Scandinavian Relative Clause Extractions

Apparent restrictions *

Fredrik Heinat & Anna-Lena Wiklund
Linnaeus University & Lund University

Abstract

This brief article investigates the restrictions on Mainland Scandinavian relative clause extraction that have figured in the literature on island constraints. The conclusion is that none of these restrictions can be regarded as constraints on relative clause extraction per se and therefore that the peripheral status standardly assigned to Mainland Scandinavian relative clause extraction cannot be maintained.

1 Introduction

It has long been recognized that the Mainland Scandinavian languages Danish, Norwegian, and Swedish are peculiar from a cross-linguistic perspective in that they fail to show the constraints on extraction usually observed for complex DPs. These languages allow relative clause extraction (RCE), traditionally described as a violation of the Complex NP Constraint or a strong island violation (Ross 1967 and Chomsky 1977, respectively), cf. (1-a) and (1-b) (from Engdahl 1997:54). Early references include Erteschik-Shir (1973), Allwood (1975), Engdahl & Ejerhed (1982); Taraldsen (1981); Maling & Zaenen, (1982); Andersson (1982).2

---

*We wish to thank the audiences at the Grammar in Focus Symposium 2014 (Feb 5-6), Lund University, and the 2014 Annual Meeting of the Linguistics Association of Great Britain (Sep, 1-5), University of Oxford, for helpful comments on this paper. We are also indebted to Eva Klingvall, Damon Tutunjian, and Johan Brandtler for comments and advice.

1Replacing who in (1-b) by that does not change the unacceptable status of the extraction.

2Only a handful of other languages have been reported to show the same exceptionality as the Scandinavian languages: Japanese and Korean (Kuno 1973b: 239-240); Akan (Saah & Goodluck 1995). However, the status of these as true RCE is debated, see e.g. Han & Kim (2004) and Cinque (2010).

(1) a. [Såna blommor], känner jag [en man [som säljer _]]. (Swe.)
    such flowers  know I  a man that sells
b.*?[Those flowers], I know [someone [who sells _]].

Whereas the English example can be improved by a resumption strategy (which
makes it indistinguishable from a left dislocation), resumption decreases ac-
ceptability in Swedish and the other Mainland Scandinavian languages (MSc):³

(2) a. Såna blommor känner jag en man som säljer (*dem). (Swe.)
    such flowers  know I  a man that sells (them)
b. *?Those flowers I know someone who sells them.

³Examples are borrowed from Engdahl (1997:54).

The relative contribution of structural and non-structural constraints on the pro-
cessing and acceptability of long-distance dependencies of this kind has been
a central issue in ongoing debates that span the research fields of both theo-
retical syntax (see Boeckx 2012 for an overview) and sentence processing (see
e.g. Sprouse & Hornstein 2013). Although the MSc data have been regularly
cited in the syntax- and processing-oriented literature, there is to date no ac-
count of RCE in Mainland Scandinavian that holds up under closer scrutiny,
see Engdahl (1997) and more recently Christensen & Nyvad (2014) and Müller
(2015) for discussion. Naturally, much of previous research on RCE in Main-
land Scandinavian has focused on identifying the right conditions for RCE from
the perspective that syntactic islands are universal. The guiding hypothesis has
been that RCE in Mainland Scandinavian is severely constrained, as expected
from theoretical considerations, but that the acceptability of the hard-to-explain
cases may be derived from either of the following: (i) discourse-organizational
factors (as in Ertechik-Shir & Lappin 1979), (ii) island obviation by way of
covered resumption (as in Cinque 1990), or (iii) structural reanalysis during pars-
ing (as in Kush et al. 2013). For arguments against a discourse-based approach,
see Boeckx (2012) and §8 below. (2-a) above seems incompatible with an ap-
proach to Swedish RCE in terms of island obviation, since overt resumption
is possible in other contexts (cf. Engdahl 1997). Finally, compelling argu-
ments against structural reanalysis during parsing are presented in Christensen
& Nyvad (2014) and Müller (2015). In other words, the problems posed by the
Scandinavian data for theories about island constraints, whether cast in terms
of processing or syntax, are far from solved. We concur with Hofmeister & Sag (2010) and Boeckx (2012) in not being satisfied with how counterexamples to island constraints have been dealt with in the literature, both empirically and theoretically. What follows is an attempt to clean up the empirical part of the Scandinavian data to deepen our understanding of the unexpected void of island effects in RCEs in these languages. We show that beyond the more general factors, like the Subject Condition and factors pertaining to discourse and processing, which are known to play a role in perceptions of acceptability of RCE also in other languages, none of the restrictions proposed to condition RCE in Mainland Scandinavian actually hold. This observation tells us two things. First, assigning a peripheral status to the exceptions is not an option. Second, there seems to be a real (although perhaps a fuzzy) difference between Mainland Scandinavian and other languages with regard to perceived acceptability of RCE, in line with the early observations. Analytical options for further exploring the exceptionality of Mainland Scandinavian in this respect are outlined in the concluding section.

