

On *wh*-questions and V2 across Norwegian dialects^{*}

A survey and some speculations

Øystein Alexander Vangsnes

Main clause *wh*-questions without V2 is geographically speaking a very widespread phenomenon in Norwegian, absent only in central eastern and southeastern Norwegian dialects. The characteristics of the phenomenon is nevertheless not uniform across the dialects—three descriptive variables can be discerned: (i) whether or not V2 is required, (ii) whether or not V2 is optional alongside non-V2, and (iii) whether or not non-V2 is allowed with short *wh*-elements only. In addition these variables are relative to a $\pm$ subjecthood distinction on the *wh*-constituent. Although the phenomenon has received considerable attention, the paper will show that it has been far from exhaustively studied. The overall goal of the paper is to give an overview of the variation across the dialects based on existing literature. In turn a tentative microcomparative account of the variation will be given, following certain theoretical conjectures. In sum this will hopefully provide useful guidelines for future investigations of the phenomenon.

1. Introduction

This paper presents a survey of main clause *wh*-questions without Verb Second (henceforth ‘V2’) in Norwegian dialects, based on existing literature. As the survey will show there is considerable variation across the dialects, and the paper furthermore presents a highly tentative account of the variation. The account is tentative not only from a theoretical point of view: as will become evident the variation pertaining to the phenomenon is empirically speaking still far from exhaustively investigated. This is so

^{*} Versions of this paper have been presented at the University of Oslo, June 2003, at the ‘Syntactic Variation’ workshop at ICLaVE 2, Uppsala June 2003, at the CASTL Kick-off conference in Tromsø October 2003, at the seminar ‘Typology and parameter theory’ in Oslo, October 2003, at the University of Padua, October 2003, at the Journée d'études: ‘La syntaxe de la périphérie gauche’ in Paris, December 2003, and at the University of Siena, March 2004. I thank the audiences on these occasions for fruitful feedback. Moreover, for comments, advice, and/or discussion I thank Tor A. Åfarli, Josef Bayer, Adriana Belletti, Valentina Bianchi, Leonie Cornips, Karen Corrigan, Helge Dyvik, Jan Terje Faarlund, Marit Julien, Svein Lie, Ora Matushansky, Nicola Munaro, Torbjørn Nordgård, Hans Georg Obenauer, Christer Platzack, Cecilia Poletto, Jean-Yves Pollock, Luigi Rizzi, and Marit R. Westergaard. This work has been supported by the Norwegian Research Council through grant no. 148440/530.

even though the lack of V2 in main clause *wh*-question is one of the few *syntactic* variables in Norwegian dialects that have received considerable attention, cf. the following, presumably incomplete, list: Iversen (1918:37), Elstad (1982), Sandøy (1987:104), Nordgård (1985, 1988), Åfarli (1986), Taraldsen (1986a), Lie (1992), Nilsen (1996), Fiva (1996), Rice & Svenonius (1998), Westergaard (2003a, 2003b), Westergaard & Vangsnes (forthcoming).

The lack of exhaustiveness is in fact the main motivation for not awaiting further empirical studies before presenting this work in progress: the hope is that both the speculations and the survey as such can be valuable for future research on the topic (and related issues). The main finding of the paper will indeed be a set of working hypotheses which further empirical investigations hopefully can answer.

As a first illustration of the topic of the paper, compare the examples in (1) and (2) from the Tromsø dialect and Standard Norwegian, respectively.

- | | |
|---|-------------------------------------|
| <p>(1) a. <i>Ka du sa?</i>
 what you said
 ‘What did you say?’
 b. <i>Ka sa du?</i>
 what said you
 ‘What did you say’</p> | <p>Tromsø</p> |
| <p>(2) a. *<i>Kva du sa?</i>
 what you said
 b. <i>Kva sa du?</i>
 what said you
 ‘What did you say?’</p> | <p>Standard (Nynorsk) Norwegian</p> |

The example in (2a) is ungrammatical: lack of V2 is not possible in Standard Norwegian. The corresponding word order is however fine for the Tromsø dialect, as illustrated by (1a).

Such a lack of V2 in *wh*-questions in Norwegian dialects is a very widespread phenomenon. In fact, judging from the existing literature, in particular Nordgård (1985) and Lie (1992), lack of V2 in main clause *wh*-questions can be encountered in all dialect areas except central eastern Norway and the adjacent southeastern Norway (‘Sørlandet’). In other

words, the phenomenon is a trait of the majority of Norwegian dialects, geographically speaking.

Notice, incidentally, that the Standard Norwegian examples in (2) are rendered in the Nynorsk standard of written Norwegian, evident by the *wh*-word (which would have been *hva* if the examples were rendered in the Bokmål standard). This is a significant fact since users of the Nynorsk standard very often are native speakers of dialects which in fact allow lack of V2 in *wh*-questions, the main “stronghold” of the Nynorsk standard being western Norway. In other words, we are dealing with a widespread syntactic phenomenon which very clearly has not made its way into the standard register, a somewhat uncommon situation for Norwegian.

Before we look more closely at dialectal Norwegian let us consider some basic properties of clausal word order in Standard Norwegian. Standard Norwegian is a “well-behaved” V2 language and exhibits subject/verb inversion both in declarative and interrogative clauses. This is exemplified in (3) and (4) for a non-subject (topic) initial declarative clause and a *wh*-question, respectively.

- (3) a. *I dag bles det veldig mykje.* Standard Norwegian
 in day blows it very much
 ‘Today there is a very strong wind.’
 b. **I dag det bles veldig mykje.*
- (4) a. *Kor ofte bles det på denne måten her?* St. Nor.
 how often blows it on this way-DEF here
 ‘How often do you have such strong winds here?’
 b. **Kor ofte det bles på denne måten her?*

Subject initial clauses such as in (5) also exhibit V2, made evident by the fact that the finite verb must precede negation and other sentence adverbs.

- (5) a. *Det bles heldigvis ikkje lenger ute.* St. Nor.
 it blows fortunately not longer outside
 ‘It’s fortunately not blowing any longer outside’
 b. **Det heldigvis ikkje bles lenger ute.*
 c. **Det heldigvis bles ikkje lenger ute*

These are well-known facts that hold for Mainland Scandinavian more generally.

There is no general V2 requirement in Standard Norwegian embedded clauses. Moreover there is no “V-to-I movement” either: the finite verb appears to the right of IP adverbs. This root/embedded asymmetry wrt. verb movement, which holds for Mainland Scandinavian in general, is exemplified by the *wh*-questions in (6) and (7) and by the declarative clauses in (8) and (9).

- (6) a. *Kva sa han egentleg?* Standard Norwegian
 what said he actually
 b. **Kva han egentleg sa?*
 what he actually said
 c. **Kva han sa egentleg?*
 what he said actually
 d. **Kva egentleg sa han?*
 e. **Kva egentleg han sa?*
- (7) a. *Eg lurur på kva han egentleg sa.* St. Norwegian
 I wonder on what he actually said
 ‘I wonder what he actually said.’
 b. **Eg lurur på kva sa han egentleg.*
 I wonder on what said he actually
 c. **Eg lurur på kva han sa egentleg.*
 I wonder on what he said actually
- (8) a. *Eg likte faktisk denne filmen.* St. Norwegian
 I liked in-fact this movie-DEF
 ‘I actually liked this movie.’
 b. **Eg faktisk likte denne filmen.*
 I in-fact liked this movie-DEF
- (9) a. *Eg trur at eg faktisk likte denne filmen.* St. Norwegian
 I think that I in-fact liked this movie-DEF
 ‘I think that I actually liked this movie.’
 b. **Eg trur at eg likte faktisk denne filmen.*
 I think that I liked in-fact this movie-DEF

With respect to embedded *wh*-questions Standard Norwegian exhibits an asymmetry between subject and non-subject questions: the element *som* obligatorily follows *wh*-subjects but is illicit after non-subject *wh*-constituents.¹ Again this is a common trait of the Mainland Scandinavian varieties in general.

- (10) a. *Eg lurer på kven *(som) kjem på konferansen.* St. Nor.
 I wonder on who SOM comes on conference-DEF
 b. *Eg lurer på kva (*som) han skal snakke om.*
 I wonder on what SOM he shall talk about
 c. *Eg lurer på kor (*som) han bur.*
 I wonder on where SOM he lives
 d. *Eg lurer på korleis (*som) du kan vite det.*
 I wonder on how SOM you can know that

The general word order pattern for Standard Norwegian (and Mainland Scandinavian) *wh*-questions can thus be summarized as in (11) and (12) for main and embedded clauses, respectively.

- (11) a. Main: *Wh*_[-subject] **Vfin** Subject (Adverb) (V) (Object)
 b. Main: *Wh*_[+subject] **Vfin** (Adverb) (V) (Object)
 (12) a. Emb.: *Wh*_[-subject] Subject (Adverb) **Vfin** (V) (Object)
 b. Emb.: *Wh*_[+subject] **SOM** (Adverb) **Vfin** (V) (Object)

Let us then turn to the Norwegian dialects which allow lack of V2 in main clause *wh*-questions.

One general property concerning main clause *wh*-questions without V2 is that their word order conforms to the word order pattern in embedded clauses in Standard Norwegian in the sense that there is *som*-insertion after *wh*-subjects and there is no “verb movement”. We illustrate this with examples from the Tromsø dialect, which is a northern Norwegian (‘nordnorsk’) dialect. The examples in (13) show that *som* can be inserted after a *wh*-subject only, the examples in (14) show that *som*-insertion is not compatible with “moving” the finite verb past a sentence adverb, and the

¹ As will be discussed in section 5, presence of *som* in such cases appears to be possible for some speakers.

examples in (15) show that the finite verb cannot “move” to a position between the subject and a sentence adverb in a non-subject *wh*-clause.

- (13) a. *Kem som kom på konferansen?* Tromsø
 who SOM came on conference-DEF
 ‘Who came to the conference?’
 b. *Ka (*som) han snakka om?*
 what SOM he talked about
 ‘What did he talk about?’
 c. *Kor (*som) han bor?*
 Where SOM he lives
 ‘Where does he live?’
- (14) a. *Kem som faktisk kom på konferansen?* Tromsø
 who SOM in-fact came on conference-DEF
 ‘Who did actually come to the conference?’
 b. **Kem som kom faktisk på konferansen?*
 who SOM came actually on conference-DEF
- (15) a. *Ka han Josef aldri fikk vite?* Tromsø
 whatART Joseph never got know
 ‘What did Joseph never get to know?’
 b. **Ka han Josef fikk aldri vite?*
 what ART Joseph got never know

Importantly, the Tromsø dialect as well as (the majority of) the other dialects which exhibit non-V2 in main clause *wh*-questions have V2 in declarative main clauses.² In other words the examples in (16)–(18) are thus representative for Norwegian dialects in general (but see note 2).

² Exceptions to this concern northern Norwegian dialects that have either replaced Finnish and/or Saami or been heavily influenced by these non-Scandinavian languages. One example of this is the dialect of the municipality Nordreisa in the northern part of Troms county, thoroughly studied by Sollid (2003)(see also Nilsen [=Sollid] 1996). Sollid (2003) investigated the language shift in the village Sappen in Nordreisa which was predominantly Finnish-speaking a few generations ago, and found that older speakers of the contemporary dialect to a significant degree accept for example non-V2 also in main clause declaratives with a non-subject initial constituent (i.e. topicalizations) and not just in main clause *wh*-questions.

- (16) a. *Æ kommer dessverre ikkje på seminaret.* Tromsø
 I come unfortunately not on seminar-DEF
 ‘I will unfortunately not come to the seminar.’
 b. **Æ dessverre kommer ikkje på seminaret.*
 c. **Æ dessverre ikkje kommer på seminaret.*
- (17) a. *Dessverre kommer æ ikkje på seminaret.* Tromsø
 Unfortunately come I not on seminar-DEF
 ‘Unfortunately I will not come to the seminar.’
 b. **Dessverre æ kommer ikkje på seminaret.*
 c. **Dessverre æ ikkje kommer på seminaret.*
- (18) a. *På seminaret kommer æ dessverre ikkje.* Tromsø
 on seminar-DEF come I unfortunately not
 ‘To the seminar I can unfortunately not come.’
 b. **På seminaret æ kommer dessverre ikkje.*
 c. **På seminaret æ dessverre kommer ikkje.*

Although the V2 requirement in declaratives is uniform across the dialects, the lack of V2 in *wh*-questions is in fact not uniform across the dialects which exhibit it, and the main objective of this paper is to give an overview—and a tentative microcomparative account—of this variation. As we will see, there appears to be three major variables: (i) first of all whether or not non-V2 is allowed in root *wh*-questions at all, (ii) then whether or not non-V2 is optional, i.e. whether V2 and non-V2 can co-exist, and (iii) whether or not non-V2 is restricted to short *wh*-elements only. Moreover, these variables can be shown to be relative to the $\pm$ subject status of the *wh*-constituent. This means that we have the following typology for (main clause) *wh*-grammars in Norwegian dialects.

	non-V2	“optional” V2	short <i>wh</i> only
<i>wh</i> [+SUBJ]	$\pm$	$\pm$	$\pm$
<i>wh</i> [-SUBJ]	$\pm$	$\pm$	$\pm$

Table 1: Typology for *wh*-grammars in Norwegian dialects

There is necessarily a dependency on the variables from right to left in this table: the issue of optional V2 cannot arise unless non-V2 is allowed in the first place, and moreover if V2 is not an option, then we would not expect non-V2 to be restricted to short *wh*-constituents only (i.e. there would not be any way for longer *wh*-constituents to be licit if V2 is not an option).

This interdependency between the variables means that the typology predicts 16 possible varieties. Less than half of these varieties can be discerned from the existing literature, and even though further empirical investigations may reveal further varieties, we will present theory internal reasons for not expecting all 16 predicted varieties to be possible.

