Subject requirement, complementizers 
and optionality in Scandinavian*

IRENE FRANCO 
University of Leiden – LUCL 
i.franco@hum.leidenuniv.nl

This paper addresses the issue of optionality in complementizer insertion/drop and related subject-object asymmetries in A'-dependencies, by focusing on the variation attested in Mainland and Insular Scandinavian. On the one hand, there is extensive literature discussing subject/object asymmetries in complementizer insertion/drop in Mainland Scandinavian A'-dependencies. On the other hand, Icelandic does not display such an asymmetry but is the only Scandinavian language with productive Stylistic Fronting, whereas Faroese displays a mixed behavior. This paper proposes a unifying analysis for Mainland Scandinavian, Icelandic and, potentially, Faroese, and argues that SF and SOM insertion are two different strategies to check one of the (at least) two features encoded on the lowest CP head Fin⁰. Such feature-checking mechanism ensures proper anchoring of the clausal semantic content (i.e. event structure and clausal arguments) to the discourse. The account is based on comparative synchronic and diachronic facts, such as recent findings on long extractions in Mainland Scandinavian, with respect to drop/insertion of the complementizer SOM, and the diachronic evolution and loss of SF in Old Swedish (Delsing 2001), by contrast to Icelandic.

1. Introduction
If we take a comparative look at the pattern of complementizer insertion and complementizer drop in Scandinavian subordinate clauses (Thráinsson 2007: 443-460), we immediately notice that Icelandic behaves differently. The following

* It is an honor for me to dedicate this paper to Luigi Rizzi, my former supervisor. Encouraged by him, I have discovered the fascinating world of linguistics. I have been so lucky to have him first as a teacher, then as a mentor and supervisor, who could patiently listen to my doubts, answer my questions and discuss theoretical and practical issues at any time. His advice was precious for my career and if I am a linguist (with a salary) now, I owe it to him. Grazie di cuore, Luigi!
A preliminary version of this paper has been presented at GIST 3, “Cartographic structures and beyond”, University of Ghent in 2011 and at the International Conference of Nordic and General Linguistics, workshop on the Left Periphery in Scandinavian, April 18th-20th 2012. I thank the audience at the above mentioned events for useful comments, specifically: Halldór Sigurðsson, Lars-Olof Delsing, David Adger, Tolli Eythórsson, and Jóhannes Gíslí Jónsson. I am also very thankful to my friend and colleague Eefje Boef, with whom I carried out part of the research that led me to achieve the results hereby presented. Comments are welcome. Usual disclaimers apply.

2012 Irene Franco
Enjoy Linguistics! Papers offered to Luigi Rizzi on the occasion of his 60th birthday
examples show that Icelandic does not allow drop of the relative particle SEM, cf. (1b), whereas Mainland Scandinavian (MSc, in this paper represented primarily by Swedish and Norwegian varieties) does, in object relative clauses: SOM is dropped in (1a). By contrast, C-drop\(^1\) in declarative complements is subject to variation in MSc. In many varieties, the complementizer AT/ATT cannot be dropped if it introduces a declarative clause whose subject has been extracted, cf. (2a), which indicates that the variety is sensitive to C-trace effects\(^2\). The situation is the opposite in Icelandic: C-drop is only marginally accepted, despite the subject extraction, cf. (2b).

(1)  
\begin{enumerate}  
\item a. *Bilen (som) vi mötte er mycket vacker.*  
(Swedish)  
Car.the SOM we met is very beautiful  
“The car we met is very beautiful”  
\item b. *Detta er maðurinn *(sem) María hitti i gær.*  
(Icelandic)  
This is man.the SEM Mary met yesterday  
“This is the man Mary met yesterday”  
[Thráinsson 2007: 447, 8.130; 407, 8.37b]  
\end{enumerate}

(2)  
\begin{enumerate}  
\item a. *Denne boka veit eg (?*at) __ vil interessere deg.*  
(Norwegian)  
this book know I that will interest you  
“This book, I know will interest you”  
\item b. *Þetta vonum við ?*(að) __ muni einhverntíma verða gert*  
(Icelandic)  
This hope we that shall sometimes become done  
“This, we hope will be done at some point”  
[Thráinsson 2007: 446, 8.129]  
\end{enumerate}

Basically, the differences between the properties of Icelandic and other Scandinavian languages, with respect to complementizer drop/insertion, can be summed up as follows:

i) C-drop in Icelandic is highly restricted, cf. (1b);

ii) Icelandic does not show C-trace effects, cf. (2b), but, rather, anti-C-trace effects.

By contrast, Mainland Scandinavian (and Faroese) do not behave uniformly with respect to i) and ii): thus, it is preferable to refer to the single languages (Norwegian, Swedish, Danish…) rather than to MSc. The problem is that even this choice would not be appropriate, given the broad dialectal variation encountered within each macrovariety. The phenomena investigated in this paper

\(^1\) In the paper I use the generic term C-drop to refer to drop of complementizer, (relative) particle and other overt subordinators.

\(^2\) There is significant variation with respect to C-trace effects in Norwegian, whereas in Swedish C-trace effects are generally shared among all the speakers. In this case the complementizer in question is the declarative AT (Norwegian) or ATT (Swedish), ScanDiaSyn fieldwork data. The presence of C-trace effects in Norwegian is highly controversial, since there is enormous variation among speakers (even of the same family, fieldwork data in Boef & Franco, in prep.). Perhaps the great variation is to be attributed to the fact that there is no such thing as spoken standard Norwegian, but rather many different coexisting grammars grouped as Norwegian. The specific parameters of variations are yet to be identified.
have been treated by previous literature as more or less uniform in each language, (Taraldsen 1986, Allan et al. 1995:193, Thráinsson 2007), although some works acknowledge the microvariation in the complementation possibility of Norwegian (Áfarli 1994) and Danish (Vikner 1991, Mikkelsen 2002).

In this respect, Icelandic stands out as a more conservative and uniform language, without significant synchronic variation. The general question, then, is:

A) which properties are related to C-drop/insertion (and C-trace effects) and how come such properties receive such a variable morphosyntactic realization in some languages (Norwegian, Swedish, Danish, but also non-Scandinavian languages, like English, etc.), but not in others (i.e. Icelandic).

For the present purposes, the answer is searched within Scandinavian languages, which offer diverse patterns of possibilities with respect to C-drop/insertion, but differ minimally in other grammatical properties. The paper does not deal specifically with AT/ATT-trace effects (discussed in Boef and Franco, in prep.), but concentrates on insertion/drop of the relative particle SOM/SUM/SEM in the various Scandinavian languages.

One significant aspect that distinguishes Icelandic from Faroese and Mainland Scandinavian is the special pro-drop status of the language (cf. Vikner 1995, Svenonius 2000, Biberauer et al. 2010, a.o.). On the one hand, this paper argues that the setting of the pro-drop parameter in Icelandic (expletive and semiargumental post-verbal pro-drop) plays a role in C-drop possibilities of this language (section 3). On the other hand, the possibility to omit the subject has been related to another peculiarity of this language, namely Stylistic Fronting (SF), cf. Maling (1980, 1990). SF is a specific type of fronting that moves various types of constituents to a preverbal position, is restricted to clauses lacking an overt preverbal subject and may occur in subordinate clauses (e.g. embedded Wh-, relative clauses; Maling 1980, 1990, Jónsson 1991, Holmberg 2000). It has been suggested that SF is an EPP-driven mechanism (Holmberg 2000, Ott 2009) or that SF actually licenses subject extraction (Rögnvaldsson & Thráinsson 1990). Poole (2007) shows that the latter argument contains a logical error. In any case the most problematic aspect for these accounts is the optionality of Icelandic SF, which is at best reduced to a difference in the EPP-checking requirement. By contrast, the optionality of SF in other non-Germanic varieties has been explained as a difference at the level of information structure (Fisher and Alexiadou 2001 for Old Catalan, Mathieu (2006) for Old French, cf. also Franco (2009, in prep.) for Old Florentine). A related question then is:

B) what is the role of SF in Icelandic grammar and whether its optionality can be related to the optionality (and variability) of C-drop/insertion in the other Scandinavian languages.

Question A) and question B) are at the basis of the investigation presented in this paper, which is structured as follows: Section 2 presents novel facts on Mainland Scandinavian long A'-dependencies, showing that previous proposals about C-drop/insertion cannot properly account for these facts; section 3 illustrates Icelandic SF and discusses the limits of some related accounts; section 4 presents the proposal: the facts illustrated up to this point are related to conditions imposed by information structure via a feature-checking mechanism that enables anchoring of the clausal semantic content to the discourse. The features involved in such mechanism are encoded on FinP (cf. Rizzi 1997, 2004, Haegeman 2006) and expressed as [finiteness] and [definiteness]. The latter spells out the nominal
counterpart of FinP, in relation to the double (N; V) nature of C (cf. Déchaine & Tremblay 2011, Muyskens 2008, Chomsky 1995). Section 5 illustrates a series of borne-out predictions following from the proposal, and section 6 is the conclusion.

