Extractions from Relative Clauses in Swedish as Self-answering Questions[1]

STEFAN HUBER, DEPARTMENT OF GERMAN, LUND UNIVERSITY

1 Introduction

Since the earliest days of generative grammar, it has been pointed out that extractions out of relative clauses (henceforth: RCs) are banned, hence the ungrammaticality of the English examples in (1):

(1) a. * Who$_i$ did you read [[$_{DP}$ an article [$_{CP}$ which is about e$_i$]]?
b. * This man$_i$, I read [[$_{DP}$ an article [$_{CP}$ which was about e$_i$]].

This restriction on RCs has been regarded as an instantiation of the so called Complex NP Constraint (henceforth: CNPC) which was originally introduced by Ross (1967). According to this constraint, no internal material must be extracted out of a DP[2]. By DP internal material, one commonly assumes attributes or subparts of attributes. DPs, therefore, are islands for movement. Regarding RCs as true postnominal attributes, the CNPC especially applies on them.

Even later theoretical stages accepted or reformulated the CNPC: Chomsky (1981) assumes the CNPC to follow from the more general Subjacency Condition. Since the fronted item in (1) has to move across both a CP and a DP, subjacency can no longer be maintained and the sentences are doomed to be ungrammatical. According to Chomsky (1986), DPs are barriers for movement of internal material.

However, it has been noted that some allow extraction out of at least a few RCs. These languages include the Mainland Scandinavian languages (see e.g. Taraldsen 1982 for Norwegian, Engdahl 1997 for Swedish, Platzack 1999 for all Mainland Scandinavian languages). In Swedish, for instance, it is possible to front an item that seems to correspond to a gap within the RC, cf. example (2):

(2) De blommorna känner jag [[$_{DP}$ en man [$_{CP}$ som säljer e].
these flowers-DEF know I a man REL sells[3]
I know a man who sells these flowers.

(= Engdahl 1997:52, ex. 2)

Apparently, the fronted item *de blommorna* corresponds semantically to the object gap within the RC *som säljer*. The most intuitive analysis would be to follow Taraldsen (1982), Engdahl

[1] Parts of this present paper have been presented at the 15th "Grammatik i Fokus" conference in Lund, in February 2001. I am especially grateful to Giuliana Giusti, Valéria Molnár and Christer Platzack for a stimulating discussion. Moreover, I would like to thank all the Swedish informants for filling in my questionnaire and answering my never-ending questions. Their intuitions are the basis of this work. All errors are of course mine.

[2] Ross (1967) put his constraint at a time when the functional D-layer of noun phrases was not yet established. Here, I will stick to the label DP since it is commonly assumed that only DPs can function as arguments at a sentence level (cf. Szabolcsi 1987, Abney 1987, Delsing 1993 among others).

[3] In the glossary, DEF symbolizes the suffixed definite article, REL the relative particle in Swedish, PRT other particles.

(1997) and Platzack (1999), among others, and describe the connection between the fronted item and the gap as antecedent-trace-dependency. However, this line of analysis does not explain why there is no violation of the CNPC in examples like (2), i.e. there should be an explanation of why the CNPC can be circumvented in these cases[4].

In this paper I will reject a movement analysis, at least for the case of Swedish. The notion "extractions from relative clauses" (ERCs) will nevertheless be kept as a pure descriptive category for this particular kind of syntactic construction. Instead of a movement analysis, I would like to suggest that ERCs are complex sentences consisting of a phonetically reduced question and an adjacent answer part. This proposal will be corroborated by the observation of some semantic, prosodic and information-structural properties observed in ERCs.

The paper is organized as follows: in chapter 2, I will present the main ERC types and discuss some basic semantic and syntactic features of ERCs. Chapter 3 contains a discussion of earlier syntactic analyses of ERCs. Here, I will particularly argue against a movement analysis. in chapter 4, I will present a new syntactic approach to ERCs that not only accounts for the syntactic facts but also takes the particular semantics of ERCs into consideration. I argue that these sentences consist both of an interrogative and a declarative sentence, i.e. they contain a yes/no-question base-generated in SpecC of the sentence and a corresponding answer part, starting out from C'. One advantage of such an approach is that the connectivity effects observed in ERCs can be easily accounted for. Secondly, it allows us to explain why languages like Swedish allow for ERCs whereas German (which shares the V2-property with Swedish) and English (which shares the VO-property with Swedish) do not. In chapter 5, I will summarize the results.

The investigation will be restricted to Swedish.

2 A descriptive analysis of ERCs in Swedish

The ERC phenomenon in Swedish appears to be rather uninvestigated. It is therefore appropriate to give a brief introduction to its empirical characteristics. In this chapter, I will present three different syntactic types of ERCs (2.1.). As it will turn out, there is a common semantic feature to all those types (2.2.). Finally, I want to discuss some further empirical aspects of and restrictions on ERCs (2.3.).

2.1 Three types of ERCs

Engdahl (1997) presents an overview over four different ERC constructions in standard and dialectal Swedish, cf.:

A. Presentational constructions (containing a "pleonastic element", i.e. an expletive element (*det* or *där*) and an existential verb (*vara* or *finnas*)):

(3)

Det	språket	finns	det	många	som	talar.
this	language-DEF	exist	there	many	REL	speak.

There are many people who speak this language.

(= Engdahl 1997:59, ex. 22a)

[4] Taraldsen (1982) and Platzack (1999) offer such structural explanations as I will show in 3.2. However, they fail to generalize their assumptions.

B. Contact clauses (without a relative item or particle):

(4) De e där ingen ved. (Scanian)
this is there nobody knows
There is nobody who knows that.

(= Engdahl 1997:61, ex. 25c)

C. Cleft constructions:

(5) Garagedörren är det bara Kalle som kan öppna.
garage-door-DEF is it only Kalle REL can open
The only one who can open the garage door is Kalle.

(= Engdahl 1997:65, ex. 22a)

D. Constructions with other thematic verbs:

(6) Den här teorin känner jag ingen som tror på.
this here theory-DEF know I nobody REL believes in
I don't know anybody who believes in this theory.