2 Relative clause extractions

It has consistently been claimed that only a subset of restrictive relative clauses allow extractions in Mainland Scandinavian and that a number of conditions have to be met for the extractions to be acceptable. Restrictions on Mainland Scandinavian RCEs that have been reported relate to the head noun, the extracted element, the extraction gap (the gap linked to the extracted XP), the matrix predicate, the matrix subject, the position of the relative clause, as well as information-structural factors. The relevant string is schematized in (4) with the verb second word order, characteristic of Scandinavian matrix clauses.

\[
(4) \text{XP}_i \rightarrow V_{\text{matrix}} \text{DP}_{\text{subj}} \ldots [\text{DP}_{\text{head noun}} [RC \text{ som } \ldots V_{\text{emb } \ldots i} ]] \]
\]

those flowers know I ... a man that ... sells

\[4\] Appositive relative clauses do not allow RCE (Engdahl 1997:58) and will therefore not be discussed here:

\[
(3) *\text{[Den teorin], känner jag en man som för övrigt tror på }\ldots. \quad \text{(Swe.)}
\]

that theory-the know I a man that by the way believes in

\[5\] The matrix V2 word order excludes a dislocation structure of RCE (cf. Engdahl 1997), since that structure does not yield surface V2 word order.
Because clefts and presentational constructions seem to be liberal with regard to extraction possibilities also in other languages, like English (cf. Kush et al. 2013), we limit our discussion to cases where we seem to find the largest difference in acceptability between Mainland Scandinavian and other languages, as in (1). We thus exclude examples like (5) below (from Engdahl 1997:57), cf. the English example in (6) (from Chung & McCloskey 1983:708).

(5) Vilket ord var det ingen som kunde stava rätt till. (Swe.)
Which word was there nobody that could spell rightly to

(6) This is a paper that we really need to find someone who understands.

3 The head noun

3.1 Definiteness, specificity, and abstractness

Definiteness and specificity of the head noun are factors that have been noted to influence the acceptability of RCEs (Erteschik-Shir & Lappin 1979; Andersson 1982; Taraldsen 1981; Engdahl & Ejerhed 1982; Huber 2002; van Valin 2005; Cinque 2010) and NP-extraction in general (Chomsky 1973; Fiengo & Higginbotham 1981; Davies & Dubinsky 2003). Contrasts like the one in (7) below have been used to claim that RCE is restricted to relative clauses following indefinite head nouns (b-example from Engdahl 1997:69).