2. Mainland Scandinavian

There is extensive literature discussing the distribution of complementizers in A'-dependencies in MSc (Taraldsen 1986, Allan et al. 1995:193, Faarlund et al. 1997:193, Teleman et al. 1999:555, Thráinsson 2007, Lohndal 2007). Specifically, some works discuss the pattern of distribution of the complementizer SOM in relative and embedded Wh- clauses. SOM is a specific form used to introduce A'-dependencies (some authors argue it is a relative operator, cf. Taraldsen 2001, Vikner 1991; others a complementizer, Thráinsson 2007: 447). In relative clauses, SOM is obligatory only with subject extraction, (3a), otherwise it can be dropped:

(3) a. Jeg kjenner mannen *(SOM) kom hit (Norwegian)
    I know man.the SOM came here
    “I know the man who came here”

b. Jeg kjenner mannen (SOM) Maria skal møte i morgen
    I know man.the SOM Maria will meet tomorrow
    “I know the man Mary will meet tomorrow”

The same pattern given in (3) holds for Swedish, whereas Danish has different possibilities related to the availability of another form, DER, which is synchretic with the subject expletive, and is used in subject extraction only (cf. Vikner 1991; Taraldsen 2001; Mikkelsen 2002):

(4) a. Vi kender en lingvist (DER/SOM) vil læse denne bog (Danish)
    We know a linguist DER/SOM will read this book
    “We know a linguist who will read this book”

b. Vi kender en bog (*DER)/(SOM) denne lingvist vil læse
    We know a book DER/SOM this linguist will read
    “We know a book that this linguist will read”
    [Vikner 1991]

Because of the existence of such lexical element (DER), the Danish system is more complex. I will briefly address the Danish cases in section 5 (but see Boef & Franco in prep.), and for the remainder of the paper I will focus on Norwegian and Swedish. Accordingly, I will refer to Norwegian and Swedish as to MSc, in this paper.

It has been proposed that SOM-insertion in Norwegian depends on the licensing of a subject feature in the subject position (Taraldsen 2001), and with a function similar to the -i morphology of the French qui, for the que/qui alternation (Kayne 1976). In the example below, a pair illustrates the subject/object asymmetry related to the complementizer morphology in French:

(5) a. Je connais l’homme qui/*que __ est venu ici (standard French)
    I know the man who is come.PART here
    “I know the man who came here”
b. Je connais l’homme que/*qui Marie va rencontrer __
   I know the man that Mary goes meet.INF
   “I know the man that Mary is going to meet”

In the GB framework, SOM-insertion would be analyzed as a strategy to enforce proper government and prevent ECP effects with subject extractions, an account that has been fairly fortunate in the literature on subject/object asymmetries (Taraldsen 1978, 1986, Rizzi 1990, Rizzi & Shlonsky 2006: 17). Taraldsen (2001: 168) analyzes Norwegian SOM as either an expletive or an argument. Specifically, he argues that in non-subject extractions (cf. 3b), SOM is optionally doubling the argument, since the head of the relative is not an operator in Norwegian. By contrast, SOM in subject relative clauses can be either an expletive or an argument, cf. (3a). The reason why Taraldsen (2001) distinguishes between expletive and argumental SOM is related to SOM-distribution in Norwegian embedded Wh-questions. Norwegian embedded Wh-questions on the subject require that the Wh-element be followed by SOM, as in (6), whereas SOM is ungrammatical in object questions, as in (7).

(6)  Vi vet ikke hvem *(SOM) oppfant ostehøvelen  (Norwegian)
    we know not who SOM invented the cheese slicer
    “We do not know who invented the cheese slicer”

(7)  Vi vet ikke hvem (*SOM) de her ansatt
    we know not who SOM they have hired
    “We do not know who they have hired”       [Taraldsen 2001: 168, 4-5]

According to Taraldsen (2001:168), the ungrammaticality of SOM-insertion in object extractions as in (7) is due to the fact that “interrogatives have operators in Spec-CP, and operators cannot be doubled by pronouns. Hence, the SOM in (4) [here (6)] can only be an expletive, and (5) [here (7)], where SOM cannot be an expletive, is ungrammatical with SOM.” A first problem with this idea is that it cannot be straightforwardly extended to other Scandinavian languages with SOM-complements, such as Swedish. In Swedish, SOM is not ungrammatical in indirect object Wh-questions, but just optional, as (8b) shows.3

(8)  a. Hon undrade vem *(som) kom
    She wondered who SOM came
    “She wondered who came”

    b. Hon undrade vem (som) Johan träffade
    She wondered who SOM Johan met
    “She wondered who John met”

So far I have mainly addressed facts concerning relative clauses: for reasons of space, I will not discuss embedded Wh-questions in detail here. The pattern of SOM-insertion in embedded Wh-questions is a bit more complex, given the

3 The same difference is reflected in long extractions. This characteristic is arguably a result of diachrony, cf. Section 5.
crosslinguistic differences between Swedish and Norwegian just outlined. However, SOM-insertion in these clauses can be properly accounted for under the hypothesis that diachronic evolution played a role in shaping the complementation system in the two languages, and will be addressed again in sections 4 and 5. Taraldsen (2001: 168) draws a parallel between Norwegian SOM, which he analyzes as an expletive, and the -i morphology on the complementizer qui in French, which would be reminiscent of the expletive pronoun il. Taraldsen compares the distribution of SOM with the French que/qui alternation, and argues that the presence of an expletive element in C is required in order to satisfy the EPP. Rizzi and Shlonsky (2006: 17), alternatively argue that –i morphology on qui checks the subject phi-features by directly C-commanding the subject position in Spec,IP. SOM arguably functions alike and, as such, enables subject extractions. However, both these accounts encounter some problems once we observe the pattern of SOM-distribution in long A’-dependencies. Consider below the examples of long extractions out of a declarative complement embedded under a verb of saying (å si = to say, for Norwegian) or a semi-factive, propositional attitude predicate (å hoppa = to hope, for Swedish).

(6) a. Jeg kjenner mannen (SOM) du sa (*SOM) kom hit. (Norwegian)
   I know the man SOM you said SOM came here
   “I know the man you said came here”

   b. Jeg kjenner mannen (SOM) du sa (*SOM) Maria skal møte i morgen.
      I know the man SOM you said SOM Mary will meet tomorrow
      “I know the man you said Mary will meet tomorrow”

(7) a. Jag känner mannen (SOM) du hoppas (*SOM) kommer hit. (Swedish)
   I know the man SOM you hope SOM comes here
   “I know the man you hope will come here”

   b. Jag känner mannen (SOM) du hoppas (*SOM) Maria ska träffa imorgon.
      I know the man SOM you hope SOM Mary will meet tomorrow
      “I know the man you hope Mary will meet tomorrow”
      [Boef & Franco, in prep.]

The subject/object asymmetry attested for Swedish SOM-insertion in relative clauses, cf. (3) above, disappears in long extractions, as the pair in (6) show. The same pattern in SOM distribution holds for Norwegian. Interestingly, neither Norwegian, nor Swedish speakers accept SOM-insertion in the most deeply embedded clause in long subject extractions, see (6a) and (7a), but they accept SOM if it is directly selected by the matrix (...mannen som du sa/hoppas...=...man SOM you said/hope...). In this case, the pattern of SOM-insertion does not depend on which argument is being extracted, because SOM-insertion in the most deeply embedded clause is ungrammatical anyway (cf. also Boef & Franco, in prep.), whereas SOM-insertion in the higher clause is always optional. I will come back to the optionality of SOM in this higher position in section 6. Compare now the pattern of French que/qui alternation for long extractions:

(8) a. L’homme que tu pense qui*/que est venu ici (standard French)
The man that you think who is come. PART here
“The man you think came here”

b. L’homme que tu pense que/*qui Marie va rencontrer
The man that you think that Mary goes meet.INF
“The man you think Mary is going to meet”

Standard French\(^4\) does not allow complementizer drop in long extractions, contrary to MSc. Moreover, the subject/object asymmetry reflected in the \(\text{que/qui}\) alternation is maintained and appears on the extraction site, i.e. on the most deeply embedded clause, on a par with French short extractions. If we consider long extractions, then, the hypothesis that the asymmetry is due to EPP or phi-feature licensing may explain the French \(\text{que/qui}\) alternation (Taraldsen 2001, Rizzi & Shlonsky 2006 and ref. therein), but not the MSc SOM-insertion. Differently from Swedish and Norwegian, Icelandic notoriously does not show any subject/object asymmetry with respect to complementizer insertion in (long) A’-dependencies. In Icelandic, the complementizer cannot be left out in object extractions, cf. (9), by contrast to Norwegian and Swedish (Thráinsson 2007: 447). The specific complementizer form that Icelandic adopts for relative clauses is \(\text{SEM}\(^5\)), but notice that this form cannot follow a Wh-element in embedded questions, in (10):\(^6\):

(9) \(\text{Báturinn *(sem) Jón á er stór}\) (Icelandic)
Boat.the SEM John owns is big
“The boat that John owns is big” [Thráinsson 2007: 410, 8.45]

(10) a. \(\text{Ég veit ekki hver (*sem) kemur}\)
I know not who.NOM SEM comes
“I do not know who will come”

b. \(\text{Ég veit ekki hvern (*sem) hann hefur séð}\)
I know not who.ACC that he has seen
“I do not know who he has seen” [Thráinsson 2007: 449, 8.134]

\(^4\) In Quebecois French the most embedded \(\text{que}\) of object extractions can be deleted. The mechanisms behind this possibility are still under debate, as it is not clear whether \(\text{que}\)-deletion is induced by contact with English, by phonological reasons (cf. Martineau 1988) or else. I cannot address this issue here, but I thank Ur Shlonsky (p.c.) for pointing that out to me.