In the remainder of the paper we will now first take a closer look at the variables that govern the crossdialectal variation, starting with the complexity constraint on the *wh*-constituent in section 2 and then moving on to the issue of optional V2 in section 3. Section 4 will present the different varieties described in the existing literature relative to the typology in table 1. In section 5 we will sketch the history and present day use of the element *som* in more detail, pointing out certain facts that will be relevant for the theoretical account of the cross-dialectal variation in root *wh*-questions. The theoretical account proper will be presented in section 6. Certain consequences of the theoretical account will then be briefly discussed in section 7 before section 8 concludes the paper.

2. The complexity constraint

In the Tromsø dialect there are only three short *wh*-constituents that allow non-V2 in non-subject (root) *wh*-questions. These are the monosyllabic *ka* ‘what’, *kem* ‘who’, and *kor* ‘where’. This “complexity constraint” on non-V2 can be illustrated with examples involving the combination of the *wh*-element *kor* ‘where’ and the locative particle *hen*³. The locative particle can be either fronted along with the *wh*-element or left in situ in the thematic position, but notice that non-V2 is only compatible with the in situ option in the Tromsø dialect. This is shown by the examples in (19) and (20): the crucial ungrammatical example is (20a).

³ The locative particle is presumably cognate with the German suffix *-hin* as in *wohin* ‘where-to’, but the Norwegian *hen* is compatible with both an illative (as in German) and an inessive meaning (unlike German).

- (19) a. *Kor du bor hen?* (20) a. **Korhen du bor?* Tromsø
 where you live LOC where-LOC you live
 b. *Kor bor du hen?* b. *Korhen bor du?*
 where live you LOC where-LOC live you
 ‘Where do you live?’ ‘Where do you live?’

The difference between the short, monosyllabic *wh*-elements and longer constituents is further exemplified by the examples in (21) and (22).

- (21) a. *Ka han Jens sa?* Tromsø
 what ART Jens said
 ‘What did Jens say?’
 b. *Kem han Jens snakka me?*
 who ART Jens talked with
 ‘Who did Jens talk to?’
 c. *Kor han Jens bor?*
 where ART Jens lives
 ‘Where does Jens live?’
- (22) a. **Korsn han Jens vesste det?* Tromsø
 how ART Jens knew that
 b. **Koffør han Jens sa det?*
 why ART Jens said that
 c. **Katti han Jens kommer?*
 when ART Jens comes
 d. **Ka slags bil han Jens har kjøpt sæ?*
 what kind-of car ART Jens has bought himself

These examples all have non-V2, and as we see the disyllabic adverbial *wh*-elements *korsn* ‘how’, *koffør* ‘why’, and *katti* ‘when’ cannot be immediately followed by the subject. The example in (22d) furthermore shows that the same holds for a fronted noun phrase with a *wh*-determiner.

The complexity constraint on non-V2 is not present in all dialects, however. The dialect of Nordmøre, spoken in northwestern Norway⁴, contrasts with the Tromsø dialect in this respect. The examples in (23) and (24) are the equivalents of (19) and (20) in this dialect, and (24a) shows that

⁴ Nordmøre is the northernmost of the three parts that make up the county Møre og Romsdal. The other two parts are Romsdal and Sunnmøre.

‘where’ and the locative particle (*hen*) can be fronted together when there is non-V2, unlike what we saw for the Tromsø dialect.

- (23) a. *Kar du bor hen?* (24) a. *Karhen du bor?* Nordmøre
 where you live LOC where-LOC you live
 b. *Kar bor du hen?* b. *Karhen bor du?*
 where live you LOC where-LOC live you
 ‘Where do you live?’ ‘Where do you live?’

The Nordmøre dialect has been studied by Åfarli (1986), and he gives the following examples to illustrate that complex *wh*-constituents can occur with both V2 and non-V2 in the dialect.

- (25) a. *Kåles bil kjøpte du deg?* Nordmøre
 what-kind-of car bought you yourself
 ‘What kind of car did you buy?’
 b. *Kåles bil du kjøpte deg?*
 what-kind-of car you bought yourself
 ‘What kind of car did you buy?’

In his study Åfarli found a preference for V2 with the disyllabic adverbs *kåfer* ‘why’ and *ka tid* ‘when’.

- (26) a. *OK Kåfer kjem du hit?* Nordmøre
 why come you here (Åfarli 1986)
 b. *?? Kåfer du kjem hit?*
 why you come here
- (27) a. *OK Ka tid kjem du?* Nordmøre
 what time come you (Åfarli 1986)
 b. *? Ka tid du kjem?*
 what time you come

This preference for V2 with the *wh*-adverbs is noteworthy, but it is also worth noticing that we are not dealing with sharp ungrammaticality with non-V2 as in the Tromsø dialect. The preference for V2 may in fact stem from variation among the informants in the study: there were (only) 10 informants of varying age from 6 different locations within Nordmøre, and Åfarli’s presentation does not clarify how the sentences were judged by the

individual informants. (See however section 6.7 for another possible way of viewing Åfarli's reported data.)

In any event there is no doubt that there is a significant difference between the Tromsø and the Nordmøre dialect with respect to the complexity constraint. One interesting question is what exactly ‘complexity’ refers to: is it phonological complexity or is it morphosyntactic complexity that matters? Earlier accounts, e.g. Elstad (1982), Taraldsen (1986a), and Rice & Svenonius (1998) have claimed that it is syllabicity that is the decisive factor in the Tromsø dialect: non-V2 is only allowed with monosyllabic *wh*-constituents.

There are however some reasons to think that the complexity constraint is morphosyntactic in nature. One piece of evidence pointed out by Merete Anderssen (p.c.) concerns the contracted form *kas* of the *wh*-determiner *ka slags* ‘what kind-of’. Both the contracted and the full form of the determiner may felicitously appear with a null noun, but a DP with the contracted form as the only overt phonological material (i.e. *kas* $\emptyset$) is nevertheless not compatible with non-V2. In other words *kas* $\emptyset$ is no different from *ka slags* $\emptyset$ in this respect. This is exemplified by the examples in (28) and (29): if the complexity constraint were purely phonological in nature we would expect (29a) to be grammatical.

- (28) a. *Kas* ___ *kjøpte han?* Tromsø
 what-kn'a bought he
 b. *Ka slags* ___ *kjøpte han?*
 what kind-of bought he
 'Which kind did he buy?'
- (29) a. **Kas* ___ *han kjøpte?* Tromsø
 what-kn'a he bought
 b. **Ka slags* _ _ *han kjøpte?*
 what kind-of he bought
 'Which kind did he buy?'

Another piece of evidence against a purely phonological account of the complexity constraint (noted by Rice & Svenonius 1998 themselves) concerns the fact that short, monosyllabic *wh*-elements can combine with swear words and still be compatible with non-V2.

- (30) a. *Ka faen du sa?* Tromsø
 what devil-DEF you said
 ‘What the hell did you say?’
 b. *Kem faen du trur du e?*
 who devil-DEF you think you are
 ‘Who the hell do you think you are?’

For the purposes of the rest of the paper we will therefore assume that the complexity constraint on non-V2 has a morphosyntactic rather than a purely phonological basis.

So far we have considered examples with non-subject *wh*-questions. Interestingly, there does not seem to be any complexity constraint on subject *wh*-questions in the Tromsø dialect. In a questionnaire survey of approximately 40 adolescent Tromsø speakers, Fiva (1996) found that subject *wh*-questions with a complex subject were judged relatively well-formed. (31) renders the relevant example (see Fiva 1996:148) with my interpretation of the overall judgment—recall that the instantiation of non-V2 in subject *wh*-questions is the insertion of *som* after the initial *wh*-constituent (plus, of course, no verb movement).

- (31) ?? *Kor mange eleva som møtte opp i dag?* Tromsø
 how many pupils SOM met up in day
 ‘How many pupils showed up today?’

Importantly, the relative acceptance of this example contrasts sharply with a clear non-acceptance of complex *wh*-DPs in non-subject *wh*-questions with non-V2 in Fiva’s study.

Fiva’s finding is partly supported by the finding reported in Nordgård (1985:12) for “Northern Norwegian”. He notes that the word order in (32a) with a complex *wh*-determiner and a null noun followed by *som* is accepted by all informants. On the other hand the example in (32b) where the complex *wh*-determiner co-occurs with an overt noun was judged as degraded by Nordgård’s informants.

- (32) a. *Kor mange som vil være med til London?* Northern Nor.
 how many SOM will be with to London
 ‘How many would like to come along to London?’
 b. *?*Kor mange kvinnelige forfattere fra den tredje verden som har fått Nobel-prisen i litteratur?*
 ‘How many female writers from the third world SOM have received the Nobel price in literature?’

Nordgård’s study suffers from certain methodological flaws, however. In his study, which altogether comprises of 25 informants, only 4 informants are from “Northern Norway” and no finer distinction is made within this vast geographical area. Moreover, the study was based on a written questionnaire of examples rendered in the Bokmål standard of Norwegian (and thus not the way I have rendered the examples in (32), i.e. in accordance with northern Norwegian pronunciation), and as we see the *wh*-phrase in (32b) is in addition quite complex, containing both an adjective and a PP in addition to the quantificational *wh*-expression. In other words, the degradedness of (32b) in Nordgård’s study may stem from (i) a too coarse a grouping of the informants (i.e. they may in fact have distinct dialects), (ii) disturbance of standard orthography, and/or (iii) simply parsing difficulties.

In fact Marit R. Westergaard, who is a native speaker of the Tromsø dialect, informs me that (32b) is acceptable for her and we can therefore conclude that there is no (clear) complexity constraint in root subject *wh*-questions in this particular variety of northern Norwegian. In turn this means that the variable as such is relative to the $\pm$ subject status of the *wh*-constituent. As we will see this is also the case for the other two variables that govern the variation across the dialects.

Let us then turn to the question of whether V2 is possible alongside non-V2.

3. Optional V2

It has been claimed (see Taraldsen 1986a, Rice & Svenonius 1998) that V2 is not allowed in the Tromsø dialect in non-subject *wh*-questions with short *wh*-element. In other words *Wh-SUB-Vfin* should be the only option in such cases. The proponents of this view claim that to the extent that the word order *Wh-Vfin-SUB* is accepted (and/or produced) by speakers of the dia-

lect, this is simply due to influence from Standard Norwegian. The main reason for advocating such a view is of course that many theoretical linguists hold, on principled grounds, that true grammar-internal optionality cannot exist.

There are however reasons to believe that this view is inadequate, and in the following I will discuss optional V2 first with respect to non-subject *wh*-questions and then with respect to subject *wh*-questions. The general picture that will emerge is that there sometimes is a choice between V2 and non-V2 in similar grammatical environments, but that the choice is governed by pragmatic factors. In other words we are not talking about true optionality in any event. Let us first consider the V2 optionality with respect to non-subject *wh*-questions.

In spontaneous speech in the Tromsø dialect short non-subject *wh*-constituents will occur with both V2 and non-V2 in root contexts. In other words, both examples in (1) can in fact be encountered in spontaneous Tromsø speech, even produced by one and the same individual.

- (1) a. *Ka du sa?* Tromsø
 what you said
 ‘What did you say?’
 b. *Ka sa du?*
 whatsaid you
 ‘What did you say’

The fact that both V2 and non-V2 occur in the Tromsø dialect has been documented by Westergaard (2003a, 2003b) in a corpus study of 1 adult Tromsø informant and 3 children acquiring the Tromsø dialect. (See also Westergaard & Vangsnes forthcoming). Table 2 gives an impression of the co-occurrence of V2 and non-V2 in the speech of the adult investigator.

<i>Wh</i> -element	V2	non-V2	Total
<i>Ka</i> (‘what’)	58 (31.9%)	124 (68.1%)	182 (100%)
<i>Kor</i> (‘where’)	38 (56.7%)	29 (43.3%)	67 (100%)
<i>Kem</i> (‘who’)	40 (78.4%)	11 (21.6%)	51 (100%)
Total	136 (45.3%)	164 (54.7%)	300 (100%)

Table 2: Occurrences of V2 and non-V2 word order in non-subject *wh*-questions of the adult Tromsø speech in the files “Ole.13-22” (Westergaard 2003a, 2003b, Westergaard & Vangsnes forthcoming)

Moreover, Westergaard finds an interesting correlation between placement of the verb in root non-subject *wh*-questions and the types of subjects that the clause contains: V2 is more frequent with full DPs and when the main verb is semantically light (typically ‘be’) whereas non-V2 is more frequent when the subject is a pronoun or an expletive and the verb is not ‘be’. Sentences (33) and (34) are thus typical examples of the V2 and non-V2 constructions with the question word *kor* ‘where’.

(33) *Kor er pingvinen henne?* (Investigator in the file “Ole.16”)
 where is penguin-DEF LOC (Tromsø)
 ‘Where is the penguin?’

(34) *Kor du har fått det henne?* (Investigator in the file “Ole.22”)
 where you have got that LOC (Tromsø)
 ‘Where did you get that?’

Westergaard’s interpretation of this correlation is that V2 is typically found with informationally *new* subjects whereas non-V2 is typically found with *given* subjects. In other words the choice is governed by pragmatic factors.⁵

This finding is to some extent corroborated by the following statement that Lie (1992:73) gives about his own Hedalen dialect, spoken in the district of Valdres in the “inner”, western part of eastern Norway: “*Hå du si?* [What you say?] is according to my intuition preferably used to have someone repeat something. If one encourages someone to say something new, I would rather say *Hå si du?* [What say you?]

⁵ An anecdote which ties in with Westergaard’s findings concerns the following self-experienced incident. Apparently, my e-mail address had been used by spammers to distribute a message saying “You try to steal” (in English). Totally unknowing of this I received a reply message from a person based in Tromsø who wrote the following two main clause *wh*-questions (and nothing more): *Kæm e du?! Ka du mene?!*, literally ‘Who are you?! What you mean?!’. The spelling leaves no doubt that this is dialectal (northern) Norwegian. Notice then that the first question has V2, and it introduces the 2nd person pronominal subject. In addition the main verb is the copula. The second question has non-V2 and the referentially identical subject is of course informationally given by now.

value it at least suggests that pragmatic factors govern the choice of V2 over non-V2.

