Extraction and Pronoun Preposing in Scandinavian

Elisabet Engdahl 1,* and Filippa Lindahl 1,2

1 Department of Swedish, University of Gothenburg, Box 200, 405 30 Gothenburg, Sweden
2 Division of Educational Science and Languages, University West, 461 86 Trollhättan, Sweden; filippa.lindahl@hv.se
* Correspondence: elisabet.engdahl@svenska.gu.se

Abstract: It has been noted that examples with extractions out of relative clauses that have been attested in Danish, Norwegian and Swedish are judged to be unacceptable in Icelandic and Faroese. We hypothesize that this may reflect whether or not speakers tend to prepose unstressed object pronouns as a way of establishing a coherent discourse. In this article we investigate to what extent pronoun preposing is used in Swedish, Icelandic and Faroese and whether there is any correlation with the acceptability of extractions from relative clauses. We show that Icelandic speakers use pronoun preposing to a very limited extent whereas Faroese speakers often prepose the VP or sentential anaphor tað. In both languages extraction from relative clauses is mainly judged to be unacceptable, with Faroese speakers being somewhat more accepting of extraction from presentational relatives. A crucial factor seems to be whether preposing is associated with a marked, contrastive interpretation or not.

Keywords: preposing; topicalization; Faroese; Icelandic; Swedish; contrastive topic; continued topic; VP ellipsis

1. Introduction

At least since Ross’ dissertation (Ross 1967), linguists have been trying to characterize the restrictions on extractions in natural languages. Ross showed that many types of extractions could be subsumed under a structural constraint, the so called Complex NP Constraint CNPC, which prohibited extraction out of a clause dominated by a noun phrase. The CNPC was soon subsumed under subjacency (Chomsky 1973) and in the following years, a number of proposals were made for what the relevant bounding nodes were in different languages in order to account for the cross-linguistic variation that was found (see, e.g., Rizzi 1990; Taraldsen 1981). However, soon Scandinavian speaking linguists such as Erteschik-Shir (1973); Andersson (1975) and Allwood (1976) pointed out that certain extractions out of relative clauses were possible in the mainland Scandinavian languages Danish, Norwegian and Swedish. In the following years several linguists approached this topic from different perspectives (see, e.g., the papers in Engdahl and Ejerhed 1982). When written and spoken corpora became available, studies based on spontaneously produced written or spoken extractions started to appear see, e.g., (Engdahl 1997; Jensen 2002; Lindahl 2010, 2011, 2017b).

The documented spontaneous extractions mainly come from mainland Scandinavian whereas the insular Scandinavian languages Faroese and Icelandic appear to behave more similar to English when it comes to extractions from relative clauses (Platzack 2014; Thráinsson 2007; Thráinsson et al. 2004; Zaenen 1985). There have been occasional reports of spontaneous extractions from relative clauses in English (see, e.g., Chung and McCloskey 1983; McCawley 1981)1 and similar examples are reported by Cinque (2013) to be at least marginally acceptable in Italian, French, and Spanish. However, they do not seem to be used productively the way they are in mainland Scandinavian, see the recent corpus investigations in Kush et al. (2021) and Müller and Eggers (2022). This suggests that the mainland Scandinavian languages have some common property which facilitates long extractions. In this article we explore the hypothesis that the common property that sets...
the mainland Scandinavian languages apart is the tendency to use preposing (topicalization) of unstressed pronouns as a way of connecting utterances in spoken language as well as sentences in texts. We test this hypothesis by comparing how pronoun preposing is used in spoken Swedish, a mainland Scandinavian language, with similar data from the insular languages Icelandic and Faroese. In this article we bring together results from previous studies that we have carried out on individual languages and apply an explicit comparative perspective.

We start in Section 2 by looking at what characterizes spontaneous extractions from relative clauses in Swedish. Since a large proportion of them involve pronouns, we turn, in Section 3, to the question how pronoun preposing is used more generally in Swedish and what similarities there are between local preposing (within a clause), long preposing (from an embedded clause) and what is often called extraction (long preposing from a relative clause). In Section 4, we look at preposing and extraction in Icelandic and in Section 5 we present relevant data from Faroese. In Section 6, we look briefly at the ways preposing is used in English, German and Dutch before evaluating our hypothesis in Section 7.

2. What Do Spontaneous Extractions Look Like?

The literature on extractions from relative clauses is largely based on constructed examples—not surprising given that most of the studies deal with languages where such extractions are not used in ordinary conversations. It is clearly a good idea to start from the types of extractions from relative clauses that are actually used. Engdahl (1997) looked at 30 naturally occurring extractions in Swedish, both written and spoken. Some examples from this article are given in (1). We mark the relative clause and the gap site of the extracted element.

(1) a. \( {}^{1}\text{det} \) finns det ingen [som kan hjälpa mig med \( {}_1^\text{-} \)]
   \( \text{it exist EXPL nobody that can help me with} \) 'There is nobody who can help me with it.' (spoken)

b. \( {}^{1}\text{det} \) var det ingen [som ville \( {}_1^\text{-} \)]
   \( \text{it was EXPL nobody that wanted} \) 'There was nobody who wanted to.' (spoken)

c. \( {}^{1}\text{där} \) har jag en moster [som bor\( {}_1^\text{-} \)]
   \( \text{there have I an aunt who lives} \) 'I have an aunt who lives there.' (spoken)

d. tapeterna\( {}_1^\text{DEF} \) var det Sven [som valde \( {}_1^\text{-} \)]
   \( \text{wall paper.DEF was EXPL Sven that chose} \) 'It was Sven who chose the wall paper.' (spoken)

e. ... ett oromantiskt namn\( {}_1^{} \) som jag då inte kände någon [som hette \( {}_1^\text{-} \)]
   \( \text{an unromantic name that I then not knew anyone who was called} \) 'an unromantic name which I did not know anyone at the time who was called.'

   (novel 1996)

In (1a,b) the pronoun \( {}^{1}\text{det} \) ‘it’ has been preposed from a relative clause which modifies the indefinite pivot in an existential sentence; we refer to this construction as a presentational relative, see Lambrecht (1988) and Chaves and Putnam (2020, p. 27). In (1c) a short deictic adverb \( {}^{1}\text{där} \) ‘there’ has been preposed from a have-construction which is very similar to presentational sentences (Keenan 1987; McCawley 1981). (1d) is a cleft construction where the object tapeterna ‘the wall paper’ has been preposed from a cleft clause; the clefted constituent Sven is definite. (1e) is a relative clause where the relativized item namn ‘name’ has been extracted from a relative clause which modifies the object of the lexical verb kännna ‘be acquainted with’.

These types of constructions, preposing out of presentational relatives or cleft constructions and preposing or relativization from an embedded relative, were most common in
Engdahl’s (1997) sample. There were a few examples involving preposing of a *wh*-phrase; they all had the form in (2) which resembles a cleft.

(2) vem1 var det ingen [som kände _1 ?]
   who was EXPL nobody that knew
   ‘Who did nobody know?’

This is not surprising given that *wh*-questions in Swedish often are clefted with the clefted constituent preposed, as in (3) (Brandtler 2019).

(3) vem1 var det _1 [som kom?]
   who was EXPL that came
   ‘Who came?’

In (2), the clefted constituent remains in situ and *vem* is linked to a position inside the cleft clause.5

The relative rareness of spontaneous examples with preposed *wh*-phrases fits well with the experimental findings in Kush et al. (2018); Kush et al. (2019) and the corpus studies in Müller and Eggers (2022).

The majority of Engdahl’s examples involved extractions out of relative clauses where the subject has been relativized, as in the examples shown so far. There were also a few cleft examples with object relatives, as in (4).

(4) matte1 var det bara pappa2 [jag kunde fråga _2 om _1]
   maths was EXPL only dad I could ask about
   ‘It was only dad that I could ask about maths.’

In some of the examples the preposed phrase was a lexical DP, as in (1d) and (4), but a large proportion of Engdahl’s examples involved either an anaphoric pronoun, as in (1a,b), or a light adverb, as in (1c). Engdahl showed that preposing of unstressed pronouns in Swedish is used as a way of establishing a coherent discourse and suggested that this might be what distinguishes the mainland Scandinavian languages from other languages, essentially the hypothesis that we are testing in this article. In the next section we look first at local preposing in Swedish and then return to extractions from relative clauses.

