

## **Relative Clause Extractions in Context\***

Elisabet Engdahl  
Department of Linguistics  
University of Göteborg

### **1. Introduction**

In the 1970s and in the early 1980s there was considerable interest in extraction phenomena in Scandinavian languages (cf e.g. Erteschik-Shir (1973), the papers in Engdahl & Ejerhed (1982), Engdahl (1980), Taraldsen (1981), Maling & Zaenen (1982). The theoretical issues concerned for instance the importance of the Scandinavian data for determining the nature of island constraints, the role of subjacency, and the issue of identifying bounding nodes in different languages. Since then not much has been written about Scandinavian extractions<sup>1</sup>. In this article, I want to take up the task of characterising one type of extractions, viz. extractions out of relative clauses. I believe that this phenomenon is central to a proper understanding of the organisation of the grammar, and in particular to a better understanding of the interaction between syntax and pragmatics. My main aim is to describe the structural characteristics of such extractions and to discuss non-structural factors which influence the actual occurrences of such extractions. A second aim is to make a larger and more representative body of extraction data available to the linguistic community who don't speak these languages. A third aim is to point out that more comparative work on Scandinavian languages is needed.

The article is organised as follows: After reviewing briefly the history of the study of relative clause extractions in section 2, I take a closer look at what types of relative clauses extractions occur naturally in section 3. In section 4, I discuss presentational sentences and clefts--two constructions which occur frequently in the extraction data--as well as some other common types. In

---

\* I have presented parts of the material in this article at the Comparative Germanic Syntax Workshop in Tromsø 1992, at seminars at the Department of Linguistics at the universities of Edinburgh, Stockholm and Göteborg, at the Department of Scandinavian languages at Lund and at a workshop at the Copenhagen Business School in February 1997. I thank the participants at these meetings for valuable I am particularly grateful to Karin Aijmer, Lars-Johan Ekerot, Staffan Hellberg and Christer Platzack for comments on an earlier draft. Some of the material in sections 3-6 is included in Engdahl 1997.

<sup>1</sup> Nordg(Erd) (1991) is an exception which adds a computational perspective.

section 5, we look at what types of NPs can head relative clauses that allow for extractions and in section 6 we consider the function of the extracted constituent. In section 7 we discuss why such extractions are possible precisely in some Germanic languages but not in others.

## 2. A brief history of the study of relative clause extractions

At least since Ross' dissertation (Ross 1967), linguists have been trying to characterise the restrictions on extractions in natural languages. By extraction we understand the configuration where a fronted constituent in for instance relative clauses, constituent questions and topicalised sentences corresponds to an empty position inside the clause. Ross pointed out that it is impossible to extract out of certain structural environments in English. He proposed the *Complex NP Constraint* (CNPC) which prohibits extraction out of a clause, dominated by a noun phrase. This constraint was subsumed under the more general *subjacency* constraint (Chomsky 1973) which prohibits movement across more than one bounding node. (1) is one of the examples Ross used to motivate the CNPC.

(A) \*The man<sub>k</sub> who I read [<sub>NP</sub> a statement<sub>j</sub> [<sub>CP</sub> which  $e_j$  was about  $e_k$ ]] is sick. (Ross 1967:4.3)

During the late 1960s and early 1970s, a lot of linguistic research was aimed at characterising *wh*-movement, and formulating appropriate restrictions on this rule. It was common to assume that subjacency was in fact a diagnostic for *wh*-movement. That is, you could tell if *wh*-movement was involved in the derivation by the fact that subjacency violations lead to ungrammaticality. Extractions out of relative clauses always involved a subjacency violation. Hence extractions out of relative clauses were excluded by the theory.

To my knowledge, the first linguist to draw attention outside Scandinavia to the fact that it is possible to extract out of relative clauses and to relate these facts to the discussion of island constraints in generative grammar was Nomi Erteschik-Shir.<sup>2</sup> In her 1973 MIT dissertation, she looked at the nature of island constraints. She argued that in order to account for when extractions out of relative clauses are possible, you have to refer to a pragmatic notion of dominance in discourse (see below, section 3.2). Soon after, Lars-Gunnar Andersson (1974) and Jens Allwood (1976) discussed extraction patterns in Swedish. The title of Allwood's 1976 paper gives a good clue to the content: *The complex NP constraint as a non-universal rule and some semantic factors*

---

<sup>2</sup> Inside Scandinavia, long-distance extractions have been discussed for around 100 years, primarily by prescriptive grammarians who admonish against using them. Such extractions are referred to in the Swedish linguistic literature as *satsflösta* ('sentence plait'), in Danish as *sætningensknode* ('sentence knot') and in Norwegian as *knutesætning*. In German, the phenomenon is referred to as *Satzverschränkung*, see e.g. Andersson and Kvam (1984).

*influencing the acceptability of Swedish sentences which violate the CNPC.* (2)  
is an example from Allwood (1976:(48)).

- (2) De blommorna känner jag en man som säljer  
*those flowers know I a man that sells*

Allwood also commented that (2) could be used as an answer to the question “Do you know where I can get some flowers like that?”, especially if prefixed by “Yes, ..”.

### 2.1. Are these real extractions?

A few years later, when I was a graduate student in Linguistics at UMass, there was some general awareness that there existed ‘weird’ relative clauses in languages like Swedish that you supposedly could extract out of. But on the east coast of the US, people were more interested in showing that these were *not* real extractions than in looking at what pragmatic or semantic factors were involved. Edwin Williams and Noam Chomsky independently suggested to me (personal communications 1979) that extractions in Swedish did not involve *wh*-movement, but instead involved an ‘empty resumptive pronoun strategy’. The term ‘resumptive pronoun strategy’<sup>3</sup> is used for constructions where an apparently fronted constituent is linked to a position that is inaccessible to *wh*-movement. Since *wh*-movement cannot apply, by assumption, there can be no trace; instead we find a resumptive pronoun coindexed with the fronted phrase, as illustrated in (3).

- (3) the man [<sub>j</sub> they think that if Mary marries him <sub>j</sub>, then everyone will be happy]

The special fact about Swedish, and presumably Danish and Norwegian as well, was that Swedish, instead of having normal resumptive pronouns with phonetic content, made use of resumptive pronouns without phonetic content, i.e. a category akin to little *pro*. The structure for (2) would, on this account, be as in (4).

- (4) De blommorna <sub>j</sub> känner jag en man som säljer *pro* <sub>j</sub>  
*those flowers know I a man that sells*

Since the initial constituent is linked not to a gap but to a pronominal category, there is no violation of subjacency. There are several objections that can be made to this proposal. Swedish does not otherwise allow empty *pro*, as the contrasts in (5) show.

---

<sup>3</sup> See Chomsky 1982, p 11 from where example (3) is taken.

(5) a Jag känner en man som säljer dem.

*I know a man who sells them*

b \* Jag känner en man som säljer *pro*.

This raises a methodological problem. If so-called empty resumptive pronouns only occur when there is a fronted constituent, it becomes impossible to distinguish them from *wh*-traces on distributional grounds, unless you take for granted that subjacency is a diagnostic for *wh*-movement.<sup>4</sup>

If examples like (2) involve pronominal copies rather than traces, one would expect these examples to behave like left dislocations. However, left dislocations in Swedish can easily be distinguished from topicalisation since the latter triggers verb second but the former doesn't. Compare the topicalisation in (2) with the corresponding left dislocation in (6).

(6) De blommorna<sub>j</sub>, jag känner en man som säljer dem<sub>j</sub>

*those flowers, I know a man that sells them*

A final reason not to analyse (2) as involving a resumptive pronoun is that this example with an overt resumptive pronoun is ungrammatical.

(7) \*De blommorna<sub>j</sub> känner jag en man som säljer dem<sub>j</sub>

*those flowers, know I a man that sells them*

In this respect, there is presumably a difference between English and the Scandinavian languages. It is often said that resumptive pronouns improve sentences which would otherwise be ungrammatical. Compare the contrast in acceptability for the English sentences in (8).

(8) a \*?Those flowers I know someone who sells

b ?Those flowers I know someone who sells them<sup>5</sup>

This is not the case in the examples involving extractions from relative clauses in Swedish. Here (2) without a resumptive pronoun is fine, but (7), with the pronoun, is not used at all.