(7) a. [Den teorin], känner jag en man som tror på __i. (Swe.)
that theory-the know I a man that believes in
b. ??[Den teorin]i, känner jag mannen som tror på __i.
that theory-the know I man-the that believes in

As noted by Engdahl, the examples provided to show these effects are, however, disfavored for pragmatic reasons. (7-b) is odd even without the extraction because of a uniqueness requirement on the object referent, cf. (8-a).6 Once this factor is controlled for, e.g. by changing the embedded predicate, the sentence is fine, (8-b), and sub-extraction from the definite head noun is impeccable in Swedish, cf. (8-c) (from Engdahl 1997:70). This is also true for Norwegian

---

6The sentence in (7-b) is fine in a (somewhat far-fetched) context where only one person is relevant as a believer of the theory under discussion, cf. Engdahl (1997).
The head noun in the examples is not only definite but also has specific reference, showing that the head noun in RCEs need not be non-specific, contra Taraldsen (1981).

(8) a. Jag känner mannen som tror på den teorin. (Swe.)
   I know man-the that believes in that theory-the
b. Jag känner mannen som kom på den här teorin.
   I know man-the that made up this here theory-the
c. [Den här teorin], känner jag mannen som kom på _i.
   this here theory-the know I man-the that made up.

(9) [Den her teori], kender jeg kvinden der fandt på _i. (Da.)
   this here theory-the know I woman-the that made up

Note that processing studies have identified a sensitivity to the referential properties of intervening nominals; there is a difference in cost between intervening indefinite nouns (less costly) and definite nouns (more costly) along a filler-gap path (Warren & Gibson 2002; Hofmeister & Sag 2010). The point that we wish to make here is that if we remove definiteness as a factor interfering with judgments, we seem to find a difference between Mainland Scandinavian on the one hand and other languages on the other with regard to perceived acceptability of RCE. That is, even though the English counterpart of (8-c) (definite head noun) may be perceived as worse than the English counterpart of (7-a) (indefinite head noun) (cf. Kluender 1992), the counterpart of (7-a) is not acceptable in English, in contrast to Swedish. Note finally that the data cannot be explained in terms of a ‘metaphysical’ concept denotation of the head noun (which enables extraction in English, too) in contrast to a physical token denotation (Davies & Dubinsky 2003). The examples above all involve concrete head nouns, which are claimed to ban extraction in English (ibid.). Despite this, extraction is possible in Swedish (see also section 6).

3.2 The relativization gap

A number of studies have made the observation that there seems to be a subject restriction on the head noun in the sense that the relativization gap has to be the
subject position of the relative clause, see Allwood (1976), Andersson (1982), Kluender (1992), Engdahl (1997), Cinque (2010), Kush et al. (2013). This restriction has been crucial in some attempts to account for RCE and other island violations, see Chung & McCloskey (1983), Chomsky (1986), Platzack (1999), Cinque (2010), Kush et al. (2013). For example, the contrast between (10) and (11) below is used by Kush et al. (2013:242) to argue in favor of a small clause analysis of RCE in Scandinavian and English.

(10)  
a. [Den teorin], finns det ingen lingvist som tror på _i. (Swe.)  
that theory-the exists it no linguist that believes in

b. *[Den här lingvisten], finns det ingen teori som _i tror på.  
this here linguist-the exists it no theory that believes in

As noted by Engdahl (1997), however, Swedish is subject to a that-trace restriction (cf. Lohndahl 2009) and (10-b) is therefore ruled out for independent reasons. If we control for the that-trace restriction, it is evident that there is no subject restriction on the relativization gap in Mainland Scandinavian RCE. Examples showing this can be construed by using a ditransitive verb, (11-a), or by using adjunct instead of argument extraction, (12-b). In these cases, the relativization gap is in the object position, yet RCE is unproblematic, cf. (12-a). The possibility of adjunct extraction in (12-b) also illustrates that Mainland Scandinavian RCE is not amenable to an account in terms of a weak rather than a strong island violation, as there is no perceived difference between argument and adjunct extraction.

