Eiríkur Rögnvaldsson
University of Iceland

ICELANDIC WORD ORDER
AND það-INSERTION

0. Introduction

The main purpose of this paper is to come up with a new analysis of the basic structure of Icelandic sentences, in many ways similar to the one proposed in Platzack (1983), but avoiding the wrong predictions of that analysis.

The paper is organized as follows: I will first review two analyses of Icelandic word order; those of Maling and Zaenen, and of Platzack. I will argue that Maling and Zaenen's approach, a surface filter which asserts that the finite verb is always in second position, is unmotivated, has no explanatory power, and occasionally makes wrong predictions. But the main part of the section is devoted to Platzack's analysis. I show that although his approach gets many important generalizations, it cannot be maintained as it stands, because it does not allow for many good sentences. In the last part of section one, I bring forth a new analysis, much in line with Platzack's, but avoiding his wrong predictions.

However, it could be regarded as a drawback of my analysis compared to Platzack's, that it does not predict the ungrammaticality of many subordinate relative and interrogative sentences with það-insertion. The second part of the paper is centered around these problems. I try to show that this is not limited to my analysis, as Platzack's analysis would also run into similar problems. I also argue that the possibility or impossibility of það-insertion in subordinate clauses is not to be explained on purely structural grounds, and we must also take notice of some semantic and functional factors.

1. Icelandic word order

1.1 V/2

During the past five years or so, Joan Maling and Annie Zaenen have

written a number of articles on Icelandic word order. The central claim of their analysis is that "Icelandic has a strict Verb-Second Constraint (V/2) that applies to all tensed clauses, embedded as well as main" (Maling & Zaenen 1978:491). Later, they admit that this is not quite correct (Maling 1980:176, Zaenen 1980:89-108); the verb is of course in first position in direct questions, imperatives, and conditionals without the conjunction ef 'if'; and sometimes even in declarative clauses (see below). It is also possible to find sentences with the finite verb in third position, as Maling (1980:176n) admits.

In Maling and Zaenen's formulation, V/2 is a surface filter, which is supposed to "explain" various things. It "explains" that we get subject-verb inversion in topicalized clauses, because otherwise the verb would end up in third position; and it also explains why the dummy það is often inserted when the subject is missing, because otherwise the verb would be in third position. But does this really explain anything? Sentences with the finite verb in first position are quite common in Icelandic written prose, especially in the "ongoing narrative" style characteristic of the Old Icelandic Sagas (see Kossuth 1981 and Halldór Ármann Sigurðsson 1983 on the conditions on V/1 word order). Haiman (1974:92) claims that this shows that Old Icelandic did not have the V/2 constraint. But subject-verb inversion accompanying fronting was just as obligatory in Old Icelandic as it is now. Thus, if Old Icelandic did not have the V/2 constraint, it is clear that this constraint cannot be used to "explain" subject-verb inversion in Old Icelandic; so why should it be used to explain the inversion in the modern language?

But the most serious weakness of Maling and Zaenen's approach is their treatment of "subject gaps". Observe the following sentences:

(1) Ég veit að $\left\{\begin{matrix}*\emptyset\\ það\end{matrix}\right\}$ hefur einhver lesið bókina

    'I know that there has somebody read the book'

(2) Ég veit að $\left\{\begin{matrix}*\emptyset\\ það\end{matrix}\right\}$ var dansað á skipinu í gær

    'I know that there was danced on the ship yesterday'

(3) Guðmundur veit ég að $\left\{\begin{matrix}*það\\ \emptyset\end{matrix}\right\}$ hefur lesið bókina

    'Guðmundur know I that has read the book'

(4) Þetta er maðurinn sem $\begin{Bmatrix} *\text{það} \\ \emptyset \end{Bmatrix}$ las bókina

'This is the man that read the book'

Zaenen (1980:95) states that "The main function of the dummy seems to be to keep the verb in second position in declarative sentences"; and Maling (1980:178) proposes that in sentences as (3) and (4), the "subject gap" is counted as first position, so that the verb is still in second position. But the question is: Why cannot all subject gaps be filled by það? And why cannot the subject gaps in (1) and (2) count as first position, but have to be filled by það? To be sure, Maling and Zaenen have an answer to the first question; það-insertion is normally not possible in "clauses which themselves inherently contain gaps, e.g., relatives, questions, and comparatives" (Maling 1980:189, cf. Zaenen 1980:105-106); or, in other words, "in binding domains það must be omitted" (Zaenen 1981:9). But this explanation is **ad hoc**, and does not follow from anything else. And to the second question, why an empty subject gap cannot always be counted as filling first position, we get no answer from Maling and Zaenen.

1.2 Platzack's analysis

The most serious attempt to explain the main features of Icelandic word order is Platzack's (1983). Working in the GB-framework (cf. Chomsky 1981), Platzack tries to explain the different word order of four Germanic languages (English, German, Swedish, and Icelandic) by different values of only two parameters; the basic order parameter and the COMP/INFL parameter. As regards Icelandic, the main features of the description are:

(5) a Both a COMP and an INFL node
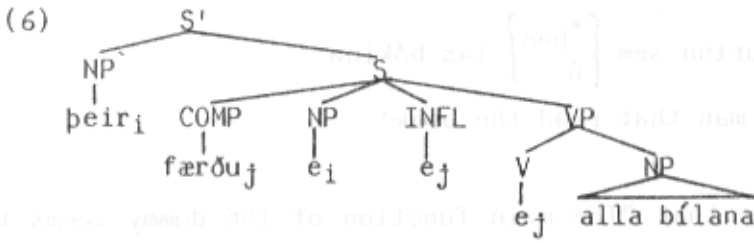
  b COMP is a constituent of S, not S'