\(^5\) On a par with MSc SOM, SEM originated as an element introducing comparative and extent clauses, cf. Faarlund (2008), Nygaard (1905).

\(^6\) Faroese does not show any subject/object asymmetry either, but the relative particle IÐ, following the Wh-element in indirect questions is optional, rather than ungrammatical:

(i) a. \(\text{Ég veit ikki hvør (ið) kemur}\) (Faroese)
I know not who.NOM IÐ comes
“I do not know who will come”

b. \(\text{Ég veit ikki, hvønn (ið) hann hefur séð}\)
I know not who.ACC that he has seen
“I do not know who he has seen” [Thráinsson 2007: 449, 8.134]

Here two issues are at stake, the first one is why the complementizer following the Wh-phrase is possibly deleted (by contrast to Icelandic, above), and the second one is how to explain the optionality of such complementizer in Faroese. This issue deserves more investigations and I leave it open for future research.
Icelandic Wh-elements are morphologically marked for case, which may suggest that these elements can check a nominative feature and enable subject extraction, à la Rizzi & Shlonsky (2006). Specifically, the nominative Wh-element would check case on its way to Spec,CP, thus allowing subject extractions, without ECP effects on the empty Spec,IP position. In Icelandic embedded Wh-questions on the subject, (10a), nominative-checking would be ensured by the Wh-element hver, whereas in object questions nominative case-checking would be done by the overt subject hann, as in (10b). This explanation, however, does not seem applicable to relative clauses, where case is not overtly marked on any Wh-operator. In relative clauses SEM is obligatory, but it seems unlikely that it can check nominative-case, given that:
(i) SEM is not inflected for case;
(ii) there is no subject/object asymmetry in SEM-insertion, which is instead expected if SEM is related to nominative licensing;
(ii) Icelandic behaves similarly to Norwegian and Swedish, with respect to long A'-dependencies: SEM is ruled out in the most deeply embedded clause and there is no subject/object asymmetry in its distribution, as well. The only difference from Swedish and Norwegian, thus, is the obligatory character of SEM-insertion in the higher clause in Icelandic:

(Icelandic)
(11) a. Ég þekki manninn *(SEM) þú sagðir (*SEM) %að kom hingað
   I know man.the SEM you said SEM that came here
   “I know the man that you said came here”

   b. Ég hata manninn *(SEM) þú sagðir (*SEM)að María ætlar að hitta á morgun.
   I hate man.the SEM you said SEM that Mary is.going to meet tomorrow
   “I hate the man that you said that Mary will meet tomorrow”

The facts in (11) show that SEM cannot be omitted when introducing the embedded clause “...þú sagðir...”, whereas it cannot be selected by this declarative predicate. By contrast, the complementizer AÐ (that) in the most deeply embedded clause is apparently optional for object clauses, but preferred for many of the interviewed speakers in the subject clauses, which shows an anti-C-trace effect (Icelandic does not have C-trace effects, see Vikner 1995). However, the discussion of the function and distribution of AÐ, although potentially relevant for the present proposal, requires further data and a separate discussion, which I leave for future research (cf. also Boef & Franco, in prep.).

2.1. Interim summary
In this section, I have illustrated some facts that are problematic for the proposal that the relative particle SOM, used in Norwegian, Swedish and Danish A'-dependencies, is inserted to check subject EPP or phi-features in C, which in turn directly c-commands the canonical subject position in Spec,IP. This idea has been adopted by comparing SOM to a complementizer that can behave either as an

---

7 Rizzi & Shlonsky refer more specifically to strategy of subject extractions based on checking the formal counterpart of subject features. I assume nominative case to fall within this category.
argument or as an expletive, on a par with the -\(i\) morphology on French \textit{qui} (Taraldsen 1986, 2001). However, the distribution of SOM in long A’-dependency disconfirms the hypothesis that SOM licenses subject-extractions via EPP, phi-feature- or nominative-case-checking, since these are all local operation, but SOM-insertion is ungrammatical in the most-deeply embedded clause containing the extraction site. Table 1 below gives an overview of the distribution of SOM/SEM in Swedish, Norwegian and Icelandic.

Table 1.

<table>
<thead>
<tr>
<th></th>
<th>Swedish</th>
<th>Norwegian</th>
<th>Icelandic</th>
</tr>
</thead>
<tbody>
<tr>
<td>short rel. clause</td>
<td>Subj OK</td>
<td>Obj (OK)</td>
<td>Subj OK</td>
</tr>
<tr>
<td>long rel. clause – high cl</td>
<td>(OK) (OK)</td>
<td>(OK) (OK)</td>
<td>OK OK</td>
</tr>
<tr>
<td>long rel. clause – low cl</td>
<td>* * * * *</td>
<td>* * *</td>
<td>* * *</td>
</tr>
<tr>
<td>short embedded question</td>
<td>OK (OK)</td>
<td>OK * *</td>
<td>* * *</td>
</tr>
<tr>
<td>long emb. question – high cl</td>
<td>(OK) (OK)</td>
<td>* *</td>
<td>* * *</td>
</tr>
<tr>
<td>long emb. question – low cl</td>
<td>* * * *</td>
<td>* *</td>
<td>* * *</td>
</tr>
</tbody>
</table>

(OK) = SOM is optional; OK = SOM/SEM is obligatory; * = SOM is ungrammatical

As pointed out above, Table 1 reveals a difference between Norwegian and Swedish in the long embedded Wh-questions of the form “she wondered who (SOM) you think came here/Mary met’. These facts reflect the optionality of SOM in Swedish, vs. its ungrammaticality in Norwegian in Wh-complements where an overt subject is spelled out (cf. (7)-(8) above). I come back to these facts in Sections 4 and 5.

3. Icelandic Stylistic Fronting

Icelandic is the only modern Scandinavian language with productive Stylistic Fronting (SF). SF is a type of movement distinct from the topicalization or the focalization of constituents that give rise to V2 word order in Germanic languages. The literature (Maling 1980, 1990; Rögnvaldsson and Thráinsson 1990; Jónsson 1991; Holmberg 2000; a.o.) agrees that the main characteristics distinguishing SF from other types of fronting are:

i) absence of an overt definite subject in Spec,IP (Spec, AgrS or Spec, SubjP in cartographic terms, cf. Cardinaletti 2004);

ii) lack of restriction to root contexts: SF is productive in embedded clauses, including A’-dependencies;

iii) SF is optional: it alternates with the subject expletive \(það\) (in impersonal clauses) or with a gap\(^9\) (in extractions, cf. Maling 1980, 1990; Rögnvaldsson 1984, 1996);

iv) the stylistically fronted element is preferably “light”: pied-piping of a large phrase is generally disallowed in SF contexts (Thráinsson 2007: 348-349; Wood 2011).

The examples below illustrate the distribution of SF, which consists of moving a

---

8 The presence of SF in Faroese seems to be restricted to formal writing (Maling 1990, Thráinsson et al. 2004).

9 Sigurðsson (1990) argues that SF may alternate with a gap, rather than the overt expletive \(það\), even in impersonal constructions, where the subject referent is easily recoverable from the discourse context.
lexical item, that previous works have analyzed as a (remnant) phrase (cf. Franco 2009a, Ott 2009), to a position preceding the inflected verb. In (12b) the fronted element is the infinitival lesa, whereas in (13a) the past participle dansað.

Subject extraction
(12) a. Þetta er maðurinn [sem lesa vildi ___ allar bækurnar] (Icelandic)
   this is man.the that read wanted all books.the

   b. Þetta er maðurinn [sem vildi lesa allar bækurnar]
   this is man.the that wanted read all books.the
   “This is the man that wanted to read all the books”
   [Thráinsson 2007: 374, 7.79, 7.81]

Subject drop/impersonal construction
(13) a. Þeir segja [að dansað verði __ í brúðkaupinu]
   they say that danced will-be in wedding-the

   b. Þeir segja [að það verði dansað í brúðkaupinu]
   they say that there will-be danced in wedding-the

   c. *Þeir segja [að það dansað/ dansað það verði ___ í brúðkaupinu]
   they say that it danced/danced it will-be in wedding-the
   “They said that people will dance at the wedding”
   [Thráinsson 2007, 355, 7.35]

In (12) SF, (12a), alternates with a subject gap, (12b). In complement clauses with impersonal constructions, as (13) shows, SF is in complementary distribution with the overt subject expletive það (the order SF – V – það is also ungrammatical). A clear distinction between SF and topicalization appears with preposing of a locative PP in subordinate clauses, where topicalization is restricted to some types of phrases (such as adverbials, locative and temporal phrases) for most Icelandic speakers (cf. Hrafnbjargsarson and Wiklund 2009). The following example shows that locative PP preposing is grammatical only if the PP is stylistically fronted in the absence of an overt subject in Spec,IP, as in (14b), whereas it is ungrammatical when undergoing the V2 type of topicalization yielding the order XP – V – subject, as in (14a)10:

(14) a. *Þeir sem í Danmörku hafði hann hitt __...
   those that in Denmark had he met
   “Those that he had met in Denmark…”

   b. Þeir sem í Danmörku hafa verið ___...
   those that in Denmark have been
   “Those who have been in Denmark…”
   [Thráinsson 2007, 380-381, 7.95-96]

10 See Franco (2009a: 29-38) for a detailed discussion on the difference between topicalization, locative inversion and SF in Icelandic.
The categories that can be probed and stylistically fronted belong to a spurious set (Maling 1980, 1990, Holmberg 2000, Jónsson 1991). Not only may past participles or locative PPs undergo SF, but also adverbs, phrasal negation, verbal particles, predicative adjectives, nominal predicates, infinitives, PPs and, allegedly (Holmberg 2000), DP arguments. These categories are arguably ranked according to an accessibility hierarchy, but the actual ordering has been object of debate (Maling 1990, Jónsson 1991, Hrafnbjargarson 2003, Franco 2009a, and Wood 2011 for experimental support). A locality condition seems to impose on such ranking, so that phrases merged higher in the structure are interveners for SF of lower constituents.