3. Swedish

The type of preposing of pronouns that we have seen in (1a,b,c) is of course not only found in extraction sentences. Local preposing, in the same clause, of object pronouns is quite common in spoken and written Swedish, as well as in Danish and Norwegian. In this section, we present data from Swedish but similar data are found in Danish and Norwegian, see Engdahl and Lindahl (2014).

3.1. Pronoun Preposing

Starting an utterance with a pronoun is very common in spoken language and Swedish is no exception to this. In most cases the pronoun functions as the subject of the sentence but in the mainland Scandinavian languages it is quite common to start with an object pronoun, or a subject pronoun from an embedded clause.6 In order to find out when this word order is used, we have conducted several studies using the Nordic Dialect Corpus (NDC) (Johannessen et al. 2009) which contains recordings and aligned transcriptions of informal conversations in all the Nordic languages (Lindahl and Engdahl Forthcoming).

The Swedish part of the NDC consists of 361 184 words, produced by 133 speakers in 37 locations. Consider the following examples, taken from the NDC. We identify the interviewer by *int* and the speakers by *s1, s2*, etc. The antecedent of the pronoun is underlined and the preposed pronoun is italicized. (.) marks a short pause, and = that the speaker continues on the next line.
In both (5) and (6), the preposed element is the third person singular neuter pronoun det ‘it’. In (5) the pronoun acts as a VP anaphor which refers to the action gå i skolan ‘go to school’, expressed as a VP in the interviewer’s question. In (6) det is a propositional anaphor, referring to the event just described by the sentence de hade slagits ‘they had been fighting’. Here det has been preposed from an embedded clause. In both utterances, det was unstressed.

Preposing of VP and propositional anaphors, as in (5) and (6), is very common in the NDC but there are also examples with a preposed pronoun which refers to a recently mentioned entity, as in (7) (somewhat abbreviated).

(7) **int:** när köpte du din första bil?
   ‘When did you buy your first car?’

   **a.** s1: den köpte jag 1980
   it bought I 1980
   ‘I bought it in 1980.’

Here the pronoun den agrees in gender and number with the antecedent bil, which is the only likely referent. The pronoun is unstressed. If it had been stressed, it would have been understood as a demonstrative, that, but this interpretation would not have been plausible in this context. Unstressed den and det are interpreted as personal pronouns whereas the stressed versions function as distal demonstrative pronouns (Faarlund 2019, p. 27). We gloss the unstressed occurrences as ‘it’ and the stressed ones as ‘that’. As we will show in Section 6.1, this type of non-contrastive preposing is hardly used in English.

In (Lindahl and Engdahl Forthcoming) we investigated how preposing of pronouns is used in the Swedish part of the NDC. We found that in a large majority of cases the preposed pronoun referred back to something that was newly introduced, an action as in (5), an event as in (6), or an entity, as in (8a). Following Erteschik-Shir (2007) we refer to this pattern as focus chaining since the antecedent is (part of) the relational focus of the preceding utterance, i.e., the new information that is asserted or questioned about the topic (Gundel and Fretheim 2004, p. 177). There were also cases where the antecedent of the preposed pronoun was the topic of the preceding utterance, often realized as a subject. This kind of topic chaining was mainly found when the antecedent was an entity, as in (8b) (we just provide the English version of the interviewer’s question).

(8) **int:** Are you in touch with anyone who did their military service with you?

   **s1:** ja (.) det var två stycken andra plutonsjukvårdare=
   yes there were two other paramedics

   **a.** dom var från Fagersta=
   they were from Fagersta
b. så dom₁ har jag ganska bra kontakt med _₁
   so them have I pretty good contact with
   ‘Yes, there were two other paramedics. They were from Fagersta, so I have pretty good contacts with them.’

The first occurrence of dom ‘they’ in (8a) is an instance of focus chaining. Once dom is realized as a subject, it becomes the topic of that utterance. Consequently the second dom in (8b) is in a topic chain.

In the examples shown above the preposed pronouns are unstressed. There are also examples where an initial pronoun is stressed and receives a contrastive interpretation, as in (9).

(9) int: jo sen finns det kontaktlinser också
    yes then are EXPL contact lenses also
    ‘Yes, then there are contact lenses as well.’

s₁: ja nä DET₁ vill jag inte ha _₁ vet du
    yes no that want I not have know you
    ‘No, THAT I do not want, you know.’

By stressing the pronoun and negating the utterance, the speaker emphasizes that he definitely does not want contact lenses.

We found that preposing of pronouns is primarily used in contexts where the antecedent and the pronoun formed a focus chain or a topic chain. In 85% of the investigated cases, the pronoun was locally preposed and in 15% we had long preposing from a subordinate clause, as in (6). One effect of the preposing is that the initial pronoun is interpreted as the aboutness topic of the utterance (Reinhart 1981) which we can show using Reinhart’s rewriting test. We here apply the test to s₁’s utterance in (8b).

(10) a. s₁ said about them (the two paramedics) that he had pretty good contacts with them.

The paraphrase (10) works well and we have a good indication that the preposed dom in (8b) acts as the aboutness topic for that utterance.

When it comes to relative frequency of object preposing in spoken Swedish, Jörgensen (1976) showed that this varies with the type of spoken interaction. Based on the material in Talbanken, he found 14% object initial sentences in interviews, and 9% in conversations and debates (Jörgensen 1976, p. 103). In a more recent study the second author investigated the word order patterns in the Swedish part of the NDC, Lindahl (under review). To obtain an unbiased sample, 1000 sentences were extracted from the corpus with the only criterion being that the sentence should have a finite verb. The sentences were manually investigated and all declarative main clauses, 712 in total, were further analyzed. In total, 46 of the sentences in that sample, or 6.5%, had a preposed object and of these, 36 were pronominal (see Table 1 in Section 4.1).

Studies of spoken Swedish show that starting a sentence with det is the most common pattern and in most cases it functions as an expletive subject (Allwood 1999; Engdahl 2012). In Engdahl and Lindahl (2014) where we specifically looked for preposed object pronouns (including subject pronouns from embedded clauses), 95% of the hits were det. The reason for this is that preposing of VP anaphors or propositional anaphors is by far more common than preposing of pronouns with entity antecedents. Mikkelsen (2015) claims that preposing of VP anaphors is actually obligatory in Danish whereas we argue that there is a strong preference for preposing in Swedish but not a grammatical constraint, Lindahl and Engdahl (Forthcoming). According to Bentzen and Anderssen (2019), preposing of VP and propositional anaphors is always an option in Norwegian.11
3.2. Extraction from Relative Clauses

As already mentioned, the first author’s investigation was limited to 30 authentic examples (Engdahl 1997). For her dissertation, the second author carried out a larger empirical study of extraction from relative clauses in Swedish based on a collection of 270 spontaneous examples, gathered between 2011 and 2016 (Lindahl 2017b). The collection contained 101 spoken examples from everyday conversations, 60 from radio and television, and 109 written examples. In a representative subset of the spoken preposing examples (Sample B), slightly over half of the examples, 56%, involved preposed pronouns. Just as in the examples with local preposing, the pronouns extracted from relative clauses typically refer back to something which has just been uttered, either by the same speaker or by an interlocutor. The extracted pronouns are thus part of a focus or a topic chain which connects the utterance to the preceding context. A few examples are given in (11)–(13). For space reasons, the preceding context is sometimes given only in English.

(11) men ingen av dom är ju varmblodiga () det finns inga insekter [som är varmblodiga] (Sw.)

‘However, none of them are warm-blooded, there are no insects that are.’ (Lindahl 2017b, p. 1)

(12) ja men det var ingen [som klagade på dem] ser du

‘Yes, but no one complained about it, you know.’ (Lindahl 2017b, p. 77)

(13) den allra största delen av befolkningen, bönderna, det var det adeln [som hade domsrätt över dem]

‘It was the nobility that had the jurisdiction over the largest part of the population, the farmers.’ (Radio Sweden, 2015) (Lindahl 2017b, p. 91)

These examples resemble the ones shown in Section 2 from Engdahl (1997). Regarding (11) and (12), these are examples of presentational relatives, cf. (1abc), and (13) is a cleft construction, cf. (1d). This example also illustrates a common use of pronoun preposing in Swedish, namely in left dislocation where the initial topic is resumed by a preposed agreeing pronoun, see Holmberg (2020) and Lindahl and Engdahl (Forthcoming) for discussion.