  c S' branches into S and X^{max}, which is the landing site for topicalized phrases and the place where það is generated
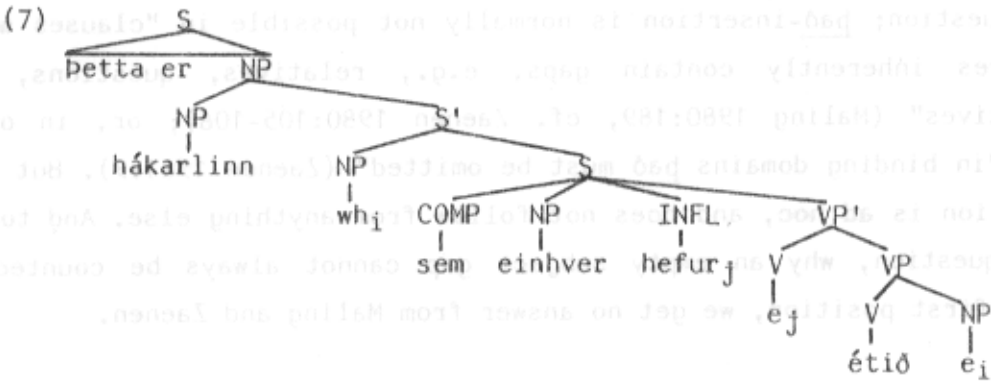
  d The finite verb can move to INFL, and sometimes from there to COMP, if necessary to prevent violations of the ECP

Let us now look at the constituent structure of four sentences, according to this description:
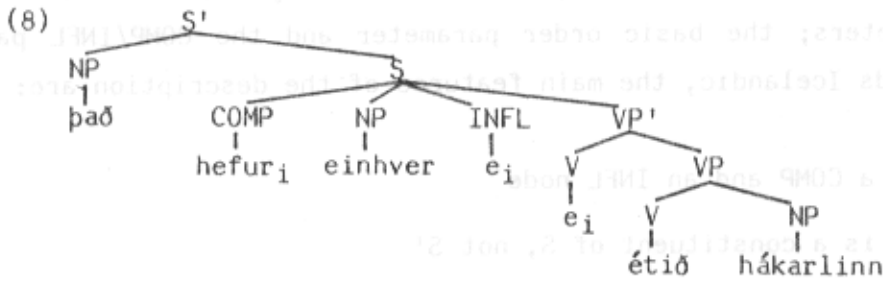
(6)

```
              S'
         ┌────┴──────┐
        NP           S
         |      ┌───┬──┴──────┐
       þeir_i  COMP NP   INFL      VP
               |    |     |      ┌──┴──┐
             færðu_j e_i  e_j    V     NP
                                 |    ╱  ╲
                               e_j  alla bílana
```

'They moved all the cars'

The finite verb has moved twice; from VP to INFL, and from there to COMP. The subject has moved to X^{max}; the result is an ordinary declarative.

(7)

```
          S'
     ┌────┴──────┐
   Þetta er      NP
              ┌──┴────────┐
             NP           S'
              |      ┌─────┴──────────────────┐
          hákarlinn NP                         S
                    |    ┌────┬─────┬──────┬────┐
                  wh_i  COMP  NP   INFL    VP
                        |     |     |    ┌──┴──┐
                       sem einhver hefur_j VP
                                        e_j ┌──┴──┐
                                           V      NP
                                           |      |
                                          étið    e_i
```

'This is the shark that someone has eaten'

Relatives are assumed to be derived by means of wh-movement; hence the wh-element in X^{max}.

(8)

```
         S'
     ┌───┴────────┐
    NP            S
     |    ┌────┬────┬─────┬──────┐
    það  COMP  NP  INFL   VP'
          |    |    |    ┌──┴──┐
        hefur_i einhver e_i   VP
                           ┌──┴──┐
                           V      NP
                           |    ┌─┴──┐
                          e_i   V     NP
                                |     |
                              étið  hákarlinn
```

'There has someone eaten the shark'

The verb moves twice, as in (6); but as það is generated in X^{max}, the subject does not get moved.

The most dubious feature in Platzack's proposals is his claim that the complementizer að has a double representation; it can either be generated under COMP, just as sem in (7), or under a special S''-level. This latter possibility is necessary as we often get topicalizations in að-clauses, and the topicalized element ends up in X^{max}, which is to the left of the COMP in (7). But in reality, topicalized elements are to the right of að; so that Platzack has to assume a structure like (9):

(9)

```
                      S''
                    /    \
                  að      S'
                        /    \
                     NP        S
                      |      / | \
                   Ólaf_i  /   |    \
                      COMP  NP  INFL   VP'
                       |     |    |    /  \
                    hafði_j Egill e_j  V    VP
                                       |   / \
                                       e_j V   NP
                                           |    |
                                          séð   e_i
```
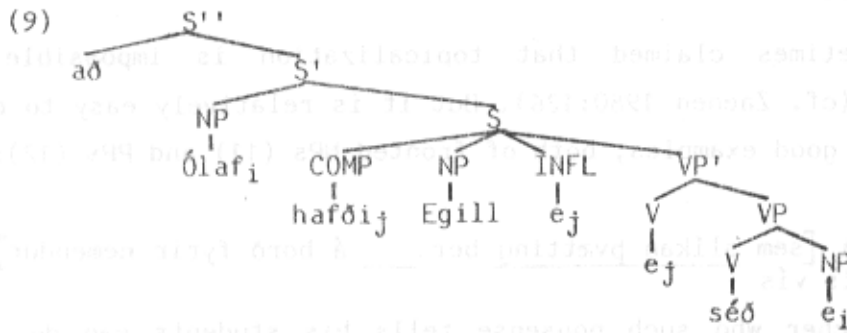
'That Ólaf (A) had Egill (N) seen'

The arguments for this behavior of að are weak, but Platzack mentions one: það is inserted under X^{max}, so að-clauses with það must have the S''-level; and S'' is considered an absolute barrier to extractions. Hence, it should be impossible to extract anything out of subordinate clauses with það. By and large this holds, but as we will see later on, there are some exceptions.