The explanations that the literature offers for Icelandic SF can be distinguished as based on either one of the following two hypotheses:

i) the P-feature hypothesis (PFH)
ii) the $\phi$-licensor hypothesis ($\phi$LH)

i) the PFH has been first proposed by Holmberg (2000) and describes SF as an EPP-driven mechanism. Along these lines, SF is preposing of any lexical category bearing a formal P(honological)-feature that can satisfy the EPP. Even assuming that Holmberg's analysis can account for the impossibility to stylistically front an auxiliary (cf. also Maling 1990), the hypothesis resorts to the EPP as a basic assumption, which is fair, as long as the function of the EPP in the grammar is clear. In the case of SF, however, the EPP is not conceptually tied to a subject requirement (cf. Chomsky 1995, Rizzi 2006, Rizzi and Shlonsky 2006), but to a generic requirement that the specifier of a functional projection be spelled out (Chomsky 2000) or the left edge of the clause be filled, as proposed by Sigurðsson (2010). In this case, then, it becomes hard to understand the real trigger behind a rather speculative notion such as the EPP. What is the EPP? Is it a universal principle and why so? These questions are still open, and this paper does not aim at solving the issue, but rather at facing the problem by resorting to different theoretical tools.

Another problem for the PFH is the restriction of some undefined spell-out conditions imposing on the stylistically fronted material. On the one hand, it is not clear which phonological or prosodic rule would impose that Icelandic SF probes only “light” elements, such as remnant VPs (verbal heads or particles, cf. Ott 2009, Franco 2009a), and cannot pied-pipe bigger chunks of structure, or even long, plurisyllabic adverbials. In this respect, the systematic study on the segmental and suprasegmental properties involved in SF of Wood (2011) gave results that do not support a purely PF-based analysis of the phenomenon. On the other hand, another challenge to the PFH hypothesis consists in the fact that phrasal categories like locative PPs must be entirely fronted when undergoing SF and P-stranding is not allowed (cf. Thráinsson 2007: 346), while SF of a verbal head cannot pied-pipe a verbal particle or an object: VP fronting is ungrammatical (cf. Thráinsson 2007: 349). It is not clear which PF condition could possibly impose such a restriction, because that would entail that PF is not blind to syntax.

ii) the $\phi$LH proposes that SF is related to the subject requirement. Ott (2009) argues that SF targets the subject position, (Spec,TP) and is an EPP-driven movement, which also partly complies with Holmberg's (2000) analysis of SF. However, Ott (2009) proposes that the optionality of SF depends on the choice.

11 This is not the case for the SF attested in Old Romance, see Franco (2009a) for a comparison between Icelandic and Old Italian SF.
between a derivation via parallel movement (Chomsky 2000) or just raising the subject to Spec, CP and delete its copy in Spec, IP under identity. The possibility offered by parallel movement is arguably available only in Icelandic, where I° has an EPP feature disjoint from φ-features (cf. also Sigurðsson 2010), thus the latter can be checked via long-distance Agree or subject raising-to-C, whereas SF satisfies the EPP. What is not clear with this type of account is, again, what determines a certain parametrization of the EPP in a grammar and what would constitute an independent motivation to make such a claim epistemologically more grounded. As it is, EPP is just a stipulation that most linguists agree not to question, although it is not clear what precisely would motivate the existence and the crosslinguistic variation of such a structural requirement. There are several alternative ways of addressing these open questions: either one finding a proper motivation for the existence of EPP, or being content with its stipulative nature, as it is, or, finally, searching for other explanatory tools that are not purely formal. Another issue that this analysis would have to deal with is what sort of mechanism induces change in a grammar: there is evidence that many Icelandic young speakers prefer það-constructions to SF, where this is an option, but why one construction should be preferred to another one is still unaccounted for, under the assumption that the two options are perfectly equivalent for information structure. Notice that this assumption is not directly undermined by the fact that SF and það-constructions have a syntactically different behavior, e.g. the latter are islands to extraction, whereas SF is not. However, in the following section it is argued that the two constructions differ in functional and interpretational terms. Identifying SF as an EPP driven-mechanism may be seen as a way to circumvent the problem of the heterogeneity of frontable items in SF. This problem consists of the fact that it is not clear how the broad range of lexical categories undergoing SF may satisfy φ-feature-checking. In order to solve the problem, EPP can be understood along the lines of Rizzi and Shlonsky's (2006) as the requirement to check the “formal counterpart of φ-features”. Thus SF would ultimately check the formal counterpart of φ-features, which enables subject extraction or drop, according to the φLH. However, the difference between a feature and its formal counterpart is not entirely clear, to me, and, as such, is at least questionable, I think. I propose an alternative answer to this puzzle, based on the observation of the complementary distribution between SF and the overt expletive subject realized as það. The feature specification of this pronoun is [3rd person singular, - referential]: assuming that 3rd person singular is [-person] (Sigurðsson 2011), það receives a default interpretation on the basis of its semantically minimal feature specification. Capitalizing on this observation and on further facts, in sections 4 and 5 I argue that such minimal feature specification on subjects may be one proper condition for SF.

4. The proposal
The facts presented in sections 2 and 3 seem to indicate that subject/object asymmetries are attested both in MSc and Icelandic but they obtain as different phenomena. In MSc, they are reflected in the patterns of C-drop (SOM-drop) and in C-trace effects, whereas Icelandic has SF which is an asymmetrical phenomenon because it can only obtain in clauses with subject extractions/drop. Moreover, both MSc C-drop and Icelandic SF are optional phenomena, but such
“optionality” remains so far ultimately unexplained.
In order to present the proposal, some assumptions and claims are made. First, the proposed account intends to reformulate the speculative notion of EPP in terms of a complex feature-checking mechanism with interpretive bearings in information structure. Rizzi & Shlonsky (2006, 2007) propose a theoretically more economic account that subsumes the notion of EPP and explains ECP effects: they refer to the “Subject Criterion” as the requirement that the features encoded in the highest IP subject position be checked. Since the lowest CP head, Fin, locally c-command the highest IP subject position, alternative strategies, checking the “formal counterpart” of these features encoded in FinP, may enable subject extraction (Rizzi & Shlonsky 2007). Elaborating on this intuition, I assume that the EPP results from the requirement of checking a combination of features encoded in the lowest CP head (cf. Rizzi 2004, Rizzi & Shlonsky 2006, 2007, Frascarelli 2007), and such feature-checking can be enforced by various grammar-specific strategies. The CP system encodes various features that enable a pragmatic anchoring of the clausal semantics, structurally provided in the vP/IP domain, to the discourse, located syntactically in CP. A vast literature discusses the structural properties of the complementizer system, primarily the seminal cartographic work by Rizzi (1997), according to which the C-domain is split into functional subheads:

(15) \[ CP \text{Force…Top*…Focus…Top*…Fin…}[IP] \]

For the present purposes, I adopt a slightly modified, updated version of the C-domain, in which the lower Topic phrase is replaced by a Modifier phrase: a non-criterial position where adverbials and other syntactic material (often related to the modal categories identified by Cinque 1999) can front and receive discourse prominence (cf. Rizzi 2004, Haegeman 2006).
Second, Sigurðsson (2011) identifies three main features in the CP domain: [Finiteness], whose interpretation provides anchoring to the speech time and location; [Λ], which identifies the speaker and the hearer; and [Topic], which can consist of an aboutness/shift topic (cf. Frascarelli 2007), of a contrastive or of a familiar topic (cf. also Bianchi and Frascarelli 2009 for English). According to Sigurðsson, then, clausal arguments are also anchored to the discourse by matching the Λ-features. Following his proposal, I assume that among the phi-features of each argument, only [Person] is interpreted in the CP. In this sense, 1st person is interpreted as [Λ: +speaker, -hearer]; 2nd person as [Λ: -speaker, +hearer] and 3rd person as [Λ: -speaker, -hearer], namely [-person] (cf. Sigurðsson 2004, 2010, 2011).
A third assumption, based on a recent proposal by Déchaine & Tremblay (2011)\(^\text{12}\), is that FinP encodes not only [finiteness] – a verbal feature enabling the pragmatic anchoring of the predicate (i.e. of the event-time) to the speech-time – but also a nominal feature, called [definiteness]. Checking of [definiteness] enables pragmatic anchoring of clausal arguments to the discourse and yields the interpretation corresponding to the Λ-features valued by the respective arguments.

