som har skrivit böcker]?
*why know.*REL *you many that have written books*
'Why do you know many people who have written books?'

a. För att jag är med i Författarförbundets styrelse.
*because I am with in Writer union.*DEF*'s board*
'Because I'm on the board of The Swedish Writer's Union.'

b. För att chockera sin samtid.
*in order to shock their.*REFLX.SG *contemporaries*
'In order to shock their contemporaries.'

(10) Av vilken anledning känner du många som har skrivit böcker?
of which reason know you many that have written books
'For what reason do you know many people who have written books?'

a. För att jag är med i Författarförbundets styrelse.
*because I am with in Writer union.*DEF*'s board*
'Because I'm on the board of The Swedish Writer's Union.'

b. För att chockera sin samtid.
*in order to shock their.*REFLX.SG *contemporaries*
'In order to shock their contemporaries.'

(11) Hur många kilo känner du många som väger?
*how many kilos know.*REL *you many that weigh?*
'How many kilos do you know many people who weigh?'

65 kg

(12) Hur många kilo känner du ingen som väger?
*how many kilos know.*REL *you no one that weighs*
'How many kilos do you know no one who weighs?'

65 kg

(13) Hur tror du att hon fick igång bilen?
*how think you that she get going car.*DEF
'How do you think she got the car started?'

Genom att byta tändstift.
through to change spark plug
'By changing the spark plug.'

(14) Hur har du aldrig träffat någon som har fått igång bilen?
*how have you never met someone that have got going car.*DEF
'How have you never met anyone that got the car going?

Genom att byta tändstift.
through to change spark plug
'By changing the spark plug.'

(15) Vad såg du någon som förgiftade?
what saw you someone that poisoned
'What did you see someone who poisoned?

a. Mitt te.
my tea
'My tea.'

b. Ett glas vin.
a glass wine
'A glass of wine.'

Scenario for (16): Three glasses of wine are standing on a table. A knows that someone has put poison in one of the glasses, but not which one. A also knows that B has seen what happened.

(16) Vilket glas vin såg du någon som förgiftade?
which glass wine saw you someone that poisoned?
'Which glass of wine did you see someone who poisoned?'

Det där. (pekar)
the there
'That one.' (pointing)

Scenario for (17): Anna and Sven have organized a party and bought 10 bottles of wine for the occasion. The following day, Anna wonders how much of the wine is left.

(17) [Hur mycket vin] var det ingen [som drack t_1]?
how much wine was there no one that drank
'How much of the wine did no one drink?'

Tre flaskor.
three bottles
'Three bottles.'

Scenario for (18) and (19): At an editors' meeting the editors are discussing which books to publish. One of the editors was not present, and asks the secretary what happened at the meeting.

(18) Hur många böcker var det flera av redaktörerna som ville ge ut igen?
*how many books was there many of editors.*DEF *that wanted give out again*
'How many book did many of the editors want to publish again?

a. Två. (Does it have to be the same two books?)
two
'Two.'

b. Två: Krig och fred och Anna Karenina.
two war and peace and Anna Karenina
'Two: War and Peace and Anna Karenina.'

(19) Hur många av böckerna på listan var det flera av redaktörerna som
*How many of books.*DEF *on list.*DEF *was there many of editors.*DEF *that*
ville ge ut igen?
wanted give out again
'How many of the books on the list did many of the editors want to publish again?'

a. Två. (Does it have to be the same two books?)
two
'Two.'

b. Två: Krig och fred och Anna Karenina.
two war and peace and Anna Karenina
'Two: War and Peace and Anna Karenina.'

Optionally:

Scenario for (20). Imagine a university system like the British, with a 0–100 grading scale, but most professors only give out marks between 40 and 75.

(20) Hur många poäng var det ingen examinator som använde?
how many points was there no examiner that used?
'How many points did no examiner use?'

a. 2/3 av poängen.
*2/3 of points.*DEF
'2/3 of the points.'

b. 65 poäng.
65 points
'65 points.'

Section C: Functional interpretations

Hur kan man svara på frågorna?

How can the questions be answered?

(21) [Vilka åsikter]$_1$ finns det många studenter [som tar avstånd från t_1]?
which opinions are there many students that take distance from
'Which opinions are there many students who condemn?'

a. Högerradikala åsikter.
right radical opinions
'Extreme right-wing views.'

b. Dom som deras föräldrar står för.
those which their parents stand for
'The ones that their parents hold'.

Om (21) är dålig:
If (21) is bad:

(22) [Vilka sorts åsikter]$_1$ finns det många studenter [som tar avstånd från t_1]?
which sort opinions are there many studenter that take distance from
'Which kind of opinions are there many students who condemn?'

a. Högerradikala åsikter.
right radical opinions
'Extreme right-wing views.'

b. Dom som deras föräldrar står för.
those which their parents stand for
'The ones that their parents hold.'

(23) [Vilken bok]$_1$ var det ingen student [som ville läsa t_1]?
which book was there no student that wanted read
'Which book was there no one who wanted to read?'

a. Krig och fred
war and peace
'War and Peace'

b. Sin mammas bok.
their.REFLX.SG *mother's book*
'Their mother's book.'

Om (23) är bra:
If (23) is good:

(24) Vilken bok vet du ingen som vill läsa?
which book know.COG *you no one that wants read*
'Which book don't you know anyone who wants to read?'

a. Krig och fred.
war and peace
'War and Peace.'

b. Sin mammas bok.
their.REFLX.SG *mother's book*
'Their mother's book.'

(25) Vilken film känner du inte nån som kan glömma?
which film know you not someone that can forget
'Which film don't you know anyone who can forget?'

a. Titanic.

b. Den första skräckfilmen hen såg.
*the first horror movie.*DEF *they.*SG *saw*
'The first horror movie they saw.'

The question in (26) was asked with both *veta* 'know.COG' and with *känna* 'know.REL'.

(26) Hurdan vet/känner du ingen som vill bli? (cf. Engdahl 1986)
*how know.*COG/*know.*REL *you no one that wants become*
'How don't you know anyone who wants to become?'

a. Gammal och trött.
old and tired
'Old and tired.'

b. Alltför lik sina föräldrar.
*too like their.*REFLX.PL *parents.*
'Too much like their parents.'

(27) Hurdana känner du många som vill bli?
*how.*PL *know.*REL *you many that wants become*
'How do you know many people who wants to become?'

a. Mer miljömedvetna.
*more environmentally conscious.*PL
'More environmentally conscious.'

b. Mer som sina förebilder.
*more like their.*REFLX.PL *role models*
'More like their role models.'