The next two examples involve the lexical main verbs veta ‘know’ and störa sig på ‘be annoyed at’. In (14), the speaker is talking about driving across the US.

(14) ja det är hättigt det vet jag en [som har gjort det] (Sw.)

‘Yes, that is cool! I know someone who has done that!’ (Lindahl 2017b, p. 126)

(15) alkoholism is not a disease however

‘People who say that annoy me.’ (Lindahl 2017b, p. 89)

In Lindahl’s Sample B, 74% were presentational relatives, 8% were clefts such as (13) and 13% were constructions with lexical verbs. In many of the presentational relatives, the relative clauses are short and contain just a finite auxiliary verb, as in (11), or the support verb göra ‘do’, as in (14). These short relative clauses help to identify or restrict the head of the relative clause. Together with the matrix, the head DP and the relative clause must form a coherent comment on the extracted item which is relevant in the sense of Grice (1975).

We end this overview of extraction from relative clauses by noting that Swedish speakers have been using such constructions for a long time. In a guide to Proper Swedish, Wellander (1939) discusses the fact that Swedish speakers say things such as (16).
3.3. Summary

In this section, we have shown that the use conditions for extraction of pronouns from relative clauses are the same as for local and non-local pronoun preposing in Swedish. The pronouns are typically part of focus or topic chains; they refer to an event, a property or an entity that has either just been introduced or is already established as a topic. This resembles a well-known strategy for cohesion in discourse, namely to start with a subject pronoun which is part of an anaphoric chain with some element in the preceding utterance (Daneš 1974; Erteschik-Shir 2007). What seems to be special about speakers of Swedish and the other mainland Scandinavian languages is that they use this strategy also for non-subject pronouns, as becomes clear when we look at spontaneous conversations. By preposing a pronoun in a focus or topic chain, the speaker ensures that it will be understood as the aboutness topic for the upcoming utterance. This may in turn explain why preposing the VP or propositional anaphor det, as in (5) and (6), is especially common. By doing so the speaker signals that det’s antecedent is what the sentence is about.

We noted earlier that preposed pronouns in Swedish are often unstressed; when a pronoun is stressed, it often invokes contrast and the presence of alternatives, see (9). Lexical DPs can also be preposed or extracted (44% in Lindahl’s Sample B). These always carry stress and normally invoke alternatives.

4. Icelandic

When Icelandic is mentioned in connection with extractions from relative clauses, it is mainly to establish that such extractions are ungrammatical (Thráinsson 2007; Zaenen 1985). There are no reports of spontaneous extractions in Icelandic in the literature. According to our hypothesis, this might indicate that preposing is not used as a method for establishing cohesion in the discourse. We start by investigating local pronoun preposing in spoken Icelandic and then turn to extractions from relative clauses.

4.1. Pronoun Preposing

Observations of informal conversations between Icelanders give the impression that they use preposing of pronouns much more seldom than Danes, Norwegians or Swedes. In order to investigate whether this impression is correct, we looked at the Icelandic part of the NDC (94 338 token, 48 speakers). We found a few examples of object preposing, all of which involve the pronoun það ‘it, that’, as shown in (17) and (18).

\[(17) \text{a. } s1: \text{maður fæ-} \text{getur ekki notað kreditkort a}l{\text{s}} s\text{tădar} \\
\text{one ge- can not used credit card all places} \\
\text{‘One cannot use credit cards everywhere.’} \]

\[(17) \text{b. } s2: {\text{það}}_1 \text{gerum við} \_1 {\text{það}}_1 \text{gerum við strákkarnir} \_1 \text{s}k\text{o} \\
\text{that do we } \_1 \text{that do } \text{we boys.PL.DEF.NOM PRT} \\
\text{‘We do, me and the boys do, you know.’} \]

\[(18) \text{a. } s1: \text{reyndar býr \text{Clinton} í *Harlem hverfinu vissir þú } \text{það?} \\
\text{actually lives Clinton in Harlem block.DAT knew you that} \\
\text{‘Clinton actually lives in Harlem. Did you know that?’} \]

\[(18) \text{b. } s2: {\text{það}}_1 \text{vissi ét ekki} \_1 \\
\text{that knew } \_1 \text{not} \\
\text{‘I didn’t know that.’} \]
In (17) the preposed það is a VP anaphor and in (18) það refers back to a proposition, the new information just conveyed. They resemble the Swedish examples we saw in (5) and (6) and are both examples of focus chaining.

This type of preposing of það can also be found in written Icelandic and is judged to be natural by many speakers in the acceptability study reported in Lindahl (2022). However, there are no examples of preposed pronouns with entity antecedents in the NDC and when asked about such examples, Icelanders tend to supply a contrastive context, see (19) from Engdahl and Lindahl (2014), provided by Halldór Ármann Sigurðsson.

(19) A: Have you seen Olaf?
   B: nei hann1 hef ég ekki séði í allan dag
   en ég sá konuna hans núna rétt aðan
   svo að hann hlýtur að vera hérna einhvers staðar
   ‘No, I have not seen him all day, but I saw his wife just now so he must be somewhere around here.’

The referent Olaf is newly introduced in A’s question so this is also a case of focus chaining. Note that B contrasts hann ‘him’ with konuna hans ‘his wife’. This suggests that preposing of entity pronouns is mainly used when the referent is contrasted with some other element.

The impression that object preposing is not very common in Icelandic is confirmed in Lindahl (under review). Using the same method as described for Swedish in Section 3.1, Lindahl extracted 1000 utterances with a finite verb in the Icelandic part of the NDC and then investigated all declarative main clauses in the sample. Lindahl categorized the initial constituents in this sample and found substantial differences with respect to the comparable Swedish sample. These are summarized in Table 1.

Table 1. Grammatical function of phrases in the prefield in NDC, from Lindahl (under review).

|                | Swedish | Percent | Icelandic | Percent |
|----------------|---------|---------|------------|---------|
| Subject/expletive | 421     | 59.1%   | 400        | 75.2%   |
| Adverbial       | 177     | 24.0%   | 87         | 16.4%   |
| V1 declarative  | 32      | 4.5%    | 37         | 7.0%    |
| Object          | 46      | 6.5%    | 17         | 0.2%    |
| Other           | 36      | 5.8%    | 7          | 1.3%    |
| Total           | 712     | 100%    | 532        | 100%    |

First we note that the proportion of subject initial utterances was much larger in Icelandic, 75.2% compared to 59.1%, whereas preposing of adverbials was more common in Swedish. Verb initial (V1) declaratives, was somewhat more common in Icelandic than in Swedish. As for preposing of objects, there were actually no unambiguous examples with a preposed object in this random sample of Icelandic utterances, compared to 46 in the Swedish sample. We interpret this as an indication that object preposing is not a particularly common strategy in spoken Icelandic.
4.2. Extraction from Relative Clauses

As mentioned in the Introduction, extraction from relative clauses in Icelandic is considered to be ungrammatical (Thráinsson 2007; Zaenen 1985). However, since these articles only mention examples with extracted lexical DPs, an interesting question is whether Icelandic speakers accept the types of extractions that are found in Swedish (see Section 3.2) even if they do not produce them spontaneously. The second author reports on such a study in Lindahl (2022). In two parallel acceptability studies, speakers of Swedish and Icelandic were asked to rate sentences with extractions, using examples similar to the spontaneously produced ones from mainland Scandinavian we have seen in this article. The extracted phrases used in all of the test sentences were *það* in Icelandic and *det* in Swedish, and they all had sentential or VP antecedents. The participants were asked to rate examples such as (20), using the scale natural, somewhat strange and unnatural.

(20) Þú getur notað gjafakortið til að kaupa bíómiða og *það* eru margir [ sem gera _1].