Another piece of evidence suggesting that pragmatic factors are involved concerns Åfarli's (1986) claim that V2 and non-V2 can be used interchangeably without significant differences in meaning in the Nordmøre dialect, but that there nevertheless is a slight preference for V2 the "heavier" the subject is. Åfarli gives the following examples to illustrate the point.

- (35) a. *Kåles gammel hattkaill du tala med i går?* Nordmøre
 which old 'hat-man' you talked with yesterday
 'Which old 'hat-man' were you talking to yesterday?'
 b. *?Kåles gammel hattkaill onkelen din fra Oslo tala med i går?*
 which old 'hat-man' uncle-DEF your from O. talked with yesterday
 'Which old 'hat-man' was your uncle from Oslo talking to
 yesterday?'

When this finding for the Nordmøre dialect is brought together with Westergaard's finding for the Tromsø dialect, one could argue that the oddness of (35b) is due to the unlikelieness of the heavy subject to qualify as informationally new.

Let us then turn to the optionality of V2 in subject *wh*-questions, i.e. where the element *som* appears in second position. First of all we should notice that *som*-insertion, and thus non-V2, appears to be obligatory in root subject *wh*-questions in both the Nordmøre and the Tromsø dialect. This is explicitly stated for the Nordmøre dialect in Åfarli (1986), who gives the examples in (36) and (37) to illustrate the point.

- (36) a. **Kåin kjem der?* Nordmøre
 who comes there
 b. *Kåin så kjem der?*
 who SOM comes there
- (37) a. **Kåles mann kjem der?* Nordmøre
 what-kind-of man comes there
 b. *Kåles mann så kjem der?*
 what-kind-of man SOM comes there

Åfarli's claim about the Nordmøre dialect is corroborated by the finding in Nordgård (1985). His informants from Møre, which subsumes Nordmøre, also require insertion of *som* in root questions with *wh*-subjects.

Nordgård (1985:11) finds the same requirement to hold also for his four "Northern Norwegian" informants. A relevant example from his study can thus be rendered as in (38).

- (38) *Kem *(som) har drukke opp all vinen?* "Northern Norwegian"
 who SOM has drunk up all wine-DEF
 'Who has drunk all the wine?'

Marit R. Westergaard (p.c.) furthermore informs me that *som*-insertion is obligatory in main clause subject *wh*-questions in her Tromsø dialect, and she reports that this holds irrespective of the complexity of the *wh*-subject. The situation can be illustrated by the examples in (39).

- (39) a. *Kem *(som) kom på forelesninga?* Tromsø
 who SOM came on lecture-DEF
 'Who came to the lecture?'
 b. *Kor mange studenta *(som) kom på forelesninga?*
 how many students SOM came on lecture-DEF
 'How many students came to the lecture?'

This (putative) obligatoriness of *som*-insertion in the Tromsø dialect may in fact shed light on the lack of a complexity constraint in subject *wh*-questions with non-V2 in this dialect (cf. section 2). If *som*-insertion, and thus non-V2, is obligatory, we should not expect complex *wh*-subjects to be barred from occurring in such structures.

On a general note we see again that the variable is relative to the $\pm$ subjecthood of the *wh*-constituent. Although neither the Nordmøre nor the Tromsø dialects optionally allow V2 with *wh*-subjects we will shortly see that there are other Norwegian dialects that do allow V2 (and lack of *som*-insertion) along with non-V2 (and insertion of *som*) in such cases.

The general finding with respect to the optionality of V2 is that whenever V2 is an option alongside non-V2 it seems to be subject to pragmatic conditions. Exactly what these pragmatic conditions are need to be further

investigated. So far Westergaard’s study provides the most detailed approximation, and we will later base some of the theoretical speculations on her findings.

In the next section we will now first summarize the properties of the Nordmøre and Tromsø dialects relative to the typology presented at the end of the introduction and then continue to present some other distinct dialect grammars that can be discerned on the basis of the existing literature.

4. Six varieties

The properties of the Nordmøre and Tromsø main clause *wh*-grammars can be summarized as in tables 3 and 4, respectively. The ‘#’ mark in the last column is meant to indicate that the question is irrelevant given the negative setting for the preceding variable.

Nordmøre	non-V2	“optional” V2	complexity constraint
<i>wh</i> [+SUBJ]	+	–	#
<i>wh</i> [–SUBJ]	+	+	–

Table 3: *wh*-grammar for the Nordmøre dialect

Tromsø	non-V2	“optional” V2	complexity constraint
<i>wh</i> [+SUBJ]	+	–	#
<i>wh</i> [–SUBJ]	+	+	+

Table 4: *wh*-grammar for the Tromsø dialect

The most important source of information about further cross-dialectal variation relevant to the phenomenon is Nordgård (1985, 1988). We noted earlier that this study suffers from certain methodological flaws, but it is nevertheless the only study of the variation which undertakes a systematic grammatical approach to the phenomenon, and the results that it reports show a highly interesting pattern. The study will therefore be of immense help to further explorations of the phenomenon.

The variables that Nordgård uses in his study are slightly different from the ones in the typology presented here, the main difference being that

he singles out a $\pm$ adjunct variable. The actual results can nevertheless by and large be translated into the present typology. We have previously noted that the Nordmøre dialect is subsumed by Nordgård’s “Møre” group of informants, which incidently indeed is reported to have the properties given in table 3. Similarly, the Tromsø dialect would be subsumed by Nordgård’s “Northern Norwegian” group of informants, comprising of just 4 individuals, and his findings for this group only partially conform to the properties given in table 4: Nordgård does not find a clear-cut lack of the complexity constraint for *wh*-subjects, although he does in fact find a quite clearcut obligatory requirement for *som*-insertion. This last point about Nordgård’s results itself suggests that his study of the “Northern Norwegian” dialect is inadequate: as discussed above we should not, on principled grounds, expect a complexity constraint on the *wh*-constituent if *som*-insertion is obligatory. The northern Norwegian dialect area is indeed one where more detailed studies are called for.

The dialect (group) with the highest number of informants in Nordgård’s study is “Mid Norwegian”. 10 of the altogether 25 informants in his study belong to this group. This dialect (type) is spoken in the Trøndelag counties, i.e in the districts around the city of Trondheim, and in the following I will refer to this dialect as the ‘Trøndelag’ dialect. Trøndelag borders with Nordmøre in the southwest, eastern Norway in the southeast, Sweden (Jämtland) in the east, and Northern Norway in the north. The properties of the Trøndelag dialect with respect to lack of V2 in root *wh*-questions are given in table 5.

Trøndelag	non-V2	“optional” V2	complexity constraint
<i>wh</i> [+SUBJ]	+	+	+
<i>wh</i> [−SUBJ]	+	+	+

Table 5: *wh*-grammar for Trøndelag (based on Nordgård 1988:32)

We see that this dialect has optional V2 in both subject and non-subject *wh*-questions, and furthermore that there is a complexity constraint in both cases. The dialect is thus different from both the (adjacent) Nordmøre dialect and the Tromsø dialect.

The “next” distinct dialect in Nordgård’s study when we continue down the coast south of (Nord-)Møre, is the Hordaland dialect, which is

spoken in the county by the same name, i.e the districts around the city of Bergen. There were only 3 Hordaland informants in Nordgård's study, and on certain points the picture is somewhat unclear for this dialect. Nordgård notes the following:

“Only one of the sentences with complex *wh*-phrases without inversion is accepted [...] It is impossible in the Hordaland dialects to have non-inverted word-order when the fronted complex determiner [e.g. *how many*] is a non-subject. However, judgements vary when subject *wh*-phrases of this sort are fronted. 50% of these constructions are accepted. [...] Fronted *wh*-adjuncts without inversion cause ungrammaticality, though one informant accepts the variant with *hvor* (=where).” (Nordgård 1985:17)

On the basis of this I therefore render the properties of the Hordaland dialect as in table 6: the question marks indicate uncertainty. Later I will reveal that I expect further studies to find that, at least for one consistent group of speakers, non-V2 is only possible in subject *wh*-questions and furthermore that for these speakers there is no complexity constraint on the subject *wh*-constituent.⁶

Hordaland	non-V2	“optional” V2	complexity constraint
<i>wh</i> [+SUBJ]	+	+	??
<i>wh</i> [-SUBJ]	-/?	#	#

Table 6: *wh-grammar for Hordaland (based on Nordgård 1985:17, 1988:32)*

⁶ Interestingly, the way I have described the (expected) Hordaland dialect is partially how Lie (1992) describes the dialect of the city Haugesund. This city lies in the northernmost part of the county Rogaland. This part, Haugalandet, is immediately adjacent to Hordaland, and Lie reports that an informant for the Haugesund dialect considers non-V2 possible only with *wh*-subjects. No information about complexity is given, however. Since what Lie conveys is not compatible with the settings for the Rogaland dialect based on Nordgård's findings, this information about the Haugesund dialect may in turn suggest a split within the Rogaland dialects (i.e. that the relevant dividing line does not follow the Rogaland/Hordaland county border). In this respect it is of some relevance to note that there are other (mainly phonological) traits that the Haugalandet dialect(s) shares with dialects in the southern part of (coastal) Hordaland (to the north) and which distinguish them from Rogaland dialects further to the south, one prominent property being that the Haugalandet and southern coastal Hordaland dialects are low tone dialects. (The majority of western Norwegian dialects are high tone varieties).

The dialect of the county Rogaland is represented by 2 informants in Nordgård’s study, and the properties of this dialect is given in table 7.

Rogaland	non-V2	“optional” V2	complexity constraint
<i>wh</i> [+SUBJ]	+	+	+
<i>wh</i> [−SUBJ]	+	+	+

Table 7: *wh*-grammar for Rogaland (based on Nordgård 1988:32)

Notice that the Rogaland dialect thus gets the same settings as the Trøndelag dialect. The two dialects are nevertheless distinct in Nordgård’s study: (bare) *wh*-adverbs are better with non-V2 in Rogaland than in Trøndelag. This possibly coincides with a difference in complexity for one of the two *wh*-adverbs (‘how’ and ‘when’) used in his investigation: the standard (Bokmål) *hvordan* ‘how’ used in the questionnaire has the form *koss* (monosyllabic) in Rogaland but *korsn* (disyllabic) in Trøndelag. Further studies should observe this important difference.

Leaving Nordgård’s study now, we observe that Lie (1992:69) notes that the dialect of Nord-Østerdalen is yet different from the varieties that we have described above. Nord-Østerdalen lies in the northeastern part of eastern Norway, adjacent to Trøndelag to the north and Sweden (Härjedalen) to the east. According to Lie a speaker of this dialect informs him that non-V2 is only possible with non-subjects, and moreover optionally so. No information is given as to whether a complexity constraint accompanies the non-V2 option. The reported properties of the Nord-Østerdalen dialect can thus be rendered as in table 8.

N.-Østerdalen	non-V2	“optional” V2	complexity constraint
<i>wh</i> [+SUBJ]	–	#	#
<i>wh</i> [−SUBJ]	+	+	?

Table 8: *wh*-grammar for the Nord-Østerdalen dialect (based on Lie 1992:69)

Finally, for the sake of completeness we should give the properties of the Norwegian dialects that just like Standard Norwegian do not allow non-V2

in root *wh*-questions at all, i.e. central eastern Norwegian and southeastern Norwegian (henceforth ‘Eastern Norwegian’ for simplicity). These dialects have the properties in table 9.

Eastern Norwegian	non-V2	“optional” V2	complexity constraint
<i>wh</i> [+SUBJ]	–	#	#
<i>wh</i> [–SUBJ]	–	#	#

Table 9: *wh*-grammar for Eastern Norwegian

Altogether we have then described seven dialects and seen how they differ from each other at a more fine-grained level with respect to the lack of Verb Second in main clause *wh*-questions. Given that two of them, Trøndelag and Rogaland, have the same (idealized) settings in the typology, we have six different varieties. This is less than half of what the typology predicts to be possible *a priori* (6 of 16), and an interesting question is of course whether the descriptive typology can be further restricted on theoretical grounds. I will later argue that this is indeed the case. We return to that issue in section 6 after the major theoretical account has been presented.

Next we will consider some facts about the history and contemporary use of the element *som*.

5. The history and use of *som*

The element *som* is used in a variety of grammatical contexts in contemporary Norwegian. We may for descriptive purposes distinguish six different uses.⁷

⁷ A seventh use, pointed out to me by Kristin M. Eide, is the use of *som* as a temporal complementizer (instead of the more common *medan/mens* ‘while’). I would be inclined to describe this use somewhat uncommon, belonging to a high register, but nevertheless not exclude the possibility that it is used to a greater extent dialectally. Interestingly, for me at least, this use of *som* is fully felicitous only when the temporal clause is preposed, cf. the following sentence pair.

(i) a. *Som/mens vi stod der, begynte det å brenne.* St. Norwegian
 as / while westood there started it to burn
 ‘As/while we were standing there it started to burn.’

- (i) comparative complementizer
- (ii) relative complementizer
- (iii) predicate marker in essive small clauses
- (iv) obligatory element in subject clefts, impossible in non-subject clefts
- (v) obligatory element in (embedded) subject *wh*-questions, (almost) impossible in non-subject embedded *wh*-questions
- (vi) obligatory element in subject exclamatives, optional in non-subject exclamatives

These uses will be exemplified below. First let us note that etymologically speaking *som* (Old Norse *sem/sum*) is a derivative of the same root as the adjective *samme* ‘same’ (Old Norse *samr*), and thus cognate with English *same*, Latin *simul*-, Greek *homos*. In that respect it is furthermore of some interest to note that older stages of Norwegian did not exhibit all these uses of the element. In fact, in the earliest Old Norse texts use (i) is the only one that we see. In later Old Norse texts *sem/sum* starts to replace other relative markers. The remaining uses seem to be absent from Old Norse (Helge Dyvik, Jan-Terje Faarlund p.c.), although it remains to be fully investigated, especially for use (iii). Contemporary Icelandic only has use (ii) and (iii).