Now let us look at the predictions of this description as regards topicalization, extractions and það-insertion. They are mainly these:

(10)a Topicalization in relative clauses should be impossible, as the X^{max}-position, which is the landing site of topicalized phrases, is occupied by a wh-element. Even if it wasn't, topicalization would still be impossible, as the topicalized element would end up to the left of sem.

b Extractions out of topicalized clauses should be excluded. The reason is that subordinate topicalized clauses must have a special S''-level, and S'' is considered an absolute barrier to extractions.

c Extractions out of subordinate clauses with það should be excluded for the same reason; these clauses must have the S''-level.

d það-insertion should be impossible in relative clauses, for the same reasons as topicalization should be excluded (cf. (10)a).

I will now look at these predictions in turn, and show that none of them holds, although counterexamples to some of them are not easy to find.

1.3 Counterexamples

It is sometimes claimed that topicalization is impossible in relative clauses (cf. Zaenen 1980:126). But it is relatively easy to come up with perfectly good examples, both of fronted NPs (11) and PPs (12):
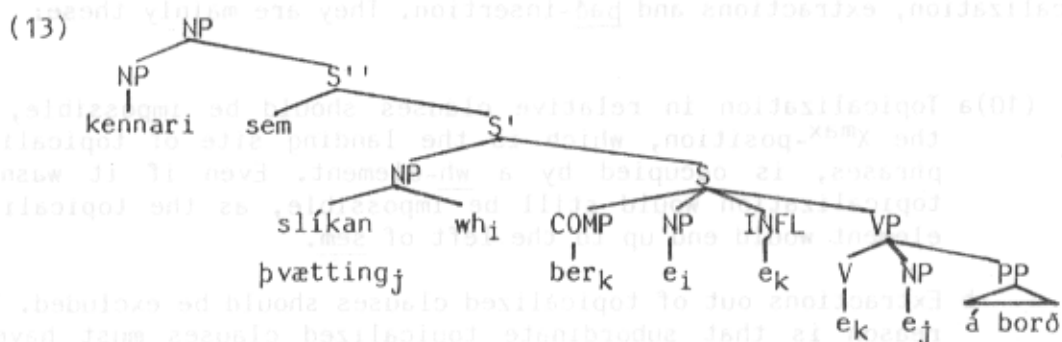
(11) Kennari [sem slíkan þvætting ber ___ á borð fyrir nemendur] er til alls vís

'A teacher who such nonsense tells his students can do any-thing'

(12) Flokkur [sem um fjögurra ára skeið hefur verið í stjórn ___] tapaði kosningunum

'A party which for four years has been in the government lost the election'

One could imagine that this could be solved by assuming a double represen-tation for sem, just as for að. Let us look at what that structure would look like:

(13)

$$
\begin{array}{c}
\text{NP} \\
\diagup \quad \diagdown \\
\text{NP} \qquad \text{S''} \\
| \qquad \diagup \diagdown \\
\text{kennari} \quad \text{sem} \quad \text{S'} \\
\qquad \diagup \diagdown \\
\text{NP} \qquad \text{S} \\
\diagup \diagdown \qquad \diagup | \diagdown \\
\text{slíkan} \quad \text{wh}_i \quad \text{COMP NP INFL VP} \\
\text{þvætting}_j \qquad \text{ber}_k \ e_i \ e_k \quad \text{V NP PP} \\
\qquad\qquad\qquad\qquad\qquad e_k \ e_j \ \text{á borð}
\end{array}
$$

As can be seen, this account does not solve the problem, because two elements have to be moved to X^{max}. If they are both moved, neither will c-command its trace, of course.

Two possible solutions to this problem come to mind. One is to assume that the wh-element ends up in COMP, not in X^{max}; and the verb stays in INFL. It is of course no new suggestion that wh-elements land in COMP, and they do so in English under Platzack's account. But if we assume that wh-elements can land in COMP, Platzack's explanation of the different word order of main and subordinate wh-interrogatives breaks down.

The second possibility is that relatives are not derived by means of wh-movement, rather by deletion in place. Then, there would only be the NP slíkan þvætting in X^{max}, and the c-command condition would not be violated. But the structure would still have some disadvantages. One is that we would expect það-insertion to be as good and usual with sem as with að, but it certainly is not. Another is that assuming a double representation for sem would be ad hoc; the extraction argument cannot be used here, because extraction out of relative clauses is (almost) always impossible anyway.

It is true that extractions out of topicalized clauses are usually bad. However, I think it is possible to find reasonably good examples (cf. also Rögnvaldsson 1982:180):

(14) Í gær$_i$ veit ég [að þar$_j$ hefur verið slegist ___$_j$ ___$_i$]
'Yesterday know I that there has been fought'

(15) Þessar bækur$_i$ hélt ég [að þér$_j$ myndi ekki nokkur maður lána ___$_j$ ___$_i$]
'These books thought I that you would not any man lend'

On the other hand, it is usually possible to extract out of subordinate wh-interrogative clauses. That is what Platzack's (1983) analysis predicts, because wh-clauses do not have the S''-level. But another explanation of the difference between topicalized clauses and wh-clauses is also available, according to Zaenen (1980:225, fn. 2):

> ... wh-elements are often secondary themes whereas topics and heads of relatives are primary themes (most likely this has to be related to the fact that wh-elements are indefinite and non-referential whereas topics and relative heads are typically definite and referential) and that one can only extract a primary theme over a secondary one, but not vice versa. Extraction of a theme over another one of equal importance is also difficult.

Assuming Kuno's (1980) definition of a **theme**, "what the rest of the sentence is about", it is clear that (14) is not about þar, nor is (15) about þér. It is also noteworthy that extractions out of clauses with Stylistic Inversion (cf. Maling 1980) are usually good (see Rögnvaldsson 1982:180-181); but it is characteristic of those elements that Stylistic Inversion moves that they do not qualify as themes according to Kuno's definition quoted above.

I will not go further into such functional matters; but it should at least be clear that other explanations than structural are available for the possibility or impossibility of extracting out of topicalized clauses.