‘You can use the voucher to buy a movie ticket, and there are many people who do.’

In brief, Lindahl found that extraction from relative clauses was rated very poorly in the Icelandic part of the study, and clearly much worse than the parallel Swedish examples. Local preposing was judged to be natural by two thirds of the participants. Long preposing from an *að*-clause received mixed ratings, but was rated worse than good fillers. In Swedish, preposing from *att*-clauses was rated on a par with good fillers. As for extraction from relative clauses, the Icelandic speakers found extraction unnatural in all of the test sentences, regardless of the type of relative clause. The Swedish speakers rated test items with a presentational relative clause as natural sounding more often than those items which contained a lexical verb. The study thus confirms previous reports that extraction from relative clauses is not acceptable in Icelandic.

4.3. Summary

Preposing of the VP or propositional anaphor *það* is sometimes used in Icelandic, primarily in local clauses, in a way that resembles the Swedish pattern, albeit much less frequently. However, when pronouns with entity antecedents are preposed, they reportedly receive a contrastive interpretation, just as in English (see Section 6.1). The low occurrence of preposed pronominal objects in Icelandic suggests that preposing is not used as a natural way of connecting utterances in the same way as in mainland Scandinavian. Since extraction from relative clauses seems to be unacceptable, we think this supports our hypothesis that there is a connection between using pronoun preposing as a way of connecting utterances and a willingness to accept long prepositions. The study reported in Lindahl (2022) confirms that there is indeed a difference between Icelandic and Swedish in how long preposing of anaphoric pronouns is judged—such prepositions are rated as natural more often in Swedish than in Icelandic, in line with our hypothesis. In addition, Lindahl found a clear effect of clause type in Icelandic. While a substantial part of the Icelandic participants rated long preposing from *að*-clauses as natural, even though such preposings are not common in spoken Icelandic, they did not do so for extraction from relative clauses. This suggests that there is also a structural constraint which blocks extraction from relative clauses in Icelandic.

5. Faroese

Turning to Faroese, very little is known about possible uses of pronoun preposing in Faroese and extraction from relative clauses has not been discussed much in previous research, apart from when it is pointed out that it is unacceptable (Platzack 2014; Thráinsson et al. 2004).
5.1. Pronoun Preposing

In the Faroese part of the NDC (64803 tokens from 20 speakers) there are more than a hundred examples with a preposed tað ‘it, that’, pronounced /tæ/, including the two following examples.

(21) a. s1: hevur tú spælt gekk nú tá ið gekkaupphaæddin var so stór
    have you played Gekk now then REL payout sum was so large

b. s2: ja (_) veitst tú hvat tað1 havi eg faktiskt _1
    yes know you what that have I actually

    ‘Do you know what, I actually have.’

(22) altso PRT tað er ordiliga hugnaligt tað1 haldi eg _1
    it is really nice it think I

    ‘It’s really nice, I think so.’

In (21) tað is a VP anaphor and in (22) tað refers back to a proposition. There are notably fewer preposed pronouns with entity antecedents, as in (23).

(23) Carl Johan Jensen hann eg veit at hann er rithøvundur (.)
    Carl Johan Jensen he I know that he is author

    hann1 veit eg einki um _1
    him know I nothing about

    ‘Carl Johan Jensen, I know he is an author. I do not know anything about him.’

Since there is so little data available on spoken Faroese, the second author carried out a controlled production study in Faroese. The study took advantage of the fact that practically all Faroese speakers know Danish well; Danish is taught in schools from third grade and used frequently in the society. In the study, 91 native Faroese speakers were asked to translate Danish sentences into Faroese. 82 of the participants were high school students, and 9 were between 38 and 69 years at the time. The Danish sentences were chosen so as to resemble spoken dialogue.

An example of the task is shown in (24). Each sentence was given in a context, also in Danish, which the participants could choose to translate if they wanted to. In the examples below, we have underlined the antecedent and italicized the pronoun, but this was not performed in the experiment.

(24) Anna spurgt os, hvad klokken var, men det1 vidste vi ikke _1. (Da.)
    Anna asked us what clock.DEF was but it knew we not

    ‘Anna asked us what time it was but we did not know.’

The Faroese speakers had no difficulties in translating this example. Their answers varied, as shown in (25). The percentages to the right indicate how often the participants used this word order in their translations.

(25) Anna spurdi okkum hvat klokkan var ... (Fa.)
    Anna asked us what clock.DEF was

    a. men tað1 vistu vit ikki _1 82%
    but that knew we not

    b. men vit vistu tað ikki 2%
    but we knew that not

    c. men vit vistu ikki Ø 14%
    but we knew not

In this example, the preposed det in the Danish version is a propositional anaphor. The Faroese participants all translated it using tað and in 82% of the answers they preposed it. The same tendency showed up in examples where a VP anaphor was preposed from an embedded at-clause. The Danish original is shown in (26) and in (27) we give the relevant parts of the Faroese translations of the example.
(26) Hvis jeg ikke tager fejl, og det tror jeg ikke, at jeg gør _1, har Ole satt kagen i køleskabet. (Da.)

‘If I’m not mistaken, and I do not think I am, Ole has put the cake in the fridge.’

(27) a. vissi eg ikki takið fel, og tað haldi eg ikki at eg geri _1 67% (Fa.)
    if I not take error and that I not that I do

b. um eg ikki taki feil, sum eg ikki haldi eg geri _1 13%
    if I not take error which I not think I do

c. vissi eg ekki taki feil, och tað haldi eg ikki _1 8%
    if I not take error and that I not

d. um eg ikki taki feil, sum eg ikki geri _1 4%
    if I not take error which I not do

e. um eg ikki taki feil, og eg haldi ikki at eg geri Ø 8%
    if I not take error and I think not that I do

In total, 92% of the Faroese participants translated the Danish sentence using a preposing strategy, either a local or long preposing of tað or relativization with sum. However, when the Danish sentence contained a preposed entity pronoun as in (28), the Faroese participants preferred a different strategy.

(28) Jeg var hjemme hos min bror i går. Han bor tæt på mig, så ham ser jeg tit _1. (Da.)

‘I was at my brother’s yesterday. He lives close to me, so I see them often.’

(29) Eg var heima hjá beiggja minum i gjar. Hann býr tætt við hjá mær ... (Fa.)

a. so hann siggi eg ofta _1 20%
    so him see I often

b. so eg siggi hann ofta
    so I see him often

c. so vit siggjast ofta
    so we see.RECIP often

For this example, only 20% of the Faroese translations replicate the preposing from the Danish original. Instead, most participants chose to start with the subject as in (29b,c). In total, 19% of the translations used an alternative formulation with a reciprocal form of the verb, as in (29c). The Faroese speakers’ preference for avoiding preposing was even more pronounced when the Danish original involved long preposing of an entity pronoun from an at-clause.

(30) Tove: Hvor er kagen? (Da.)

Mette: Den tror jeg, at Ole satte _1 i køleskabet.
    it think I that Ole put in fridge.DEF

‘I think Ole put it in the fridge.’

(31) Hvar er kakan? (Fa.)

‘Where is the cake?’
a. Hana₁ haldi eg, at Óli setti _₁₁ í køliskápi
   her think I that Óli put in fridge.DEF
   4%
b. Eg haldi at Óli setti hana₁ í køliskápið
   I think that Óli put her in fridge.DEF
   79%
c. Eg haldi at Óli setti Ø í køliskápi
   I think that Óli put in fridge.DEF
   17%

Only 4% of the translations replicated the preposing of an entity pronoun. The Faroese participants clearly preferred to leave the pronoun in the verb phrase or leave it out. Given the different strategies used by the Faroese participants, it seems that preposing of the VP or propositional anaphor tað is a common strategy in Faroese, but it does not extend to entity pronouns.

There are both advantages and disadvantages with using a translation method. It can be a useful method for finding examples that are hard to locate in corpora and it can elicit natural sounding Faroese exchanges, especially if the materials to be translated come from natural dialogue. One disadvantage is that participants may transfer some aspects of the Danish originals into Faroese. However, when the Faroese translations deviate from the Danish originals, this is most likely an indication that the participants do not recognize this as an acceptable word order in Faroese. For example, only a few of the translations retained the preposed word order when the pronoun referred to an entity, see (29a) and (31a).