In other words, there has been a spread in the use of *som* as a functional element throughout the history of Norwegian. It started out as a comparative complementizer, then spread to relative clauses, and is now used in an even wider set of grammatical contexts. Several of these uses, and in particular the latter three, may presumably be manifestations of the same grammatical conditions, but I keep them apart here for descriptive purposes. For a fuller treatment of the history of *som* see Stroh-Wollin (1997).

Let us then for completeness exemplify the various uses. The examples in (40) illustrate the use of *som* as a comparative complementizer.

- (40) a. *Jon er like gammel som Marit.* St. Norwegian
Jon is just-as old SOM Marit.
'Jon is the same age as Marit.'
- b. *Så mange stillingar (som) de har, vil vi aldri få.*
so many positions SOM you have will we never get
'We will never get as many positions as you have.'
- c. *Som eg sa, er etymologien til som interessant.*

b. *Det begynte å brenne mens/??som vi stod der.*

SOM I said is etymology-DEF to *som* interesting
 ‘As I said the etymology of *som* is interesting.’

In this use *som* corresponds to English ‘as’.

The use of *som* as relative complementizer is possibly the use most often referred to by linguists and grammarians, and therefore also the most “famous” one—Norwegian school children are taught that *som* should be classified as a “relative pronoun”, clearly an inadequate doctrine for a number of reasons. Examples to show this noun phrase internal use of *som* is given in (41).

- (41) a. *fondet* **(som) har finansiert prosjektet* St. Norwegian
 fund-DEF SOM has financed project-DEF
 ‘the fund which has financed the project’
 b. *fondet* *(som) vi søkte penger frå*
 fund-DEF SOM we sought money from
 ‘the fund that we applied for money from’
 c. *fondet* *(som) vi oppretta*
 fund-DEF SOM we established
 ‘the fund that we established.’

Notice that *som* is obligatory when the correlate for the relative clause is the subject and that it is optional otherwise. This is so only in restrictive relatives clauses, however—in non-restrictive relatives *som* must be present. As a relative complementizer one may say that *som* corresponds to English (complementizer) ‘that’.

The examples in (42) illustrate the use of *som* as a predicate marker in essive small clauses.

- (42) a. *Jon jobba der fire år som lærar.* St. Norwegian
 Jon worked therefour year as teacher
 ‘Jon worked there for four year as a teacher.’
 b. *Vi bør betrakta problemet som løyst.*
 we should consider problem- DEF as solved
 ‘We should consider the problem solved.’

Again the English equivalent of *som* is ‘as’.

Some cleft sentences illustrating the use of *som* in that context are given in (43).

- (43) a. *Det var Ola *(som) utløyste brannalarmen.* St. Norwegian
 it was Ola SOM triggered fire-alarm-DEF
 ‘It was Ola who triggered the fire alarm.’
 b. *Det var brannalarmen (?*som) Ola utløyste.*
 it was fire-alarm-DEF SOM Ola triggered
 ‘It was the fire alarm that Ola triggered.’
 c. *Det var då (*som)/(?at) Ola forstod at han hadde dumma seg ut.*
 it was then SOM/that Ola realized that he had dumber REFL out
 ‘It was then Ola realized that he had made a fool out of himself.’

Notice that *som* cannot be used when the clefted constituent has a temporal denotation: in such cases the declarative complementizer *at* can optionally be used.⁸

The use of *som* in embedded *wh*-questions has been exemplified earlier, but we give some additional examples in (44) (all examples still Standard Norwegian).

- (44) a. *Petter undersøkte kven *(som) hadde gitt Ola fyrstikken.*
 Petter inquired who SOM had given Ola match-DEF
 ‘Petter inquired about who had given Ola the match.’
 b. *Petter undersøkte kven (*som) Jon hadde gitt fyrstikken.*
 Petter inquired who SOM Jon had given match-DEF
 ‘Petter inquired about who Jon had given the match.’
 c. *Petter undersøkte kva (*som) Jon hadde gitt Ola.*
 Petter inquired what SOM Jon had given Ola.
 ‘Petter inquired about what Jon had given Ola.’
 d. *Petter undersøkte kvifor (*som) Jon hadde gitt Ola fyrstikken.*
 Petter inquired why SOM Jon had given Ola match-DEF
 ‘Petter inquired about why Jon had given Ola the match.’

⁸ In this respect Swedish differs from Norwegian. As is fairly well-known, Swedish uses *som* also when a temporal constituent is clefted.

(i) *Det var då som Olof förstod vad som hänt.*
 it was then that Olof understood what SOM happened

According to these examples *som* is obligatory in embedded subject *wh*-questions and impossible otherwise. Not all speakers of Norwegian seem to agree on this confinement of *som* to subject *wh*-questions only. Åfarli & Eide (2003:257f) claim that *som* can be used if a non-subject *wh*-constituent is a complex DP, and they give the following example to illustrate this.

- (45) *Eg veit kva mann (?som) du burde treffe.*
 I know whatman SOM you should meet
 ‘I know which man you should meet.’

Moreover, in a search in the Oslo Corpus of Tagged Norwegian Texts I found the following four examples of embedded non-subject *wh*-questions with *som* present after the *wh*-constituent. Personally I find all the examples ungrammatical, but other speakers I have consulted accept them.

- (46) a. *Når påtalemyndigheten ber retten foreta oppnevning av sakkyndige, skal den skriftlig angi hva som i tilfelle ønskes undersøkt og **hva som det** ønskes uttalelse om.*
 ‘When the authorities ask the court to appoint experts, it shall specify in writing what should be investigated and *what SOM there* is-wanted a statement about.’
- b. *... medmindre skadelidte mottar ytelser fra ordninger som skal gi en bestemt brutto uansett **hva som folketrygden** yter.*
 ‘... unless the offended receives benefits from arrangements that should give a certain gross number no matter *what SOM the public insurance* extends.’
- c. *... tar han opp nasjonens problemer og legger ut om sine forslag til løsninger, eller skryter av **hva som hans regjering** angivelig alt har oppnådd.*
 ‘... he brings up the problems of the nation and outlines his suggestions for solutions or brags about *what SOM his government* purportedly already has achieved.’
- d. *Nå får dei vite **kva som de** har bestemt to månader etterpå.*
 ‘Now they will know *what SOM you-PL* have decided two months later.’

Notice incidently that these examples from the Oslo Corpus all involve simple *wh*-constituents, thus not conforming to Åfarli & Eide's claim.⁹

The most important fact, which seems to be clearcut enough, is nevertheless that *som* is obligatory in embedded subject *wh*-questions. What the situation is for the optionality of *som* in the non-subject cases is yet another issue that deserves to be further investigated: there clearly seems to be some variation among speakers, and possibly also across dialects.

The final use of *som* that we listed above is the use in exclamatives.¹⁰ Some examples are given in (47).

- (47) a. *Så mange mennesker *(som) spring rundt her, då!*
 so many people SOM runs around here then
 ‘How many people there are running around here!’
 b. *Så mange mennesker (som) det var her, då!*
 so many people SOM it was here then
 ‘How many people there are here!’
 c. *Så fin kjole (som) du har kjøpt deg!*
 so nice dress SOM you have bought yourself
 ‘What a nice dress you have bought!’

As we see, *som* is obligatory if the exclamated constituent is the subject and optional otherwise.

⁹ Åfarli & Eide (2003:257) remark that in Swedish *som* is allowed in embedded non-subject *wh*-questions also with simple *wh*-constituents. On the same issue Teleman et al. (1999:556) say that *som* is usually absent from embedded non-subject *wh*-questions in written Swedish. In spoken Swedish on the other hand, they claim that the presence of *som* is subject to regional and contextual variation. As a main rule of thumb they say that *som* is more often present the more complex the *wh*-constituent is (conforming to Åfarli & Eide's claim for Norwegian) and the more complex the (non-*wh*) subject is.

¹⁰ Norwegian exclamatives can be introduced by the degree element *så* ‘so’ as in (47). In some varieties of Norwegian they can also be introduced by *wh*-elements, in particular the *wh* degree element *kor* ‘how’ (which is homophonous to, but has a different etymological origin than, the locative *wh* word *kor* ‘where’). Interestingly, the *wh* option appears to be impossible in central eastern Norwegian dialects and in the Bokmål written standard—it is at least not possible without yielding an archaic flavor. This means that there is a potential geographical correlation between the dialects that allow non-V2 in main clause *wh*-questions and the ones that can let *wh* elements introduce exclamatives. Whether this correlation has a deeper grammatical source remains to be investigated, as do the exact distribution of *wh*-exclamatives in Norwegian dialects.

All in all *som* is fairly robustly subject related. Setting the “*as*-uses” of *som* aside (comparative complementizer, essive predicate marker) we see that *som* is always obligatorily present in the context of subject extraction/relativization, i.e. in its “*that*-uses”. In non-subject contexts *som* is either optional or impossible: it is optional in non-subject relatives and non-subject exclamatives, but less felicitous in non-subject clefts and embedded *wh*-questions. Further studies of the optionality in non-subject contexts are most welcome as there appears to be variation among speakers with respect to that issue.

After this exposition of the overall use of the element *som* we will now return to the main topic of the paper and present an analysis of the dialectal variation pertaining to lack of V2 in main clause *wh*-questions.

6. Analysis

6.1 Preliminaries

The analysis to be presented here will largely build on the analysis developed in Westergaard & Vangsnes (forthcoming). A first crucial assumption is that of a split CP along the lines of Rizzi (1997, 2001), Benincà (2001), Benincà & Poletto (forthcoming), however with certain modifications. The specific split-CP that we will assume is sketched in (48).

(48) [_{CP} Int(errogative) Top(ic) Foc(us) Wh Fin(iteness) [_{IP} T

The main (micro)parametric tool of the analysis will furthermore be a parameterized requirement for lexicalized heads. V-to-C movement, and thus V2, will be a way to meet this requirement for left peripheral heads (at least in the standard cases). The notation for the requirement will be [X°_{EPP}] where X is a variable referring to the various heads.

As an immediate illustration of this parametric tool, consider table 10 which indicates the assumed properties of the left peripheral heads in Standard Norwegian, the Tromsø dialect, the Nordmøre dialect, and English.

	Int	Top	Foc	Wh	Fin
St. Norwegian	[Int [°] _{EPP}]	[Top [°] _{EPP}]			
Tromsø	[Int [°] _{EPP}]	[Top [°] _{EPP}]			
Nordmøre		[Top [°] _{EPP}]			
English	[Int [°] _{EPP}]				

Table 10: *Requirements on C[°] heads in three Norwegian dialects and English*

In standard Norwegian there is a requirement for lexicalization of both the Int[°] and the Top[°] heads, and there is therefore V2 both in (*wh*) interrogatives and in declaratives: a necessary condition for this line of reasoning to hold, is to say that all declaratives have a default topic, and that the subject fills the topic position (Spec-TopP) if no other constituent does.

In the Nordmøre dialect, on the other hand, there is a requirement for lexicalization of Top[°] but not of Int[°]. This is the reason why this dialect has V2 in declaratives but not (necessarily) in *wh*-questions.

In English Int[°] requires lexicalization whereas Top[°] does not, and this is the reason why V2 is encountered in *wh*-questions but not in declaratives.

The question then is the Tromsø dialect which has the same settings as Standard Norwegian. Earlier we have seen that the Tromsø dialect is one where we do in fact find lack of V2 in *wh*-questions. This leads to another main assumption for our account of the cross-dialectal variation in Norwegian: I will assume that simple/short *wh*-elements are non-projecting, clitic in some dialects and projecting, non-clitic in other dialects. The non-projecting, clitic *wh*-elements have the capacity to meet the lexicalization requirement on Int[°], and this will essentially speaking be the account of why some dialects allow lack of V2 with short *wh*-elements only. A similar idea was entertained in Taraldsen (1986a).

A third and similar assumption concerns variation pertaining to the element *som*. I will assume that this element is a head (X[°]) in some dialects and a specifier (XP) in other dialects. This means that *som* can meet the lexicalization requirement on Left Peripheral heads in some dialects (where it is an X[°]) but not in others (where it is an XP). In turn this will account for why the complexity constraint can be “lifted” in subject *wh*-questions with non-V2, as for example in the Tromsø dialect.

I will shortly discuss these assumptions more thoroughly. Let me first point out that there will be just three factors that govern the variation across the Norwegian dialects: (i) a “parameter” pertaining to the properties of Int° , (ii) a lexical difference pertaining to some *wh*-elements, (iii) a lexical difference pertaining to *som*.

Before we proceed it is important to stress that the Int° head assumed here is similar, but not identical, to the Force° head in Rizzi (1997). In fact, I will take clause typing more generally not to be an effect of properties of a single head in the Left Periphery, but rather to follow from the total composition of the whole CP domain: the presence of Int° yields a clause with interrogative force, its absence does not, and instead we get a declarative clause, provided that Top° is present.

Moreover, the root/embedded asymmetry in Norwegian *wh*-questions (and Mainland Scandinavian in general) with respect to verb movement will fall out on the assumption that Int° is not present in embedded *wh*-questions. The assumption is rooted in the view that embedded *wh*-questions are not, illocutionary speaking, real questions, but rather reports of (real) questions. (See Wechsler 1991 for an enlightened discussion of this in view of (embedded) Verb Second in Swedish in particular and Germanic in general.)

Notice in this respect the presence of the left peripheral head *Wh*, adapted from Rizzi (2001). This will be the domain in which *wh*-related features other than illocutionary force are processed, and a crucial assumption is that *wh*-constituents move to this domain—and then on to IntP in root clauses.

Let us then return to the claim about the varying status of short *wh*-elements, i.e. as $\pm$ clitic.

6.2 The complexity issue on *wh*-elements

The disyllabic *wh*-elements which do not allow non-V2 in the Tromsø dialect clearly have a complex morphological structure. At least diachronically, and maybe even synchronically, they may be analyzed as indicated in (49).