---

<sup>4</sup> See Cinque 1991 for a detailed discussion of how to distinguish empty resumptive pronouns from *wh*-traces in Italian. Only some of Cinque's tests carry over to Swedish.

<sup>5</sup> (8b) is of course indistinguishable from a left dislocation structure in English.

## 2.2. An extraposition analysis

Taraldsen (1981, 1982) brings up some interesting contrasts in Norwegian concerning extractions from relative clauses, which he takes to show that there is a structural restriction on acceptable extractions that cannot be explained in semantic or pragmatic terms. Taraldsen argues that there is a contrast in acceptability as shown in (9) and (10). The relative clauses are underlined.

- (9) a \*Her er en bok<sub>j</sub> som ingen som leser e<sub>j</sub> blir lykkelig (Taraldsen  
*here is a book that nobody that reads becomes happy*'81:(52))
- b Her er en bok<sub>j</sub> som ingen blir lykkelig som leser e<sub>j</sub>  
*here is a book that nobody becomes happy that reads* \_
- (10) a \*Rødsprit slipper vi ingen som har drukket e<sub>j</sub> inn. (Taraldsen  
*red spirit let we nobody that has drunk in* '82:(9))
- b Rødsprit slipper vi ingen inn som har drukket e<sub>j</sub>.  
*red spirit let we nobody in that has drunk*

In the unacceptable (a)-examples, the relative clause is adjacent to the head NP, whereas in the acceptable (b) clauses, the relative clause has been extraposed across some lexical material<sup>6</sup>. Taraldsen suggests that this kind of extraposition occurs in all cases, i.e. also in examples like (2) where extraposition would be string vacuous. After extraposition, the relative clause is no longer dominated by NP and hence no subjacency violation occurs. This was an interesting proposal which had the merit of explaining a distributional fact, i.e. that relative clauses with extractions always occurred clause finally, with the help of a grammatical mechanism thought to be generally available, viz. extraposition. However, it is not so easy to express this idea using current assumptions. For instance, Kayne (1994) proposes that all movement is leftward, which means apparent rightward extraposition must be reanalysed as resulting from leftward movements of all other constituents. I don't know if anyone has attempted to

---

<sup>6</sup> The contrast between (9a) and (9b) is noticeable in Swedish as well. Since verb particles normally precede the NP in Swedish, the only possible order is as in (i).

(i) Rødsprit slöpper vi inte in någon som har druckit.  
*red spirit let we not in somebody that has drunk*

recapture Taraldsen's analysis with these restrictions, but it doesn't seem like a plausible way to go.

The role of bounding nodes changed considerably with the *Barriers* approach (Chomsky 1986). Bounding nodes were seen as a relational notion, dependent on the configuration. Whether a node is a barrier or not depends on whether it is L-marked, roughly whether it is lexically selected. As far as I understand, relative clauses are never L-marked, and hence will always introduce a barrier.

Returning to Taraldsen's proposal, we note that it does capture what seems to be a characteristic property of relative clause extractions, namely the fact that the relative clause tends to be clause final. Extractions out of non-clause final relative clauses only seem to be possible in the case of parasitic gaps, as in (11), from Lars-Johan Ekerot. (For discussion of such examples, see e.g. Kayne 1983, Engdahl 1985, and Cinque 1991.)

- (11) Den avhandlingen<sub>j</sub> blir ingen som läser <sub>-p</sub> upphetsad av *e<sub>j</sub>* .  
*that dissertation gets noone who reads -excited by*

### 2.3. A restriction on subjects

There is one further restriction that needs to be mentioned. The fronted phrase is never extracted from the (highest) subject position in the relative clause. Allwood (1976) proposes a *Complex NP Subject Constraint* to block ungrammatical extractions as in (12).

- (12) \*Den här lingvisten<sub>k</sub> finns det [<sub>NP</sub>ingen teori<sub>j</sub> [<sub>CP</sub> som e<sub>k</sub> tror på e<sub>j</sub> ]  
*this linguist                    there is no theory that                    believes in*

I suspect that (12) is ungrammatical for two reasons. It involves an intersecting extraction and these are in general unacceptable in Swedish (cf Engdahl 1982b). Furthermore it involves extracting a subject after a lexically filled Comp, a context which requires a resumptive pronoun in standard varieties of Swedish (Engdahl 1985). In Norwegian, both these constraints can be relaxed (Christensen 1982, Nordgård 1991). Still, subjects cannot be extracted out of normal headed relative clauses as in (12) in Norwegian either. However, it is possible to construct acceptable Norwegian examples where the subject of a *free* relative has been extracted as in (13), (Nordgård, p.c.).

- (13) Denne kunstneren<sub>j</sub> kjøper jeg [ hva<sub>k</sub> [ e<sub>j</sub> (enn) produserer e<sub>k</sub> ]]  
*this artist                    I buy                    what                    ever produces*

There are however restrictions on lexical material in the head NP. I have not come across any natural examples of this kind. More research is clearly needed in order to establish the exact conditions on subject extractions in these contexts.

We have now established two syntactic restrictions on relative clause extractions. We now turn to characterising the discourse conditions that seem to be relevant for when an extraction is used. But first we will extend the range of examples.

### 3. Natural extraction data

Consider the spontaneously produced examples in (14).<sup>7</sup>

---

<sup>7</sup> Most of the examples in this article are taken from a collection of naturally occurring extractions, both written and spoken, that I have gathered over the past 20 years. When no reference to a written source is given, the example was produced in conversation.

(14) a Den teorin<sub>k</sub> känner jag [<sub>NP</sub> ingen<sub>j</sub> [<sub>CP</sub> som e<sub>j</sub> tror på e<sub>k</sub>]]  
*that theory know I nobody that believes in*

- b Där<sub>k</sub> har jag [<sub>NP</sub> en moster<sub>j</sub> [<sub>CP</sub> som e<sub>j</sub> bor e<sub>k</sub>]]  
*there have I an aunt that lives*
- c Det<sub>k</sub> finns det [<sub>NP</sub> ingen<sub>j</sub> [<sub>CP</sub> som e<sub>j</sub> kan hjälpa mig med e<sub>k</sub>]]  
*that there is nobody that can help me with*

The examples in (14) all involve topicalisation. Examples where a constituent has been relativised out of a relative clause also occur, preferably in sentences that have a presentational function as in (9b) and (15a). Questioning out of relative clauses is, as far as I'm aware, less common, but utterances like (15b,c) do occur. Note that these are both cleft constructions.

- (15) a Här är en fråga<sub>j</sub> som jag inte känner någon som kan svara på e<sub>j</sub>  
*here is a question that I don't know anybody that can answer*
- b Vem<sub>j</sub> var det ingen som kände e<sub>j</sub> ?  
*who was there nobody that knew*
- c Vilket ord<sub>j</sub> var det ingen som kunde stava rätt till e<sub>j</sub>?  
*which word was it nobody that could spell correctly*

So far all examples have involved extractions out of relative clauses where the subject has been relativised. These are by far the most common. However, extractions out of other types of relative clauses occur. (16a) is an example with an infinitival relative clause and in (16b) the object has been relativised.<sup>8</sup>

- (16) a Det<sub>j</sub> har jag inte haft någonting<sub>k</sub> att göra e<sub>k</sub> med e<sub>j</sub>  
*that I have not had anything to do with*
- b Matte<sub>j</sub> var det bara pappa<sub>k</sub> (som) jag kunde fråga e<sub>k</sub> om e<sub>j</sub>  
*maths it was only dad that I could ask about*

Referentiality is sometimes brought up in discussions about extractions (cf Rizzi 1990, Cinque 1991). In (17) we find examples of fronted non-referential constituents.

---

<sup>8</sup> In order to make the English paraphrases sound more natural, I will from now on not render the verb second order in the translations. The reader can easily check the Scandinavian original for the position of the verb.