(11)  
a. Jag vet tre saker som han vill ge Lisa. (Swe.)  
I know three things that he wants give Lisa

I know two things that you should do in Paris

(12)  
a. Lisa; vet jag tre saker som han vill ge _i.  
Lisa know I three things that he wants give

b. [I Paris;] vet jag två grejer som man bör göra _i.  
in Paris I know two things that you should do
3.3 Clause function

Erteschik-Shir & Lappin (1979) observe that the head noun must be a direct object of the matrix verb in cases of RCE. This is correct but, as well known, extraction from a (displaced) subject is constrained for independent reasons, regardless of whether it involves a relative clause or not: this is the so-called Subject Condition (Chomsky 1973; Huang 1982); see Kluender (2004), Boeckx (2012), and Haegeman et al. (2014) for recent discussion.

4 The matrix predicate

Along with the subject restriction on the relativization gap, shown to be wrong in the above section, Kush et al (2013) list choice of embedding verb as a key structural factor that influences acceptability of RCE, their so-called predicate restriction. The claim is that RCE is acceptable only with matrix verbs that also select small clauses, because in these cases the parser can reconstruct the complex noun phrase as a small clause (from which extraction is not blocked). The predicate restriction has recently been investigated in detail in acceptability judgment experiments on both Swedish and Danish speakers, see Müller (2015) and Christensen & Nyvad (2014), respectively. Neither of these studies found any statistically significant differences between small-clause selecting vs. non-small clause selecting verbs. Müller (2015) provides a number of examples from the literature where RCE occurs with verbs that cannot select small clauses, all perceived as acceptable by native speakers of Swedish, cf. (13) (from Teleman et al. 1999[4]:423).

(13) a. Akupunktur, brukar det delta en läkare som kan _i vid våra seminarier.

b. [Piratdelar till Volvo]_i har jag tagit reda på en som säljer _i.

7 The contrast between Swedish and English is proposed to derive from differences with regard to the relative pronoun/complementizer. See Christensen & Nyvad (2014) and Müller (2015) for counterexamples to that proposal.
Even though acceptability arguably varies to a certain degree with properties of the intervening verb (Ertechik-Shir 1973; Kothari 2008; Hofmeister & Sag 2010), the examples in (13) refute the suggestion that the matrix verb in cases of RCE is semantically light by necessity, as has been claimed by Allwood (1976) and Ertechik-Shir & Lappin (1979:57). Matrix verbs in RCE clearly may have complex event structures and may also be rather specific with regard to the manner component of the event referred to.

5 The matrix subject

Erteschik-Shir & Lappin (1979:57) claim that RCE in Danish is subject to a person restriction such that extraction is only possible when the matrix subject is 1st person. This observation seems incorrect. As the examples from Danish in (14) show, RCE is also possible with 2nd and 3rd person matrix subjects (Ken Ramshøj, p.c.).

(15) a. [Den slags musik]i kender du vist ingen der kan lide _i.
   that kind music know you PRT nobody who can like
b. [Den slags musik]i kender hun vist ingen der kan lide _i.
   that kind music know she PRT nobody who can like
c. [Den slags musik]i kender Mille vist ingen der kan lide _i.
   that kind music know Mille PRT nobody who can like

6 The extracted element

Putative restrictions on the extracted element can be derived from well-known factors that influence the possibility to front constituents in general. That demonstratives extract more easily than indefinite noun phrases (Allwood 1976:11) is because the former front more easily than the latter, see Engdahl (1997) for discussion. Note that even non-referential constituents are extractable (a-example from Engdahl 1997:57):

---

8Norwegian (Marit Julien, p.c.) and Swedish are similar in this respect:

(14) [Såna blommor]i känner du/hon/Lisa väl en man som säljer _i.
    such flowers knows you/she/Lisa PRT a man that sells (Swe.)
Davies & Dubinsky (2003) propose that extraction from NP in English is restricted to elements that count as participants in the lexical conceptual structure of the head noun, which is why extraction from NPs involving concrete nouns is never possible; these do not have an argument structure, nor do they imply participants. The possibility to extract from NPs involving concrete nouns (noted in §3 above) shows that the lexical conceptual approach to possibility of NP-extraction cannot be applied to Mainland Scandinavian NP-extractions and that no participant restriction holds for the extracted element in Mainland Scandinavian RCE.