- | | | | | | | |
|---------|---------------|--------|---|-------------------------------|------------------|--------|
| (49) a. | <i>korsn</i> | ‘how’ | □ | <i>kor</i> + <i>-leis(en)</i> | ‘how + ADV.SUFF’ | Tromsø |
| b. | <i>koffør</i> | ‘why’ | □ | <i>kor</i> + <i>før</i> | ‘how + for’ | |
| c. | <i>katti</i> | ‘when’ | □ | <i>ka</i> + <i>tid</i> | ‘what + time’ | |

As for the short *wh*-elements no morphological complexity can be discerned, but we know for certain that these elements stem from lexical items that were inflected in earlier stages of Norwegian. The pronouns *ka* ‘what’ and *kem* ‘who’ stem from Old Norse *hva-t* and *hver-r*, respectively, which both were inflected for number and case (and gender)—the indicated forms are the nominative singular ones. The Old Norse equivalent of the locative *wh*-adverb *kor* on the other hand was equipped with directional affixes which were since lost: *hvar* ‘where’, *hva-ðan* ‘where-from’, *hvar-t* ‘where-to’.

The idea then is that the loss of case, number, and directional affixes on the (now) simple *wh*-words made it possible to reanalyze them from being projecting (morphosyntactically complex) to being non-projecting (morphosyntactically simplex). There are various ways to be more specific about this, and we may cast the idea in terms of Cardinaletti & Starke’s (1997) tripartite distinction between strong, weak, and clitic forms: some Norwegian dialects have clitic short *wh*-elements that can occupy X^0 positions, others do not. All Norwegian dialects will be taken to have weak short *wh*-elements (which occupy specifier positions, although they have a deficient internal structure) as well as strong ones (which occupy specifier positions but have a fuller internal structure). We can assume that stressed short *wh*-elements belong to the ‘strong’ category, and in that respect it is important to notice that stressed short *wh*-elements are only compatible with V2 in the Tromsø dialect. Consider the relevant minimal pair in (50).

- (50) a. *KA sa han Ola?* Tromsø
 what said ART Ola
 b. **KA han Ola sa?*
 what ART Ola said
 ‘WHAT did Ola say?’

6.3 Optional V2, information value, and the licensing of left peripheral heads

Consider then the fact that V2 can be optional alongside non-V2 in the *wh*-questions of some of the dialects, as for instance in the non-subject *wh*-questions of the Tromsø dialect. A set of important facts related to the $\pm$ V2 distinction concerns the relative placement of non-*wh* subjects and sentence

adverbs. In non-subject *wh*-questions with V2 the subject may appear either before or after a sentence adverb as illustrated by the examples in (51) from the Tromsø dialect.

- (51) a. *Ka mente egentli han Ola med det der?* Tromsø
 what meant really ART Ola with that there
 b. *Ka mente han Ola egentli med det der?*
 what meant ART Ola really with that there
 ‘What did Ola really mean by that?’

In cases with non-V2 on the other hand, the subject *must* precede sentence adverbs. This is illustrated in (52).

- (52) a. *Ka han Ola egentli mente med det der?* Tromsø
 what ART Ola really meant with that there
 b. **Ka egentli han Ola mente med det der?*
 what really ART Ola meant with that there
 ‘What did Ola really mean by that?’
 c. **Ka han Ola mente egentli med det der?*
 what ART Ola meant really with that there

This difference between V2 and non-V2 cases can be viewed in the light of combining Westergaard’s (2003a, 2003b) finding about the role of the information value of the subject with Nilsen’s (1997, 2003) general claim that subjects that follow sentence adverbs are focused. Recall that Westergaard found the subjects of V2 (non-subject) *wh*-questions to be (predominantly) informationally new. Given that ‘new information’ involves focalization (in some sense), Westergaard’s finding squares with Nilsen’s claim.

In turn, we can invoke the idea that new/focalized subjects always occupy a syntactic position to the right of (certain) sentence adverbs. In Westergaard & Vangsnes (forthcoming)(henceforth ‘W&V’) the claim is that new subjects occupy the IP-internal position Spec-TP, whereas given subjects occupy the higher Spec-AgrSP position, still within the IP. Here we will follow the basic idea that given and new subjects occupy distinct positions, but we will assume that the higher position is Spec-FinP, i.e. the specifier position of the lowest left peripheral head, and that there is no AgrS head. Moreover, the properties assumed to be encoded by AgrS° in

W&V (and derivatively Vangsnes 1999, 2002), notably that of providing an event denotation, are here taken to be properties of Fin° .

We then have the following picture so far: New subjects stay within the IP, given subjects leave it. Given that sentence adverbs (such as ‘really’) mark the boundary between CP and IP, it follows from the current understanding of the facts in (51) and (52) that given subjects *may* occur in V2 constructions whereas new subjects may *not* occur in non-V2 constructions. This conclusion is inevitable given that V2 is the only grammatical option in some dialects and some contexts: the given/new distinction must exist in these dialects and contexts also.

The question is nevertheless why V2 does not necessarily obtain with given subjects and why non-V2 even more strongly correlates with given subjects. I will suggest that this in part pertains to the licensing conditions on Fin° . I follow the general licensing requirement on (functional) heads advocated in Vangsnes (1999, 2001) which in short says that some relevant material must be merged within the projection of a head in order for the head to be *identified* and its properties activated—if identification does not obtain, the semantic properties of the head in question will not be part of the interpretation of the phrase structural object.¹¹ The licensing requirement on Fin° can be met either by the subject (if it moves to Spec-FinP) or by the (finite) verb (if it moves to Fin°). If the subject does not move, as is the case when it is informationally new, the other option must be resorted to, and the verb moves past both the (new) subject in Spec-TP and the sentence adverb (if present) to Fin° .

This then accounts for Westergaard’s (op.cit.) finding that new subjects occur with V2 in non-subject *wh*-questions with short *wh*-elements. The verb is not attracted by Int° , but rather by Fin° , and this yields V2 surface word order. Conversely, given subjects will move to Spec-FinP and render V-to-Fin movement unnecessary. Since Int° will be lexicalized by the short *wh*-element this head does not attract the verb either and the result is non-V2 surface word order. The account so far can be summarized as in

¹¹ Notice that the notion of *identification* is different from the parameterized lexicalization requirement that we have postulated as our main microparametric tool. The former is an omnipresent visibility condition, whereas the latter bears more resemblance to the poorly understood (and often arbitrary) notion of ‘checking’ as the labelling indicates.

(53)—notice that we assume the finite verb to occupy a position at least as high as T(ense).

- (53) a. $[_{CP} \text{ } Wh_{cl}\text{-Int}^\circ \dots \text{ } Subj_{given} \text{ } Fin^\circ \text{ } [_{IP} \text{ } Adv \text{ } t \text{ } V_{fin}\text{-T}^\circ \dots$
 b. $[_{CP} \text{ } Wh_{cl}\text{-Int}^\circ \dots \text{ } V_{fin}\text{-Fin}^\circ \text{ } [_{IP} \text{ } Adv \text{ } Subj_{new} \text{ } t\text{-T}^\circ \dots$

We are still left with the case in (51b), though, i.e. V2 with a given subject. There could be several answers to why this is a viable surface order. One possibility is that the *wh*-element in this case is the weak form rather than the clitic, in which case verb movement takes place to lexicalize Int° . Another solution is that the verb is attracted by some other left peripheral head than Int° , for instance Foc° .

The latter solution is entertained in W&V and does in fact play a more general role there in explaining the V2 option. W&V argue that the verb always is attracted to Foc° when the sentence contain a new subject, and that this applies in order to identify Foc° : Foc° hosts a focus operator in its specifier position, but this phonetically empty operator does not serve to identify the projection, and V-to- Foc is therefore called for. Above we gave a different account for the V2 option by reference to the identification of Fin° , but a way of combining the two ideas could be to say that the given/new distinction is not a subvariety of focus after all. What new and focused subjects have in common is that they occupy the IP-internal Spec-TP position, but only *focused* subjects relate to the left peripheral Foc° head. When this head is present V-to- Foc is indeed called for and crucially so also when something else than the subject is focalized. The latter point entails that V-to- Foc , and V2 surface order, may obtain even in the context of a given subject (in Spec- $FinP$), provided that something else is focalized.

In other words, V2 obtains in the context of new (unfocused) subjects as a result of V-to- Fin movement, triggered by the licensing requirement on Fin° , and V2 may obtain in the context of a given subject as a result of V-to- Foc movement, triggered by the licensing requirement on Foc° . The latter situation is sketched in (54).

- (54) $[_{CP} \text{ } Wh_{cl}\text{-Int}^\circ \dots \text{ } Op \text{ } V_{fin}\text{-Foc}^\circ \text{ } Subj_{given} \text{ } t\text{-Fin}^\circ \text{ } [_{IP} \text{ } Adv \text{ } t \text{ } t\text{-T}^\circ \dots$

Needless to say, these issues deserve further consideration, but we will leave them for now.

So far we have only taken into consideration non-subject *wh*-questions. Let us now turn to subject *wh*-questions and thus issues that involve the element *som*.

6.4 Subject *wh*-questions and the status of *som*

We recall from previous sections that the presence of *som* is obligatory in embedded subject *wh*-questions in Norwegian, as well as in main clause subject *wh*-questions in the Tromsø, Nordmøre, and Hordaland dialects. In general, the presence of *som* is always what instantiates non-V2 in root subject *wh*-questions in the Norwegian dialects.

The account for the obligatory insertion of *som* will crucially hinge on the following assumption: *wh*-subjects move from Spec-TP directly to Spec-WhP (and on to Spec-IntP/Int°)(cf. the structure in (48)), i.e. not via Spec-FinP.¹² The claim is then that in such a situation *som* will be inserted within FinP in order to identify (license) Fin°.

The obvious question then is why the verb does not move to Fin° instead of *som* being merged. The answer is that *som* is an expletive or, in the sense of Vangsnes (1999, 2001), a ‘functional element’. Such elements have the privileged status of being accessible for postlexical insertion, and on the further assumption that merge is always preferred over move and (remerge) whenever possible, merger of *som* is preferred over moving the verb.¹³

In other words, Fin° triggers the merger of *som* rather than movement of the finite/tensed verb in the case where the subject skips Spec-FinP. However, in the case where the verb is attracted to the left periphery by some higher head, for example Int° or Foc°, we take it that the verb must headmove cyclically through Fin°, thereby obviating the need for merger of *som*. Notice that this accounts for the root/embedded asymmetry in Standard Norwegian on the assumption that Int° is not present in embedded *wh*-questions (cf. section 6.1). In root contexts the attraction of the verb by Int°

¹² At a certain level of abstraction a similar view is taken by Holmberg & Hróarsdóttir (2003): in their C-T-v-V system they claim that (Icelandic) *wh*-subjects move directly from Spec-v to Spec-C.

¹³ A necessary condition in this respect is that functional elements are not part of the numeration, contrary to Chomsky’s (2001) proposal. In other words, economy does not evaluate two distinct numerations but rather two (hypothetical) derivations, one including the postlexical insertion of the functional element.

will give identification of *Fin*^o as a by-product, but in embedded contexts there will be no attraction of the verb and *som* must be merged.

One may ask on what grounds one can argue that *som* has the special status of being a functional element/expletive. One piece of evidence comes from Danish where *som* in fact can be replaced by the regular expletive *der* in (embedded) *wh*-questions and relatives. This is pointed out by Taraldsen (1986b:151ff) who gives the following pair of examples to illustrate the parallel use of *som* and *der*.

- (55) a. *Vi ved hvem *(der) taler med Margrethe.* Danish
 we know who there talks with Margrethe
 b. *Vi ved hvem *(som) taler med Margrethe.*
 we know who SOM talks with Margrethe

Importantly, *der* can only replace *som* when the *wh*-constituent is the subject. This is best shown for relative clauses where *som* is optional (as in Norwegian) when the relativized constituent is a non-subject.

- (56) a. *manden *(som/der) Margrethe taler med* Danish
 man-DEF SOM/there Margrethe talks with
 b. *manden (som/*der) Margrethe taler med*
 man-DEF SOM/there Margrethe talks with

In a way this ties in with the very strong relation between *som* and extracted/relativized subjects that we have seen earlier, in particular in section 5. As we recall *som* must always appear in such cases. In cases where non-subjects are extracted or relativized *som* is either only optionally present or not licit at all.

A further assumption about the properties of *som* is that it is anaphoric. In other words it is an *anaphoric expletive*. This is in fact reminiscent of how Taraldsen (1986b) treats the element. It then follows that *som* can only be merged in cases of extraction/relativization where a constituent within the left periphery (i.e. its local domain) c-commands it.

At this point we can present the announced assumption about the varying head/specifier status of *som* across the Norwegian dialects. The general point of this idea is that in the dialects where *som* is a head, it will be able to meet the lexicalization requirement on *Int*^o whereas it will not in the dialects where it is a specifier. More specifically, this will account for

why *som* is obligatory in main clause subject *wh*-questions in some dialects, e.g. the Tromsø and Hordaland dialects: Int° can target *som* rather than the verb as a constituent to meet the $[\text{Int}^\circ_{\text{EPP}}]$ feature.

In the dialects where *som* is a specifier, on the other hand, it cannot meet the lexicalization requirement on Int° . In such dialects main clause *som*-insertion can only occur if the dialect otherwise possesses clitic short *wh*-elements, and thus exhibit non-V2 in subject *wh*-questions with a complexity constraint, such as the Trøndelag and Rogaland dialects.¹⁴

The facts that the etymological origin of *som* is an adjective ('same') and that it can be replaced (in certain contexts) by an ordinary expletive in Danish suggest that the idea of treating *som* as a specifier is not far-fetched.¹⁵ On the other hand the lightness, both structural and semantic, of this element suggests that it is a head. The current proposal relates to both sets of arguments: the phrase structural status of *som* simply varies across the dialects, and with noticable effects. In the next sub-section we will see how this varying status of *som* interacts with the other proposed variables so as to explain the cross-dialectal variation sketched in section 4.