- (17) a Så<sub>j</sub> känner jag ingen som kan måla e<sub>j</sub>.  
*so I know nobody that can paint*  
*'I don't know anybody who can paint like that'*
- b ...ett fyrkantigt, oromantiskt namn<sub>j</sub> som jag då inte kände  
*a square unromantic name that I then didn't know*  
 någon som hette e<sub>j</sub>. (Å.Moberg (1996) *Simone*  
*somebody that was called och jag, p206)*

In (17a) the fronted manner adverbial modifies the verb in the relative clause. The extraction in (17b) corresponds to *namn* (name) which is presumably non-referential.

Note furthermore that extractions are impossible out of non-restrictive relative clauses as shown in (18), which is similar to (14a). The adverbial *för övrigt* (by the way) forces the relative clause to be understood as a non-restrictive one.<sup>9</sup>

- (18) \*Den teorin<sub>j</sub> känner jag en man, som för övrigt tror på e<sub>j</sub>.  
*that theory I know a man, who by the way believes in*

If we look at what the naturally occurring examples of extractions have in common, we see that most of the examples given so far involve some kind of *presentational* or *cleft* construction. The main verb is usually some form of *vara* (be), *ha* (have) or *finnas* ('there is'). We have also seen that the extraction most often consists of a topicalisation. In the next section we will look closer at the connection between presentational constructions and extractions. Examples like (14a) with main verbs such as *känner* (know, be acquainted with) are also common and will be discussed in section 4.4.

## 4. Common construction types

### 4.1. Presentational constructions

---

<sup>9</sup> See Platzack (forthcoming) for a proposal which makes a structural distinction between restrictive and non-restrictive relative clauses which also can account for the difference in extraction possibilities.

Presentational constructions typically introduce a new referent into a situation. In Scandinavian languages, the introduction of novel referents often takes the form of a there-insertion sentence. The form of the pleonastic element is *det* (it) or *där/der* (there), mainly in Danish and southern Swedish dialects. The choice of verbs that can appear in presentational constructions is quite wide (see e.g. Sundman 1980, Anward 1981 and Sveen 1996). The role of the situation into which the new referent is introduced can be illustrated by the following pair of examples.

- (19) a    Det springer en hund på gatan.  
          *there runs    a dog    on the street*

- b Det springer en hund.  
*there runs a dog*
- (20) a Vad hände sen?  
*what happened then*
- b Det kom en man.  
*there came a man*

(19a) is a common presentation utterance whereas (19b) sounds odd. In my view, (19b) can only be used in a context where the speaker and the hearer already agree on which situation or scene they are talking about. For instance, (19b) could be felicitously uttered if you were studying a picture closely. Similarly, (20b) is an OK, somewhat elliptical utterance in the context introduced by the question in (20a). Presentational utterances thus seem to require something more than a novel noun phrase (cf. Ekberg 1990). The general existential predicate *finns* (exist) often introduces an indefinite NP followed by a relative clause which characterises the new referent, as in (21).

- (21) Det finns många som talar det språket.  
*there are many who speak that language*

The NP head often lacks descriptive content. Typical NPs are *någon*, *några* (some), *inga* (no) and *flera* (several). The discourse relevant property of the NP you are talking about is introduced in the relative clause. According to Prince (1981), the ‘informational centre’ of the utterance lies in the relative clause. What happens when you extract out of such an utterance is that part of the informational centre is fronted to a prominent topic position. (In section 6 we will look closer at the intonation patterns associated with the fronted constituent.)

- (22) a Det språket<sub>j</sub> finns det många som talar *e<sub>j</sub>*.  
*that language there are many that speak*
- b Här är ett språk<sub>j</sub> som det finns många som talar *e<sub>j</sub>*.  
*here is a language that there are many that speak*

- c Grovbröd<sub>j</sub> finns det bara en ost som är god att ha på *e<sub>j</sub>*.  
*coarse bread there is only one cheese that is good to have on*  
 (Hagström (1976) p 143.)

Erteschik-Shir (1973, 1982) uses the term ‘dominant’ to characterise the domains out of which you can extract and uses a number of operational tests to identify dominant subordinate clauses.

The observation that the good extraction domains are informational centres gives us a clue why it is impossible to extract out of non-restrictive relative clauses such as (18) above. A non-restrictive relative clause does not provide information that is essential for identifying or characterising the head NP. Consequently, such a clause cannot be an informational centre or dominant in the sense we are using these terms. The contrast between (14a) with a restrictive relative clause and (18) with a non-restrictive one supports the notion that extracted material must be informationally salient.

Another type of presentational relative construction is illustrated in (22’). Consider the English utterance in (22’a). This would most naturally be translated into Swedish as in (22’b). This presentational construction is often the source of a fronting, especially of direct quotes, as in (22’c).

- (22’) a Someone said that we got the wrong evening.
- b Det var någon som sa att vi tagit fel kväll.  
*there was someone that said that we got the wrong evening*
- c ["Vi kanske kommer fel kväll"]<sub>j</sub>, var det en liten pojke, som sa *e<sub>j</sub>*.  
*we maybe got the wrong evening was there a little boy that said*  
 (A.Lindgren: *Pippi har julgransplundring*)

## 4.2 Contact clauses

In some dialects, the relative complementiser can be omitted in presentational relative clauses and we get what Jespersen (1929) called a *contact clause*. Some examples are given in (23).<sup>10</sup> \_ marks the position of the omitted relative complementiser.

---

<sup>10</sup> I heard (14 a,b) on Radio Scotland. (14c-e) are from Jim Miller’s ECOSSE corpus of spoken Scottish. (14f,g) come from Judy Delin and (14h) comes from Quirk et al. (1972:14.29), who note that this kind

- (23) a There is nothing \_ can be done about it.  
 b I have a daughter \_ stays in Leith.  
 c My friend's got a brother \_ used to be in the school.  
 d There's a man in our street \_ has a Jaguar.  
 e We had this French girl \_ came to stay.  
 f It was him \_ did it.  
 g Who was it \_ let the cat out?  
 h There is something \_ keeps upsetting him.  
 j \*I killed a man \_ had brown eyes

Such contact clauses are most common in presentational constructions with *be* and *have* but they also occur in clefts as shown by (23f,g). The natural examples can be contrasted with (23j) which is totally impossible according to people who use contact clauses.

In her detailed article from (1963) on dialects spoken in the Danish islands and in Scania, Inger Ejkskjær gives several examples of contact clauses in Danish. The other examples are taken from Brink (1987).

- (24) a Der er et tog \_ går klokken seks (Ejkskjær p 113)  
*there is a train leaves (at) six o'clock*
- b Der er ikke ret mange \_ kan spinde. (n.w.Sjæl.) (Ejs. p 113)  
*there are not very many can spin*
- c Jeg har to døtre \_ bor i København. (n. Sjæl.) (Ejs. p 151)  
*I have two daughters live in Copenhagen*
- d Di har båda \_ ligger ute på Atlannten. (Scania) (Ejs. p 151)  
*they have boats lie out on the Atlantic*
- e Hvem er det \_ går nede på vejen? (Brink)  
*who is it walks down on the road*
- f Så var der en mand deroppe \_ tog imod det. (Brink)  
*and then there was a man up there received it*

---

of subject omission is only possible in relatives in existential sentences. Lambrecht (1988) has several examples from American English.

Note that (24c) is an almost exact parallel to the Scottish example in (23b).

The contact clauses are particularly interesting for the theme of this article, since many of Ejskjær's examples also involve extractions.

- (25) a De<sub>j</sub> e där ingen \_ ved e<sub>j</sub>. (Scania) (Ejs. p 114)  
*that there is nobody knows*
- b Det<sub>j</sub> var der ikke noget \_ hed e<sub>j</sub>. (n. Sjæl.) (Ejs. p 114)  
*that there was not anything was called*
- c Dét<sub>j</sub> er der vist ingen andre \_ synes e<sub>j</sub> end du, Marie. (Da.)(Brink)  
*that there is nobody else thinks than you*  
*'Nobody other than you think so, Marie'*

Ejskjær also notes that in some dialects the relative complementiser coincides with the conjunction, as in (26). (26d) involves an extraction.