It is also not easy to find good examples of extractions out of clauses with það-insertion, but I think that most speakers will not find these too bad:

(16) <u>Hvenær</u> heldur þú [að það geti allir keypt sér íbúð ____]?

'When think you that there can everyone by himself a flat'

(17) <u>Hvaða víntegund</u> heldur þú [að það sé drukkið mest af ____ á Íslandi]?

'Which sort of wine think you that there is drunk most of in Iceland'

(18) <u>Þessa mynd</u> finnst mér nú [að það verði allir að sjá ____]

'This picture think I that there must everybody see'

It must be stressed that such examples are hard to find; but the main point is that Platzack's (1983) analysis predicts that they should not occur at all, and that is undoubtedly wrong.

Now let's turn to það-insertion in relative clauses. There it is still harder to find reasonably good examples, but I think that many speakers would accept these two:

(19) Þetta er maðurinn [sem það var talað við ____ í sjónvarpinu í gær]

'This is the man that there was talked to on the TV yesterday'

(20) ?Þetta er stelpa [sem það eru margir skotnir í ____]

'This is a girl that there are many in love with'

Although these sentences would not win any beauty contest, there is a clear difference between them and (21):

(21) *þetta er stelpa [sem það ___ elskar alla]
'This is a girl that there loves everybody'

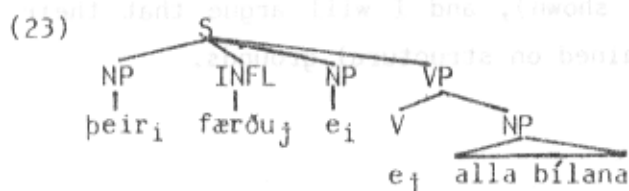Under Platzack's analysis, however, they should all be equally bad.
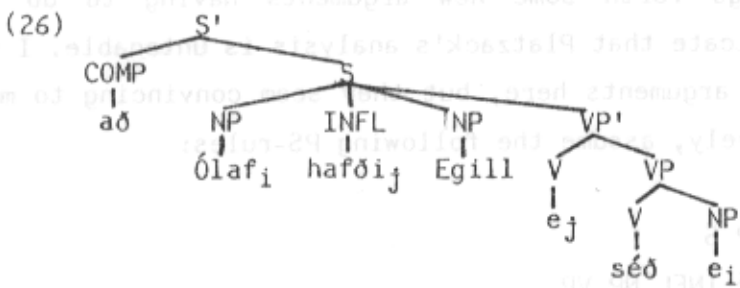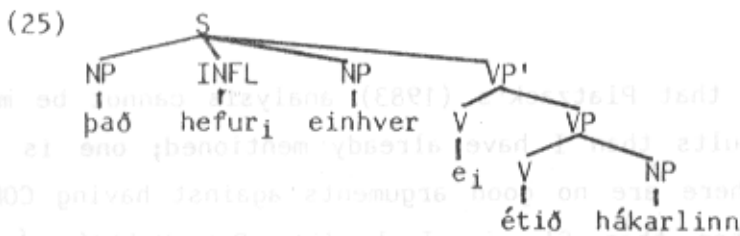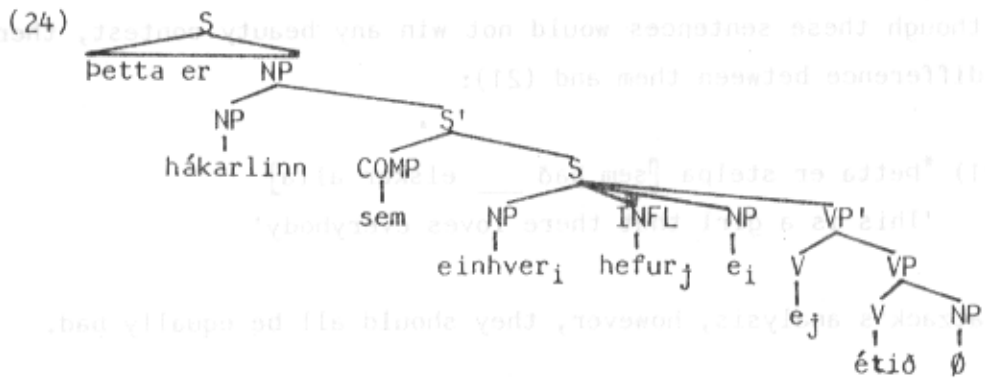

1.4 A new analysis

It is thus clear that Platzack's (1983) analysis cannot be maintained. It has more faults than I have already mentioned; one is that Platzack claims that there are no good arguments against having COMP a constituent of S, rather than S', in Icelandic. But Halldór Ármann Sigurðsson (1983) brings forth some new arguments having to do with coordination, which indicate that Platzack's analysis is untenable. I will not review Sigurðsson's arguments here, but they seem convincing to me. I will therefore, tentatively, assume the following PS-rules:

(22)a S' $\longrightarrow$ COMP S
  b S $\longrightarrow$ X^{max} INFL NP VP

These rules are made up so as to capture all of Platzack's relevant generalizations, while at the same time avoiding the wrong predictions his rules make. I think my rules succeed in this task, provided (i) that no movement is involved in relativization (on this point I accept Maling's (1977, 1978) and Zaenen's (1980) arguments); and (ii) that main clauses have only S, not S', so that wh-phrases will end up in X^{max} instead of COMP in main clauses; this explains the different word order between main and subordinate wh-interrogative clauses (Halldór Ármann Sigurðsson (1981:11) has argued that there is no reason to assume that COMP is present in main clauses in Icelandic).