6.5 Pulling things together

There are now basically three grammatical variables that, by hypothesis, govern the variation that we see across the Norwegian dialects with respect to word order in main clause *wh*-questions. These are:

¹⁴ A consequence of this line of reasoning is that *som* must be a specifier in those Mainland Scandinavian varieties which do not allow non-V2 in root *wh*-questions whatsoever, thus in all of the standard Mainland Scandinavian varieties. This is not compatible with the analysis of *som* (in relative clauses) advocated by Platzack (1998, 1999). Platzack explicitly argues that Mainland Scandinavian *som* is a head (complementizer) and that the corresponding Icelandic element *sem* a specifier, a claim which among other things explains why long extraction from an embedded relative clause is possible in Mainland Scandinavian but not in Icelandic and other Germanic and Romance languages more generally (i.e. why Mainland Scandinavian does not obey the Complex NP Constraint). Hopefully, further studies may find a way of reworking Platzack's results in a way which allows *som* to be analyzed as a specifier also in Mainland Scandinavian.

¹⁵ Another fact which indirectly supports the view that *som* can be a specifier concerns the Icelandic equivalent *sem*. As pointed out by Platzack (1999) Icelandic *sem* can be followed by the general complementizer *að* 'that'.

(i) *Jón hitti konuna sem að þú þekkir.* Icelandic
 Jón met woman-DEF SOM that you know

- (57) (i) The properties of Int° : whether it must be lexicalized or not.
 (ii) The properties of *som*: whether it is an X° or an XP.
 (iii) The properties of simple *wh*-elements: whether they are clitic or not.

These variables are all strictly speaking lexical, but the first is special in that it concerns a(n abstract) functional head rather than a lexical item or a lexical class. The first variable thus comes closest of the three to what we may call a ‘syntactic parameter’. Let us now go through each of the dialects presented in section 4 and see how the syntactic properties of their *wh*-questions fall out from the proposed analysis.

In the Tromsø dialect the variables will have the properties given in table 11.

$[\text{Int}^\circ_{\text{EPP}}]$	$\text{som} = X^\circ$	short <i>wh</i> = cl°
+	+	+

Table 11: *Lexical properties of the Tromsø dialect*

There is a lexicalization requirement on Int° , but non-V2 always arises in subject *wh*-questions since *som* is a head: *som* may meet the lexicalization requirement and obviate V-to- Int movement. In fact, *som* will be able to lexicalize any left peripheral head, and V2 (= verb movement) should therefore not be possible in any context where Fin° fails to be identified by the subject.

Non-V2 may also arise in non-subject *wh*-questions in the Tromsø dialect, but only with the short *wh*-elements, which are clitic and therefore may lexicalize Int° . In subject *wh*-questions on the other hand there is no complexity constraint: Int° can be lexicalized by *som* and any *wh*-constituent may move into Spec- IntP .

In the Nordmøre dialect the variables have the properties in table 12.

$[\text{Int}^\circ_{\text{EPP}}]$	$\text{som} = X^\circ$	short <i>wh</i> = cl°
–	+	?

Table 12: *Lexical properties of the Nordmøre dialect*

In this dialect there is no lexicalization requirement on Int° , and therefore non-V2 obtains without any complexity constraint on the *wh*-constituent.

Moreover, V2 is not possible in subject *wh*-questions since *som* is a head and therefore may lexicalize/identify any left peripheral head, barring verb movement. V2 may optionally arise in non-subject *wh*-questions however, as a reflex of V-to-Foc movement. As for the status of the short *wh*-elements it is not possible to decide whether they are clitic or not since the effect will be overridden by the other relevant properties of the dialect.

In the Trøndelag and Rogaland dialects, which we recall are basically alike with respect to the syntax of *wh*-questions, the grammatical variables have the properties in table 13.

$[\text{Int}^{\circ}_{\text{EPP}}]$	$\text{som} = \text{X}^{\circ}$	short <i>wh</i> = cl°
+	–	+

Table 13: *Lexical properties of the Trøndelag and Rogaland dialects*

There is a lexicalization requirement on Int° , and since *som* is a specifier and thus is not able to lexicalize Int° , non-V2 can only obtain with the short *wh*-elements, which are clitic elements in these dialects. At the same time V2 may optionally obtain in both subject and non-subject *wh*-questions since *som* cannot lexicalize/identify other left peripheral heads either.

In eastern Norwegian dialects, as well as in Standard Norwegian, the variables have the properties in table 14.

$[\text{Int}^{\circ}_{\text{EPP}}]$	$\text{som} = \text{X}^{\circ}$	short <i>wh</i> = cl°
+	–	–

Table 14: *Lexical properties of eastern Norwegian (and Standard Norwegian)*

The most difficult kind of situation to account for is the kind of “complete separation” that has been reported for the Hordaland and Nord-Østerdalen dialects: as we recall the Hordaland dialect allows non-V2 in subject *wh*-questions only whereas the Nord-Østerdalen dialect allegedly only allows non-V2 in non-subject *wh*-questions.

Recall that there were some uncertainty with respect to the complexity constraint for the Hordaland dialect. The proposed theory can however only account for a situation where the Hordaland dialect does not have a complexity constraint (in the subject *wh* cases). The reason is the following. If the Hordaland dialect only allows non-V2 in subject *wh*-questions, this must mean (i) that *som* is a head and (ii) that short *wh*-elements are not

clitics. If the short *wh*-elements were clitics we would of course expect non-V2 to obtain also in non-subject questions, contrary to facts.¹⁶ On the other hand, given that *som* lexicalizes Int° (since it is a head) also complex *wh*-constituents should be able to precede *som* (i.e. in Spec-IntP).

I will therefore conjecture that the uncertainty in the data stems from the Hordaland dialect being in a transition phase, targetting the properties in table 15. These properties entail that non-V2 obtains in subject *wh*-questions, and obligatorily so, and with no complexity constraint on the *wh*-constituent.

$[\text{Int}^\circ_{\text{EPP}}]$	$\text{som} = \text{X}^\circ$	short <i>wh</i> = cl°
+	+	–

Table 15: (*Targetted*) lexical properties of the Hordaland dialect

The variable concerning *som* in effect relates directly to subject *wh*-questions, but no variable that relate directly to non-subject *wh*-questions has been offered. Accordingly, the present theoretical approach cannot account for the reported properties of the Nord-Østerdalen dialect, i.e a situation where non-V2 is allowed in non-subject *wh*-questions only. If such cases in the dialect are due to the clitic nature of the short *wh*-elements, there is no reason to expect that the short *wh*-elements should not also be able to lexicalize Int° and in turn be followed by *som*, which then identifies Fin° (as it always does in embedded subject *wh*-questions). Recall that the information about the Nord-Østerdalen dialect given in Lie (1992) was provided by one single speaker, and not integrated in a study of the kind reported by Nordgård (1985, 1988). We may therefore await further studies of the Nord-Østerdalen dialect before we conclude that it represents a counter-example to the analysis developed here.

Lie does not provide information as to whether the Nord-Østerdalen dialect has a complexity constraint or not. In any event, neither a version with a complexity constraint nor one without can be accounted for by the

¹⁶ Recall that Nordgård (1985) reports that one speaker allows non-V2 in non-subject *wh*-questions with the short locative *wh*-element *kor* ‘where’. This could in fact mean that this short *wh*-element but not the other ones has the status of being a clitic for this individual speaker. In turn similar “splits” in the class of short *wh*-elements would then be expected among other speakers (of all dialects). Future investigations should address this issue.

present theory. This immediately reduces the number of predicted varieties from 16 to 14. Also other varieties predicted by the descriptive typology are ruled out by the present theoretical approach. We do for instance not expect any variety to exhibit a complexity constraint in subject non-V2 *wh*-questions and not in non-subject ones: no complexity constraint in the non-subject cases can only obtain if Int^o does not need to be lexicalized, in which case the requirement does not apply to subject *wh*-questions either. Moreover, the theoretical account does not leave open the possibility for obligatory non-V2 in the non-subject *wh*-questions, and varieties with this property have so far not been encountered either. All in all this means that we, according to the current approach, may expect further empirical investigations to reveal at the most 5-6 additional varieties of Norwegian *wh*-grammars to exist.

6.6 *Diachronic speculations*

Setting the Nord-Østerdalen dialect aside we are then left with only five distinct types of "*wh*-grammars" across Norwegian dialects. These may be ordered on a scale as follows with respect to how "liberal" their lack of V2 in main clause *wh*-questions is.

(58) Nordmøre >> Tromsø >> Tr.lag/Rogal. >> Hordal. >> Eastern Nor.

The Nordmøre dialect (and presumably Møre more generally) is the most liberal one since it allows non-V2 in both subject and non-subject *wh*-questions, and it does so without a complexity constraint on the *wh*-constituent. The Tromsø dialect is slightly less liberal in that it exhibits the complexity constraint in non-subject *wh*-questions, though not in subject *wh*-questions. In turn the Trøndelag and Rogaland dialects exhibit the complexity constraint in both subject and non-subject *wh*-constructions. The Hordaland dialect on the other hand only allows non-V2 in subject *wh*-questions. Finally, the eastern (and southeastern) Norwegian dialects do not allow non-V2 at all in *wh*-questions.

The following map shows the geographical locations of the various dialects. Bear in mind that the Tromsø dialect is a city dialect, thus a point, not an area, but presumably the adjacent dialects have similar properties as described here for the Tromsø dialect.



Map 1: Map of Norway with indication of some districts/counties and cities/towns.

A highly interesting question is how the cross-dialectal variation has arisen. Given that obligatory V2 is the most widespread situation in Scandinavian main clause *wh*-questions in general, and given that this was the situation in Old Norse too, it appears well-founded to assume that the properties exhibited by the eastern Norwegian dialects were the original ones in all dialects. Notice that this is furthermore plausible on the assumption that *som* has originated as a specifier (cf. its etymological origin as an adjective).

Starting from a grammar of the eastern Norwegian type we can then envisage the following “chains of events” leading to the current cross-dialectal situation. If *som* is reanalyzed as a head, we will get the Hordaland dialect (as idealized above), i.e. a dialect with non-V2 in subject *wh*-constructions only, and without the complexity constraint. If on the other hand the short *wh*-elements are reanalyzed as clitic (rather than weak) elements, we will get a dialect of the Trøndelag/Rogaland type, i.e. with non-V2 in both subject and non-subject *wh*-questions, but with a complexity constraint in both contexts.

If both of the reanalyses take place we will get the Tromsø dialect, i.e. a dialect type where the complexity constraint only obtains in non-subject *wh*-questions. If on the basis of a Tromsø dialect type we get reanalysis of Int° as not having the $[\text{X}^\circ_{\text{EPP}}]$ feature, we get the Nordmøre dialect type, i.e. where non-V2 is required in subject *wh*-questions and possible in non-subject *wh*-questions, and where there is no complexity constraint. In fact, the reanalysis of Int° can also proceed from a grammar type of the Hordaland dialect—as we noted earlier it is not possible to tell whether the short *wh*-elements of the Nordmøre dialect are clitics since the overt evidence is overridden by the properties of the other variables.¹⁷

These speculations can be summarized as in (59) and (60).

- (59) i. *som* is reanalyzed as $\text{X}^\circ \sqsubset$ ‘Hordaland’
 ii. Int° is reanalyzed as not having the $[\text{X}^\circ_{\text{EPP}}]$ feature $\sqsubset$ ‘Nordmøre’
- (60) i. short *wh* are reanalyzed as clitic $\sqsubset$ ‘Trøndelag/Rogaland’
 ii. *som* is reanalyzed as $\text{X}^\circ \sqsubset$ ‘Tromsø’
 iii. Int° is reanalyzed as not having the $[\text{X}^\circ_{\text{EPP}}]$ feature $\sqsubset$ ‘Nordmøre’

This account predicts one particular unattested dialect type, namely one which is minimally different from the Nordmøre dialect in that *som* is (still) a specifier. In such a dialect type V2 should be possible alongside non-V2 in subject *wh*-questions, for example if the verb is attracted by some medial

¹⁷ Given the geographical adjacency of Nordmøre and Trøndelag, and lack of such between Nordmøre and Hordaland, one might consider it likely that the Nordmøre dialect has gone through a stage where it was like the Trøndelag dialect (before subsequent reanalysis of *som*) and thus that the short *wh*-elements indeed are clitic in this dialect.

left peripheral head, e.g. Foc° . Diachronically speaking we could get this dialect type if the Int° -variable is reanalyzed on the basis of a Trøndelag/Rogaland grammar type, i.e. without an intermediate reanalysis of *som* as a head.

Let us call this so far unattested dialect type ‘Norwegian X’. The question of whether Norwegian X actually exists relates to an even more interesting question: is it possible to change directly from a grammar type of the eastern Norwegian kind to one of the Nordmøre or Norwegian X type? In other words, is it possible to have reanalysis of the properties of Int° (loss of the $[\text{Int}^\circ_{\text{EPP}}]$ feature) without a preceding reanalysis of either *som* or the class of short *wh*-elements? At a more general level this is a question of whether the properties of a(n abstract) functional head can change without being preceded by more low level changes pertaining to the morphosyntax of lexical items or classes.

On purely conceptual grounds we may argue that the change concerning Int° at least is more likely to take place if it is subsequent to changes in the $\pm\text{XP}$ status of *som* and/or short *wh*-elements. The principled reasons should be obvious, and are the following. Reanalysis of either *som* and/or the short *wh*-elements as heads in one generation of individuals yields the presence of non-V2 word order in (some) main clause *wh*-questions in the *primary linguistic data* (‘PLD’) that serve as input for parameter setting in subsequent generations of speakers. As we have noted earlier the variable pertaining to Int° is the only one that can be regarded as a syntactic (micro)parameter, and the line of reasoning here is that whereas the reanalysis of the lexical variables can happen more or less “at will”¹⁸, the re-setting of the parameter is dependent on overt evidence in the PLD. In a situation where the language acquiring child does not detect for instance a complexity constraint in the PLD, the way is open to conclude that there is no lexicalization requirement on Int° in the target grammar.