\(^{12}\) Déchaine & Tremblay (2011) refer to Chomsky’s (1995) identification of C as (+V;+N) and to Muyskens (2008: 249) who interprets definite/indefinite determiners as markers of discourse status. Whence the hypothesis that [definiteness] be interpreted in the information structure.
according to their person-feature specification. In this respect, the fact that 3rd person is interpreted as [-person] is crucial for the productivity of SF. In section 3 it was illustrated how SF is in complementary distribution with the overt expletive það, which specified as [-person]. A is indeed [-person] in expletive constructions, where either SF or það are possible, and in other 3rd person constructions as well, e.g. in relative clauses modifying an NP/DP, where SF (but generally not það) can occur, cf. (12a)\(^\text{13}\). Moreover, I will assume that SF may target a position in the C-domain (cf. Rögnvaldsson & Thráinsson 1990, and Franco 2009a for evidence).

A final assumption, crucial for the analysis proposed here, follows Sigurðsson’s (2011) idea that Scandinavian, and, more generally, Germanic V2 languages license argument drop configurationally, by contrast to, say, Romance languages where verbal morphology has pronominal properties and may license argument drop via long-distance agreement. Configurational languages such as the Scandinavian ones, then, require that the dropped argument be in a local configuration with the licensing CP head where it must be interpreted. For this reason, topic drop requires an empty left edge:

(16) a. _____ kommer tillbaks imorgon
   come.Ø-AGR back tomorrow
   “I/we/she/etc. will be back tomorrow”

   b. Imorgon kommer *(jag/hon/ …) tillbaks.
      tomorrow come.Ø-AGR *(I/she/ …) back
      [Sigurðsson 2011: (3)-(4)]

In other words, Germanic topic drop, which includes subject-drop, requires that the dropped argument locally match [A] in CP, which explains the V2 restriction (cf. Sigurðsson 2011). From this perspective, a subject can be dropped or extracted if (i) it is also a Topic in the CP or (ii) an alternative strategy ensures that the relevant person features are interpreted in the CP, i.e. [A]-values properly match with an element in a local configuration with them. With respect to (ii), such local position could in principle be Spec,IP, but this position is specific for subjects, so only specific elements such as (null) expletives pronouns, could be merged there, and this option is actually available in Icelandic clauses where the subject is extracted and there is no SF. Otherwise elements that can locally match [A] values in CP may occupy FinP. This second possibility finds a concrete realization in það-expletive insertion, which allows post-position or extraposition of subjects, and, arguably, in SF (cf. Wood 2011: (41)).

In view of distribution and syntactic properties of SF, I argue that SF is a mechanism that checks the verbal counterpart of the feature combination encoded in FinP, namely [Finiteness]. As will become clearer in the following section, the nominal counterpart [Definiteness] either receives a default interpretation, i.e. [A: -person], or is checked by another element (e.g. a Wh-OP), in SF constructions. The core idea of this proposal is then stated in (17) below:

---

\(^{13}\) In the examples in (13a) and (14b) SF occurs in a relative clause modifying a 3rd person plural subject pronoun, but given that only [person] is interpreted in the CP, 3rd yields [-person] anyway, and [number] is then irrelevant here.
(17) SOM-insertion in MSc and SF in Icelandic are both strategies to check one of the [Finiteness/Definiteness] complex of features encoded in FinP, in constructions where the syntactic subject cannot do that.

4.1. Icelandic SF

An important aspect that distinguishes this proposal from the hypothesis that SOM is an element checking the subject EPP or phi-features concerns the nature of the structural projection hosting SOM (and SF, in Icelandic). Differently from Spec, IP, the lexical realization of FinP is not merely dependent on syntax, but is imposed by pragmatics. This entails that FinP may not be spell-out in case, for instance, the arguments that must be anchored to the discourse are recoverable from the discourse pragmatics. Of course, overt subjects in Spec, IP can be easily anchored to the discourse, because they are in a local configuration with FinP. However, there are cases where the subject can be dropped simply because it is recoverable from previous context, i.e. A is identified by a matching relation with the features valued in the CP of a preceding clause. This can be the case for Italian, among other pro-drop languages, where the features identifying the dropped subject are “inherited” by a preceding matrix CP (cf. Frascarelli 2007).

In Scandinavian languages, verb agreement has no pronominal properties: thus subject-drop is restricted to configurations in which its reference is either locally recoverable or interpreted by default, as in the case of Icelandic non-referential 3rd person singular. In this respect, the basic difference between Icelandic and MSc is that Icelandic allows [Λ: -person, -referential] to be morphologically realized as Ø: compare (18) and (19b) with (19a) (below, (19) repeats (13) for convenience).

(18) ?Deir segja [ad ___verði dansað í brúðkaupinu] (Icelandic)
    they say that there will-be danced in wedding-the
    “They say that people will dance at the wedding”

(19)    a. Deir segja [ad dansað verði ___ í brúðkaupinu]
        they say that danced will-be in wedding-the

    b. Deir segja [ad það verði dansað í brúðkaupinu]
        they say that there will-be danced in wedding-the

    c. *Deir segja [ad það dansað/ dansað það verði ___ í brúðkaupinu]
        they say that it danced/danced it will-be in wedding-the
        “They said that people will dance at the wedding”
        [Thráinsson 2007, 355, 7.35]

The lack of an overt expletive in (18) is not completely ungrammatical, because the features that the pronoun spells out (in (19b)) may be interpreted as default (-person, non-referential). Thus the reason of the alternation between a gap and SF cannot be related to checking subject-specific features. I propose that the reason is pragmatic, and it amounts to spelling out a position whereby some clausal content can be anchored to the context, i.e. FinP.

The stylistically fronted element in (19a) conveys some semantic content (in this case, aspectual information concerning the predicate dansað, i.e. the result of an event, cf. Franco, in prep.) that is fronted to the CP. The event-time is thus
interpreted as anchored to the discourse (speech-time, cf. Third assumption, section 3 above). Such interpretation is arguably given from the feature-checking mechanism driving SF: the fronted element (a remnant) phrase, in (19) above, cf. Franco 2009a, Ott 2009, Wood 2011), checks the verbal counterpart on FinP: [Finiteness\textsuperscript{14}]. At this point we can just assume that the nominal counterpart of FinP, in SF constructions, is either specified as [-Definiteness], which corresponds to a [-referential, -person] subject, or as [+ Definiteness], in which case it may be checked by a (Wh-)OP undergoing successive cyclic movement to a higher position, as in (12), repeated here as (20) for convenience:

(20) a. Þetta er maðurinn
this is man.the
[OP\textsubscript{rel} sem [SF < OP\textsubscript{rel} > lesa] vildi allar bækurnar]
that read wanted all books.the

b. Þetta er maðurinn
this is man.the
[OP\textsubscript{rel} sem < OP\textsubscript{rel} > vildi lesa allar bækurnar]
that wanted read all books.the

“This is the man that wanted to read all the books”
[Thráinsson 2007: 374, 7.79, 7.81]

The feature-checking mechanism involved in (20a) is given in (21) below. In this case, the remnant phrase (i.e. vP) containing the subject-relative OP is moved to Spec,FinP\textsuperscript{15}. In (21) the OP matches the nominal counterpart of FinP, whereas SF can check [Finiteness].

(21) … Spec, FinP [<Rel/Wh-OP> [-Fin, +De] SF [+Fin, -De] ] Fin\textsuperscript{\circ} [-Fin, +De]: [SubjP \varnothing [-pn]]

The [Definiteness] feature on FinP may be arguably checked in an alternative way, by expletive insertion, given that Icelandic það is only preverbal and occurs exclusively in a [+definite] subject position (cf. Thráinsson 2007, Bentzen 2007). For (19b), the feature-checking mechanism would then be as in (22) below:

(22) … Spec, FinP [það [-Fin, +De]] Fin\textsuperscript{\circ} [-Fin, +De]: [SubjP \varnothing [-pn]]

Notice that in expletive-constructions the string það-SF, contrary to relative OP – SF in (21), is ungrammatical because it is syntactically underviable and semantically incompatible: það-merger applies only at the CP/IP interface. Moreover, það is semantically vacuous and creates a chain with an indefinite associate in a rhematic position carrying the relevant semantic information. By

\textsuperscript{14} Notice that also other elements commonly undergoing SF, such as adverbs, verbal particles, predicative adjectives, nominal predicates, generally categorize at least as [+V]. Franco (2009a), however, argues that nominal predicates, predicative adjectives and potentially also adverbs (i.e. all the categories that may also undergo V2 topicalization) target a non-criterial position in CP, where they receive discourse prominence: ModP (cf. Rizzi 2004, Haegeman 2006). Along these lines, adverbs may be cyclically independent from Wh-movement to Spec,FinP and move directly to ModP, so they do not intervene. Cf. fn. 17 below.

\textsuperscript{15} For a derivation of SF in subject-relative clauses see Franco, in prep.
contrast, SF contributes to information structure in that it fronts the relevant semantic content to a thematic position (which is to be understood in terms of the functional opposition between theme and rheme). Thus expletive constructions and SF-constructions are minimally different strategies to check the feature-complex encoded in FinP, which leads to some predictions discussed in the following section.

Before doing so, I wish to discuss how the present proposal applies to MSc.