(= Engdahl 1997:67, ex. 41b)

As for Standard Swedish, I would like to reorganize those four groups and reduce them into three. The rather colloquial or dialectal class of contact clauses, I will not address here. Even those, I suppose, contain an RC and can therefore semantically and syntactically be subsumed under the three other types. For instance, sentence (4), must also be a "presentational construction" in Engdahl's (1997) sense, since it contains an expletive and an existential verb.

The term "presentational construction", however, is somewhat misleading. The aim of these sentences is not to introduce a new entity into the ongoing discourse but rather to determine a set of individual elements quantitatively. In (3), for instance, the speaker describes the number of people speaking a specific language by "many". The matrix clause of such an ERC construction always contains an existential verb and a quantified RC head. Therefore, I would like to label this group existential ERCs. As for the RC head, many different quantifiers are allowed:

(7) a. Det språket finns det många/bara en/en hel kontinent/ingen
this language-DEF exist there many just one a whole continent nobody
som talar.
REL speak.
There are many/is just one/is a whole continent/is nobody who speak(s) this language.

b. Denna olycka var det bara en som överlevde.
this accident was there only one REL survived
There was only one who survived this accident.

The verb *vara* (be) can be used in the existential sense like in (7). In other contexts, *vara* can also be a specificational copula[5]. In this reading, it is used in cleft constructions and related copular constructions. Clefts can also allow for the ERC phenomenon (cf. ex. 5). This specificational type forms the second group of ERCs.

[5] On the difference between predicational, identificational and specificational copulas see Higgins (1973) and Huber (forthcoming).

A last group contains other thematic verbs as already shown in (6). Further examples are given in (8):

(8) a.

Det	har	jag	sett	många	som	gjorde.
this	have	I	seen	many	REL	did.

I have seen many who did this.

b.

Det	vet	jag	ingen	som	kan.
this	know	I	nobody	REL	can

I don't know anybody who can (do) this.

Usually, even in this thematic type, we find a quantified RC head. The range of the quantifying elements probably is as extensive as with the "existential" type.

2.2 A remark on the semantics of ERCs

What do all these types of ERCs have in common? It seems that we find two semantic strategies within ERCs: the first strategy is that a set is determined quantitatively. This is the case in existential and thematic ERCs. For instance, in an existential sentence like (3), the speaker declares that the number of the set of people speaking a particular language is large. Likewise, in the thematic type like the one in (6), the speaker describes a set of individuals believing in a particular theory by uttering his/her relationship to the amount of this set - s/he does not know any individual of this set. The element that has been "extracted" from the RC in these sentences is a part of the set that the speaker has commented on. Thus, quantification seems to be a necessary condition for "extraction" from RCs.

The second strategy which one can discover in specificational (cleft) sentences does not determine the quantity of a set. As I will point out in Huber (forthcoming), the characteristics of a specification rather extends to the referential determination of a set. Consider e.g. the cleft sentence in (3), repeated here for the sake of convenience:

(3)

Garagedörren	är	det	bara	Kalle	som	kan	öppna.
garage-door-DEF	is	it	only	Kalle	REL	can	open

The only one who can open the garage door is Kalle.

A paraphrase of (3) would be *The only one who can open the garage door is Kalle* (cf. the translation). By saying so, the speaker presupposes that there is a person who is able to open the garage door (cf. also Delin 1992). This person is specified as *Kalle*. Furthermore, *Kalle* is supposed to be the only one who is able to open the garage door. The set of people who can open the garage door is, thus, restricted to one, i.e. it is a closed set. As the cleft constituent (*Kalle*) determines the presupposed set, one might say that also specificational (cleft) constructions semantically operate on a set. The difference to the existential and the thematic type, respectively, is only manifested by the kind of the determination of the set (referential vs. quantificational).

Therefore I find it plausible to assume that the predication within an ERC concerns the determination of a set: whereas in the existential and in the thematic cases, the strategy is a quantificational one, in the specificational type, it is a referential one.

2.3 Further empirical properties of ERCs

A typical ERC construction contains a fronted phrase that corresponds to a gap in the RC further down in the sentence. Note that the RC therefore contains two gaps, one corresponding to the head and the other corresponding to the fronted item, cf. the internal structure of ex. (9):

(9) De blommorna$_2$ känner jag [[$_{DP}$ en man$_1$ [$_{CP}$ som **e$_1$** säljer **e$_2$**].
these flowers-DEF know I a man REL sells
I know a man who sells these flowers.

Keep in mind that, when talking about the "gap" within the RC, I will not refer to the one that can be analysed as a trace of the relativized head (e_1), but to the one that semantically corresponds to the fronted item (e_2). The relevant gap e_2 is not a gap that arises from relativization.

As for the fronted item itself, it always seems to contain given material in information-structural terms, which means that it refers to a referent in the co(n)text of the speech situation. Often, deictic elements like pronouns are used in this position:

(10) [Co(n)text: De frågade var Peter var.]
they asked where Peter was
They asked where Peter was.

Men **det** var det ingen som visste.
but this was there nobody REL knew.
But there was nobody who knew this.

Since Swedish is a V2-language, the finite verb follows the fronted constituent and precedes the subject. The subject itself is also mostly given in the utterance situation. Often, we find proper names or pronouns in this position. However – as can be seen in (4) – the subject may also be non-referential. The only restriction seems to be that the subject must not consist of an indefinite DP:

(11) a. Lax känner **han** ingen som gillar.
salmon knows he nobody REL likes
He does not know anybody who likes salmon.

b. Lax känner **Kalle** ingen som gillar.
salmon knows Kalle nobody REL likes
Kalle does not know anybody who likes salmon.

c. * Lax känner **en** **man** ingen som gillar.
salmon knows a man nobody REL likes.
A man does not know anybody who likes salmon.

I have no explanation on this definiteness restriction on the matrix subject. Possibly, there are some pragmatic factors involved here, too.