As for the frequency of object preposing, we can get at least an idea about it from a smaller sample of 300 sentences from the Faroese part of the NDC. The sample contained 192 declarative main clauses, and out of these, 15 started with a preposed object, in 12 cases tað, which was the only pronoun found. Although the material is smaller, the rate of object preposing, 7.8%, is comparable to what Lindahl found using the same type of search in Swedish, 6.5%, see Table 1. Object preposing thus seems to be fairly common in spoken Faroese and definitely more common than in Icelandic.

5.2. Extraction from Relative Clauses

According to the literature, extraction from relative clauses is not acceptable in Faroese. Platzaek (2014) gives the following example, which is a translation of a Swedish example in Allwood (1976).

(32) * Sliðar blómur₁ kenni eg ein mann [sum selur₁ ]
   such flowers know I a man who sells
   (Fa)
   (Platzaek 2014, p. 10)

However, we have noticed a handful examples in Faroese newspapers, two of which are shown here.

(33) Tað ber eisini til at koma til viðgerð saman við fyrirverandi makenum, og
   it bears also to come to treatment together with former spouse.SG.DEF and
   tað₁ eru tað nógv [sum gera₁]
   that are there many who do
   (Fa.)
   ‘It is also possible to undergo treatment together with your former spouse, and there are many people who do.’ (Dimmalætting, 24 April 2015)

(34) “Kjakokrati”₁ var tað onkur, [sum rópti tað₁ ], tá fólk á ymsum
   debate-ocracy was there no one who called it when people on various
   internetsíðum viðmerkja evnír, ella geva sína meining til kennar.
   web pages comment topics, or give REFL meaning to known
   (Fa.)
   ‘No one called it “debateocracy” when people on various web pages commented on topics or let their options be known.’ (Dimmalætting, 1 March 2019)

In order to investigate whether such constructions are used productively, Lindahl included extractions from relative clauses in her translation study. The 91 Faroese speakers were
asked to translate five examples with relative clauses. Three of them were similar to (35) where the relative clause modifies the object of a lexical verb, here ved ‘know’.

(35) Jeg synes sommeren er den bedste tid på året, og det ved jeg mange, I think summer.DEF is the best time on year.DEF and it know I many [som er enige med mig om _1]. (Da.) that are in agreement with me about

‘I think that summer is the best time of year, and I know many people who agree with me about that.’

(36) Eg haldi at summarið er besta tíðin á árinum, (Fa.) I think that summer.DEF is best time.DEF on year

a. og tað1 veit eg nógv [ sum er einigur við meg um _1 ] [1 inf.] and that know I many that are in agreement with me about

b. og tað1 veit eg, at tað eru nógv, [ sum eru samd við mær í _1 ][1 inf.] and that know I that there are many that are same with me in

c. og tað1 veit eg at nógv eru samd um _1

and that know I that many are in agreement about

d. og eg veit at fleiri eru samd við mær

and I know that many are in agreement with me

e. og eg veit fleiri, sum er samd við mær

and I know many who are in agreement with me

f. og eg veit nógv ið er einig við mær

and I know many who are in agreement with me

g. og eg veit nógv ið er einig við mær í tí

and I know many who are in agreement with me in that.DAT

Only two of the Faroese participants (2%) replicated the preposing in the Danish original and produced the versions in (36a,b). The rest of them used a variety of strategies to convey the content of the Danish sentence.25 In (36c), tað is fronted, but the sentence has been reformulated using an at-clause instead of a relative clause. (36d,e,f) use slight reformulations leaving out the pronoun, and (36g) keeps the relative clause, but leaves the pronoun tí (the dative form of tað) in situ.

For the two sentences where the Danish original was a presentational relative, as in (37), the Faroese speakers did produce some extractions from relative clauses in their translations, see (38).

(37) Ole undrede sig over, om det ville regne i morgen, men det1 var der ingen, Ole wondered REFLEX over if it would rain tomorrow but that was there no one [der troede _1]. (Da.) who believed

‘Ole wondered whether it would rain tomorrow, but no one thought so.’

(38) Óli ivåést f, um tað fer at regna í morgin ... (Fa.) Óli doubted in whether EXPLAIN go to rain tomorrow

a. men tí1 var tað eining [ íð trúði _1 ] [1 inf.] but that.DAT was there no one who believed

b. men tað1 var tað ongin [ sum helt _1 ] [16 inf.] but that was there no one who thought

c. men tað var ongin [ sum helt tað ]

but there was no one who though that

d. men tað1 helt ongin _1

but that thought no one
A total of 17 participants (19% of the replies) retained the Danish structure, as shown in (38a,b). One participant used the relativizer *ið* as in (38a) and 16 participants used the relativizer *sum*. The other 74 informants reformulated the sentence either by not using a relative clause in the translation, by leaving the object in situ, as shown in (38c–g), or both, or using other reformulations, as in (38h). In the two examples where the matrix verb was *vera* ‘be’, 10–20% of the translations retained the Danish extraction structure, compared to 2–4% when the verb was *vita* ‘know’ or *kende* ‘be acquainted with’. This suggests that this particular type of extraction sentence is less degraded in Faroese, and may even be acceptable for some speakers, while extraction from other relative clauses is unacceptable. Faroese could thus potentially be a language that shows evidence for the “pseudo-relative” hypothesis put forth by McCawley (1981). He proposed that the relative clause in existential sentences in English is not a true relative clause; instead it has a somewhat reduced structure. This could explain why extraction in such an environment is sometimes marginally acceptable in English (see also Chung and McCloskey 1983; Kush et al. 2013).

5.3. Summary

There are clearly similarities between Faroese and Icelandic. In both languages the VP or propositional anaphor *tað,tuð* can be preposed locally and from *að/at*-clauses, as shown in the acceptability and translations studies summarized above. However, the impression is that this is more common in Faroese. Preposing of entity pronouns seems not to be a natural strategy given the alternative translations provided in (29) and (31). Given that preposing of *tað* is fairly common in Faroese, one might expect speakers to use preposing of *tað* from relative clauses as well. The translation study showed that this is not the case, except for preposing in presentational relative clauses, as in (38). This is also the type of preposing that Faroese speakers are most likely to hear and read since this is the most common type in Danish (Müller and Eggers 2022). In this respect, the Faroese speakers differ from Icelandic speakers who judge extraction from presentational clauses to be unnatural to the same extent as extractions from other types of relative clauses, as discussed in Section 4.2.

6. Comparisons with Other Languages

We have suggested that the frequent use of preposing of unstressed anaphoric pronouns in Swedish is a crucial factor for explaining the presence of spontaneous extractions from relative clauses in this language. Since extraction from relative clauses is not found in English and the continental Germanic languages German and Dutch, it becomes relevant to investigate whether local and long preposing is used in these languages.

6.1. English

In English, preposed (topicalized) phrases are normally stressed and are understood to imply contrast. Exceptions to this are the light adverbs *then* and *there* which do not require stress in initial position when they serve to connect the utterance to the preceding context. Attempts to prepose an unstressed personal pronoun sound very strange. Compare the options for answering the question in (7) in English.

(39) When did you buy your first car?

a. *It₁ I bought _₁ in 1980.
b. I bought it in 1980.
c. #THAT₁ I bought _₁ in 1980.

The preposed version in (39a) is clearly unacceptable whereas the version with the pronoun in situ in (39b) is fine. Preposing a stressed demonstrative is grammatical, (39c), but hardly appropriate in this context.27

When we look at the Swedish examples involving the VP anaphor det, we find that they are best rendered in English using VP ellipsis (Ø) (Bentzen et al. 2013; Hankamer and Sag 1976; Hardt 1999). Consider the question in (40).

(40) Have you been to Oslo?
   a. Ja, det har jag _₁.
      yes it have I
   b. Ja, jag har det.
      yes I have it
   c. Yes, I have Ø.