Now let us look at the structure of (6)-(9) under my analysis:

(23)
```
              S
      _____|_____
     NP    INFL   NP    VP
     |      |     |    __|__
   þeir_i  færðu_j e_i  V     NP
                       __|____
                     e_j  alla bílana
```

(24)

```
              S
         _____|_____
      Þetta er        NP
                 _____|_____
               NP             S'
                |        _____|_____
           hákarlinn   COMP          S
                        |       _____|_____
                       sem    NP   INFL  NP    VP'
                               |    |    |   ___|___
                          einhver_i hefur_j e_i V     VP
                                            |   |   __|__
                                           e_j  V    V   NP
                                                     |    |
                                                    étið  Ø
```

(25)

```
              S
      _____|_____
    NP    INFL     NP         VP'
     |     |        |      ____|____
    það  hefur_i einhver   V       VP
                           |    ___|___
                          e_i   V     NP
                                |      |
                               étið hákarlinn
```

(26)

```
              S'
       _____|_____
    COMP              S
     |        _____|_____
    að      NP   INFL    NP            VP'
             |    |       |        _____|_____
          Ólaf_i hafði_j Egill     V         VP
                                   |      ___|___
                                  e_j     V     NP
                                          |      |
                                         séð    e_i
```

In conditional clauses without the conjunction _ef_, the finite verb is assumed to move to COMP. The adverb _kannski_ is also assumed to be base generated under COMP in sentences like (27), where a movement analysis is untenable (cf. Svavarsdóttir 1982):

(27) Kannski ég komi á morgun

'Maybe I come tomorrow'

The only case where Platzack's analysis would appear to be preferable concerns _það_-insertion in relative clauses and _wh_-clauses; my analysis does not predict the normally low acceptability of these sentences. However, they are not all ungrammatical, as Platzack's analysis would predict (and as I have already shown), and I will argue that their low acceptability is not to be explained on structural grounds.

2. Conditions on það-insertion

In this section, I will look at the conditions on the use of the dummy það in subordinate clauses. I will show that all the existing descriptions of its use are in some way misleading; and although my description in section 1 does not answer all questions on the use of það, it allows for its use where necessary. I admit that it must be supplemented in some way to exclude some of the possibilities that it allows for, but I try to show that exactly the same goes for the other proposals, especially Platzack's (1983).

2.1 Thráinsson's (1979) rule

Thráinsson (1979:478) suggests that það is inserted into empty subject slots by the rule (28):

(28) $[_{NP} \triangle] \ V \implies [_{NP} \text{ það}] \ V$

The empty subject slot can either be base generated as such, or the subject has been moved away by some movement rule. The V in the structural description is to ensure that the rule only applies to subject NPs, and that it does not apply if something else has been fronted (because empty NPs are assumed to invert with the verb just as lexical NPs if something is fronted).

But as Zaenen (1980, 1981) and Platzack (1983) have pointed out, this rule is much too general. It would predict that, e.g., both (29) and (30) should be good:

(29) *Ég spurði hvern það hefðu allir séð
     'I asked whom there had everybody seen'

(30) ?*Þetta er maður sem það sáu allir
     'This is a man whom there saw everybody'

This rule would also predict that (31) and (32) should be grammatical:

(31) *Þetta er maður sem það hefur lesið bókina
     'This is a man who there has read the book'

(32) *Guðmundur veit ég að það hefur lesið bókina
     'Guðmundur know I that there has read the book'

Notice that in these latter sentences, the empty subject slot is the
result of subject deletion (31) and subject extraction (32), whereas in
the first two sentences, the subject has been moved to the right. We will
see later that this is an important distinction.

2.2 Zaenen's (1980) rule

Zaenen (1980:10$_1$) gives the following rule of það-insertion:

(33) X   COMP $_S[$ V   W
      1   2   3   4
      1   2   það # 3   4

Here, það is not assumed to be inserted into subject position. This rule
is supplemented by one condition: "in binding domains það must be omitted"
(Zaenen 1981:9). A "binding domain" is defined as the stretch between a
binder (which can be a fronted element or the relative particle sem) and a
bindee (i.e., a gap). Hence, this condition would exclude, quite correct-
ly, all of the sentences (29)-(32). But it would also exclude quite a
number of reasonably good sentences, such as:

(34) Hvenær heldur þú að það geti allir keypt sér íbúð ___ ? (= (16))

(35) Hvaða víntegund heldur þú að það sé drukkið mest af ___ á
     Íslandi? (= (17))

(36) Þessa mynd finnst mér nú að það verði allir að sjá ___ (= (18))

(37) Daginn, sem það rigndi mest ___, var ég ekki heima
     'The day that there rained most, was I not home'

It is quite clear that we must somehow allow for such sentences, while
still excluding at least (31) and (32).

2.3 Platzack'z explanation

Platzack's (1983) proposals make essentially the same claims on the possibilities of það-insertion as Zaenen's. There is, however, an important difference: Platzack's proposals follow directly from his analysis of the basic structure of Icelandic sentences. Hence, he does not have to stipulate when það must not be inserted. According to his proposals, það is excluded in relative clauses because the X^{max}-position, where það is generated, is occupied by a wh-element (besides, the X^{max}-node is to the left of COMP, where sem is generated). In subordinate wh-interrogative clauses, það is bad for the same reason; the X^{max}-position is occupied by the wh-phrase.

Platzack's account is obviously superior to Zaenen's, because the former gives a principled explanation of the impossibility of það-insertion in certain contexts. But at the same time, Platzack's proposal suffers from the same weaknesses as Zaenen's; i.e., it excludes reasonably good sentences such as (34)-(37). Hence, his analysis must also be abandoned, because it does not make the right predictions in all cases. However, it would obviously be desirable to have some principled explanation of these problems. My analysis is an attempt to catch a part of the explanatory power of Platzack's analysis, while at the same time avoiding its wrong predictions.

2.4 Predictions of the new analysis

Now turn to my analysis. It is clear that it allows for all of the sentences in (34)-(37), because the X^{max}-position is assumed to be to the right of the COMP, and both sem and wh-phrases are in COMP. As explained in section 1, no movement is assumed to be involved in relative clauses, so that no wh-element occupies the X^{max}-position there. But the obvious problem with my analysis is, of course, that it not only allows for the good sentences in (34)-(37); it also allows for the bad sentences (29) and (30), which Platzack's analysis correctly excludes. So the score appears to be 1:1.