The matrix predicate can either be an existential verb like *vara* (be, exist) or *finnas* (exist), a specificational copula or a thematic verb (cf. 2.1). In case of a thematic matrix verb, there is a clear preference for two place predicates. One or three place predicates are judged as more or less marked by my informants:

(12) a. <u>One place predicates</u>:

1. ?? Det kunde ingen sova som hade glömt.
   this could nobody sleep REL had forgotten
   Nobody who had forgotten this could sleep.

2. */??? Det sprang (det) ingen som hade sett.
   this ran there nobody REL had seen
   Nobody who had seen this ran.

b. <u>Three place predicates</u>:

1. ??Det skrev Peter Kalle ett brev som innehöll.
   this wrote Peter Kalle a letter REL contained
   Peter wrote a letter to Kalle which contained this.

2. ?? Det gav Kalle Peter en bok som innehöll.
   this gave Kalle Peter a book REL contained
   Kalle gave Peter a book which contained this.

I am not sure what the reason for this two place predicate restriction is. I suppose that there are different restrictions on the thematic matrix verb. As we have seen, one of those restrictions is the predication upon a set. The choice of a two place predicate, then, might be an epiphenomen of this lexical constraint. However, I will not go deeper into this question.

As for the head of the RC, we already noticed that it has to be quantified when a thematic matrix verb is involved. As a consequence of this, only indefinite RC heads – which can be construed as quantified nouns in Swedish – but not definite RC heads can occur in ERCs:

(13) Blommorna känner jag en man/*mannen som säljer.
flowers-DEF know I a man man-DEF REL sells
I know a man who sells the flowers.

Note further that the RC involved must be a restrictive one. Thus, items like *för övrigt* (by the way) that indicate appositiveness are banned:

(14) Denna teorin finns det ingen lingvist som (* för övrigt)
this theory exists there no linguist REL by the way
tror på.
believes in
There is no linguist who believes in this theory.

This fact might depend on semantic reasons: after quantified nouns like the one in (14), a non-restrictive RC is never possible (* *nobody who by the way believes in that*).

Within the RC, the most preferred predicate structure also seems to be a two place one. In case of a three place predicate the subcategorization has to be of the kind [__ DP, DP, PP] (15a). The order [__ DP, DP, DP], on the other hand, is odd (15b):

(15) a1. Matte var det ingen som hjälpte mig med e.
maths was there nobody REL helped me with

a2. Mig var det ingen som hjälpte e med matte.
me was there nobody REL helped with maths
There was nobody who helped me with maths.

b1.	??? Boken	var	det	ingen	som	gav	e	Maria.
	book-DEF	was	there	nobody	REL	gave		Maria
b2.	* Maria	var	det	ingen	som	gav	boken	e.
	Maria	was	there	nobody	REL	gave	book-DEF	

There was nobody who gave Maria the book.

Engdahl (1997) and Platzack (1999) formulate another constraint on the internal structure of the RC: whereas it is possible to front an item that corresponds to the direct (or possibly the indirect) object gap within the RC, the fronted item must not correspond to an RC internal subject gap:

(16)	* Denne	lingvist$_1$	finns	det	ingen	teori$_2$	som e_1	tror	på e_2.
	this	linguist	is	there	no	theory	REL	believes	in

There is no theory which this linguist believes in.

In (16), the relativized element in the RC is the object (e_2). The fronted item, however, corresponds to the subject gap in the RC. Regard that the "non-extracted" variant is completely fine:

(16')	Det	finns	ingen	teori	(som)	denne	lingvist	tror	på.
	there	exists	no	theory	REL	this	linguist	believes	in

There is no theory which this linguist believes in.

Whereas Platzack (1999) assumes the subject constraint to be a consequence of the structural configuration and the special requirements on restrictive RCs, I would like to propose that sentence (16) is bad because of parsing reasons: it automatically gets the reading that evokes a correspondence between the fronted constituent and the object gap. This reading, however, is out due to semantic conditions – a theory cannot believe in a linguist. The fact that sentence (16) is parsed semantically incorrectly might result from the strong subject feature in Swedish: it appears that in Swedish, the topmost subject position is located within the C-domain (cf. Platzack & Rosengren 1998). Let us call this projection FinP, for the sake of simplicity. In order to satisfy the EPP condition, FinP must be visible at PF. Hereby, it is sufficient that either the specifier or the head of FinP is spelled out. I take it that *som* can serve as a functional head base-generated in Fin° and thus being able to fulfil the EPP requirement (see also Huber forthcoming). Therefore, *som* cannot be omitted when a subject is relativized since the EPP would be violated. When a non-subject is relativized, *som* can be realized, but it cannot stay within FinP since every projection within the C-domain must not be filled doubly, i.e. either the specifier or the head might be spelled out (cf. Platzack 1998). Therefore, *som* can be omitted (cf. 16') or must be raised to a higher functional head position, which I here will call Force°. FinP, on the other hand, is spelled out by the lexical subject, thus satisfying the EPP. Now regard sentence (16): since a non-subject is relativized, *som* cannot stand in Fin°, but must stand in Force°. On the other hand, there is no lexical subject that is placed within FinP. FinP, therefore, is not visible at PF and the EPP violated. The only way to circumvent this difficulty is to parse the sentence so that the fronted item corresponds to the object gap within the RC. However, this reading is bad due to semanic reasons.

Taraldsen (1982) and Engdahl (1997) point out that the RC within an ERC construction tends to be right-peripheral. Generally speaking, I suppose this to be a correct assumption. However, it does not seem to be an absolute constraint: if an ERC contains a comparative clause like the one in (17a), both orders are possible (CP_{comp} > RC and RC >

CP_{comp}). The same can be observed in stacked RC constructions like (17b). Even here, the "extracted" RC and the additional RC can be inverted:

(17) a1. Det fanns det fler kvinnor [$_{RC}$ som påstod]
this existed there more women REL claimed
[$_{CPcomp}$ än jag trodde].
than I thought
There were more women who claimed this than I thought.

a2. (?) Det fanns det fler kvinnor [$_{CPcomp}$ än jag trodde]
this existed there more women than I thought
[$_{RC}$ som påstod].
REL claimed
There were more women than I thought who claimed this.

b1. Det var det ingen [$_{RC1}$ som visste]
this was there nobody REL knew
[$_{RC2}$ som inte hade läst medicin].
REL not had studied medicine
There was nobody who knew this who had not studied medicine.

b2. Det var det ingen [$_{RC2}$ som inte hade läst medicin]
this was there nobody REL not had studied medicine
[$_{RC1}$ som visste].
REL knew
There was nobody who had not studied medicine who knew this.