- (26) a Der er nogen og \_ kan (Sjæl.) (Ejs.p 149)  
*there is someone and knows (how to)*
- b Vi har nogle senge og \_ står oppe på loftet (s.w.Sjæl.) (Ejs. p 151)  
*we have some beds and stand in the attic*
- c Det er hende og \_ har gjort det (Fal.) (Ejs. p 151)  
*it is her and has done it*
- d Ham<sub>j</sub> er der ingen og \_ kan lide e<sub>j</sub> (e.Sjæl.) (Ejs. p 151)  
*him there is nobody and can stand*

This may suggest that contact clauses don't involve real subordination but should rather be analysed as a type of coordination where the conjunction and the subject in the following clause is deleted, under identity with the preceding NP. The example in (23b) could be analysed as in (27).

- (27) I have a daughter (and she) stays in Leith

If this is the correct analysis, contact clauses would be better seen as an instance of conjunction reduction. There are however arguments against this analysis of contact clauses. Note that not all examples can be reanalysed in this way. For instance, if the head NP consists of a negative quantifier, then a conjunction analysis is implausible on semantic grounds.

(28) ?There is nothing (and it) can be done about it

For the Danish examples, it is possible to show that the clause introduced by the apparent conjunction nevertheless behaves as a subordinate clause when it comes to word order. In Danish, as in Norwegian and Swedish, sentence adverbials follow the finite verb in main clauses, but precede the finite verb in subordinate clauses. Compare (29a) and (29b).

(29) a De kommer aldrig i kirke (Da) (Main clause)  
*they come never to church*

b at de aldrig kommer i kirke (Da) (Subordinate clause)  
*that they never come to church*

(30 a,b) shows that *og* introduces clauses with typical subordinate word order. The same holds in (30c) where the complementiser is missing.

(30) a der er mange og \_ aldrig kommer i kirke (Fyn) (Ejs. p  
*there are many and never come to church* 126)

b der er en anden en og \_ pænt må blive hjemme (s.w.Sjæl.)  
*there is another one and just has to stay home* (Ejs. p 127)

c Der er ingen \_ rigtig ved noget (Loll.) (Ejs.  
*there is nobody really knows anything* p 126)

From this I conclude that contact clauses are some form of relative clauses.

The fact that there are many examples of extractions out of contact clauses might lead one to think that they are two aspects of the same phenomenon. However, despite this convergence, it is not possible to equate the structural and/or pragmatic conditions that trigger relative pronoun omission in contact clauses with the structural and/or pragmatic conditions that determine when extractions are used. A comparison between the English and the Danish dialects

may yield that contact clauses are used in very similar contexts. Nevertheless, extractions are common in the Danish dialects, but not so in the English ones<sup>11</sup>.

### 4.3. Cleft constructions

Many naturally occurring extractions involve some form of cleft construction with a subordinate clause with the relative complementiser *som*. These clauses are rather similar to relative clauses and it seems appropriate to discuss them in this paper.<sup>12</sup> I will refer to these clauses as *c-clauses*. Delin (1989) and Paggio (1996) distinguish three types of cleft constructions, viz. *it-clefts*, *wh-clefts* and *reverse wh-clefts*. For Scandinavian languages, it seems appropriate to divide *wh-clefts* into two types: morphological *wh-clefts*, which are introduced by a *wh*-phrase, often corresponding to a CP, and demonstrative clefts (*dem-clefts*), introduced by a demonstrative pronoun. The different types of clefts are shown in (31). The *c*-clause contains a gap which is interpreted with reference to the clefted constituent, which may be an NP, PP, AdjP, AdvP or CP. Examples of the different types are given in (32).

---

<sup>11</sup> Ejskj34r has some examples of contact clauses from SkCene (Scania), but this construction does not seem to be used in other dialects in Sweden. Further north, in dialects spoken around lake MŠlaren, a similar construction exists, as discussed by Nordberg (1977) who gives the examples below.

- (i) Jag har cykeln \_stCER pCE gCERden  
*I have the bike stands in the yard*
- (i) PCE bordet<sub>i</sub> hade hon en massa Špplen \_ lCEg e<sub>j</sub> .  
*on the table had she a lot of apples lay*

The matrix verb is always a form of *ha*, (have). The NP can be definite or indefinite. The following verb agrees in tense with the matrix *ha*. Fronting out of the subjectless clause is possible as shown in (ii).

<sup>12</sup> The relation between the head NP and the *som* clause is of course not the same in relatives and in clefts. Restrictive relative clauses can normally not modify proper names, but *c*-clauses often cooccur with proper names. C.Platzack (pers.comm.) suggests that the clefted constituent is moved into Spec-CP of the *c*-clause. Support for this analysis comes from the fact that the case of the clefted constituent is determined by the position in the *c*-clause.

- (i) Det var han som e kom.  
*it was he-NOM that came*
- (ii) Det var honom (som) jag trŠffade e.  
*it was him-ACC (that) I met*

- (31) a **det-cleft:** *det* copula {NP,AP,PP,CP} c-clause  
 b **wh-cleft:** *wh* c-clause copula CP  
 c **dem-cleft:** *dem* c-clause copula NP  
 d **reverse wh:** CP copula *wh* c-clause  
 e **reverse dem:** NP copula *dem* c-clause

(32) **det-cleft:**

- a Det var Kalle [som *e* slog sönder fönstret.]  
*it was Kalle who broke the window*

**wh-cleft:**

- b [Vad jag inte förstår *e*] är [CP hur vi ska hinna med allting]  
*what I don't understand is how we will get everything done*

**dem-cleft:**

- c [Den som *e* bakar de godaste bullarna] är mormor.  
*the (one) who bakes the best buns is granny*

**reverse-wh:**

- d Att flyga är [vad jag alltid har drömt om *e*.]  
*to fly is what I have always dreamt of*

**reverse-dem:**

- e Mormor är [den som *e* bakar de godaste bullarna].  
*Granny is the (one) who bakes the best buns*

*det*-clefts are very common and reverse *dem*-clefts occur relatively frequently, whereas *wh*-clefts are not used very often in Swedish (cf. Johansson 1996). We have already seen some examples of extractions out of *det*-clefts in Swedish in (8b) and (9b) and additional examples can be found in (33). (33c) is an authentic Norwegian example which was pronounced with stress on the pronominal head of the relative clause which I have marked by SMALL CAPS. Note that whenever there is an extraction, the copula will occur in second position.

- (33) a Mästerdetektiven Blomkvist<sub>j</sub> var det två som klädde ut sig till *e<sub>j</sub>*  
*detective Blomkvist it was two that dressed up as*  
 på kalaset. (MC 6 years)

*at the party*

- b En sådan häst; är det nog ingen mer än jag som har  $e_j$ .  
*such a horse there is probably nobody else but me what has*  
(A.Lindgren: *Mio min Mio*)

- c Dette er det ikke JEG<sub>j</sub>, som har funnet på e<sub>j</sub>  
*this it is not I who has invented*

It is important to distinguish genuine *det*-clefts introduced by the unstressed pleonastic *det* (it) from other utterances with a referential *det* subject, corresponding more or less to *that* in English or *das* in German. This *det* is also used as a personal or demonstrative pronoun, referring to neuter entities (34a) and events (34b), and may be stressed in certain contexts.

- (34) a DET är bordet.  
*that is the-table (NEUT)*
- b Dricka kaffe, DET är gott.  
*drink coffee, that is nice*

The referential *det*<sup>13</sup> is also used with non-neuter antecedents in certain deictic contexts as shown in (35). I will refer to these kinds of copula constructions as *identificational*.

- (35) a Vem är DET?  
*who is that*
- b DET är Kalle.  
*that is Kalle*

The need for distinguishing *det*-clefts from other utterances introduced by a referential *det* becomes clear in the following minimal pair.

- (36) a Det är bara Kalle som kan öppna garagedörren.  
*it is only Kalle who can open the garage door*
- b Garagedörren<sub>j</sub> är det bara Kalle som kan öppna e<sub>j</sub>  
*the garage door it is only Kalle who can open*

---

<sup>13</sup> Some colleagues have remarked that it is somewhat unusual to call this *det* 'referential'. I suspect that this depends on whether you take 'referential' implicitly to mean reference to an individual. I use the term in a wider sense which includes reference to properties, states, situations, beliefs and events. Regardless of what you call *det* in (34) and (35), it is clearly distinct from the pleonastic *det* in cleft-constructions.