At first, it could appear that my analysis also predicts that (31) and (32) should be good. But they are ruled out in a principled manner. In (32), we may assume that we have the trace of Guðmundur in the X^{max}-position, which can therefore not be filled by það. In (31), we can either assume that both the X^{max}-node and its content have been deleted (cf.,

e.g., Chomsky & Lasnik 1977), or we can say that the X^{max}-position is occupied by the trace of the subject. Such sentences are thus explicitly ruled out, and I have not been able to find any examples of them which come near to being acceptable. My analysis thus predicts a principled distinction betwen such sentences on one hand and sentences like (29)-(30) and (34)-(37) on the other hand; Platzack's and Zaenen's analyses predict that they should all be equally bad. As I have already shown, many sentences of the latter type are in fact acceptable, so that the distinction seems to be motivated.

2.5 Troubles with the new analysis?

If my analysis is accepted, it follows that a new explanation must be found for the ungrammaticality of sentences like (29) and (30). Platzack (1983) mentions that this problem can be avoided "if we assume that Indefinite Subject Postposing for some reason is impossible in relative clauses or subordinate wh-interrogatives, or we could propose a filter that rules out sentences with það after sem and after a wh-word". But these solutions are undesirable, Platzack claims, as they "would not explain [my stress] why það does not occur in these contexts".

Now it is clear (as Platzack's examples indeed show) that these sentences are also bad without það, if Indefinite Subject Postposing has applied. This could, however, be attributed to violations of the V/2 constraint. But fortunately, it is possible to test whether it is the það-insertion or the postposing by itself which makes the sentence bad. This is because það-insertion is not the only way to save sentences with subject postposing form violating V/2; they can also by saved by a fronting process which Maling (1980) calls **Stylistic Inversion**. Stylistic Inversion can always apply if það-insertion can apply, but not vice versa (cf. Maling 1980:189); but it needs an independently created subject gap. This means that Stylistic Inversion cannot cause subject-verb inversion by itself; it can only apply if the subject has been deleted or moved to the right by an independent rule such as Indefinite Subject Postposing. Hence, if Indefinite Subject Postposing is independently motivated in relative and subordinate interrogative clauses, Stylistic Inversion should be good there. However, (38)b is at least as bad as (38)a:

(38)a ??Þetta er bók sem það hafa allir lesið

'This is a book that there has everybody read'

b ??Þetta er bók sem lesið hafa allir

'This is a book that read has everybody'

The same goes for interrogatives:

(39)a ?*Hann spurði hvað það hefði einhver étið

'He asked what there had somebody eaten'

b ?*Hann spurði hvað étið hefði einhver

'He asked what eaten had somebody'

This appears to indicate that it is the postposing, but not the það-insertion, which makes such sentences bad. Hence, they pose no problem for my analysis.

However, some problems still remain. My analysis thus predicts that (40) should be good with það, because there we have a base generated empty subject position:
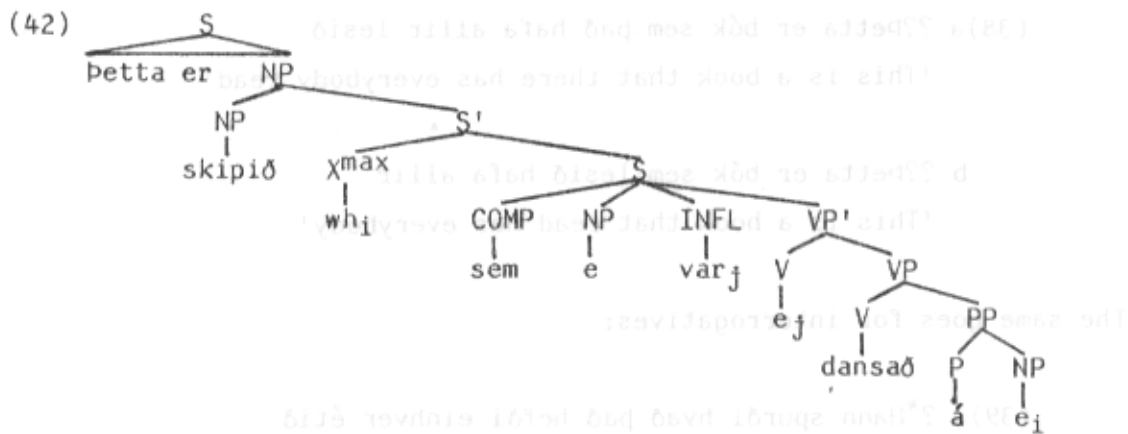
(40) Ég vissi ekki hvar $\begin{Bmatrix} *það \\ \emptyset \end{Bmatrix}$ hefði verið dansað

'I knew not where (there) had been danced'

We could of course claim that a wh-element somehow binds an adjacent empty X^{max}-position, but that move would be **ad hoc**. I do not have an explanation to offer, but I must point out that Platzack's (1983) analysis runs into similar problems with sentences like (41):

(41) Þetta er skipið sem $\begin{Bmatrix} ?það \\ \emptyset \end{Bmatrix}$ var dansað á ___ í gær

'This is the ship that (there) was danced on yesterday'

Most people would prefer the það-less version. Let's look at the structure of this sentence, according to Platzack's analysis:

(42)

```
                    S
        ┌───────────┴───────────┐
    Þetta er              NP
              ┌────────────┴────────────┐
            NP                          S'
            │              ┌────────────┴────────────┐
          skipið        X^max                        S
                          │          ┌───────┬───────┬────────┐
                        wh_i        COMP    NP    INFL      VP'
                                     │       │      │      ┌──┴──┐
                                    sem      e    var_j    V     VP
                                                           │   ┌──┴──┐
                                                          e_j  V     PP
                                                               │   ┌──┴──┐
                                                            dansað P    NP
                                                                   │     │
                                                                   á    e_i
```
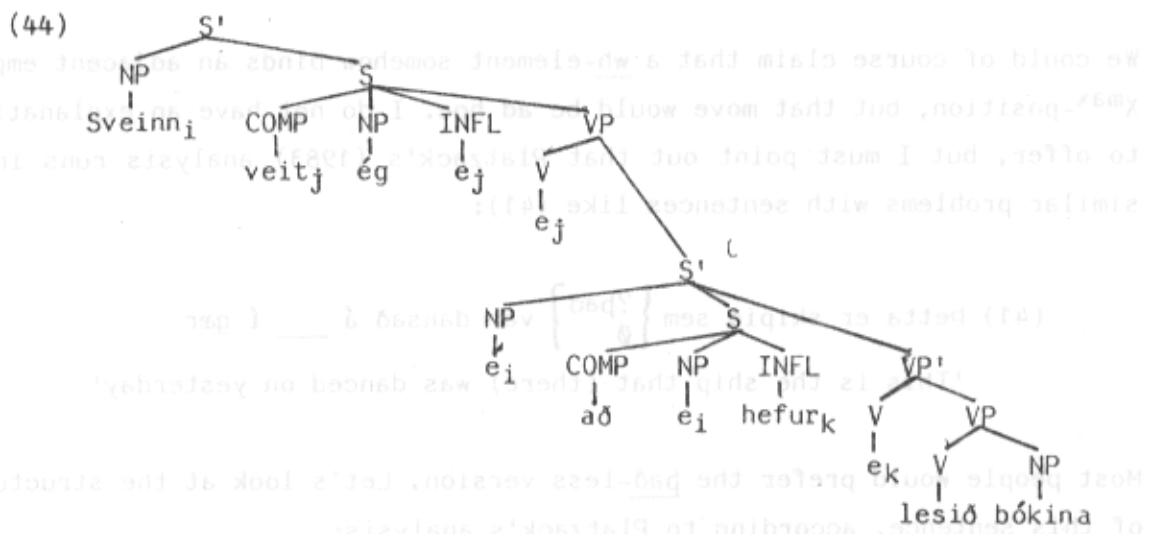
Where is the governor of the subject NP in this sentence? I don't know;
but I presume that if a governor can be found for the empty category in
(42), it is also possible to find one for the empty category in (40).
Note, however, that Zaenen's (1981) prohibition against það-insertion in
"binding domains" would predict that both (40) and (41) should be ungram-
matical with það.

Similar problems will also face Platzack's analysis in the descrip-
tion of sentences like (43):

(43) <u>Sveinn</u> veit ég [að ___ hefur lesið bókina]
'Sveinn know I that has read the book'

The structure of this sentence, assuming Platzack's analysis, must be:

(44)

```
                    S'
        ┌───────────┴───────────┐
      NP                        S
        │          ┌─────┬──────┬──────┐
    Sveinn_i      COMP   NP   INFL    VP
                   │      │     │      │
                 veit_j  ég   e_j      V
                                       │
                                      e_j
                                         └──────────────┐
                                                        S'
                                            ┌───────────┴───────────┐
                                          NP                        S
                                           │          ┌─────┬──────┬──────┐
                                          e_i        COMP   NP   INFL    VP'
                                                      │      │     │     ┌──┴──┐
                                                      að    e_i  hefur_k V    VP
                                                                         │  ┌──┴──┐
                                                                        e_k V    NP
                                                                            │
                                                                        lesið bókina
```

The explanation of the ungrammaticality of the English sentences which
correspond to (43) is usually assumed to be that the empty subject has no
proper governor. Now these sentences are clearly grammatical in Icelandic
(cf. Maling & Zaenen 1978). I don't know how Platzack's analysis would

proper governor. Now these sentences are clearly grammatical in Icelandic (cf. Maling & Zaenen 1978). I don't know how Platzack's analysis would explain that; but the same explanation should also be able to handle the það-less version of (40) and (41).

2.6 The obligatoriness of það-insertion

At last, let us look a bit at the obligatoriness or impossibility of það-insertion. Maling & Zaenen (1978) claim that "það is not optional in Icelandic; it is either required or simply impossible". And Zaenen (1981:9) says that "það can only occur if nothing else prededes the tensed part of the verb, and in that case it is obligatory both in main and in embedded clauses". This holds for declarative main clauses (with the exception of the "ongoing narrative" style mentioned in section 1); but in certain types of subordinate clauses, það-insertion appears to be optional, for many speakers at least. Let us look at some examples:

(45) Mér leiðist alltaf þegar $\left\{\begin{matrix} \text{það} \\ \emptyset \end{matrix}\right\}$ rignir

   'I am bored always when (there) rains'

(46) Ég vissi að $\left\{\begin{matrix} \text{það} \\ \emptyset \end{matrix}\right\}$ væri ekið vinstra megin í Ástralíu

   'I knew that (there) were driven on the left side in Australia'

Note especially (46); it may be contrasted with an example from Maling (1980:189):

(47) Ég veit að $\left\{\begin{matrix} \text{það} \\ *\emptyset \end{matrix}\right\}$ er ekið vinstra megin í Ástralíu

   'I know that (there) is driven on the left side in Australia'

Since það is never optional in main clauses, it should be possible to find something which those subordinate clauses where það is obligatory have in common with main clauses; subordinate clauses where það is optional (or where it is impossible, even if there is an empty X^{max}-slot available) should, on the other hand, differ from main clauses in some way.

Now it is a well-known fact that main clause features are more likely to be found in those subordinate clauses that are assertions than in other subordinate clauses (cf. Hooper & Thompson 1973, Platzack 1983). We would thus expect það to be obligatory in asserted subordinate clauses, just as it is in main clauses; but where we have instances of optional það, we would not expect the clauses to be asserted. The "minimal pair" in (48)-(49) gives some support to this proposal, I think:

(48) Veistu að $\begin{Bmatrix} það \\ ??\emptyset \end{Bmatrix}$ er rigning úti?