However, in most ERCs, the RC concludes the sentence.

2.4 Summary

There exist three major ERC types in Swedish, an existential, a specificational and a thematic type. They all appear to semantically predicate upon a set, either in a quantitative (existential or thematic ERCs) or a referential (specificational ERCs) way. As for the internal ERC structure, both the fronted item and the postverbal subject often contain given information. The matrix verb is a two place one, if thematic[6]. The RC itself must be restrictive. Its subject cannot correspond to the fronted item. The RC tends to be sentence-final. ERCs, therefore, can be described as a highly restricted sentence type obeying a lot of different syntactic, semantic and information-structural restrictions.

3 Earlier accounts for the ERC phenomenon

The obvious correlation between the fronted phrase and the (second) gap within the RC in ERC constructions has traditionally been described as a movement out of the RC, thereby violating the CNPC. However, Engdahl (1997) discusses an alternative line of analysis. In this chapter, I want to discuss both approaches in order to show that they cannot be correct.

[6] Note that the predicates in existential and specificational ERCs also can be described as two place predicates.

3.1 Empty pronouns?

Engdahl (1997:53f.) argues against an account of ERCs that involves empty pronouns. As advocates of such an theory, she refers to "personal communication 1979" with E. Williams and N. Chomsky. According to them, a *pro*-like element is placed in the RC of an ERC:

(18) Detta känner jag en man som vet *pro*.
this know I a man REL knows
I know a man who knows this.

As Engdahl correctly states, such an approach can hardly be maintained. Empty pronouns do exist in Swedish, for instance in the sentence initial position (i.e. in cases of so called "topic" or "subject drop"). But it is always possible to replace these empty pronouns by overt ones:

(19) [Context: Kalle, du vet.] Honom/e har jag träffat igår.
Kalle, you know him have I met yesterday.
Kalle, you know. I met him yesterday.

Within ERCs, the gap corresponding to the fronted item must never be filled by an overt pronoun, even if the distance between the two positions is long:

(20) a. Honom är det ingen som känner (* honom).
him is there nobody REL knows him
There is nobody who knows him.

b. Det trodde jag länge att det inte kunde vara
this believed I for-a-long-time that it not could be
möjligt att det fanns en enda människa
possible that there existed one single person
på hela jorden som hade bara
on whole earth-DEF REL had only
ett svagt intresse för (* det).
a slight interest for this
I believed for a long time that it could not be possible that there was one single person on the whole earth who only had a slight interest for that.

Thus, if there exists a silent pronoun within the "extracted" RC, one should be able to explain why it cannot be realized overtly. Another problem would be where the moved item originates from. Is it base-generated in the topmost position or moved there from elsewhere?

I therefore agree with Engdahl in assuming that an empty pronoun approach is not the correct way to explain the internal structure of ERCs.

3.2 Movement?

Taraldsen (1982), Engdahl (1997) and Platzack (1999) all argue for a movement account. This means that they have to explain why there is no violation of the CNPC. Taraldsen (1982), for instance, assumes the RCs in ERCs to be extraposed. In extraposition, he further argues that the CNPC does not apply any more. However, this assumption does not prove to be correct since it makes the wrong predictions: if any extraposed RC allowed for extraction, there should be no difference between a definite and an indefinite RC head:

(21) a. Detta känner jag en man som kan.
this know I a man REL can
I know a man who can (do) this.

b. * Detta känner jag mannen som kan.
this know I man-DEF REL can
I know the man who can (do) this.

The same problem arises within Platzack's (1999) system. He claims that the special syntactic configuration in restrictive RCs in Swedish paves the way for a possible extraction. He, too, cannot explain why the extraction from every restrictive RC in Swedish is not correct, as can be seen again in example (21).

However, the movement analysis is based upon some empirical facts that I would like present and discuss in the following section.

3.2.1 Arguments for a movement analysis: connectivity

The assumption that an element has been moved presupposes that the element has another basic position which is posited deeper in the sentence. Movement always leaves a trace in the original position. The moved element must end up in a position where it can bind its trace. The relevant fact, now, is that the moved item and the trace display the same lexical and morphological features (according to the so called "projection principle", cf. Chomsky 1981). Therefore, the moved item can be called "connective" with the sentence, i.e. it appears to be part if the deep structure of the sentence (cf. Higgins 1973).

As can be shown easily, ERCs display connectivity: note first that the fronted items can be realized within the RC of an ERC construction without changing the meaning of the sentence:

(22) Blommor känner jag en man som säljer.
flowers know I a man REL sells
→ Jag känner en man som säljer blommor.
I know a man REL sell flowers
I know a man who sells flowers.

A second test for the connectivity between the fronted item and the gap within the RC concerns case marking: a moved item should keep its case throughout the derivation. In Swedish, case marking can only be observed with pronouns. A fronted pronominal constituent gets the case of the gap corresponding to it:

(23) Honom/* han är det ingen som gillar.
him he is there nobody REL likes
There is nobody who likes him.

Finally, even lexical anaphora can be fronted in ERCs:

(24) Sig själv är det ingen här som gillar.
themselves is there nobody here REL likes
There is nobody who likes himself.

Since anaphora are regarded to be bound within their minimal binding domain, the fronted anaphor in (24) should behave like any other moved anaphor, i.e. it is assumed to be moved back to its original positions on LF (by virtue of reconstruction). The only position in which

sig in (24) can be bound on LF seems to be the gap within the RC[7]. Therefore, it should be its basic position.