In Swedish, this is a context where the VP anaphor det typically would be preposed. Leaving it in situ, (40b), is not ungrammatical but the preposed version is preferred (Lindahl and Engdahl Forthcoming).28 In English, repeating the finite verb and leaving out the rest of the VP is common and this is also what we have done in the English translations of the examples with VP anaphors, Swedish (5), Icelandic (17) and Faroese (21). Note that the VP ellipsis strategy can also be used in embedded clauses in English where Swedish uses preposing.

(41) Do you think I should go to Oslo?
   a. Ja, det tycker jag _₁.
      yes it think I _₁.
   b. Yes, I think you should Ø.

There are clearly similarities between VP ellipsis in English and VP anaphor preposing in Swedish; in particular both require that the antecedent VP is available in the immediate context.

In contexts where a preposed det in Swedish refers back to a recently expressed proposition, the English version can sometimes have the anaphor so, as in (42) from the NDC.

(42) a. int: tycker du det är roligt med små barn?
       think you it is fun with small children
       ‘Do you think small children are fun?’
   b. sp1: ja det₁ tycker jag faktiskt _₁
       yes it think I actually
       ‘Yes, I actually think so.’

Given that VP ellipsis can also be used inside embedded clauses in English, as in (41b), one might ask whether this strategy can be used in English translations of extraction clauses. We have actually done so in (11) and repeat the translation here in (43).

(43) However, none of them are warm-blooded, there are no insects that are Ø.

This long distance VP ellipsis appears to be acceptable at least to some English speakers. However, the rest of the Swedish extraction examples, which involve preposed propositional or entity pronouns, are best translated into English with the pronoun in situ.29
6.2. German and Dutch

German and Dutch are both verb second languages and local preposing is quite common, but there are differences compared with Swedish. Consider the German version of the car dialogue.

(44) Wann hast du dein erstes Auto gekauft?
   a. Das habe ich in 1980 gekauft.
   (that have I in 1980 bought)
   b. Ich habe es in 1980 gekauft.
       (I have it in 1980 bought)
   c. *Es habe ich in 1980 gekauft.
       (that have I in 1980 bought)

The unmarked answer would be with a preposed unstressed demonstrative pronoun, a so-called d-pronoun, as in (44a), or an unstressed personal pronoun in situ, (44b). Preposing of the neuter pronoun es is very restricted, see the discussions in Frey (2006); van Craenenbroeck and Haegeman (2007) and Theiler and Bouma (2012). Corpus studies reveal that preposing of personal pronouns is uncommon compared to preposing of d-pronouns. In a newspaper corpus, it was much more common for object d-pronouns to be preposed than to appear in situ whereas no personal object pronouns were preposed (Bosch et al. 2007).

In Dutch, preposing of personal pronouns is also uncommon compared to preposing of demonstratives (d-pronouns) (van Kampen 2007).

(45) Wanneer heb jij je eerste auto gekocht?
   a. Die heb ik in 1980 gekocht.
       (that have I in 1980 bought)
   b. Ik heb hem in 1980 gekocht.
       (I have it in 1980 bought)
   c. #Hem heb ik in 1980 gekocht.
       (it have I in 1980 bought)

As in German, the unmarked answers have a preposed d-pronoun or a personal pronoun in situ. Preposing an unstressed personal pronoun is not ungrammatical but pragmatically odd. Bouma (2008, p. 112) found 4 preposed object personal pronouns in a 9 million words corpus of spoken Dutch (Corpus Gesproken Nederlands), compared to 2723 preposed d-pronouns.

When it comes to non-local preposing, both German and Dutch restrict this to contrastively stressed phrases. This has led some German linguists to propose that there are two types of preposing. Werner Frey distinguishes what he calls formal movement from true A-bar movement (Frey 2004, 2006, 2007). Formal movement involves local preposing to Spec,CP of the highest constituent in the middle field (IP). This accounts for preposing of subjects, both referential and expletive. Objects can also be preposed by formal movement when they are topics. On Frey’s analysis this means that they have first moved to a topic position above the subject position in the middle field; this way they become the highest constituent in IP. True A-bar movement accounts for all other movement into Spec,CP and is linked to a marked interpretation, typically involving contrast. A similar proposal is made by Fanselow (2016) who distinguishes unrestricted V2 where there are no pragmatic effects of the preposing and restricted V2 which comes with pragmatic effects and induces contrast.

This distinction seems to capture the situation in German. All examples of spontaneously produced long preposings in German that we are aware of are contrastive, as predicted by Frey and Fanselow, see Andersson and Kvam (1984) and Lühr (1988). Two examples from Andersson and Kvam (1984) are shown here, involving preposing of contrastive adverbial phrases from a daß ‘that’ clause and an ob ‘if’ clause.
That could perhaps differ regionally. For instance, in Hannover, I would doubt that anyone says that.

Tuesday, I’m not sure if he will come, but Wednesday, he will certainly be there.'

To some extent, the distinction between formal movement and true A-bar movement is relevant for Swedish as well since expletive subjects can only be preposed in the local clause. However, we find no evidence for a topic position above the subject in the middle field in Swedish which means that preposing of unstressed object pronouns cannot be handled by formal movement (Lindahl and Engdahl Forthcoming). Instead, it seems that A-bar movement in Swedish is not correlated with contrast or a marked interpretation but can be used in order to connect utterances involving focus and topic chains.  

Highly proficient L2 speakers of German with Swedish L1 sometimes extend this type of preposing into spoken and written German with the result that “the preposed item is understood to be highlighted as very important even when it is not”, as a native speaker of German recently commented.

7. Concluding Remarks

In this article, we have explored the hypothesis that frequent use of local preposing of object pronouns increases the likelihood that speakers will also use long preposing from embedded clauses as well as from relative clauses. For Swedish, it seems clear that preposing of both subject and non-subjects pronouns is a common strategy when the pronoun refers to a recently mentioned event, property or entity, that is when they are part of a focus or topic chain. This preposing strategy can be extended to subordinate clauses and to relative clauses, provided that the rest of the utterance is a relevant comment on the preposed item in the context. Acceptable long prepossings and extractions are thus doubly context dependent; first, the preposed item must be part of a focus or topic chain with respect to the previous utterance and second, it must be possible to interpret the preposed item as the aboutness topic of the utterance with the rest providing a comment that meets Grice’s Maxim of Relevance. The second property presumably lies behind the fact that short presentational relative clauses are most common. When both these pragmatic conditions are met, long extractions can be used and do not cause problems for the listeners/readers.

This does not mean that there are no syntactic constraints on extractions in Swedish. For instance, the Coordinate Structure Constraint seems to apply, see, e.g., the discussion in Lindahl (2017b).

The situation in Icelandic is clearly different. Local preposing of the VP or propositional anaphor það occurs but is used much more seldom than in Swedish. Preposing of entity pronouns is unusual and seems to have a contrastive effect instead of being used for cohesion. Furthermore the pragmatic conditions that are important in Swedish do not seem to play any role when it comes to extraction from relative clauses. Not even when the sentences were pragmatically very plausible did the Icelandic participants in Lindahl’s (2022) study judge them to be natural. We conclude that there must be a structural constraint operating in Icelandic which prevents such extractions and which is not affected by pragmatic conditions.

Faroese presents a more mixed picture. On the one hand, Faroese speakers often prepose the VP or sentential anaphor tað, similar to Danish. This suggests that preposing is employed as a cohesive device, but note that it does not seem to extend to entity pronouns.
On the other hand, Faroese speakers react more similar to Icelanders when it comes to extractions from relative clauses; they try to avoid using them when they translate such sentences from Danish. This suggests that extractions from relative clauses are unacceptable in Faroese, similar to Icelandic, presumably due to a structural constraint. However, there is one interesting difference between the Faroese and Icelandic speakers present in our studies. The Faroese speakers seem to accept, and occasionally produce, extractions from presentational relatives, whereas the Icelanders find these unnatural. This may be due to influence from Danish where this kind is the most common type of relative clause extraction (Müller and Eggers 2022). It may also be that these relatives have a reduced structure in Faroese which makes extraction more acceptable. More research on Faroese is clearly called for.