4.2. MSc SOM

As proposed in (17) above, SOM also enters the feature-checking mechanism imposed on FinP in order to license subject-extractions. How so? Looking back at Table 1, we can see that SOM is only obligatory with short subject extractions, i.e. when the subject of the relative clause is otherwise impossible to recover. In long subject extractions, SOM in the lower clause is ungrammatical altogether, despite the lack of an overt subject in the most deeply embedded clause, contrary to the expectations naturally emerging from Taraldsen’s (2001) or Rizzi & Shlonsky (2006) accounts (cf. Section 2). Below, (6a) is repeated as (23) for convenience.

(23) Jeg kjenner mannen [*(som) du sa [(*som) kom hit]]. (Norwegian)

"I know the man you said came here"

The difference between (23) and short subject-extraction is explained, in the current proposal, as an interpretive difference. While [EPP] or subject [\(\phi\)] always need to be checked locally (as is the case for the –i morphology on qui, in French), the function of MSc subject-specific SOM is that of anchoring the extracted argument to the discourse. This analysis is in line with the idea that SOM insertion is a strategy to anchor an argument to the discourse, in order to satisfy the requirement of Full Interpretation (Chomsky 1986, 1995).

Crucially, discourse-anchoring is a mechanism that determines the interpretation of that argument and, as such it occurs in the higher clause, i.e. the clause containing the probe for the relative OP movement. The fact that the extracted argument is anchored in the higher, rather than in the most deeply embedded clause, is due to the fact that its reference is identified by the entire scope of the OP movement. In other words, SOM insertion in (23) is only grammatical where the relative OP is interpreted, i.e. in the highest embedded CP. This is precisely what is expected, given that the relativized argument in (23) is interpreted as “the man [YOU SAID [came here]]”, which is not (necessarily) coreferent with “the man [that came here]”. In order to account for this difference, I propose that subject-specific SOM checks the nominal counterpart of FinP, locally c-commanding the Spec, IP subject position. In short A’-dependencies (e.g. simple relative clauses), this feature-checking mechanism allows subject-extraction to occur without triggering ECP effects, or, in Rizzi’s (2004) words, criterial freezing of the subject in the highest IP position.

(24) a. …manner [FinP[\(\mathsf{De}\)] SOM[\(\mathsf{De}\)] [IP ___ kom hit]

man.the SOM came here

“The man who came here”

In this perspective, ECP effects are provoked by an unsatisfied requirement on C,
rather than on I, which in the present case corresponds to [Definiteness] checking. Because such requirement is not imposed on the most deeply embedded clause of A’-dependencies, for the reasons discussed above, SOM is not required in the embedded CP.

A related question, at this point, concerns the optionality of SOM in object extractions\(^\text{16}\) and in the higher clause of long A’-dependencies. If the function of SOM is to check [Definiteness] on FinP, why is it at all possible to have SOM in clauses when an overt subject in Spec,IP is present? Postulating that SOM may or may not check such feature, depending on the argument that is being extracted, is of course a completely \textit{ad hoc} solution that remains unexplained in terms of acquisition.

This optionality puzzle can be faced from a diachronic perspective. Following a seminal proposal in Franco (2009a), also on the basis of facts analyzed in Delsing (2001), the proposed account explains the optionality of SOM in short object clauses and higher clauses of long A’-dependencies vs. its obligatoriness in short subject clauses as the result of an overlap between two grammars introduced at different diachronic stages. This proposal is discussed in greater detail in the following section, where diachronic evidence is provided. In a nutshell, the analysis is that the first grammar had just a non-subject specific SOM, which was distributed symmetrically over subject and object clauses, and there is a later reanalysis of SOM to subject-specific in a second grammar. As a relic of the first system, “optional”, non-subject specific SOM still coexists with the second, subject-specific SOM. This difference can be accounted for in structural terms, by assuming that the “old” and the “new” SOM occupy two distinct positions in a split-CP (Boef & Franco, 2010, in prep.). Figure 1 illustrates optional, non-specific SOM, which is arguably merged on a high CP head, specifically as a relative subordinator. By contrast, Figure 2 illustrates that subject-specific SOM is arguably merged in Fin\(^0\), where it checks [Definiteness] and it locally C-commands the subject position in IP.

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig1}
\caption{Short object extractions/Higher clause of long extractions}
\end{figure}

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{fig2}
\caption{Short subject extractions}
\end{figure}

\(^{16}\) As mentioned in the introduction, SOM is optional in object-relative clauses both in Norwegian and Swedish, but while it is optional in object-indirect-Wh- questions in Swedish, it is ungrammatical in the same context in Norwegian (cf. Table 1). This difference is tentatively explained below.

167
The analysis proposed here for both Icelandic SF and MSc SOM gives rise to a series of predictions that are discussed in the following section.

5. Predictions
In this section, I discuss some predictions following from the proposal, by distinguishing synchronic comparative facts, from the diachronic facts supporting the idea in (17) above.

5.1. Comparative support
A first prediction following from (17) is that, if subject-specific SOM-insertion and SF are both mechanisms that check the features specified on FinP, they should be attested, at least in the same syntactic contexts, as alternatives for the grammars in which they are respectively productive. Looking at the respective languages, we can observe that Icelandic, which has productive SF, does not have any subject-specific particle introducing relative clauses: the distribution of SEM does not display any subject-object asymmetry (cf. Table 1). By contrast, MSc grammars do not have SF but derive subject extractions by means of subject-specific particles. This is not only true of Norwegian and Swedish, which use SOM, but also of Danish, where the subject expletive DER is reanalyzed as relative particle in subject extractions, and alternates with SOM in subject relative clauses (25a and b). Crucially, Danish allows for multiple complementizers, but only cases where at least either SOM or DER are present in the order SOM – (AT) – DER are grammatical as subject relatives (cf. Vikner 2001, Mikkelsen 2002 for more data):

(25)  
  
  Vi kender de lingvister  
  we know the linguists  
  (Danish)
Mikkelsen (2002) observes that (i) DER occurs only in subject relative clauses; (ii) SOM occurs in subject and non-subject relative clauses; (iii) when SOM AT DER are all present they must appear in that order; (iv) either SOM or DER must be present in the subject relative clause (Mikkelsen 2002: 2).

These facts are explained under the hypothesis that Danish has a mixed system where two alternative strategies, subject-specific SOM or DER-insertion, apply, cf. (25a) and (25b) and (iv) above. The co-occurrence of DER and SOM in (25c) is here explained with the hypothesis that the double nature of SOM, cf. (iv), is due to an overlap of the new complementation system with subject specific SOM with a diachronic relic of the old MSc complementation where SOM is non-subject specific, in Danish as well as in other MSc languages. By contrast to SOM, however, Danish DER occurs in the most deeply embedded clause in a subject Wh-extraction recalling the distribution of French qui: compare (26) below with Table 1 and (8a).

(26)  Jeg ved ikke  
I know not  

a. hvem du tror DER har gjort det  
who you think there has done it  

b. *hvem du tror SOM har gjort det  
who you think SOM has done it  

“I do not know who has done it”  
[Mikkelsen 2002: 7, (20)]

The fact that DER is only subject-specific and behaves more similarly to French qui than to MSc SOM can be explained with the idea that Danish has a complex complementation structure, making use of different strategies: expletive insertion (DER), subject-specific particle, and non-subject specific relic (SOM). Moreover Danish, on a par with other MSc, does not allow for SF.

Let us now consider a borderline case: Faroese. In Faroese, relative clauses are introduced by the relative particle SUM, whose distribution matches that of SOM.

SUM is not used in embedded Wh-questions, where the complementizer associated to the Wh-XP is ID: this complementizer is optional and has a symmetric distribution (see fn. 6). Albeit I cannot discuss this in the present paper, these facts do not represent an obstacle to the present
in Norwegian and Swedish, as is illustrated in (27) below.

(27)  a. **Báturin *(SUM) Jón eigur er stórrur**  (Faroese)
     boat.the that John owns is big
     “The boat that John owns is big”

     b. **Ofta eru tað konurnar *(SUM) koma fyrst**
     Often are there women who come first
     “Often there are women who come first”
     [Thráinsson 2007: 448, (8.132)]

In addition, let us take into consideration another syntactic context, i.e. non-restrictive relative clauses. In MSc as well as in Faroese, SOM/SUM can never be omitted in non-restrictive relative clauses (cf. Thráinsson 2007: 447), namely, non-restrictive relative clauses do not display any subject-object asymmetry with respect to SOM/SUM insertion.

(28)  a. **Den här bilen, *(SOM) jag aldrig har sett förut, er vacker**  (Swedish)
     This here car SOM I never have seen before, is beautiful
     “This car here, which I have never seen before, is beautiful”

     b. **Tær konurnar, *(SUM) _ skulu vera heima, eru burtustaddar**  (Faroese)
     They women.the who shall be home are away
     “The women, who shall be home, are away”
     [Thráinsson 2007: 447, (8.131), (8.132b)]

The obligatoriness of SOM/SUM in all non-restrictive relative clauses\(^{18}\) can be tentatively explained in terms of the different structural properties of non-restrictive and restrictive relative clauses. I will tentatively argue here that, differently from restrictive relative clauses, the CP of non-restrictive relative clauses constitutes a full phase-edge, subject to the Phase Impenetrability Condition (PIC, Chomsky, 1999). The CP of restrictive relative clauses, instead, is transparent for inheritance of the features that permit the recoverability of the extracted noun. In non-restrictive relative clauses, the reference of the extracted noun is not recoverable, unless the C-position onto which it is interpreted is overtly spelled out (by SOM, in this case). Given that both subject and object non-restrictive relative clauses are subject to the PIC, this requirement expectedly holds symmetrically for coreference of subject or object OPs with the head noun. The lack of symmetry in non-restrictive relative clauses, on a par with long extractions, is then analyzed as a consequence of the fact that SOM/SUM is not a EPP/\(\phi\) - licensor, but an element that checks the nominal features (i.e.