- (37) a Vad är det som händer?  
*What is happening? (What is that?)*
- b Det är bara Kalle som öppnar garagedörren.  
*it is only Kalle who is opening the garage door*

- c \*Garagedörren är det bara Kalle som öppnar.  
*the garage door it is only Kalle who is opening*

(36a) is a normal *det*-cleft and the extraction in (36b) is fine. *det* in (37b) is most naturally interpreted as referring to whatever noise prompted the question in (37a) and can either be stressed or unstressed. Note that it is impossible to extract out of the relative clause on this interpretation (37c). Just like relative clauses in presentational constructions, c-clauses in cleft constructions tend to provide the informational centre of the utterance. This is clear in (36a) where the focalising adverb *bara* (only) highlights the fact that Kalle is the only person who has the property provided by the c-clause. Fronting an element from the c-clause further emphasises that the discourse relevant property predicated of Kalle is one that involves precisely the garage door. In (37a) on the other hand, *det* presumably refers to the event of Kalle opening the garage door. The purpose of the utterance seems to be to identify which event *det* refers to. In this identificational use, extractions are impossible.<sup>14</sup>

Extractions out of *reverse dem-clefts* sometimes occur, especially if the demonstrative is strengthened by a focalising adjective like *ende* (only).

- (38) Surströmming<sub>j</sub> är Fredrik den ende som tycker om e<sub>j</sub>.  
*fermented herring is Fredrik the only one who likes*

I have not come across any naturally occurring examples of extractions out of a *wh-cleft* or *dem-cleft*. Attempts to construct examples sound very bad.

- (39) \*Surströmming<sub>j</sub> är [den ende som tycker om e<sub>j</sub>] Fredrik,  
*fermented herring is the only one who likes Fredrik*

Note that in *wh-clefts* and *dem-clefts*, the c-clause is internal to the sentence and the reason extractions out of such clefts don't occur may well be that the c-clause isn't clause final. Compare the discussion in (2.2.). In *det*-clefts, reverse *dem-clefts* and reverse *wh-clefts*, on the other hand, the c-clause is always sentence final. Note that there is still a clear contrast in acceptability between the naturally occurring examples of extractions out of *det*-clefts in the

---

<sup>14</sup> There are interesting parallels between examples like (37) and the so-called pseudo-relative constructions in Romance languages, see e.g. Cinque (1995) and Strudsholm (1996). I'm grateful to Nigel Vincent for bringing the latter's thesis to my attention. Taraldsen (1986) refers to the corresponding French construction as *predicative qui-clauses*.

Scandinavian languages and constructed English parallel examples. Compare for instance the English counterparts to the Swedish examples in (9b) and (22b).

(40) a \*Maths<sub>k</sub> it was only dad<sub>j</sub> I could ask  $e_j$  about  $e_k$ .

b \*The garage door<sub>j</sub> it is only John who can open  $e_j$

#### 4.4. Other main verbs

Not all naturally occurring examples of extractions involve presentational or cleft constructions. We have already seen one common type in (14a), repeated here in (41a). Extractions also occur out of NPs that are complements to verbs like *behöva* (need), *känna till* (know of), *se* (see), *hitta på* (make up), *beundra* (admire).

(41) a Den teorin<sub>k</sub> känner jag ingen<sub>j</sub> som tror på  $e_k$   
*that theory know I nobody that believes in*

b Det<sub>j</sub> behöver vi någon som tar hand om  $e_j$ .  
*that we need someone who takes care of*

c En sådan frisyr<sub>j</sub> har jag aldrig sett någon som ser snygg ut i  $e_j$ .  
*that kind of hairstyle, I have never seen anyone who looks good in*

d Latex är det viktigt att känna någon som kan hjälpa till med  $e_j$ .  
*Latex it is important to know someone who can help with*

e Ja, det<sub>j</sub> kan vi hitta på en sång som heter  $e_j$ . (Vardagsgruppen:  
*yes, that we can make up a song that is called 'Mamma Mu')*

f Den här teorin<sub>j</sub> beundrar jag dem som förstår  $e_j$ .  
*this theory I admire those that understand*

What do these verbs have in common? Allwood (1976) discusses examples with *känna* (know), *veta om* (know about) and *se* (see) and suggests that the relevant property is that these verbs designate relations both to singular objects and to state of affairs and that this is important in the interpretation of the utterance. In particular, it is important that the remainder of the clause, i.e. the

sentence minus the initial topicalised constituent, can be interpreted as a reasonable predication (comment) with respect to the topic (Allwood 1976 [1982 p28]).<sup>15</sup> This is an important observation which correctly highlights the context sensitivity of many of the examples. If the context is one in which whatever follows the initial constituent can be interpreted as a relevant predication given what the speaker and the hearer are attuned to at that point in the conversation, then an extraction is often possible. In fact, I believe that Allwood's 'relevant predication' is another way of identifying what Prince refers to as the 'informational centre' of the utterance. This may also explain why extractions out of cleft constructions are so common. Clefting is a way of informationally structuring an utterance so that the new and the presupposed information can more easily be identified. In presentational relative constructions such as (9b), the relative clause clearly is part of a coherent predication.

## 5. Properties of the head NP

In this section we will look closer at what types of NPs can head relative clauses which allow for extractions. In the linguistic literature, one sometimes comes across the claim that extractions out of relative clauses are only possible if the head NP is indefinite. As we have already seen several examples with definite heads this cannot be correct. It seems rather that the form of the head NP depends on the type of sentence and the type of verbs used, both in the matrix sentence and in the relative clause. Looking first at sentence type, we note that in presentational sentences in Scandinavian languages, the head must be indefinite<sup>16</sup>. This also holds for presentational relative clauses, regardless of whether there is an extraction or not. Compare the wellformed pair in (42) with the illformed pair in (43).

- (42) a Det finns många som talar det språket.  
*there are many who speak that language*

---

<sup>15</sup> As we have seen above, the extracted constituent need not be topicalised, but can be relativised or questioned. However the requirement that the remainder of the clause should form an informationally coherent predication still holds.

<sup>16</sup> The claim that the NP must be indefinite is too strong. Nimb (1993) notes that focus adverbials may cooccur with definite NPs in presentational constructions as in (i) and (ii).

- (i) Det Šr bara jag kvar.  
*there is only I left*  
 (ii) Det finns bara MtEverest kvar att bestiga.  
*there is only MtEverest left to climb.*

- b Det språket<sub>i</sub> finns det många som talar  $e_j$  .  
*that language there are many that speak*
- (43) a \*Det finns kvinnan som talar det språket.  
*there is the woman who speaks that language*
- b \*Det språket<sub>j</sub> finns det kvinnan som talar  $e_j$  .  
*that language there is the woman that speaks*

In cleft constructions, on the other hand, the head can be either definite (a proper name (44a) or a definite NP (44b)) or indefinite (44c), depending on the context. Extractions are possible in all types, as shown.

- (44) a Tapeterna<sub>j</sub> var det Sven som valde  $e_j$  .  
*the wallpaper it was Sven who chose*
- b Detta<sub>j</sub> är det bara presidenten som kan avgöra  $e_j$  .  
*this it is only the president who can decide*

- c Lax<sub>j</sub> var det många som ville ha e<sub>j</sub> .  
*salmon it was many who wanted*

Looking next at extractions out of sentences with matrix predicates such as *känner* (know), it seems at first glance that a definite head makes the extraction worse. Compare (45 a and b).

- (45) a Den teorin<sub>k</sub> känner jag ingen som tror på e<sub>k</sub>  
*that theory I know nobody that believes in*
- b ??Den teorin<sub>k</sub> känner jag mannen som tror på e<sub>k</sub>  
*that theory I know the man that believes in*

On closer inspection it turns out that the reason for the illformedness is not due to the extraction but that the whole sentence is strange.<sup>17</sup> Note that the same acceptability pattern obtains in the corresponding sentences without extractions.