One important aspect of pronoun preposing in the mainland Scandinavian languages is that it often involves unstressed pronouns. If the antecedent of the pronoun has just been introduced or already is a topic in the previous utterance, it is often natural to continue with an unstressed pronoun. (This holds for both subject and non-subject pronouns). In English, object preposing tends to induce a contrastive interpretation and consequently the preposed item must be stressed; this holds both for local and non-local preposing. In Dutch and German, unstressed demonstrative pronouns can be preposed from the same clause but only contrastively stressed items can be preposed from embedded clauses. In English, one way of connecting utterances without invoking contrast is to use VP ellipsis. This often turns out to be the best translation of Swedish utterances with preposed pronouns, but there are limitations, especially with regard to relative clauses.

Our comparative investigation has shown that preposing in some languages can be used to connect the utterance to the preceding context through anaphoric chains. This means that (long) preposing by itself is not necessarily associated with a contrastive interpretation. Whether a contrastive interpretation is plausible depends rather on the context of use and the type of preposed item. This should have consequences for psycholinguistic investigations of extractions in languages such as Swedish. If, as has been common so far, the experimenter uses materials with preposed lexical DPs, then very often a contrastive reading emerges. However, if the context calls for a cohesive continuation, a focus or topic chain, then a sentence with a preposed pronoun might be more natural.

Author Contributions: E.E. is mainly responsible for writing Sections 1, 2, 3.1, 6 and 7. The remaining Sections were written together. F.L. carried out the corpus studies reported in Sections 3–5 as well as the experimental studies in Sections 4–6. All authors have read and agreed to the published version of the manuscript.

Funding: Part of this research was supported by an International Postdoc grant (Dnr. 2017-06139) from the Swedish Research Council, awarded to F.L.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Informed consent was obtained from all subjects involved in the study.

Data Availability Statement: Nordic Dialect Corpus http://www.tekstlab.uio.no/nota/scandiasyn/. The Swedish data concerning extraction from relative clauses are available at https://svn.spraakbanken.gu.se/sb-arkiv/pub/lindahl/2017/. The Icelandic and Faroese data presented in this study are available on request.

Acknowledgments: Earlier versions of parts of this paper have been presented at the workshop on Word Order in the Scandinavian languages, University of Konstanz, December 2018, the Linguistics Research Seminar at the University of Gothenburg, May 2021, and at the 12th International Conference of Nordic and General Linguistics, University of Oslo, June 2021. We would like to thank Maia Andréasson, Miriam Butt, Nomi Erteschik-Shir, Benjamin Lyngfelt, Peter Sells and Annie Zaenen for discussion and comments. We are grateful to Kristin Hagen and Anders Nøklestad from the Text laboratory at the University of Oslo for help with using the Nordic Dialect Corpus. We also thank Victoria Absalon, Gunnvor Hoydal Brimnes, Zakaris Svabo Hansen, Jógvín Í Lón, Johannes Gislí Jónsson, Ida Larsson, Solveig Malmsten, Hjalmar Petersen, Sigriður Sigurjónsdóttir, and Annika
In Object shift contexts, such as (18b), the gap could also be located before the negative adverbial, but nothing hinges on this for Lindahl (2017b, p. 146f.) also discusses examples where an extracted stressed item is understood as the focus of the utterance, A reviewer pointed out that examples with preposed object pronouns can be found in the large corpus of written Icelandic, the Swedish data are discussed in Sundman (1980, p. 59) and Teleman et al. (1999, volume 3, p. 423f.).

Establishing what is a relevant comment in a particular context is difficult, see Allwood (1976); Andersson (1982) and Engdahl (1982) for some attempts. Recently Chaves and Putnam (2020, p. 120) have introduced the term Relevance Islands for contexts which are sensitive to Gricean maxims. The earliest mention of this type of extraction that we are aware of is in Mikkelsen (1894), an early grammatical description of Danish. See Lindahl (2017a, p. 27) for some of his examples.

Swedish is a verb second language and we assume that both subjects and non-subjects are preposed to a Spec position in the C domain in declarative clauses, which we refer to as the prefield. In (Lindahl and Engdahl Forthcoming) we develop the syntactic analysis further, adopting the bottleneck hypothesis in Holmberg (2020).

The sound files along with transcripts with word by word translations into English can be accessed on the web page of the Text Laboratory at the University of Oslo (http://www.tekstlab.uio.no/nota/scandiasyn/ accessed on 1 March 2022). You can search for the utterances, listen to the sound files and see the examples in context. To facilitate the search, the transcripts use standard orthography which we have retained in the examples cited, but added underlining, italics and gap locations. The Norwegian and Eldfdalian transcripts in addition contain a simplified phonetic transcription.

On VP anaphora in Scandinavian, see, e.g., Ørsnes (2011); Ledrup (2012) and Mikkelsen (2015).

Personal pronouns include first, second and third person referential pronouns. The pronoun det also functions as a non-referential expletive, in which case it cannot be stressed.

The relative frequency for lexical verbs, 13%, is much higher in Lindahl’s spoken collection than what Müller and Eggers (2022) find in their corpus study of written Danish (7/940). We suspect that a similar study of written Swedish would also find a higher proportion of presentational relatives.

Establishing what is a relevant comment in a particular context is difficult, see Allwood (1976); Andersson (1982) and Engdahl (1982) for some attempts. Recently Chaves and Putnam (2020, p. 120) have introduced the term Relevance Islands for contexts which are sensitive to Gricean maxims.

The earliest mention of this type of extraction that we are aware of is in Mikkelsen (1894), an early grammatical description of Danish. See Lindahl (2017a, p. 27) for some of his examples.

The same probably holds for extraction from embedded interrogatives. In Lindahl’s (2022) acceptability study, the participants rated sentences where the VP or propositional anaphor det had been extracted from wh-clauses as natural 65.7% of the time. Lindahl (2017b, p. 146ff.) also discusses examples where an extracted stressed item is understood as the focus of the utterance, typically in answers to questions.

The propositional or VP anaphor það is usually glossed as ‘that’ in the literature on Icelandic and we follow this in the examples below.

See Sigurðsson and Wood (2020) for an analysis of the use of conjoined subjects as in við strákarnir.

In Object shift contexts, such as (18b), the gap could also be located before the negative adverbial, but nothing hinges on this for our analysis.

A reviewer pointed out that examples with preposed object pronouns can be found in the large corpus of written Icelandic, the Gigaword Corpus (https://malheildir.arnastofnun.is/, accessed on 28 April 2022). However, if these sentences had been spoken, the pronoun would have “a regular main-clause initial stress”, according to the reviewer. This may be a relevant difference with Swedish. In Lindahl and Engdahl (Forthcoming) we include Praat analyses (Boersma and Weenink 2020) of Swedish examples with preposed unstressed pronouns. These analyses show that the initial pronoun often does not form a separate prosodic phrase but is incorporated into the verb, see Myrberg and Riad (2015). Similar investigations of the prefield in Icelandic are needed as well as more informant studies.