Evidently, the PLD produced by speakers who have a grammar with *som* = X° will, in principle, provide the strongest “support” for re-setting of the Int° parameter since such a grammar will allow complex *wh*-constituents in subject *wh*-questions. It is thus also interesting to ask whether the reanalysis of *som* as X° by itself is a necessary prerequisite for the re-

¹⁸ The reanalysis of the *wh*-elements must of course be preceded by reduction of morphology as argued for in section 6.2.

setting of the Int° parameter, and that in fact the Nordmøre dialect type cannot evolve directly from the Trøndelag/Rogaland dialect type without passing through a stage where it is like the Tromsø dialect type. If this holds true, Norwegian X should not exist.

If a comprehensive investigation of Norwegian dialects turns out to show that Norwegian X is never encountered we have reasons to assume that the change of the Int° parameter indeed requires the presence of complex *wh*-constituents in root non-V2 *wh*-questions in the PLD (in effect subject questions allowed by *som* being a head). If such an investigation moreover also comprise of dialects that are currently undergoing change in a more “liberal direction” with respect to lack of V2 in *wh*-questions there are all the more reasons to believe in the above.

6.7 Some notes on clefts

Lie (1992) proposes that the non-V2 *wh*-questions in Norwegian dialects have evolved from clefts. As evidence for this he notes the existence of a possible intermediate stage in his own Hedalen dialect. Consider the examples in (61) from Lie (op. cit.:72) which according to him all are used in the dialect without any significant meaning difference.

- (61) a. *Hå e de du si?* Hedalen
 what is it you say
 b. *Hå e du si?*
 what is you say
 c. *Hå du si?*
 whatyou say

The example in (61a) is a clefted *wh*-question. (61c) is a non-V2 *wh*-question of the kind we have encountered numerous times above, i.e. without the characteristics of a cleft, lacking both the finite copula and the expletive. The example in (61b) on the other hand has a finite copula along with the finite main verb, but lacks the expletive. This is then the intermediate stage between a cleft and a non-V2 non-cleft.

Lie also mentions diachronic evidence for the intermediate stage, citing the following example from 1776 Vågå dialect (Gudbrandsdalen, northwestern eastern Norwegian, west of Nord-Østerdalen) as the oldest he has encountered.

- (62) *Naa kaa va du saag?* Vågå (18th century)
 now whatwas you saw
 ‘Now, what did you see?’

Lie furthermore points out a correlation between clefts and non-V2 constructions when it comes to types of *wh*-constituents allowed. In section 2 we saw that the *wh*-adverb *kåfer* ‘why’ in the Nordmøre dialect was somewhat degraded with non-V2 word order. Moreover, in other dialects this and other *wh*-adverbs are not allowed whatsoever because of the complexity constraint. A similar degradedness can be found with *wh*-adverbs in standard Norwegian clefts. Consider the examples in (63).

- (63) a. *Kven er det som bankar på mi dør?* St. Norwegian
 who is it SOM knocks on my door
 b. *Kva er det som plagar deg?*
 what is it SOM bothers you
 c. *Kor er det du bur?*
 where is it you live
 d. *Når er det du står opp om morgonen?*
 when is it you stand up in morning-DEF
 e. *??Korleis er det du vet det?*
 how is it you know that
 f. *??Kvifor var det du sa det?*
 why was it you said that

In addition to providing support for Lie’s suggestion that the non-V2 word order has evolved from a cleft structure, these facts could suggest that the awkwardness/ungrammaticality of non-locative/-temporal *wh*-adverbs is not an effect of complexity but rather of semantics/pragmatics—their meaning is “uncleftable”, as it were.

Although this is a relevant point to keep in mind there should be little doubt that there exists such a thing as the complexity constraint in some of the dialects. The following complex argumental *wh*-constituent is fine in a Standard Norwegian cleft, but could nevertheless not appear in a (non-subject) non-V2 structure in varieties like the Tromsø dialect.

- (64) *Kor mange studentar er det du vanlegvis underviser?* St. Nor.
 how many students is it you usually teach
 ‘How many students is it you usually teach?’

Another problem with the cleft hypothesis is that in the non-V2 dialects also clefts may exhibit non-V2. In other words, both of the structures in (65) are fine in for example the Tromø dialect.

- (65) a. *Kem e det du skal møte?* Tromsø
 who is it you shall meet
 b. *Kem det e du skal møte?*
 who it is you shall meet

Lie recognizes this problem and his answer to it is straightforward enough: once the non-V2 structure has been established as an independent construction, the rule that governs it can also affect other structures, including clefts.

On a general note it is not immediately clear how Lie’s suggestion about a cleft origin for non-V2 *wh*-questions can be brought together with the speculations that I have put forward here—at first sight the two lines of reasoning appear incompatible. Even if my suggestions concerning the development of *som* and short *wh*-elements are taken as guidelines for future research, the cleft hypothesis is worth keeping in mind, and it would be especially interesting to investigate the status of the “intermediate” structure more closely (both from a grammatical and a geographical point of view).¹⁹

6.8 Questions for future research

All in all there are a number of interesting challenges for future research on the dialectal variation pertaining to lack of V2 in *wh*-questions across Norwegian dialects. The most important challenge of all is to perform a

¹⁹ Another general point concerning the status of clefts pertains to methodology and the use of questionnaires (either written/oral). Given that construing a cleft alternative is very often an option, one should choose the example sentences with non-V2 carefully: the speaker may dismiss an example simply because the cleft version is felt as being more felicitous. A (partial) way out is make sure that $\pm$ V2 is tested (also) by the use of cleft sentences.

comprehensive and methodologically sound study of a substantial number of Norwegian dialects, preferably evenly distributed across the country. Such a study is called for both to control and adjust the picture that has been presented in this survey and to search for additional, distinct varieties. Although the existing literature provides certain pointers as to the distribution of various varieties we are far from a good understanding of where the *isosyntagmas* go and moreover if there are subvarieties within the areas that have been identified by for example Nordgård (1985, 1988).

What ‘methodologically sound’ means for the study of dialect syntactic variation is a subject of debate itself. I will not specify here how the design of a future and comprehensive study of this particular phenomenon should be, but it is quite clear that none of the studies referred to in this survey meets the methodological standards of contemporary dialect syntax investigations such as the Italian ASIS-project or the Dutch SAND-project. See Cornips & Poletto (2004) for some relevant discussion of methodology for dialect syntactic investigations, including references to these two projects.

The present paper has developed and pursued a certain theoretical approach which in turn has posed several questions for future research. Let me mention two of the more particular ones which pertains to individual dialects:

- *What are the properties of the Nord-Østerdalen dialect?*
The present analysis is not compatible with how it is reported to be in Lie (1992), i.e. with non-V2 only in non-subject *wh*-questions.
- *What are the properties of the Hordaland dialect?*
The present analysis predicts that the uncertainty reported in Nordgård (1985, 1988) should be resolved in certain ways: the dialect, or at least a variety of it, should have non-V2 in subject *wh*-questions only, and without a complexity constraint on the *wh*-subject.

Another, more general issue, concerns the nature of the complexity constraint. In this paper we have argued for and assumed that the constraint is sensitive to the morphosyntactic rather than phonological status of the *wh*-element. We have however seen that on the basis of reports in Nordgård (1985, 1988) we can establish a few quirks that indirectly make reference to a phonological side of the complexity constraint. One quirk concerns our

idealization of the Rogaland dialect: we noted that the relative acceptance of bare *wh*-adverbs with non-V2 might be due to the fact that *koss* ‘how’ is monosyllabic in the Rogaland dialect whereas the corresponding form in other dialects most often is disyllabic (*korsn*, *kåles*, etc.). This was the basis for our claim that the Rogaland dialect is of the same type as the Trøndelag dialect. In turn, if correct, this at least suggests that phonology plays an important role with respect to allowing reanalysis of *wh*-elements as clitic.

On this particular issue further investigations may reveal that for instance a $\pm$ adjunct (or $\pm$ argument) variable is also relevant for the variation. This is indeed what Nordgård (op.cit.) suggests, but interestingly it is only with respect to the somewhat unclear data from the Rogaland dialect that it makes a difference in his study. Remember in this respect that Nordgård had only two informants from Rogaland. It therefore seems well-founded to await further studies before we adjust the general theoretical approach followed here. In any event, future studies should of course make it possible to decide whether a $\pm$ adjunct/argument variable plays a role for the variation.

A related issue concerns whether or not the class of short *wh*-elements is constant. Recall that Nordgård (1985:17) reports that in the Hordaland dialect one informant (out of the three) allows non-V2 in non-subject *wh*-questions with the short locative *wh*-element *kor* ‘where’. This could in fact mean that this short *wh*-element but not the other ones has the status of being a clitic for this individual speaker. In turn similar “splits” in the class of short *wh*-elements would then be expected among other speakers (of all dialects). Future thorough investigations should also control for this issue.

It is not immediately clear how the diachronic speculations in the previous section may be addressed from an empirical point of view. Some historical dialect material from the last hundred years or so does exist, but we may expect the frequency of main clause *wh*-questions to be quite low in this material (as in any free speech material). On the other hand it should be possible to observe the principles of the proposed evolution by looking at contemporary dialects that are undergoing change, manifested by age differences. A dialect (area) which seems to be undergoing a change towards a more “radical” non-V2 grammar (of the Nordmøre type) is the one spoken in Sogn, a district in western Norway to the immediate north of Hordaland and south of Møre (including Nordmøre). This dialect is currently being investigated by the present author, and the present survey is

in effect a result of the preparation for that investigation. Hopefully the investigation can shed some light on the mechanisms governing this particular language change in Norwegian and in turn language change more widely.

In the next section we will now return to syntactic theory and point out some general consequences of the theoretical assumptions that underlie the analysis.

7. Clause typing and the status of complementizers

We noted in section 6.1 that a crucial point in the analysis is the assumption that Int° is not present in embedded *wh*-questions. This is a necessary condition for explaining the root/embedded distinction in Standard Norwegian *wh*-questions: there is V2 in root questions, but not in embedded ones. The examples given to illustrate this asymmetry is repeated here.

- (6) a. *Kva sa han egentleg?* Standard Norwegian
 what said he actually
 b. **Kva han egentleg sa?*
 what he actually said
 c. **Kva han sa egentleg?*
 what he said actually
 d. **Kva egentleg sa han?*
 e. **Kva egentleg han sa?*

- (7) a. *Eg lurer på kva han egentleg sa.* St. Norwegian
 I wonder on what he actually said
 ‘I wonder what he actually said.’
 b. **Eg lurer på kva sa han egentleg.*
 I wonder on what said he actually
 c. **Eg lurer på kva han sa egentleg.*
 I wonder on what he said actually

The rationale given for the assumption was that the Int° head encodes interrogative force, and since embedded questions are not questions in and of themselves (i.e. “(explicit) requests for an answer”), but rather reported questions, it appears well-founded to assume absence of Int° .

Along the same lines we may note that the general assumption about a parameterized lexicalization requirement on left peripheral heads captures ‘embedded V2 topicalization’ (cf. Vikner 1995) and the difference between Norwegian/Scandinavian and English in this respect. In Norwegian/Scandinavian Top° is equipped with the $[\text{Top}^\circ_{\text{EPP}}]$ feature whereas in English it is not, and only Norwegian/Scandinavian has embedded V2 topicalization. Compare the following examples

- (66) *Ole fortalte oss at ...* (Standard) Norwegian
 Ole told us that
 a. ... *sønnen hans skulle gifte seg i morgen.*
 son-DEF his would marry REFL in to-morrow
 b. *i morgen skulle sønnen hans gifte seg.*
 in to-morrow would son-DEF his marry REFL
 c. **i morgen sønnen hans skulle gifte seg.*
 in to-morrow son-DEF his would marry REFL
- (67) Oliver told us that ... English
 a. his son was getting married tomorrow.
 b. ?tomorrow his son was getting married
 c. *tomorrow was his son getting married.

At this point a problem concerning clause typing arises. If Int° partially equals Rizzi’s (1997) Force° and there is no $\text{Int}^\circ/\text{Force}^\circ$ in embedded clauses since embedded clauses in general does not have an illocutionary force of their own, what is then the general status of complementizers? From a purely syntactic point of view this problem is highlighted by the fact that *at/that* precedes the embedded topic in (62) and (63): what is the position of the complementizer if not for example Force° ?

A tentative solution to this is to say that embedded clauses are closed off by a prepositional/Case-related layer of functional heads which establishes a relation to the matrix clause and determines a ‘force’ indirectly through this relation. This would in effect mean that embedded clauses are either PPs (non-argumental/adverbial) or KPs (argumental/nominal), and we would hold that complementizers license these P/K heads—some are merged as heads, others as specifiers (cf. *if* vs. *whether* and extraction facts).

Such a solution sheds light on the many correlations, synchronic as well as etymological, between complementizers ~ prepositions ~ determiners (cf. English *that, for, to*). Moreover, *wh*-elements parttake in the licensing of P/K heads. This is why they (typically) move to the left of embedded clauses and this is why they can turn into complementizers (cf. Romance *que/che*, Norwegian *hvis* ‘if’ (from ‘whose’)).

8. Conclusion

The main purpose of this paper has been to present an overview of what is known about the geographical variation pertaining to lack of V2 in main clause *wh*-questions across Norwegian dialects. On the basis of this overview I have allowed myself to put forward some fairly detailed speculations to account for the variation both from a synchronic microparametric and from a diachronic point of view. The specific theoretical proposal has been that there are three factors that govern the variation. These grammatical variables are repeated here:

- (57) (i) The properties of *Int*^o: whether it must be lexicalized or not.
- (ii) The properties of *som*: whether it is an *X*^o or an *XP*.
- (iii) The properties of simple *wh*-elements: whether they are clitic or not.