'Know you that (there) is rain outside'

(49) Ég vissi ekki að $\begin{Bmatrix} það \\ \emptyset \end{Bmatrix}$ væri rigning úti

'I knew not that (there) were rain outside'

By uttering (48), the speaker is asserting that it rains; making the fact known to his audience. In (49), on the other hand, the rain is presupposed; what the speaker is making known to the audience is that he is aware of the rain, but not that it rains. I think that a similar contrast can be found between (46) and (47), although it is perhaps not as obvious.

It is interesting in this connection to note that it appears to be easiest to find good examples of optional það in adverbial clauses. This agrees well with what Hooper & Thompson say (1973:494): "Some adverbial subordinate clauses, such as those beginning with when, before, and after, ... are always presupposed ..." They also argue that restrictive relative clauses are always presupposed - a claim which agrees well with the low acceptability of relative clauses with það.

It is also possible that assertion is involved in the difference between (50) and (51):

(50) ?Þetta er saga sem það er einhver að lesa í útvarpinu núna

'This is a story which there is somebody reading on the radio now'

(51) ??Það er mjög góð saga sem það er einhver að lesa í útvarpinu núna

'It is a very good story which there is somebody reading on the radio now'

Neither sentence is perfect, but many speakers find (50) considerably

better than (51). The former could be paraphrased as "There is someone reading this story on the radio now", but the latter might be paraphrased as "It's very good, this story which someone is reading on the radio now". Thus, (50) presupposes that someone is reading a story, whereas (51) does not. The relative clause in (50) is therefore like a main clause in many respects, and I find it likely that this is the reason for the greater acceptability of <u>það</u> there.

3. Conclusion

What I am suggesting, then, is that the conditions on the use of <u>það</u> in subordinate clauses are the results of an interplay between formal and functional (or semantic) factors. The syntactic prerequisites for the use of <u>það</u> are that an empty X^{max}-position is available; but the functional factors are more complex, and I will not try to work them out here.

That functional or semantic factors are involved should come as no surprise, of course. These factors govern the use of <u>það</u> in main clauses; <u>það</u> can only be used in a clause with a postposed subject if the subject can be interpreted as new information (cf. Rögnvaldsson 1983). This is further complicated by the fact that if the postposed subject is definite, <u>það</u> cannot be used if the subject is only moved immediately to the right of the verb; it must be moved to the end of the clause if <u>það</u> is to be accepted. I do not know the reason for this; and I do not know either the reason for the fact that Indefinite Subject Postposing and <u>það</u>-insertion appear to be more acceptable in relative clauses if the relativized NP is the object of a preposition than if it is the object of a verb, as (52)-(53) show:

(52) ?*Þetta er stelpa sem það elska ___ margir
'This is a girl who there love many'

(53) ??Þetta er stelpa sem það eru margir skotnir í ___
'This is a girl who there are many in love with'

I will not try to explain this difference; but my conclusion is that a purely structural account of the possibilities of <u>það</u>-insertion is bound to fail. Hence, it presents no special problem for my analysis that it does not account for all the peculiarities of <u>það</u>-insertion; no structural analysis can hope to do that.

References

Chomsky, Noam. 1981. **Lectures on Government and Binding**. The Pisa Lectures. Foris Publications, Dordrecht.

Chomsky, Noam, & Howard Lasnik. 1977. Filters and Control. **Linguistic Inquiry** 8:425-504.

Haiman, John. 1974. **Targets and Syntactic Change**. Mouton, The Hague.

Hooper, Joan B., & Sandra A. Thompson. 1973. On the Applicability of Root Transformations. **Linguistic Inquiry** 4:465-497.

Kossuth, Karen C. 1981. Unmarked Definite NPs and Referential Cohesion in Old Icelandic Narrative. **Íslenskt mál** 3:85-100.

Kuno, Susumu. 1980. Functional Syntax. Edith A. Moravcsik & Jessica R. Wirth (eds.): **Syntax and Semantics**, Vol. 13, Current Approaches to Syntax, pp. 117-135. Academic Press, New York.

Maling, Joan. 1976. Old Icelandic Relative Clauses: An Unbounded Deletion Rule. **Proceedings of the Seventh Annual Meeting of the North Eastern Linguistic Society**, pp. 175-188.

_____. 1978. An Asymmetry with Respect to Wh-Islands. **Linguistic Inquiry** 9:75-89.

_____. 1980. Inversion in Embedded Clauses in Modern Icelandic. **Íslenskt mál** 2:175-193.

Maling, Joan, & Annie Zaenen. 1978. The Nonuniversality of a Surface Filter. **Linguistic Inquiry** 9:475-497.

Platzack, Christer. 1983. Germanic Word Order and the COMP/INFL Parameter. **Working Papers in Scandinavian Syntax** 2.

Rögnvaldsson, Eiríkur. 1982. **Um orðaröð og færslur í íslensku**. M.A.-thesis, University of Iceland, Reykjavík (mimeographed).

Sigurðsson, Halldór Ármann. 1981. Dukunarögn um íslenskar aukatengingar. Unpublished paper, University of Iceland, Reykjavík.

_____. 1983. **Um frásagnarumröðun og grundvallarorðaröð í forníslensku**. M.A.-thesis, University of Iceland, Reykjavík (mimeographed).

Svavarsdóttir, Ásta. 1982. Lokaskýrsla um "Ferðasögu Árna Magnússonar frá Geitastekk 1753-1797". Unpublished paper, University of Iceland, Reykjavík.

Thráinsson, Höskuldur. 1979. **On Complementation in Icelandic**. Garland Publishing, New York.

Zaenen, Annie. 1980. **Extraction Rules in Icelandic**. Unpublished Doctoral Dissertation, Harvard University.

_____. 1981. Characterizing Syntactic Binding Domains. **Occasional Paper** #17, MIT Center for Cognitive Science.