Notes
1. See Chaves and Putnam (2020, p. 67) for additional references.
2. We use the term preposing rather than the commonly used term topicalization in order to distinguish the syntactic positioning from any information structural effects this may have, see Ward (1985) and the discussion in (Lindahl and Engdahl Forthcoming).
3. Swedish data are discussed in Sundman (1980, p. 59) and Teleman et al. (1999, volume 3, p. 423f.).
4. Löwenadler (2015) suggests that these common types should be seen as conventionalized constructions.
5. Softeeland (2013) calls this type a presentational cleft since the clefted constituent ingen ‘nobody’ is indefinite, as in presentational sentences, see (1b). She distinguishes them from it-clefts where the clefted constituent is normally definite, as in (1d).
6. Swedish is a verb second language and we assume that both subjects and non-subjects are preposed to a Spec position in the C domain in declarative clauses, which we refer to as the prefield. In (Lindahl and Engdahl Forthcoming) we develop the syntactic analysis further, adopting the bottleneck hypothesis in Holmberg (2020).
7. The sound files along with transcripts with word by word translations into English can be accessed on the web page of the Text Laboratory at the University of Oslo (http://www.tekstlab.uio.no/nota/scandiasyn/ accessed on 1 March 2022). You can search for the utterances, listen to the sound files and see the examples in context. To facilitate the search, the transcripts use standard orthography which we have retained in the examples cited, but added underlining, italics and gap locations. The Norwegian and Eldfdalian transcripts in addition contain a simplified phonetic transcription.
8. On VP anaphora in Scandinavian, see, e.g., Ørsnes (2011); Ledrup (2012) and Mikkelsen (2015).
9. Personal pronouns include first, second and third person referential pronouns. The pronoun det also functions as a non-referential expletive, in which case it cannot be stressed.
10. Talkbanken (96,346 words) was collected in Lund in the 1970s. The materials are available in Språkbanken and can be searched using the search engine Korp (https://sprakbanken.gu.se/korp, accessed on 1 March 2022.)
11. Mikkelsen (2015); Bentzen and Anderssen (2019) and Lindahl and Engdahl (Forthcoming) also discuss interactions with Object shift.
12. See Lindahl (2017b, p. 45ff.) for details about the data collection.
13. The relative frequency for lexical verbs, 13%, is much higher in Lindahl’s spoken collection than what Müller and Eggers (2022) find in their corpus study of written Danish (7/940). We suspect that a similar study of written Swedish would also find a higher proportion of presentational relatives.
14. Establishing what is a relevant comment in a particular context is difficult, see Allwood (1976); Andersson (1982) and Engdahl (1982) for some attempts. Recently Chaves and Putnam (2020, p. 120) have introduced the term Relevance Islands for contexts which are sensitive to Gricean maxims.
15. The earliest mention of this type of extraction that we are aware of is in Mikkelsen (1894), an early grammatical descriptions of Danish. See Lindahl (2017a, p. 27) for some of his examples.
16. The same probably holds for extraction from embedded interrogatives. In Lindahl’s (2022) acceptability study, the participants rated sentences where the VP or propositional anaphor det had been extracted from wh-clauses as natural 65.7% of the time.
17. Lindahl (2017b, p. 146ff.) also discusses examples where an extracted stressed item is understood as the focus of the utterance, typically in answers to questions.
18. The propositional or VP anaphor það is usually glossed as ‘that’ in the literature on Icelandic and we follow this in the examples below.
19. See Sigurðsson and Wood (2020) for an analysis of the use of conjoined subjects as in við strákarnir.
20. In Object shift contexts, such as (18b), the gap could also be located before the negative adverbial, but nothing hinges on this for our analysis.
21. A reviewer pointed out that examples with preposed object pronouns can be found in the large corpus of written Icelandic, the Gigaword Corpus (https://malheildir.arnastofnun.is/, accessed on 28 April 2022). However, if these sentences had been spoken, the pronoun would have “a regular main-clause initial stress”, according to the reviewer. This may be a relevant difference with Swedish. In Lindahl and Engdahl (Forthcoming) we include Praat analyses (Boersma and Weenink 2020) of Swedish examples with preposed unstressed pronouns. These analyses show that the initial pronoun often does not form a separate prosodic phrase but is incorporated into the verb, see Myrberg and Riad (2015). Similar investigations of the prefield in Icelandic are needed as well as more informant studies.

Conflicts of Interest: The authors declare no conflict of interest.
This involved excluding questions, imperatives, embedded clauses and tags. The Icelandic sample contained more embedded clauses and more tags such as þú veist ‘you know’ than the Swedish sample with the result that there were fewer relevant clauses in Icelandic.

A more detailed presentation of this study in Swedish is forthcoming, Lindahl (to appear).

The percentages do not always add up to 100, because there are a few cases where a participant did not contribute a translation. The participants were asked to use the word order they found natural for spoken or informal written Faroese. We are showing the answers exactly how they were written by the participants, including spelling errors and any informal/non-standard spelling. For example, many speakers have chosen to leave out -ð in the definite suffix, which is silent in spoken language.

Since we are mainly interested in the extraction cases here, we only give the number of informants who produced such translations.

Note the interesting translation in (36b) where the informant inserts an additional at-clause which permits him/her to reformulate it as a presentational relative, thereby avoiding having a relative clause embedded under vide as in the Danish original.

(39c) would have been appropriate if the question had involved a narrow focus, When did you buy your FIRST car?.

It is possible that the in situ order is used more frequently in Norwegian than in Danish and Swedish, cf. Bentzen and Anderssen (2019). More comparative research is required in order to establish if this is the case and why. There is also an issue whether the VP anaphor can precede negation in Object shift, see Mikkelsen (2015); Ørsnes (2013) and Engdahl and Zaanen (2020).

Similarities and differences between movement and ellipsis have been much discussed, see, e.g., Johnson (2001) and Aelbrecht and Haegeman (2012).

The examples and judgments in (45) were supplied by Gerlof Bouma.

The examples in Zifonun et al. (1997) are all taken from Andersson and Kvam (1984). There is considerable discussion concerning long wh-movement, especially the so called was was construction, see the articles in Lutz et al. (2000).

See Lindahl and Engdahl (Forthcoming) for a detailed discussion of the structure of the C-domain in German and Swedish.

Professor Christiane Andersen, personal communication.

We have looked at extended contexts for the attested examples we have investigated and have not seen any evidence that they are difficult to produce, for the speaker, or to understand, for the addressee. There are no clarification requests or other signs of comprehension problems from the interlocutors.

References

Aelbrecht, Lobke, and Liliane Haegeman. 2012. VP-Ellipsis Is Not Licensed by VP-Topicalization. Linguistic Inquiry 43: 591–614. [CrossRef]

Allwood, Jens. 1976. The complex NP constraint as a non-universal rule and some semantic factors influencing the acceptability of Swedish sentences which violate the CNPC. In University of Massachusetts Occasional Publications in Linguistics II. Edited by Justine Stillings. Amherst: University of Massachusetts, pp. 1–20.

Allwood, Jens. 1999. The Swedish spoken language corpus at Göteborg University. In Gothenburg Papers in Theoretical Linguistics. Göteborg: Göteborg University, p. 81.

Andersson, Lars-Gunnar. 1975. Form and Function of Subordinate Clauses. (Gothenburg Monographs in Linguistics 1). Gothenburg: Department of Linguistics, University of Gothenburg.

Andersson, Lars-Gunnar. 1982. What is Swedish an exception to? In Readings on Unbounded Dependencies in Scandinavian Languages. Edited by Elisabet Engdahl and Eva Ejerhed. Stockholm: Almqvist & Wiksell International, pp. 33–45.

Andersson, Sven-Gunnar, and Sigmund Kvam. 1984. Satzverschränkung im Heutigen Deutsch. Volume 24 of Studien zur deutschen Grammatik. Tübingen: Gunter Narr.

Bentzen, Kristine, and Merete Anderssen. 2019. The form and position of pronominal objects with non-nominal antecedents in Scandinavian and German. Journal of Comparative Germanic Linguistics 22: 169–188. [CrossRef]

Boersma, Paul, and David Weenink. 2020. Praat: Doing Phonetics by Computer [Computer Program, Version 6.1.16]. Available online: http://www.praat.org/ (accessed on 6 June 2020).

Bosch, Peter, Graham Katz, and Carla Umbach. 2007. The non-subject bias of German demonstrative pronouns. In Anaphors in Text: Cognitive, Formal and Applied Approaches to Anaphoric Reference. Edited by Monika Schwarz-Friesen, Manfred Consten and Mareile Knees. Amsterdam: John Benjamins, pp. 154–64.

Bouma, Gerlof. 2008. Starting a Sentence in Dutch. Ph.D. thesis, University of Groningen, Groningen, The Netherlands.

Brandlt, Johan. 2019. The question of form in the forming of questions. The meaning and use of clefted wh-interrogatives in Swedish. Journal of Linguistics 55: 755–54. [CrossRef]

Chaves, Rui P., and Michael T. Putnam. 2020. Unbounded Dependency Constructions: Theoretical and Experimental Perspectives. Oxford: Oxford University Press.

Chomsky, Noam. 1973. Conditions on transformations. In A Festschrift for Morris Halle. Edited by Steven R. Anderson and Paul Kiparsky. New York: Holt, Rinehart and Winston, pp. 232–86.

Chung, Sandra, and James McCloskey. 1983. On the interpretation of certain island facts in GPSG. Linguistic Inquiry 14: 704–13.