I have shown how these variables can be taken to account for the variation within a coherent grammatical framework.

A more general picture that has emerged from the exposition is that microparametric variation to a large extent can be viewed as an effect of lexical variation, i.e. variation in the inventory of and properties of lexical elements. The last two variables in (57) arguably concerns lexical variation in this sense. The first variable on the other hand is more syntactic in nature in the sense that it concerns a particular part of the clausal structure, not a lexical element as such (although also abstract parts of clausal structure presumably are represented in the mental lexicon). Accordingly we can characterize the first variable as a syntactic parameter whereas the latter two are lexical parameters.

An important possible restriction on a change in the syntactic parameter that was discussed in section 6.6 is that such a change must be pre-

ceeded by changes in at least one of the lexical parameters. The main conceptual argument for this is that changes in the lexical parameters in turn yield the presence of non-V2 structures in the material which subsequently serve as input for the next (or later) generation(s) of speakers—the basic line of reasoning is that changes in the lexical parameters pave the ground for a more radical (truly) syntactic change.

Whether this last speculation as well as all the other speculations presented in this paper can be maintained is strictly speaking an empirical matter. As should be evident from the presentation, although the lack of V2 in main clause *wh*-questions in Norwegian dialects has received considerable attention among both dialectologists and generative linguists, the phenomenon is still far from exhaustively studied. It is my sincere hope that this survey has provided some useful guidelines as to how to proceed with future investigations.

References

- Åfarli, Tor: 1986, 'Some syntactic structures in a dialect of Norwegian', *University of Trondheim Working Papers in Linguistics* 3, 93-111.
- Åfarli, Tor A. & Kristin Mehlum Eide: 2003, *Norsk generativ syntaks*. Novus, Oslo.
- Benincà, Paola: 2001, 'The position of Topic and Focus in the left periphery', in Guglielmo Cinque & Giampaolo Salvi (eds.), *Current Studies in Italian Syntax*, Elsevier, Amsterdam, pp. 39-64.
- Benincà, Paola & Cecilia Poletto (forthcoming): 'Topic, Focus and V2: defining the CP sublayers', to appear in Luigi Rizzi (ed.), *The CP Layer*. Oxford University Press, Oxford/New York.
- Cardinaletti, Anna & Michal Starke: 1999, 'The typology of structural deficiency: A case study of the three classes of pronouns'. In Henk van Riemsdijk (ed): *Clitics in the Languages of Europe*. Mouton de Gruyter, Berlin/New York, pp. 145-233.
- Chomsky, Noam. 2001: 'Derivation by phase', in Michael Kenstowicz (ed.): *Ken Hale: a life in language*. The MIT Press, Cambridge, Massachusetts, pp. 1-52.
- Cornips, Leonie & Cecilia Poletto: 2004, 'On standardising syntactic elicitation techniques (part 1)', in press in *Lingua*.

- Elstad, Kåre: 1982, 'Nordnorske dialekter', in Tove Bull & Kjellaug Jetne (eds.): *Nordnorsk: Språkarv og språkforhold i Nord-Noreg*. Samlaget, Oslo, pp. 11-100.
- Fiva, Toril: 1996, 'Spørresetninger i Tromsø-dialekten', in *Nordica Bergensia* 9, 139-155.
- Holmberg, Anders & Thorbjörg Hróarsdóttir: 2003, 'Agreement and movement in Icelandic raising constructions', in *Lingua* 113:10, pp. 997-1019.
- Iversen, Ragnvald: 1918, *Syntaksen i Tromsø bymaal*. Bymaals-lagets forlag, Kristiania.
- Lie, Svein: 1992, 'Ka du sei?', *Maal og Minne*, 62-77.
- Nilsen, Hilde: 1996, 'Koffør dæm sir det? Spørresetninger i nordreisadialekten', Cand. Philol. Thesis, University of Tromsø.
- Nilsen, Øystein: 1997, 'Adverbs and A-shift', *Working Papers in Scandinavian Syntax* 59, pp. 1-31.
- Nilsen, Øystein: 2003, *Eliminating positions: syntax and semantics of sentential modification*, doctoral dissertation, University of Utrecht, HIL dissertation 73.
- Nordgård, Torbjørn: 1985, *Word order, Binding and the Empty Category Principle*, Cand. Philol. thesis, University of Trondheim.
- Nordgård, Torbjørn: 1988, 'Omkring ordstilling i hv-spørsmål i norske dialekter', *Skriftserie fra Institutt for fonetikk og lingvistikk, Universitetet i Bergen* 33, pp. 26-37.
- Platzack, Christer: 1999, 'Satsfläta med relativsats', in Inger Haskå & Carin Sandqvist (eds.): *Alla tiders språk. En vänskrift till Gertrud Pettersson november 1999*. Lunds universitet, Lund, pp. 189-199.
- Platzack, Christer: 2000, 'A Complement-of-N° account of restrictive and non-restrictive relatives: the case of Swedish', in Artemis Alexiadou, Paul Law, André Meinunger & Chris Wilder (eds.), *The Syntax of Relative Clauses*. Amsterdam, Benjamins, pp. 265-308.
- Rice, Curt & Peter Svenonius: 1998, 'Prosodic V2 in Northern Norwegian', ms., University of Tromsø.
- Rizzi, Luigi: 1997, 'The fine structure of the Left Periphery', in Liliane Haegeman (ed.): *Elements of Grammar: Handbook of Generative Syntax*, Kluwer, Dordrecht, pp. 281-337.

- Rizzi, Luigi: 2001, 'On the position "Int(errogative)" in the Left Periphery of the clause', in Guglielmo Cinque & Giampaolo Salvi (eds.), *Current Studies in Italian Syntax*, Elsevier, Amsterdam, pp. 287-296.
- Sollid, Hilde: 2003, 'Dialektsyntaks i Nordreisa: Språkdannelse og stabilisering i møtet mellom kvensk og norsk'. Doctoral dissertation, University of Tromsø.
- Stroh-Wollin, Ulla: 1997, 'Framväxten av obligatoriska *som*', in Patrik Åström (ed.): *Studier i svensk språkhistoria 4*, pp. 171-178. Stockholm.
- Taraldsen, Knut Tarald: 1986a, 'On Verb Second and the functional content of syntactic categories', in Hubert Haider and Martin Prinzhorn (eds.), *Verb Second Phenomena in Germanic Languages*. Foris, Dordrecht, pp. 7-25.
- Taraldsen, Knut Tarald: 1986b, 'SOM and the Binding Theory', in Lars Hellan & Kirsti Koch Christensen (eds.), *Topics in Scandinavian Syntax*, Reidel, Dordrecht, pp. 149-184.
- Vangsnes, Øystein Alexander: 1999, 'The identification of functional architecture'. Doctoral dissertation, University of Bergen.
- Vangsnes, Øystein Alexander: 2001, 'On noun phrase architecture, referentiality, and article systems. In *Studia Linguistica* 55, pp. 249-299.
- Vangsnes, Øystein Alexander: 2002, 'A verb's gotta do what a verb's gotta do! On Scandinavian infinitivals and the AGR parameter', in C. Jan-Wouter Zwart and Werner Abraham (eds.) *Studies in Comparative Germanic Syntax*, John Benjamins, Amsterdam, pp 205-219.
- Vikner, Sten: 1995, *Verb movement and expletive subjects in the Germanic languages*, Oxford University Press, New York and Oxford.
- Wechsler, Stephen: 1991, 'Verb Second as illocutionary force', in Katherine Leffel and Denis Bouchard (eds.): *Views on Phrase Structure*. Kluwer, Dordrecht, pp. 177-191.
- Westergaard, Marit R.: 2003a, 'Word order in *wh*-questions in a North Norwegian dialect: some evidence from an acquisition study', in *Nordic Journal of Linguistics* 26.1, pp. 1-29.
- Westergaard, Marit R.: 2003b, 'On the acquisition of word order in WH-questions in the Tromsø dialect, in *Nordlyd* 31:3, pp. 589-605. (Published online at <http://www.ub.uit.no/munin/nordlyd/viewissue.php?id=3>.)

Westergaard, Marit R. and Øystein A. Vangsnes: forthcoming, ‘WH-questions, V2, and the left periphery of three Norwegian dialect types’, to appear in *Journal of Comparative Germanic Linguistics*.

Appendix: Variation in the properties of *som*

In this appendix I will present some further information on crossdialectal variation pertaining to the element *som*. I have not exploited these differences in the analysis of the cross-dialectal variation in Norwegian *wh*-questions, but further investigations may reveal that the differences are indeed relevant. The information is therefore given here in order to give as broad a basis as possible for further studies.

First of all, Nordgård (1985, 1988) reports that in all of the dialects in his study that allow non-V2 in *wh*-questions *som* can replace *at* ‘that’ in cases of extraction of a *wh*-subject from an embedded clause. In this respect these dialects contrast with the eastern “obligatory V2” dialects (and Standard Norwegian). This is exemplified by the examples in (1) and (2), and the relevant important contrast is between (1c) and (2b).

- (1) a. *Jens tror at Jon er i baren nå.* eastern Norwegian
 Jens thinks that Jon is in bar-DEF now
 ‘Jens thinks that Jon is in the bar right now.’
 b. **Jens tror hvem som er i baren nå.*
 Jens thinks who SOM is in bar-DEF now
 c. *Hvem tror Jens (?at/*som) er i baren nå?*
 who thinks Jens that/SOM is in bar-DEF now
 ‘Who does Jens think is in the bar right now?’
- (2) a. *Han Jens tror at han Jon er i baren no.* “non-eastern” Nor.
 ART Jens thinks that ART Jon is in bar-DEF now
 ‘Jens thinks that Jon is in the bar right now.’
 b. *Kem hanJens tror (som) er i baren no?*
 who ART Jens thinks SOM is in bar-DEF now
 ‘Who does Jens think is in the bar right now?’

Second, in many of the “non-eastern” Norwegian dialects *som* is homophonous to the consequential complementizer *så* (‘so, therefore’) and the

measure adverb *så* ('so'). I exemplify this with my own native Leikanger dialect from the district Sogn in western Norway (immediately to the north of Hordaland)—the a.-example is a subject *wh*-question, the b.-example contains a consequential adverbial clause, and the c.-example illustrates the measure adverb.

- (3) a. *Kem so kjemme í brüdleupe?* Leikanger (Sogn)
 who SOM comes in wedding-DEF
 'Who's coming to the wedding?'
 b. *Eg ha føsove meg so dú má koyra meg.*
 I have overslept myself so you must drive me
 'I have overslept so you'll have to drive me.'
 c. *Fisken va só stóre.*
 Fish-DEF was so big
 'The fish was this big.'

The Nordmøre dialect displays the same homophony (Tor A. Åfarli, p.c.), but the Tromsø dialect does not (Marit R. Westergaard, p.c.)—in the Tromsø dialect the consequential complementizer and the measure adverb are homophonous, but in turn these are distinct from *som*.

- (4) a. *Kem som kommer i bryllupet?* Tromsø
 who SOM comes in wedding-DEF
 'Who's coming to the wedding?'
 b. *Æ har førsove mæ så du må kjøre mæ.*
 I have overslept myself so you must drive me
 'I have overslept so you'll have to drive me.'
 c. *Fisken va så stor.*
 Fish-DEF was so big
 'The fish was this big.'

The situation in the Tromsø dialect is in this respect the same as in Standard Norwegian.

A third point indirectly involving *som* concerns exclamation. As mentioned in footnote 10 the element *så* can introduce exclamation in Norwegian. In the "non-eastern" dialects *wh*-elements, in particular the *wh* degree element *kor* 'how', can also be used to introduce exclamation. This

is however not possible in eastern Norwegian and not in the Bokmål written standard either, at least not without yielding an archaic flavor. This fact concerning eastern Norwegian and Bokmål is illustrated in (5) and (6).

- (5) a. *Så stor du har blitt!* eastern/Bokmål Norwegian
 so big you have become
 ‘How tall you’ve become!’
 b. *Så mange presanger dere har fått!*
 so many presents you have received
 ‘How many presents you have received!’
 c. *Så fin kjole du har!*
 so nice dress you have
 ‘What a nice dress you have!’
- (6) a. **Hvor stor du har blitt!* eastern/Bokmål Norwegian
 how big you have become
 b. **Hvor mange presanger dere har fått!*
 how many presents you have received
 c. **Hvor fin kjole du har!*
 how nice dress you have

As an example of a dialect which allows the degree *wh*-element *ko(r)* ‘how’ to introduce exclamatives I again use my own native Leikanger dialect. As we recall from above the non-*wh* degree element is homophonous to *som* in this dialect, and this is then the reason why this variation pertaining to exclamative is potentially relevant with respect to *som*.

- (7) a. *Ko/Só stóre dú e vorten!* Leikanger (Sogn)
 how/so big you are become
 ‘How tall you’ve become!’
 b. *Ko/Só mánge presanga di ha fått!*
 how/so many presents you have received
 ‘How many presents you’ve received!’
 c. *Ko/Só fíne kjóle dú ha!*
 how nice dress you have
 ‘How nice a dress you have!’

Eastern Norwegian seems to be the only Scandinavian variety which cannot use *wh*-elements to introduce exclamatives.

In sum we can give the following table to illustrate how Norwegian dialects differ with respect to lexical and syntactic differences that relate to the element *som*.

	<i>som</i> may replace <i>at</i>	<i>som</i> = <i>så</i>	<i>wh</i> exclamatives
eastern Norw.	–	–	–
western Norw.	+	+	+
Tromsø	+	–	+

Table 16: *Variation directly and indirectly related to som across Norwegian dialects*

In what ways this variation can be exploited to account for the variation pertaining to word order in main clause *wh*-questions is an issue for future research.

Author's address:
 Department of Linguistics
 Faculty of Humanities
 University of Tromsø
 NO-9011 Tromsø

oystein.vangsnes@hum.uit.no