3.2.2 Wh-movement?

Engdahl (1997:56ff.) indirectly argues for a movement analysis by giving an example that could indicate that independent (*wh*-) extractions out of RCs are possible in Swedish:

(25)

Vem	var	det	ingen	som	kände?
who	was	it	nobody	REL	knew

Whom did nobody knew?

Within the classic Minimalist Program (Chomsky 1995), movement is triggered by morphological marking. A *wh*-element must be raised to the C-domain in order to check its *wh*-feature. If a *wh*-word is merged within an RC, however, a problem arises since RCs are subject to the CNPC. Therefore, a *wh*-word should not be allowed to leave the complex DP in which it is base-generated. As example (25) shows, *wh*-movement seems to be possible. Thus, movement from RCs should be allowed *per se*. If one assumes that even syntactic topicalization in declarative ERCs resembles *wh*-like movement[8], it should be plausible that movement in ERCs indeed takes place.

Taking a closer look at interrogatives like (25), one can notice, however, that they only can serve as echo questions. A proper context for (25) would be (25'):

(25')

Speaker 1:	[Unparseable noise]	var	det	ingen	som	kände.
		was	there	nobody	REL	knew.

Speaker 2:	Vem	var	det	ingen		som	kände?
	who	was	it	nobody	REL	knew	

What did nobody knew?

On the other hand, it is not possible to use interrogative ERCs as proper questions:

(26)

[Context:	Tapeterna	var	det	Emil	som	valde,
	wallpapers-DEF	was	it	Emil	REL	chose
	möblerna	var	det	Hans	och	spisen Greta.]
	furniture-DEF	was	it	Hans	and	stove-DEF Greta

It was Emil who chose the wallpapers, Hans who chose the furniture and Great who chose the stove.

#	Och	vad	var	det	Peter	som	valde?
	and	what	was	it	Peter	REL	chose
OK	Och	vad	var	det	som	Peter	valde?
	and	what	was	it	REL	Peter	chose?

And what did Peter choose?

It is also impossible to embed interrogative ERCs:

[7] As I will show in 3.2.3 and 4.2.2, respectively, such an anaphor is bound by a silent operator in the C-domain of the RC.

[8] Note though that there is some debate upon the nature of syntactic topicalization, cf. Platzack (1996), Büring (1999) and many others.

(27) a. * Jag undrade vad det var ingen som visste.
I wondered what it was nobody REL knew

b. OK Jag undrade vad det var som ingen visste.
I wondered what it was REL nobody knew
I wondered what nobody knew.

I therefore assume that interrogative ERCs are not proper interrogatives. Being echo questions, I regard them as structural copies of a preceding sentence in which one item is replaced by a *wh*-word. Thus, they involve no proper syntactic derivation. Example (25) can therefore not be considered as an argument for movement within ERCs.

3.2.3 Arguments against a movement analysis: the CNPC

The CNPC was formulated at a time when a uniform movement analysis was not yet given. The transformation rules at the stage of the so called Standard Theory were first simplified and minimalized by Chomsky's (1981) command "move α". According to this principle, any constituent can be moved into any position as long as no other principles are violated. In the light of the historical development, one could argue that the CNPC was put forward at a much earlier stage of the theory and therefore cannot be valid any more. However, the basic assumption of this constraint is still compatible with a more modern theory of movement (e.g. Kayne 1994, Chomsky 1995). As I want to show, Swedish RCs are subject to the CNPC, whereby it does not matter which exact internal structure of RCs one bases the own analysis upon.

As for the syntax of Swedish RCs, there are two different RC-approaches. The first one – which I will call the standard analysis – implicates that the general Swedish relative particle *som* or its silent counterpart[9] is regarded as a head, merged in the C-domain of the clause. The gap after the relativized element is bound by a silent operator in a specifier position within the C-domain (cf. Platzack 2000, Huber forthcoming and others). An RC like (28a), thus, has the structure in (28b):

(28) a. en man som Maria såg
a man REL Maria saw
a man who Maria saw

b. $[_{DP}$ a man $[_{CP}$ op_i $[_{C'}$ som $[_{IP}$ Maria såg $t_i]]]]$[10]

An alternative approach to the structure of RCs was formulated by Kayne (1994). Roughly speaking, he proposed that the relativized element is merged within the RC and then raised to SpecC. The derived RC is selected by a determiner head. Thus, there exists no silent operator, and the dependency between the gap and the relativized element is that of a movement. Applying this approach to the case of Swedish, one could argue that *som* still is merged as a head within the C-domain:

(28) c. $[_{DP}$ $[_{D'}$ a $[_{CP}$ man$_i$ $[_{C'}$ som $[_{IP}$ Maria såg $t_i]]]]]$

---

[9] Like the English *that*, *som* can be omitted when a non-subject is relativized, cf.:

(i) mannen *(som) kommer
man-DEF REL comes

(ii) mannen (som) jag såg
man-DEF REL I saw

[10] I will restrict the discussion here to a simple, i.e. non-split CP (cf. also the next footnote).

The Kaynean approach has been widely debated under the last part of the 1990's (for a detailed discussion see Alexiadou et.al. 2000). It seems that it faces a lot of empirical and theoretical problems which not only arise when trying to transfer the analysis to other languages than English but also can be noticed for the case of English itself.

For the analysis of ERCs, it does not play a role which approach one applies: ERCs are not compatible with either view. The crucial point is that in an RC of an ERC, there are two gaps, one for the relativized and one corresponding to the fronted item (cf. ex. 9). A movement of the fronted item to the sentence initial position is subject to the principle of cyclicity since it extends over a minimal sentence. The notion cyclicity stems form the subjacency theory (Chomsky 1973, 1981). In order not two cross two boundary nodes at the same time, Chomsky (1973) suggested that movement should be cyclic, thus taking place in several steps. The fronted item in a ERC should successively be moved from its original position via SpecC of the RC into the topmost SpecC-position:

(29) $[_{CP}$ det $[_{C'}$ vet $[_{IP}$ jag $[_{DP}$ ingen $[_{CP}$ t_i' $[_{C'}$ som $[_{IP}$ kan t_i.]]]]]]]
this know I nobody REL can
I don't know anybody who can (do) this.