\(^{18}\) SOM/SUM-deletion is ungrammatical also when introducing long non-restrictive relative clauses, as the Norwegian example below shows:

(i)  a. **Johan, *(SOM) du sa (*SOM) jobber hele tida, gikk til en konferanse i går**
     John, SOM you said SOM work all time went to a conference yesterday
     “John, who you said works all the time, went to a conference yesterday”

     b. **Johan, *(SOM) du sa (*SOM) du ikke har mott, synger i et kor**
     Johan, SOM you said SOM you not have met sings in a choir
     “Johan, who you said you have not met, sings in a choir”
     [Boef & Franco, in prep.]
[Definiteness]) required for the interpretation of the extracted arguments in CP. Differently from MSc, Faroese allows for SF, although this phenomenon is less productive than in Icelandic and seems to be restricted to literary texts or written language (cf. Thráinsson et al. 2004). The fact that the Faroese grammar has both a SUM distribution that matches that of MSc SOM, and SF, like Icelandic, does in fact comply with the claim that SOM/SUM insertion and SF are two alternative strategies. The following example indeed shows that an ungrammatical deletion of SUM, cf. with (28b) above, can be rescued by SF (of heima), as Thráinsson (2007: 447) points out:

(29)  Tær konurnar, (SUM) heima skulu vera ___, eru burtustaddar  (Faroese)
They women.the who home shall be are away
“The women who shall be home are away”
[Thráinsson 2007: 447, (8.132c)]

The peculiarity of Faroese, then, is that the two alternatives are coexistent in one system, which actually seems to result from the competition between two grammars (i.e. the MSc-like and the Icelandic-like) also with respect to other syntactic properties, such as V-to-I (cf. Heycock et al. 2010, a.o.). A further expectation concerns the feature specification of Icelandic clauses with SF, in contrast to clauses with expletive insertion or with a subject gap. Arguably, SF checks the verbal counterpart of FinP, whereas expletives are nominal elements, thus checking the nominal counterpart of FinP. Accordingly, it is predicted that það-clauses, namely expletive constructions, are islands for the extraction of an argument (e.g. subject relativization), because such an extraction would trigger minimality effects with the feature specification of the expletive in Spec, FinP. This prediction is borne out by well-known facts in the literature on Icelandic: in (30c) below, það intervenes with the movement of the subject relative OP, whereas SF in (30b), bearing [+Finiteness], does not.

(30)  a. þetta er mál sem __ hefur verið rætt  (Icelandic)
this is issue that has been discussed

b. þetta er mál sem hefur verið
this is issue that discussed has been

c. *þetta er mál sem það hefur verið rætt
this is issue that there has been discussed
“How is an issue that has been discussed”
[Thráinsson 2007, 353, 7.28-31]

The (lack of) intervention effects in (30b) and (30c) is schematically represented in (31a) and (31b) respectively:

(31)  a. Rel/Wh-OP [+Fin, +De] SF [+Fin, -De] → OK


The fact that SF does not create an island for Wh- movement can be accounted as follows. SF of “items from the verbal complex” (Maling 1990) such as past participles, infinitives, verbal particles, or predicative adjectives is phrasal
movement of a (remnant) vP or PP to Spec,FinP (cf. Franco 2009a, Ott 2009). The moved phrase also contains the probed Wh- argument that is extracted, once the phrase that contains it reaches the CP. Following Franco (2009a), I propose that there are different kinds of SF: when the fronted element is a sentential adverb or a negation, there is no remnant movement involved, but XP movement to Spec,ModP, a projection immediately above FinP where adverbials receive discourse prominence (Rizzi 2004, Haegeman 2006). Adverbial movement is thus cyclically independent from Wh-movement to Spec,FinP (see Franco 2009a for details). The feature specification of the adverbial phrase is such that no minimalitiy with the Wh-OP is triggered. Moreover, the accessibility hierarchy of SF (Maling 1990, Holmberg 2000, Hrafnbjargarson 2004, Franco 2009°, Wood 2011) is explained by the fact that the Wh-OP cannot “pied-pipe” the (remnant) vP, if the latter is within the scope of a phrasal adverb or a negation, for semantic reasons. Movement of a (remnant)-vP across a phrasal adverb or a negation would indeed give rise to scope interaction effects at the level of information structure. As a consequence, the sentence would become uninterpretable.

What remains to be explained is the alternance of a gap with SF, given in (30a) and (30b). In section 4, I have proposed that SF contributes to information structure by anchoring the event-time to the speech-time, namely anchoring some semantic content to the discourse. In this perspective, (30b) above would receive a slightly different interpretation from (30c). Where the past participle is fronted, as in (30b), an aspectual feature of the predicate is anchored to the speech-time: the information concerning the result of an action/event (rætt) acquires discourse prominence, by being “thematized” (cf. Franco, in prep. and references therein). By contrast, the semantic content concerning the action/event encoded in the predicate structure remains in rhematic position. Thus clauses without SF have a different information structure: no semantic information is given discourse prominence, but is instead conveyed in a presentational way, by leaving it in rhematic position, with either and expletive or a gap in thematic position. The contrast between functional theme and rheme is given in Table 2 below:

<table>
<thead>
<tr>
<th>Construction type</th>
<th>Theme</th>
<th>Rheme</th>
</tr>
</thead>
<tbody>
<tr>
<td>SF</td>
<td>SF: semantically full</td>
<td>Semantically vacuous / residual</td>
</tr>
<tr>
<td>Subject gap (e.g. relative clause)</td>
<td>Ø: inheriting feature via OP-A’-binding</td>
<td>vP: semantically full</td>
</tr>
<tr>
<td>Expletive construction</td>
<td>það: semantically vacuous</td>
<td>Associate: semantically full</td>
</tr>
</tbody>
</table>

Table 2

The alleged subtle interpretive difference between cases of subject extractions where nothing fronts and cases with SF, cf. (30a) and (30b), seems corroborated by further facts. If the head noun of a subject relative clause with SF is definite, the sentence is degraded for some speakers. Halldór Sigurðsson (p.c.) provides the following judgments:

19 But not for others: at a first look it seems that younger speakers (20-30+) are not sensitive to such definiteness effect, whereas older speakers (45-50+) are. There seems to be an even sharper definiteness effect (irrespectively of speakers’ age) in Wh-clauses where the Wh-Op is a definite subject. This point deserves further investigation.
Notice that, if the head noun is a pronoun (Peir) or a generic/indefinite noun (menn/allir), cf. (32c), the sentence is totally acceptable, whereas it is less acceptable with a definite head noun (maðurinn, cf. (32b)). This judgment is reflected in a frequency test: by searching the exact string “sem lesið hafa” (=who read have.3PL) on a search-engine (Google search, 27.04.2012, 11:30), the great majority of the results (total of about 67.300 entries) display an indefinite/generic head noun. On the first 10 Google pages, the most frequent head nouns are Peir (=they), allir (=all), Íslendingar (=Icelanders), margir (=many), flestir (=most), or bare nouns modified by these elements. The trigger for the marginal definiteness effect in (32b) vs. (32a) indicates that there is a pragmatic, rather than a syntactic difference between clauses with and without SF. The degradation effects in (32) are still an open issue. From the present perspective, this difference could be tentatively explained in terms of the feature specification on FinP. On the one hand, it is true that SF does not create an island for Wh- extraction (cf. (30) and the discussion above). On the other hand, FinP has a double specification, namely [+Finiteness; ±Definiteness]. A working hypothesis is that the nominal/verbal features of FinP (cf. Chomsky 1995) may not be simultaneously valued (+) and be both morphologically spelled out. A strict interpretation of this requirement is that a [+Finiteness] specification on FinP is incompatible with a [+Definiteness] specification because that clause may not have both an overt element checking [Finiteness], i.e. SF, and an overt element checking [+Definiteness]. Given that Icelandic SEM is not subject-specific and cannot check the [+Definitenes] imposed for the interpretation of the extracted definite subject, in cases like (32b), the requirement cannot be fulfilled, and the interpretation is partially compromised. For what is presently known in Icelandic, it seems that valuation of the verbal counterpart of FinP, i.e. [+Finiteness], pairs with lack of valuation of the nominal counterpart, i.e. [-Definiteness]. However, this hypothesis requires further testing.

5.2. Diachrony
In this subsection I provide some support for my proposal, which is based on diachronic comparative facts. Given that relative clauses are the ideal context for observing both the distribution of SOM in MSc and of SF in Icelandic, I will concentrate on those. The complementation structure of Scandinavian relative clauses has undergone some diachronic changes: both Old Eastern Scandinavian (here represented by Old Swedish) and Old Western Scandinavian (here, Old Icelandic) used similar particles and strategies to derive relative clauses (cf.
Faarlund 2008, Delsing 2001, Nygaard 1906), but, around 1350-1400, the Swedish system underwent some major changes that resulted in a basic difference from the Icelandic system (cf. Franco 2009a, ch. 4). The history of the various relative particles in (Old) Swedish and (Old) Icelandic is illustrated in Table 3 below.