- (46) a Jag känner ingen som tror på den här teorin.  
*I know nobody who believes in this theory*
- b ??Jag känner mannen som tror på den här teorin.  
*I know the man who believes in this theory*

The reason that (46b) (out of context) is odd is presumably that *tro på* (believe in) denotes a many-to-one relation, i.e. there are normally many believers in one theory. Saying that you know *the man* who believes in the theory is pragmatically odd, unless the context warrants the conclusion that there is exactly one man who believes in the theory, in which case the relative clause in (46b) is sufficient for identifying the man. If we replace *tro på* (believe in) with a verb that denotes a one-to-one relation, like *uppfinna* (invent), we would expect that the preference would now be for a definite subject and hence a definite head of the relative clause. This expectation is borne out as can be seen in (47). (47a) with a definite head sounds more natural than (47b) with an indefinite head. The same judgments hold for the pair with extractions, as shown in (48).

---

<sup>17</sup> The observation that it's to a large extent the naturalness of an utterance without an extraction that determines the felicity of an extraction was first made by Andersson 1982. The discussion in the text owes a lot to his account.

(47) a Jag känner mannen som uppfann den här teorin.  
*I know the man who invented this theory*

- b ??Jag känner ingen som uppfann den här teorin.  
*I know nobody who invented this theory*
- (48) a Den här teorin<sub>j</sub> känner jag mannen som uppfann  $e_j$  .  
*this theory I know the man that invented*
- b ??Den här teorin<sub>j</sub> känner jag ingen som uppfann  $e_j$  .  
*this theory I know nobody that invented*

We have seen that some semantic property like the distributivity of the predicate in the relative clause is largely responsible for determining whether the head should be definite or indefinite. Similar properties of the main predicate, i.e. the predicate that takes the NP + relative clause as an argument, are relevant for choosing e.g. the appropriate number of the head NP. Ultimately, the choice of the main predicate reflects contextual factors, in particular what can be construed as a relevant predication at that point in the conversation. Consequently it is not possible to make a general claim about what type of NP allows for extractions. It is necessary to look at the sentence type, whether it imposes indefiniteness or definiteness requirements on the head, and at the entailment properties of the verbs involved.

## 6. Function of the extracted constituent

In this section we will look briefly at the function of these extractions in dialogue and coherent texts. We will look in particular at the discourse functions of the fronted constituents and their prosodic realisation.

### 6.1 Focus preposing

For the purpose of this paper it will be sufficient to distinguish focus constituents, topic constituents and ground material<sup>18</sup>. The focus part of an utterance corresponds to those constituents that express the new information that the speaker wants the hearer to add to his/her information state and will be represented with underline. The ground represents material that the speaker assumes the hearer already knows about or can reconstruct from the utterance context or general knowledge. Topics will be defined below.

---

<sup>18</sup> See Vallduv' and Engdahl (1996) for a more systematic presentation of these terms with cross-linguistic illustrations.

Fronted constituents can be the focus of an utterance, as shown with the simple question-answer pair in (49). (49 a) elicits a narrow focus answer. The focus is stressed, which is represented by SMALL CAPS.

- (49) a Vad tycker du om att dricka?  
*what do you like to drink?*

- b KAFFE tycker jag om, men inte TE.  
*coffee I know many that like*

The narrow focus KAFFE can be contrasted with an alternative, TE, which the speaker doesn't like. This type of focus preposing is quite common, and if other circumstances permit, may well take place out of a relative clause as in (50).

- (50) 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*

## 6.2. Contrastive topics

The most common function of fronted constituents is probably to be a topic. There are (at least) two types of topics in the Scandinavian languages. There are contrastive topics as illustrated in (51). The fronted constituents have to be stressed. I will represent stressed contrastive topics by **boldface**<sup>19</sup>.

- (51) a Vad tycker du om pojkarna?  
*what do you think about the boys?*
- b **Kalle** tycker jag OM, men **Pelle** AVSKYR jag.  
*Kalle I like but Pelle I can't stand*

The question introduces a contextually relevant set of boys and the answer picks up on individual boys in this set. An English version of (51) might use topicalisation as well. Not surprisingly, a lot of the naturally occurring extraction data in Swedish are of this type.

- (52) Brandmännen frågade mamma och pappa, men  
*the firemen asked mummy and daddy but*

---

<sup>19</sup> In English, contrastive topics are distinguished by a special intonational tune called the B-accent in Jackendoff (1972) and the L+H\* tune in Steedman (1991). Contrastive topics in Swedish may be realised with the same accent as used for focus. In so-called contrastive dislocations as in (i), a distinct accent and lengthening pattern is often used on the first NP.

(i) **Kalle** honom tycker jag OM.  
*Kalle him I like*

**Pelle**<sub>j</sub> var det ingen, som frågade *e*<sub>j</sub>. (G.Knutsson: *Pelle Svanslös*)  
*Pelle there was no one who asked*

(53) Men **Italien**<sub>j</sub> var det ingen som hade tänkt på *e*<sub>j</sub>. (GK 97)  
*but Italy there was nobody who had thought of*

In (52), the fronted constituent **Pelle** is accented, and is interpreted contrastively as can be seen from the context. The effect of the utterance is to highlight the fact that although the firemen asked the parents, nobody thought of asking Pelle, the cat, who had in fact witnessed the fire. The reader may want to test his/her ability to construct a plausible context for (53). The correct answer is given in footnote (20).<sup>20</sup>

I believe that when the term *topicalisation* is used in the linguistic literature, it is normally this kind of contrastively stressed topic that is meant. However, we also find fronted topics which are not stressed.

### 6.3. Continuous topics.

Consider the discourse in (54).

- (54) a Igår köpte jag en ny jacka.  
*yesterday I bought a new jacket*
- b **Den**<sub>j</sub> ska jag ha *e*<sub>j</sub> på mig ikväll.  
*it I'll wear tonight*

(54b) has a fronted pronominal topic. One way of uttering (54b) is with no particular stress on the fronted pronoun *den*. In this use, *den* resumes the NP *en ny jacka* (a new jacket) introduced in the previous sentence. There is no sense of contrast between this jacket and other articles of clothing. Rather the speaker uses the fronting as a way of establishing cohesion between the two utterances. I will refer to this type of initial topic as a *continuous topic*. Andersson (1974, 1982) uses the term *topic movement* for the process which fronts a non-contrastive pronominal topic.

---

<sup>20</sup> (53) was produced in the coffee room of the English Department at the University of Göteborg during a discussion about the 1997 Nobel Prize in literature.

The presence or absence of an accent on the topic constituent determines whether it is to be interpreted as a contrastive or continuous topic. Compare (55b) and (55c), both uttered as follow-ons to (55a).

- (55) a Igår köpte jag en ny jacka.  
*yesterday I bought a new jacket*
- b **Den** ska jag ha på mig ikväll, inte **jeansjackan**.  
*that I'll wear tonight not the jeans jacket*
- c Den ska jag ha på mig ikväll, (#inte jeansjackan)  
*it I'll wear tonight not the jeans jacket*

The stressed fronted *den* in (55b) is clearly a contrastive topic, as shown by the possibility of overtly alluding to an alternative. The unstressed pronoun *den* in (55c) on the other hand does not invoke a contrast set and hence the continuation *inte jeansjackan* (not the jeans jacket) is inappropriate in this context.

I believe that this type of fronting of a non-contrastive continuous topic is not used in English or the Romance languages. Notice that a literal translation of (55c) is impossible in English. Germanic languages use this option to various degrees. German also uses topic fronting of unstressed personal pronouns<sup>21</sup>, but this seems to be restricted to arguments of the matrix clause. In the mainland Scandinavian languages, continuous topic movement of an unstressed pronoun is common also out of complement clauses, i.e. long-distance movement, as shown in (56).

- (56) Den<sub>j</sub> borde tvillingarna ringa till Per och be att få se e<sub>j</sub> på hans  
*it ought the twins phone Per and ask to watch on his*  
 färg-tv. (Hagström 1976 p 145)  
*colour TV*

Not surprisingly, continuous topics can also be fronted out of relative clauses.

- (57) Jag undrade om någon hade en KANOT, och  
*I wondered if anybody had a canoe and*

---

<sup>21</sup> Except for the neuter singular *es* which is replaced by the demonstrative *das*. S-G Andersson (1988) investigates syntactic and pragmatic conditions on dislocations in German.

det<sub>j</sub> var det en flicka som hade e<sub>j</sub> .  
*it it was a girl who had*  
*'and there was a girl who had one'*

There is no sense of contrast between the initial *det* in (57) and some other thing or property. The utterance functions as confirmation that there was indeed a girl who had a canoe. Just as in (55c) and (56), the weak initial pronoun in (57) resumes some topic introduced in the previous utterance.