The intermediate trace t_i' is the indicator for the cyclic movement which guarantees that subjacency is satisfied.

However, considering the analyses of RCs in Swedish presented above, the structure in (29) cannot be correct. Consider fist the standard approach: on this view, a silent operator is placed in SpecC of the RC. Therefore, the fronted item cannot pass through this position since one position cannot be filled by two elements at the same time:

(30) * $[_{CP}$ det $[_{C'}$ vet $[_{IP}$ jag $[_{DP}$ ingen $[_{CP}$ **t_i' op$_k$** $[_{C'}$ som $[_{IP}$ t_k kan t_i]]]]]]]]^{11}

Likewise, within Kayne's system, another element (the RC head) stands in SpecC thereby making this position unavailable for an intermediate trace:

(31) $[_{CP}$ det $[_{C'}$ vet $[_{IP}$ jag $[_{DP}$ $[_{D'}$ $_{CP}$ **t_i' ingen$_k$** $[_{C'}$ som $[_{IP}$ t_k kan t_i]]]]]]]

Any argumentation in favour of a movement analysis should be able to fully explain why the CNPC in these cases can be circumvented. As far as I can see, no such convincing proposal has been made. On the other hand, the CNPC can still be regarded as fully compatible with a modern movement theory.

Summing up: besides the connectivity effects seen in 3.2.1, there seems to be no evidence for a movement analysis. The presumed independent *wh*-movement argument can be disregarded from since it appears that those sentences are echo questions and thus involve no proper derivations. From a more theoretical point of view, movement out of an RC never meets the cyclicity requirement. The CNPC has still explanatory force.

[11] Note that the same problem arises when one assumes a split CP with a Force-Fin-structure: As a *wh*-like element, the silent operator should be placed in SpecForce, i.e. higher than the functional subject category FinP. The fronted element would have to move through SpecForce which is occupied. Subjacency, thus cannot not be maintained.

4 A new proposal

4.1 ERCs as self-answering questions

In 2.3, I showed that the fronted item within ERCs always contains given information. From a pragmatic perspective, this item often serves as a so called topic (see also Engdahl 1997:70). As a topic, the fronted constituent is the starting point of the pragmatic predication (cf. Molnár 1991). We often find a special intonation contour in ERCs which is commonly called "I-" or "hat contour". The characteristics of such an intonation contour are a falling-rising tone at the beginning and a falling tone at the end of the sentence (cf. Büring 1997, Jacobs 1996, Molnár & Rosengren 1996 et. al.):

(32)	√KAFfe	känner	jag	\MÅNga	som	gillar.
	coffee	know	I	many	REL	like
	I know many people who like coffee.					

The item that is marked with the falling-rising accent tone (*kaffe* in 32) is usually called a "marked " or "contrastive topic". The falling-rising ("root") accent on the fronted constituent gives the impression of a question that is answered within the same sentence. Thus, (32) could be understood as a complex sentence like the one in (32'):

(32') Kaffe? – känner jag många som gillar.

I would like to propose this basic structure for ERCs. Note, however, that the question in (32') is elliptical. I suppose that it is the remainder of a monologue-like question like the one in (32''):

(32'')	Känner	jag	X	som	gillar	kaffe?	
	– känner	jag	många	som	gillar.		

Thus, ERC constructions are divided into two separate sentences: an interrogative and a declarative one. I assume that the basic position of the interrogative clause is the SpecC position of the matrix clause[12]. The answer part begins in C':

(32''')	[$_{CP}$ [$_{CP}$	Känner	jag	X	som	gillar	kaffe?]
	[$_{C'}$	känner	jag	många	som	gillar.]]	

Furthermore, I assume a polarity quantifier (*X*) to be the head of the RC within the question part. This element has no phonetic features, but it semantically resembles very much to the polarity item *any* in English. It is licensed by the interrogative operator in the question part of the sentence. In 4.2, I will show that the assumption of such a silent element makes sense both syntactically and semantically.

Although (32''') is a fairly detailed analysis of the underlying structure of (32), it is still not complete. As one easily can see, the answer part lacks a full RC. I would like to suggest that in ERCs, the question and the answer part are built phonetically parallelly which is supported by the verb raising in Swedish. Because of the structural parallelism, all double elements are deleted on PF. Hereby, even pragmatic factors play a certain role: since the sentence should be able to display a topic and it seems that there is a constrain on topics not

[12] Even here, I dispense with a split CP.

to be preceded by other material (cf. Molnár 1991), everything in front of the potential topic must be deleted. So, if *kaffe* in (32) is meant to be a topic, every phonetic material in front of it must be eliminated:

(32’’’’)		~~Känner~~	~~jag~~	X	~~som~~	~~gillar~~	kaffe?
	–	känner	jag	många	som	gillar	~~kaffe~~

Note further that *kaffe* cannot be doubled and therefore it must be deleted at PF within the RC. The polarity quantifier X is deleted on PF, as all empty elements are (cf. Chomsky 1995). The structure in (32’’’’), thus, would be the complete structure of the ERC construction in (32).

Before I demonstrate the advantages of my proposal, I want to stress that the pragmatic form still must be understood as independent of the syntactic form by means of modularity. Thus, even if an item like *kaffe* in (32) often may serve as a pragmatic topic, it does not necessarily have to. As Engdahl (1997:70ff.) points out, there are at least three different kinds of pragmatic functions observable with ERCs. The reduction to just one “marked topic” type with an “I-intonation” would be too simplified. Nevertheless, the proposed structure is assumed to hold for all ERC types.

4.2 Advantages of the proposed analysis

4.2.1 The explanation of the types of ERCs

ERCs are a very restricted construction type. As I already pointed out several times, it is not possible to “extract” an item from every restrictive RC in Swedish. In 2.2, I described the predication upon a set as a common feature of all ERCs. Why is this?