Table 3. Relative particles in Old Swedish and Old Icelandic

<table>
<thead>
<tr>
<th>1200</th>
<th>1300</th>
<th>1400</th>
<th>1500</th>
<th>1600</th>
<th>1700</th>
<th>1800</th>
<th>1900</th>
</tr>
</thead>
<tbody>
<tr>
<td>SWEDISH</td>
<td>OV</td>
<td>--------</td>
<td>OV/VO</td>
<td>--------</td>
<td>VO</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>Ø</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>som</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ær</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ther</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>hvilkin (som)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>ICELANDIC</td>
<td>OV/VO</td>
<td>--------</td>
<td></td>
<td></td>
<td>VO</td>
<td>--------</td>
<td></td>
</tr>
<tr>
<td>sem</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>er</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Table 3 shows that Old Swedish had several possibilities for introducing a relative clause: the particle ær, a Ø-complementizer and som, which was in fact an unspecific form, on a par with Old and Modern Icelandic sem, given that it displays a symmetric distribution between subject and object clauses. In Old Swedish (i.e. up to 1350), SF is practically always found in subject extractions, according to Delsing (2001). Around 1350-1400, however, a major change occurred in the Swedish grammar involving word order change (from OV to VO) among other properties. At this stage, the particle ær disappears, while ther is introduced, and, crucially, a Wh- pronoun, hvilkin (=which) is adopted for introducing relative clauses. As Delsing (2001) observes, object relative clauses are usually introduced by hvilkin alone, whereas subject relative clauses are introduced by the sequence hvilkin som or hvilkin ther (cf. (33)). The fact that som/ther begins to have an asymmetric distribution suggest that it becomes subject-specific.

(Old Swedish, 1350-1400)
(33) Brudhgöma HULKIN SOM är åronna konungir
Groom which SOM is honour king.GEN
“The groom who is the king’s honour”
[Delsing 2001, 159, 48, BU]

Under the hypothesis of such a reanalysis, the prediction is that SF is in

The Old Swedish part of the table is taken from Delsing (2001).

The fact that ther alternates with som in subject extractions seems to indicate that they are both subject specific. However, further investigations with respect to the diachrony of ther are necessary. For the moment I might just speculate that ther is the ancestor of an expletive, originally a locative, der, which is indeed reanalyzed as a subject specific element in Modern Danish (cf. section 5.1), a language developed from Old Eastern Scandinavian, on a par with Modern Swedish.
complementary distribution with SOM in subject relative clauses, once SOM becomes subject-specific. In other words, the Wh- element HVILKIN should be followed either by SOM or by SF, but not by both. As expected, strings of the type HVILKIN – SOM- SF or HVILKIN- SF- SOM have not been attested so far, and when the Wh- is not followed by SOM, it is obligatorily followed by SF, in subject relative clauses. Observe, in (34) below, that forms of HVILKIN (hulke, hulkan) are not followed by SOM but by a stylistically fronted adverb (adrigh, (34a)) or particle (fram, (34b)).

(34) a. som grymasto diwr       HULKE  adrigh kunna ___ mättas 
   like most.cruel anymals  which never can be.satisfied
   “Like the most cruel animals which can never be satisfied”

b. HULKIN  framgik til sanctum gregorium
   which forth.went to saint Gregory
   “Which went forth to saint Gregory”
   [Delsing 2001, 163, 54, Greg]

Therefore, this prediction is borne out. The explanation for such a reanalysis of SOM and the eventual loss of SF in Swedish and the other MSc languages onto which this system spread could be based on an economy principle. Arguably, the old complementation structure, in which generic SOM combines with SF in subject extractions, competes with the new system, after the introduction of the Wh- element. The combination “Wh- SF” attested in the Early Modern Swedish period loses the competition (in the sense of Lightfoot 1999, Lightfoot & Westergaard 2007) with the newer and more economic strategy “Wh-SOM”. In the latter option, SOM is reanalyzed by the grammar as subject specific. This strategy is more economic because it involves just one case of movement (the Wh-element) and one case of merge (subject-specific SOM, rather than movement of the stylistically fronted phrase, cf. Franco 2009a).

The question, at this point, is why Icelandic SEM has such a different distribution from SOM, i.e. it is not subject-specific. I analyze this difference as a result of the fact that the Icelandic system is far more conservative and has never undergone the Swedish reanalysis.

If this is the case, a further prediction is that (Old) Icelandic subject specific clauses, e.g. restrictive relative clauses on the subject, should never be introduced by a Wh- element followed by SEM, by contrast to the Swedish HVILKIN SOM cases. In order to verify this prediction I have carried out a corpus search on the Old Norse data reported in Faarlund (2008), and on 44 Old Icelandic sagas, from the online archive in the IcePahC. It resulted that the only cases where the string Wh- SEM is attested are free relatives, in which SEM has a subject/object symmetric distribution and the Wh- element is interpreted as a quantifier. Moreover, these sentences display SF in a Wh- SEM- SF string, in case of subject extraction, as in (35), contrary to Old Swedish where HVILKIN – SOM – SF is unattested:

(35) a. Til  HVERS  SEM  draga vill               (Old Icelandic)
   To each.GEN  SEM  pull wants
   “To everyone who wants to pull”
b. Er þú hefðin jafnan uppi látið  
That you have always up let each.NOM SEM asked have
“... that you have always told whoever has asked”

This means that Icelandic maintains the same strategy for checking the FinP features in subject extractions, and the complementation system does not undergo a reanalysis of the Old Swedish kind.

6. Optionality explained
On the one hand, the diachronic facts illustrated in section 5.2 provide further support to the analysis of SOM distribution in MSc. It has been shown that SOM undergoes a diachronic reanalysis from non-subject specific to subject specific. Arguably this change in the grammar occurred as the result of a competition between two different systems. In the older system, which was similar to Icelandic, SF was productive, but eventually it disappeared, due to the fact that the old system lost the competition against the new one, in which SOM can actually check the nominal feature on FinP. In cartographic terms, I analyze the old SOM as a bare relative subordinator, on a par with Icelandic SEM, and the “new” SOM as an element that merges in FinP. In a split CP structure, the generic subordinator SOM/SEM is then merged in SubP (cf. Haegeman 2006, a.o.), whereas the subject-specific SOM merges in FinP, on a par with SF, as is schematically represented in (36)\(^\text{22}\).

\[
\begin{array}{cccccc}
\text{Sw./No.} & \text{SubP} & \text{Sub^o} & \text{ModP} & \text{FinP} & \text{Fin^o} \\
\text{Ic.} & \text{Wh/OP} & \text{(SOM)} & \text{SEM} & \text{Adv/Neg SF} & \langle \text{Wh/OP} \rangle \text{SOM}! \\
\end{array}
\]

The different, but compatible, functions of the old and the new SOM and the fact that they do not occupy the same position allow them to be both retained in the same grammar, where the competition between the two systems is stabilized at this specific point (cf. Lightfoot 1999). The optionality of SOM in object and long \(A^\prime\)-dependencies is explained under the hypothesis that SOM, in those cases, is just a diachronic relic of the generic subordinator of the old complementation system, where relative clauses could be introduced by a morphologically null subordinator as well, cf. Table 3 above. In terms of language change, this hypothesis is compatible with the idea that the competition between the old and the new system may reach different points of stability, depending on the specific grammar. For this reason, crosslinguistic variation and microvariation among the Swedish and the Norwegian dialects is expected.

On the other hand, the optionality of SF in Modern Icelandic, vs. its higher frequency, arguably obligatoriness, in Old Icelandic is explained both in light of its interpretive properties (cf. Sections 4, 5) and as a consequence of a diachronic change. The different interpretation of extraction clauses with and without SF may account for the optionality of SF in relative clauses, cf. (30) above, and discussion. On the contrary, the alternation between the Icelandic expletive \(það\)

\(^{22}\) Brackets indicate that SOM is optional, exclamation mark that SOM is obligatory.
and SF in other clause types (e.g. impersonal constructions) results from a diachronic change: as Falk (1993) already observed, Old Scandinavian has no overt expletive form, i.e. það is not present in the lexicon with an expletive function. However, once það begins to function as an expletive, the productivity of SF drops dramatically (Falk 1993, Rögnvaldsson 1984), arguably because það-insertion is in competition with SF and is a more economic strategy, under the hypothesis that merge wins on move.

To conclude, I have argued that SF and SOM are both strategies to check the features encoded on FinP. SF checks the verbal counterpart, [Finiteness], whereas the obligatory SOM checks the nominal counterpart [Definiteness]. This account generates some predictions concerning the syntax and the interpretation of the clauses in which these elements appears, and such predictions are borne out by comparative synchronic and diachronic facts. The optionality of non-subject specific SOM, as well as of SF is explained in the light of the respective discourse-configurational properties related to information structure and as a result of a complex diachronic reanalysis that created a fundamental division in the Scandinavian complementation system.

References


Bentzen K. *Order and Structure*. PhD Dissertation, University of Tromsø.


Lund University.


Sigurðsson H. Á. 2010. On EPP effects Studia Linguistica, 64, 2, 159–189.


http://www.linguist.is/icelandic_treebank