The kind of long-distance topic movement we have seen in (56) and (57) has no counterpart in English, or in German or Dutch, to my knowledge. But in Swedish, Danish and Norwegian, these fronting operations are used quite frequently, both in spoken language and in narratives, including literature for children.<sup>22</sup> In fact, all types of frontings are probably more common in the Scandinavian languages than in e.g. English.<sup>23</sup>

---

<sup>22</sup> Falster Jakobsen (1996) shows that continuous topics are common in long-distance topic movement in Danish. Kristiansen (1996) discusses the links between accent, topic continuity and referent tracking in Danish.

<sup>23</sup> Many Scandinavian linguists have noted that fronting is 'easier' in Scandinavian languages than in English, but these observations have, to my knowledge, not yet been confirmed by large-scale corpus based contrastive analyses. Cf. Aijmer et al (1996) where a number of interesting contrastive studies are reported based on a corpus of parallel English and Swedish texts.

## 7. Why is Scandinavian different?

In this section we will discuss the intriguing question how come extractions out of relative clauses are possible in the Scandinavian languages but not in related languages such as Dutch, German and English or in the Romance languages. In the preceding section we found that fronted constituents have at least two different information structures, one contrastive and one cohesive. Most languages have some way of fronting contrastive topics, either by topicalisation, as in English and Germanic, or dislocation as in Romance. The Germanic languages also allow fronting of unstressed continuous topics. Whereas this kind of fronting seems limited to arguments and adverbials of the main clause in Dutch and German, the mainland Scandinavian languages also uses long-distance fronting out of complement clauses. This appears to be the fact that distinguishes the situation in the Scandinavian languages from the other languages. It thus seems plausible to connect this fact with the possibility of extracting out of a wide range of clauses, including relative clauses<sup>24</sup>.

If this is correct, the fact that it is possible to extract out of (some) relative clauses in Scandinavian does not point to some specific structural difference between Scandinavian and e.g. the other Germanic languages. Rather the difference seems to have to do with how the languages organise coherent discourses, both dialogue and monologue. It seems that the Scandinavian languages favour an utterance structure with fronted topics, contrastive as well as continuous topics. This means that part of learning these languages involves learning when to use a contrastive topic or a continuous topic and how to realise these, syntactically and prosodically.<sup>25</sup> I believe that these are the important factors that determine whether a particular extraction is wellformed or not. Whether the fronted constituent originates in the same clause, in a complement clause, in an interrogative complement<sup>26</sup>, or in a relative clause is less important. In this paper, we have looked mainly at extractions out of relative clauses. I believe that the same factors are responsible for distinguishing natural, coherent extractions out of adverbial clauses from illformed ones. Two examples of extractions from adverbial clauses are given in (58).<sup>27</sup>

---

<sup>24</sup> Thanks to Lars-Johan Ekerot for prompting me to make this connection more explicit.

<sup>25</sup> Many of the authentic examples in this paper come from children's literature. It would be particularly interesting to look at when e.g. English speaking children and Scandinavian speaking children begin to use frontings. Such a comparative study should clearly include looking at parental input.

<sup>26</sup> See Engdahl 1986 for a discussion of extractions out of embedded questions.

<sup>27</sup> (39b) comes from Hagström (1976 p 149). This short article, written in Swedish, contains a number of authentic examples of extractions. The author discusses these extractions from the point of view of

---

text and discourse strategy, and also makes an interesting psycholinguistic conjecture on the basis of incremental processing. He conjectures that those extractions will be easiest to process where the nature of the fronted element makes it impossible to associate with any of the predicates in the intervening clauses.

- (58) a    Då måste man arbeta på en deskription<sub>j</sub> som det kan ta  
           *then you must work on a description that it may take*  
           flera decennier innan man kommer så långt  
           *several decades before you get so far*  
           att man kan börja formalisera  $e_j$ .                    (KN 97)  
           *that you can start to formalise*
- b    Hennes stämma<sub>j</sub> dröjde det en bra bit in i programmet innan  
           *her voice it took a big chunk of the programme before*  
           man kom underfund med  $e_j$ .  
           *you could figure out*

## 8. Concluding remarks

In this article, I have presented a range of naturally occurring examples involving extractions out of relative clauses in Swedish, Danish and Norwegian. I have tried to identify what the relevant properties are that determine when an extraction is possible or not. As a point of departure we took Ross' structurally defined Complex NP Constraint, reformulated as the subjacency condition by Chomsky. We found that subjacency is not a particularly useful diagnostic since it does not make the right distinctions between acceptable and unacceptable extractions. Instead we looked at what constructions are involved in the acceptable extractions. We found that many examples involved presentational and cleft constructions, where the relative clause provided the informational centre of the utterance.

We also looked at the function of the fronted constituent. A crucial property of the acceptable examples seemed to be that the remainder of the clause was interpreted as a coherent and relevant comment on the fronted constituent in the utterance context. I take this to show that the acceptability of an extraction depends primarily on two pragmatic factors: whether the fronting is motivated in the context and whether the information structure of the clause (i.e. the particular focus-background partitioning that it imposes) fits in with the information states of the participants.

We then addressed the wider question how come there is such a difference between extractions in Swedish, Danish and Norwegian on the one hand and other languages, in particular closely related languages such as Dutch, German and English on the other hand. I have not been able to identify any structural difference. The difference that seems most relevant is the fact that the mainland

Scandinavian languages tend to use fronting both for contrastive and cohesive purposes whereas fronting in English is presumably reserved for contrastive topics. Long-distance fronting of continuous topics was found to be the property that distinguished the mainland Scandinavian languages from Dutch and German. There is probably also a difference in how frequently fronting is used in the different languages. The mainland Scandinavian languages tend to use fronting a lot. However, more comparative research on the discourse functions of fronting and the relative frequencies of different word order patterns is needed. In addition it seems very important to look at the acquisition of extractions in various languages to see when and where differences show up.

In this context, it would be particularly relevant to look at the insular Scandinavian languages Icelandic and Faroese. These languages share many syntactic properties with Danish, Norwegian and Swedish, including the possibility to extract out of embedded questions. Surprisingly enough, extractions out of relative clauses don't seem to occur spontaneously in these languages at all.<sup>28</sup> It would thus be interesting to investigate whether there are differences in the way fronting is used in insular languages compared with the other Scandinavian languages. For instance, is continuous topic movement used in Icelandic and Faroese in the same way as in the mainland languages? If it isn't, then we have further confirmation that topic movement somehow is a precondition for a wider range of extractions. If Icelandic and Faroese use topic movement in the same way as Swedish, Danish and Norwegian, we would have to look for alternative explanations. We could then look at diachronic data to see when and in what text types extractions out of relative clauses started to emerge and try to trace the source of the diverging judgments for these closely related languages.

## References

Aijmer, Karin, Bengt Altenberg and Mats Johansson (eds) (1996) *Languages in Contrast. Papers from a Symposium on Text-based Cross-linguistic Studies*. Lund Studies in English 88, Lund University Press.

---

<sup>28</sup> At the CGSW meeting in Tromsø, Eiríkur Rögnvaldsson, Halldor Sigurdsson and Håskuldur Thrainsson told me that they didn't like any of the examples, apart from possibly (16a). Hjalmar Petersen from the Faroe Islands judges (14a) as strange and (14c) as acceptable, but doesn't like any of the other examples in (14)-(16). Note that none of these linguists have seen the full range of data in this paper. More research is clearly needed.