My analysis of ERCs as self-answering questions provides an explanation for the different types of ERCs. Remember that I assumed a polarity quantifier in the question part that resembles English *any*. In English, the yes/no-question that corresponds to (32) is *Do I know any people that like coffee?*. The basic property of yes/no-questions is that they can simply be answered by just confirming or negating the open proposition. Another, more detailed answer is achieved by specifying the presupposed set.

There are basically two alternative options: the first is to describe the quantity of the individuals that the predication is about. The question *Do I know any people that like coffee?* could thus be answered by e.g. *(Yes), many* or *(No), nobody*. In existential ERCs and thematic ERCs, this strategy is chosen:

(33) a. <u>Existential ERCs</u>:

Kaffe	finns	det	bara	en/ många/ingen	som	gillar.
coffee	exists	there	only	one many nobody	REL	like(s)

There is only one/are many/is nobody who like(s) coffee.

b. <u>Thematic ERCs</u>:

Kaffe	känner	jag	bara	en/ många/ingen		som	gillar.
coffee	know	I	only	one many	nobody	REL	like(s)

I know only one/many/nobody who like(s) coffee.

The answer part does not consist of a simple *yes* or *no*, but elaborates further on the amount of individuals for whom the predication holds.

The second strategy which is chosen in the case of specificational ERCs consists of referentially mentioning the individuals that the predication is about:

(33) c. Specificational ERCs:

Kaffe	är	det	bara	Kalle	som	gillar.
coffee	is	it	only	Kalle	REL	likes

It is only Kalle who likes coffee.

Here, the speaker answers the question *Are there any people that like coffee?* by *(Yes,) Kalle.* It is therefore no coincidence that semantic predications of this kind are involved within ERCs. The predication upon a set which we noticed in 2.2 can be understood to emerge from the need of a proper answer to the fronted yes/no-question. Other types of predication are not able to form correct answers to the corresponding yes/no-questions and hence banned from ERCs.

4.2.2. Explaining the connectivity effects

The main argument for a movement analysis consisted of the observable connectivity effects demonstrated in 3.2.1. By applying my interrogative-declarative-analysis, there is no need for any additional assumptions: since the question part in the prefield is self-contained, case marking and binding of anaphora can take place within this domain:

(34) a.

Honom	är	det	ingen	som	gillar.
him	is	it	nobody	REL	likes

There is nobody who likes him.

= ~~Är det~~ X ~~som gillar~~ honom? – är det ingen som gillar ~~honom~~.

b.

Sig själv	är	det	ingen	här	som	gillar.
themselves	is	there	nobody	here	REL	likes

There is nobody here who likes himself.

= ~~Är det~~ X ~~här~~ op ~~som gillar~~ sig själv? – är det ingen här som gillar ~~sig själv~~.

I assume that it is the silent operator that binds the anaphor *sig själv* in (34b). Likewise, case is assigned by the verb *gilla* in the question part of (34a). The two subclauses, thus, are independent of each other.

4.2.3 No ERCs in German and English

A interrogative-declarative-complex should be possible to formulate even in other languages like in German or English. However, my analysis predicts correctly that this cannot be the case.

Let us first consider German: like Swedish, German is a V2-language. The V2-allowed structural parallelism between the question and the answer part of the sentence. Parallel structures of this kind could be achieved in German as well since even here, a yes/no-question is formulated by a verb-initial structure. Note further that I assumed a silent polarity quantifier *X* as the RC head of an ERC construction. Such a silent element could also be assumed for German. However, as I pointed out in 3.2.3, it is very plausible to assume that the gap in a Swedish RC is bound by a silent operator in SpecC. The empty character of the operators in the question and the answer part, respectively, allows a further phonetic parallelism between the two parts of the sentence, cf. the detailed analysis of example (32):

(32’’’’’)		~~Känner~~	~~jag~~	X	**op**	~~som~~	~~gillar~~	kaffe?
	–	känner	jag	många	**op**	som	gillar	~~kaffe~~.

Since the two involved RC heads differ, the following operators might be differ, too, thereby rendering two non-parallel structures. For instance, one might argue that the operator following the polarity quantifier has a singular form and the operator after *många* a plural form. In this case, phonetic deletion would be out due to the formal differences. I suppose that using silent relative operators, Swedish does not face this problem, i.e. the silent operators can be coordinated in the proposed way although they might differ in form.

German, on the other hand, generally does not allow for silent relative operators. Instead, German uses (lexical) relative pronouns. If we assume a silent polarity quantifier in German as well, we face two problems: first, it is not clear which the proper pronominal form after a silent RC head would be since this quantifier might be underspecified for number and gender. Due to this underspecification, it is not possible to choose the right verbal form within the RC of the question part. Therefore, it is impossible to formulate a correct question. The second problem is that the two different RC heads involved could evoke two different relative pronouns:

(35)	* Kenne	~~ich~~	X	**?der$_{\text{MASK-SING}}$/?die$_{\text{FEM-SING}}$/?das$_{\text{NEUTR-SING}}$/?die$_{\text{PL}}$**			
	know	I		REL			
	Kaffee	**?mag/?mögen**?					
	coffee	like					
	– kenne	ich	viele,	**die$_{\text{PL}}$**	~~Kaffee~~	mögen.	
	know	I	many	who	coffee	like	

Therefore, ERCs in German are not possible to formulate.

English faces several problems with potential ERC constructions. Note first that English is not a V2-language, which means that the verb can be placed in third position in case of a syntactic topicalization. Furthermore, the structure of English yes/no-questions depend on the type of matrix verb: with thematic verbs, an English speaker uses the so called *do*-support. In an ERC, this *do*-form could not be omitted since it is not doubled:

(36)	* **Do** I know X op	that like ?(es)	coffee?
	– I know many people	that like	coffee.

Remember now that *do* in (36) precedes the presumed sentence topic *maths* which is illicit. Therefore, thematic ERCs are out in English.

A second problem with a construction like (36) is that the speaker cannot decide which verbal form to choose: if there is a silent polarity quantifier *X* like in Swedish, this quantifier could be underspecified for number and, hence, the speaker is not able to decide which agreement form on the verb is required. Swedish does not have this problem, since the verbal morphology makes no difference between number and personal endings.