7 The relative clause

According to Taraldsen (1981:486), the relative clause from which extraction has taken place must be clause-final (cf. Teleman et al. 1999[4]:423):

(17) a. *[Såna böcker] tar hon en kompis som läser _i med sig. 

   such books  takes she a friend that reads  with herself

b. [Såna böcker] tar hon med sig en kompis som läser _i. 

   such books  takes she with herself a friend that reads

The clause final restriction is not specific to RCE but seems to hold also for other extractions barring clefts and presentational constructions (Kuno 1973a). In view of example (13-a) above however, the restriction is not categorical. In that example, the sentence-final PP belongs to the matrix predicate delta ‘attend’, not to the relative clause. Despite this, extraction is possible.

8 Information-structural factors

Before concluding this excursion in reported restrictions on RCE, we wish to make a note on claims about information structural factors, reported to play a role for RCE and for extraction more generally (Erteschik-Shir 1973; 1982;
Erteschik-Shir & Lappin 1979; van Valin 2005; Goldberg 2013). The claim is that back-grounded constituents are extraction islands. Back-grounded constituents are those which are not interpreted as pragmatically dominant in discourse in the terminology of Erteschik-Shir (1973) and Erteschik-Shir & Lappin (1979). The operational test adopted to determine the relevant dominance relations among constituents involves:

"[...] placing the entire complex sentence in a context of direct discourse and denying first the matrix sentence and then the embedded sentence. If it is not possible to deny the complement this indicates that the environment defined by the matrix excludes the possibility of interpreting the complement as dominant"

(Erteschik-Shir & Lappin 1979:46)

The relative clause in (18-a) cannot be interpreted as pragmatically dominant, according to this test, cf. (18-c). In this respect, Swedish is not different from English.

(18)  
(a) Jag känner dom som utvecklade den teorin. 
I know them that developed this theory-the

(b) Nej, det gör du inte. 
no, that do you not
‘No, you don’t.’

(c) #Nej, det gjorde dom inte. 
No, that did they not
‘No, they didn’t.’

Still RCE is possible in Mainland Scandinavian, but not in English, see (19). An information-structural account of the difference between Mainland Scandinavian and English is therefor not likely to be on the right track. For discussion and arguments against discourse-based accounts of islands more generally, see Boeckx (2012:28-29).

(19)  
(a) [Den teorin], känner jag dom som utvecklade _i. 
this theory-the know I them that developed

(b) *[This theory], I know the guys that developed _i.
9 Conclusion

We have shown that the restrictions on Mainland Scandinavian RCE that have figured in the literature vanish under closer scrutiny. Some can be reformulated as preferences derivable from semantic, pragmatic, and processing factors, which influence the acceptability of complex structures more generally (cf. Klunder 1992; 2004; Hofmeister & Sag 2010; Sprouse & Schütze 2013) and thus cannot be regarded as constraints on RCE per se. When we remove these factors, the MSc languages stand out with regard to acceptability of RCE, in line with the early observations. That is, the difference between the Mainland Scandinavian languages and languages like English with regard to acceptability of relative clause extraction seems real although may appear fuzzy in the presence of the above-mentioned factors.

One possibility is that Swedish RCEs, although intuitively acceptable, pattern more like island structures in terms of processing. Tutunjian et al. (submitted) address this hypothesis in an eyetracking while reading study. They conclude that Swedish RCEs pattern closer to non-island structures (that-clause extraction) than to island structures (non-restrictive relative clause extraction), in terms of processing. This leaves us with two possibilities. The first is that Swedish RCEs do not involve island structures, in which case we need to look harder to find a tenable account of the phenomenon in terms of structure. The second possibility is that Swedish relative clauses are exempted from island constraints, in which case there is true variation in the island constraints themselves (see Phillips 2013, for discussion).

References

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, 1–20. Amherst, Ma.