- 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 J.Stillings (ed.) *University of Massachusetts Occasional Papers in Linguistics II*, Amherst, Ma. (reprinted in Engdahl & Ejerhed 1982)
- Andersson, Lars-Gunnar (1974) Topicalization and relative clause formation. Gothenburg Papers in Theoretical Linguistics 25.
- Andersson, Lars-Gunnar (1982) What is Swedish an Exception to? Extractions and Island Constraints. In Engdahl & Ejerhed (eds).
- Andersson, Sven-Gunnar (1988) Zur Grammatik und Pragmatik von Herausstellungsstrukturen. Sprache und Pragmatik 1, Lund.
- Andersson, Sven-Gunnar and S.Kvam (1984) *Satzverschränkung im heutigen Deutsch*. Studien zur Deutschen Grammatik 24, G.Narr, Tübingen.
- Anward, Jan (1981) *Functions of Passive and Impersonal Constructions: A Case Study from Swedish*. Doctoral dissertation, Department of Linguistics, Uppsala University.
- Brink, Liisa Theilgaard (1987) Relative clauses. ms University of Iceland.
- Chomsky, Noam (1973) Conditions on transformations. In S.Anderson & P.Kiparsky (eds) *A Festschrift for Morris Halle*. Holt, Rinehart and Winston.
- Chomsky, Noam (1986) *Barriers*. MIT Press, Cambridge, Ma.
- Christensen, Kirsti. K. (1982) On Multiple Filler-Gap Constructions in Norwegian. In Engdahl & Ejerhed (eds).
- Cinque, Guglielmo (1991) *Types of A-bar Dependencies*. MIT Press, Cambridge, MA.
- Cinque, Guglielmo (1995) The pseudo-relative and ACC-ing constructions after verbs of perception. In G.Cinque *Italian Syntax and Universal Grammar*, Cambridge University Press.
- Delin, Judy (1989) *Cleft Constructions in Discourse*. Ph.D. Thesis, University of Edinburgh.
- Ejlskjær, Inger (1963) Nogle ejendommeligheder ved relativkonstruktioner i danske ømål og i skånsk, *Acta Phil. Scand.* 26(3), 109-152.
- Ekberg, Lena (1990) Theta Role Tiers and the Locative PP in Existential Constructions. Working Papers in Scandinavian Syntax 46, Lund.

- Engdahl, Elisabet (1980) WH-Constructions in Swedish and the Relevance of Subjacency. *Proceedings of NELS 10*, Cahiers Linguistiques d'Ottawa 9, 98-108,.
- Engdahl, Elisabet (1982a) Restrictions on Unbounded Dependencies in Swedish. In Engdahl and Ejerhed (eds.) pp. 151-174.
- Engdahl, Elisabet (1982b) Interpreting sentences with multiple filler-gap dependencies. *Working Papers in Linguistics and Phonetics 24*, Lund University, pp 21-48.
- Engdahl, Elisabet (1985) Parasitic Gaps, Resumptive Pronouns, and Subject Extractions, *Linguistics*, 23(1), 3-44.
- Engdahl, Elisabet (1986) *Constituent Questions: The Syntax and Semantics of Questions with special reference to Swedish*, Reidel, Dordrecht.
- Engdahl, Elisabet (1997) Scandinavian Relatives Revisited. In D.Faber and F.Sørensen (eds) *From Words to Utterances in LSP*. Copenhagen Studies in Language 20, Samfundslitteratur, Copenhagen.
- Engdahl, Elisabet and Eva Ejerhed (1982) *Readings on Unbounded Dependencies in Scandinavian Languages*, Almquist & Wiksell International, Stockholm.
- Erteschik-Shir, Nomi (1973) *On the nature of island constraints*. PhD dissertation, MIT.
- Erteschik-Shir, Nomi (1982) Extractability in Danish and the Pragmatic Principle of Dominance. In Engdahl & Ejerhed (eds), 175-192.
- Falster Jakobsen, Lisbeth (1996) Sentence Intertwining. In Betty Devriendt, Louis Goossens, and Johan van der Auwera (eds) *Complex structures: a functionalist perspective*, Mouton de Gruyter, Berlin.
- Hagström, Björn (1976) Om satsflätans förutsättningar. En skiss. In Svensson, Lars, A-M Wieselgren and Å.Hansson (eds) *Nordiska studier i filologi och lingvistik. Festskrift tillägnad Gösta Holm*. Studentlitteratur, Lund, 138-150.
- Jackendoff, Ray (1972) *Semantic Interpretation in Generative Grammar*. MIT Press.
- Jespersen, Otto (1974 [1929]) *A Modern English Grammar on Historical Principles*, Allen & Unwin, London and Munksgaard, København.

- Johansson, Mats (1996) Contrastive data as a resource in the study of English clefts. In Aijmer et al (eds).
- Kayne, Richard (1983) Connectedness. *Linguistic Inquiry* 14, 223-249.
- Kayne, Richard (1994) *The Antisymmetry of Syntax*. MIT Press.
- Kristiansen, Erik (1996) Topic continuity and prosody. In E. Engberg-Pedersen, M. Fortescue, P. Harder, L. Heltoft & L. Falster Jakobsen (eds) *Content, Expression and structure. Studies in Danish Functional Grammar*. John Benjamins, Amsterdam.
- Lambrecht, Knud (1988) There was a farmer had a dog: Syntactic amalgams revisited. *BLS 14, Parasession on Grammaticalization*.
- Maling, Joan and Annie Zaenen (1982) A Phrase Structure Account of Scandinavian Extraction Phenomena. In P.Jacobson and G.Pullum (eds) *The Nature of Syntactic Representation*. Reidel.
- Nimb, Sanni (1993) En formel semantisk analyse av danske 'der'-sætninger. Specialopgave i datalingvistik, Københavns universitet.
- Nordberg, Bengt (1977) "Har du Dahlstedt ligger på nattduksbordet?" En syntaktisk egendommelighet i mellansvenskt talspråk. I C.C.Elert, S.Eliasson, S.Fries and S.Ureland (eds) *Dialectology and Sociolinguistics*, Umeå Studies in the Humanities 12, 113-136.
- Nordgård, Torbjørn (1991) *A GB-related parser for Norwegian*, Doctor artium thesis, Univ. of Bergen.
- Nølke, H. (1984) Clefting in Danish. *Nydanske studier* 14, 72-111.
- Paggio, Patrizia (1996) *The Treatment of Information Structure in Machine Translation*. Ph.D, thesis, University of Copenhagen and Centre for Language Technology.
- Platzack, Christer (1974) An exception to the complex NP constraint. *Papers from the first Scandinavian Conference of Linguistics*, Dept of Linguistics, University of Göteborg.
- Platzack, Christer (forthc.) A Complement-of-N<sup>0</sup> Account of Restrictive and Non-Restrictive Relatives: The Case of Swedish.
- Prince, Ellen (1981) Towards a taxonomy of given-new information. In P.Cole (ed) *Radical Pragmatics*, Academic Press, p 223-255.
- Quirk, Randolph, Sidney Greenbaum, Geoffrey Leech and Jan Svartvik (1972) *A Grammar of Contemporary English*. Longman.

- Rizzi, Luigi (1990) *Relativized Minimality*. MIT Press.
- Ross, J.R. (1967) *Constraints on Variables in Syntax*. Ph.D. dissertation, MIT.
- Steedman, Mark (1991) Structure and intonation. *Language* 67, 260-296.
- Strudsholm, Erling (1997) *Relative situazionali in italiano moderno*. Ph.D. thesis, Department of Romance Philology, Copenhagen University.
- Sundman, Marketta (1980) *Existentialkonstruktionen i svenska*. Åbo Akademi.
- Sveen, Andreas (1996) *Norwegian Impersonal Actives and the Unaccusative Hypothesis*. Dr.Art. thesis. Department of Linguistics, University of Oslo.
- Taraldsen, Knut Tarald (1981) On the Theoretical Interpretation of a Class of "Marked" Extractions. In A.Belletti, L.Brandi & L.Rizzi (eds) *The Theory of Markedness in Generative Grammar*, Scuola Normale Superiore, Pisa, p. 475-516.
- Taraldsen, Knut Tarald (1982) Extraction from relative clauses in Norwegian. In Engdahl & Ejerhed (eds).
- Taraldsen, Knut Tarald (1986) *som* and the Binding Theory. In L.Hellan & K. K. Christensen (eds) *Topics in Scandinavian Syntax*, Reidel, Dordrecht.
- Vallduví, Enric and Elisabet Engdahl (1996) The linguistic realisation of information packaging *Linguistics*. 34, pp.459-519.