In yes/no questions with a non-thematic verb (for instance with existential verbs or copular constructions), there is an inversion of the (formal) subject and the verb. This inversion is responsible for the fact that there cannot be a structural parallelism between the two parts of the sentences in an ERC and phonetic deletion cannot take place:

(37)	**Is there**	X	that like ?(es)	coffee?
	– **there is**	nobody	that likes	coffee.

Since the *is there* part cannot be deleted in the question, it precedes the topic at PF which leads to unacceptability. Besides, the quantifier X again seems to be underspecified for number and it is therefore impossible to choose the correct verbal form in the question RC. Like in German, ERCs are not possible to be formulated.

5 Concluding remarks

In this paper, I argued that there is no need for a movement analysis for ERCs in Swedish. Movement not only violates the still relevant CNPC, but it also cannot be restricted to the proper ERC types. The relevant types have been described as involving a semantic predication upon a set. The description fitted very well with the syntactic analysis that was proposed here: since a yes/no-question can be answered by elaborating on the quantity or the reference of a set of individuals, the approach seems to be the correct one. Moreover, I have shown that the only arguments for a movement analysis – the so called connectivity effects – can be explained without further assumptions. Finally, it could be shown how Swedish, but not German and English meet the necessary conditions for a language licensing ERCs.

Summing up, we find that a language must fulfil the following criteria, in order to allow for ERCs:

a) The question and the answer part of the ERC must be built structurally parallel. This is best achieved with a V2-language that has subject-verb-inversion in yes/no-questions, but no *do*-support.
b) There must be no lexical relative operators such as relative pronouns.
c) The verbal morphology of the must not show any marking for person and number on the verbal stem.

Mainland Scandinavian languages like Swedish meet these preconditions.

6 References

Abney, S.P. 1987. *The English Noun Phrase in its Sentential Aspect*. Unpublished Ph.D. dissertation, MIT, Cambridge, Mass.

Alexiadou, A., Law, P., Meinunger, A., Wilder C. 2000. (eds), *The syntax of relative clauses*, John Benjamins, Amsterdam/Philadelphia.

Büring, D. 1997. *The meaning of topic and focus: the 59th Street Bridge accent*, Routledge (= Routledge Studies in German Linguistics 3), London.

Büring, D. 1999. 'Topic', in Bosch, P. & van der Sand, R. (eds), *FOCUS. Linguistic, cognitive, and computational perspectives. Part II: Semantic interpretation of focus phenomena*, Cambridge, Cambridge University Press, pp 142-165.

Chomsky, N. 1973. 'Conditions on transformations', in Anderson, S. & Kiparsky, P. (eds), *A Festschrift for Morris Halle*, Holt, Rinehart and Winston. New York, pp 232-286.

--- 1981. *Lectures on government and binding*, Foris, Dordrecht.

--- 1986. *Barriers*, The MIT Press, Cambridge, Mass.

--- 1995. *The minimalist program*, The MIT Press, Cambridge, Mass.

Delin, J. 1992. 'Aspects of cleft constructions in discourse', Arbeitspapier Nr. 19 des Sonderforschungsbereichs 340, Stuttgart/Tübingen.

Delsing, L.-O. 1993. *The internal structure of noun phrases in the Scandinavian languages*. A comparative study. Ph.D Dissertation, University of Lund.

Engdahl, E. 1997. 'Relative Clause extractions in context', in *Working Papers in Scandinavian Syntax 60*, pp 51-79.

Higgins, F.R. 1973. *The pseudo-cleft construction in English*. Ph.D Dissertation, MIT, Cambridge, Mass.

Huber, S. Forthcoming. *Es-Clefts und det-Clefts. Spaltsätze im Deutschen und Schwedischen*, Ph.D. Dissertation, University of Lund.

Jacobs, J. 1996. 'Bemerkungen zur I-Topikalisierung', in *Sprache und Pragmatik* 41, pp 1-48.

Kayne, R.S. 1994. *The antisymmetry of syntax*, The MIT Press, Cambridge, Mass.

Molnár, V. 1991 *Das TOPIK im Deutschen und Ungarischen*. Ph.D Dissertation, University of Lund (= Lunder germanistische Forschungen 58).

Molnár, V. & Rosengren, I. 1996. 'Zu Jacobs' Explikation der I-Topikalisierung', in *Sprache und Pragmatik 41*, pp 49-88.

Platzack, C. 1996. 'Germanic verb second languages. Attract vs. Repel: On optionality, A-bar movement and the symmetrical/asymmetrical verb second hypothesis', in Lang, E. & Zifonun, G. (eds), *Deutsch – typologisch*, Walter de Gruyter, Berling & new York, pp 92-120.

--- 1998. 'A visibility condition for the C-domain', in *Working Papers in Scandinavian Syntax 61*, pp 53-100.

--- 1999. *Satsfläta me relativsats*, Ms, University of Lund

--- 2000. 'A complement-of-N° account of restrictive and non-restrictive relative clauses: The case of Swedish', in Alexiadou et. al. (eds), pp 265-308.

Platzack, C. & Rosengren, I. 1998. 'On the subject of imperatives: A minimalist account of the imperative clause', in *The Journal of Comparative Germanic Linguistics 1*, pp 177-224.

Rizzi, L. 1997. 'The fine structure of the left periphery', in Haegeman (ed.), *Elements of grammar. Handbook in generative syntax*, Kluwer, Dordrecht, Boston, London, pp 281-337.

Ross, J.R. 1967. *Constraints on variables in syntax*. Ph.D Dissertation, MIT, Cambridge, Mass.

Szabolcsi, A. 1987. 'Functional categories in the noun phrase', in *Approaches to Hungarian*, Vol. II, pp 16-189.

Taraldsen, T. 1982. 'Extraction from relative clauses in Norwegian', in Engdahl . E. & E. Ejerhed (eds), *Readings on Unbounded Dependencies in Scandinavian Languages*, Almqvist & Wiksell International, Stockholm.