This is a post-print version reflecting changes made in the peer review and editing process, but it is not the publisher’s PDF. This article appeared as


When citing, please use the page numbers given there. The publisher should be contacted for permission to re-use or reprint the material in any form.
Word order and verb movement in Norwegian wh-questions: A comparison of production and judgment data

Marit Westergaard
UiT – The Arctic University of Norway

Abstract (120 words)
This paper compares production and judgment data collected from several North Norwegian dialects, focusing on word order variation in wh-questions (V2 vs. non-V2). The findings show that non-V2 is widespread in questions with monosyllabic wh-elements and also fully accepted in questions with phrasal wh-elements in language contact areas. Furthermore, the acceptability of non-V2 correlates with the complexity of the wh-element, confirming previous accounts of the historical development. Age differences suggest that this development is currently being reversed. However, there is a major discrepancy between the production and judgement data, in that non-V2 is hardly ever produced with complex wh-elements. The findings are discussed in relation to what frequencies in diachronic data can(not) reveal about the grammar that produced historical texts.

Keywords: Verb second, frequency, language contact, dialect variation, information structure, language change

1. Introduction

This article discusses the microvariation found with respect to word order in wh-questions in several North Norwegian dialects. As is well known from the literature (e.g. Vangsnes 2005, Westergaard 2009a), certain wh-questions in these dialects require strict verb second (V2) word order, others require non-V2, while others again allow both word orders. The syntactic variation is generally dependent on the function and length of the wh-element, in that e.g. subject questions require non-V2 and non-subject questions with disyllabic wh-elements (phrases) require V2. With short (monosyllabic) wh-elements there is word order variation which has been shown to be linked to information structure, more specifically whether the subject is given or new and/or focused information. The patterns are thus not unlike what is found in declaratives in the history of English (see e.g. Bech 2001, Westergaard 2009b, van Kemenade & Westergaard 2012).

Much of the previous literature on wh-questions in North Norwegian dialects has focused on the dialect of Tromsø. The present paper is based on data collected at

* The data collection in Northern Troms reported on in this paper was supported by small-scale funding (Småforskningsmidler) from the Research Council of Norway, granted to the author of this paper. I am grateful to Øystein A. Vangsnes and an anonymous reviewer for helpful comments.
a number of locations in Troms county\(^1\), on the island Senja (to the south of Tromsø) and three different locations in Northern Troms (to the northeast). The main issue is the word order found in wh-questions introduced by long (disyllabic and phrasal) wh-elements, which have been found to optionally display non-V2 in some of the areas in Northern Troms, argued to be due to language contact with Saami and Kven/Finnish, both languages without the V2 property (Nilsen 1996, Sollid [=Nilsen] 2003).

The data collected include acceptability judgments as well as semi-spontaneous production, which turn out to yield surprisingly different results: While the speakers from Senja neither produce nor accept non-V2 with phrasal wh-elements, speakers from Northern Troms fully accept them, but produce them only very rarely (0.7%). The judgment data from Northern Troms also indicate that the variation found with the long wh-elements is not dependent on information structure, as is the case with the monosyllabic wh-words. These findings are discussed in relation to what frequencies in diachronic data can (and cannot) tell us about the underlying grammar that produced historical texts. They also corroborate findings from late Middle English (ME), showing that patterns of information structure break down before the (almost) complete loss of V2 (van Kemenade & Westergaard 2012).

2. Background

The word order of wh-questions in Norwegian dialects has been the focus of many studies, e.g. Åfarli (1985, 1986), Nordgård (1985), Taraldsen (1985, 1986), Lie (1992), Nilsen (1996), Sollid (2003), Vangsnes (2005), Westergaard (2003, 2009a), Westergaard & Vangsnes (2005), Rognes (2011), and Westergaard, Vangsnes & Lohndal (2012, 2017). The reason for this is that, while Norwegian is generally assumed to be a V2 language, certain wh-questions also either require or allow non-V2 in many dialects, illustrated in (1).\(^2\) In comparison, this word order is not grammatical in either of the two written standards of Norwegian, as shown in (2).\(^3\)

\[(1) \quad \text{Ka du sir?} \quad \text{(Tromsø dialect)}
\]

\[
\begin{align*}
& \text{what you say} \\
& \text{‘What are you saying?’}
\end{align*}
\]

\(^1\) Troms is the name of one of three counties in North Norway.

\(^2\) In traditional accounts of V2 in Germanic languages, this word order is generally assumed to be the result of verb movement to the C position of the clause (den Besten 1977, Vikner 1995), while the variable word order in Norwegian wh-questions has been argued to involve a lower head position in the C domain, Top\(^a\) (Westergaard 2009a). For earlier stages of English, it has been argued that the verb either moves to C (e.g. in questions) or to a lower position (in most declaratives), typically referred to as the head of FP (e.g. Haeberli 1999, Fischer et al. 2000, van Kemenade & Westergaard 2012).

\(^3\) The two written standards of Norwegian are bokmål and nynorsk (lit. ‘book language’ and ‘new Norwegian’). Examples illustrating Standard Norwegian in this article are given in bokmål. See Venås (1993) or Vikør (1995) for more information on the language situation in Norway.
This article discusses the nature of the variation between V2 and non-V2 in North Norwegian dialects. The main focus in the previous literature has been on the variety spoken in Tromsø. Speakers of most northern dialects, including Tromsø, make a clear distinction between the word order of questions introduced by short (monosyllabic) wh-words on the one hand and questions introduced by disyllabic wh-elements or full wh-phrases on the other: While the latter only allow V2 word order, the former display variation between V2 and non-V2. This is shown in examples (3)-(4) and (5) respectively.

(3)  Koffer *sir dem* det? / *Koffer dem* sir det?  (Tromsø)
    why say they that / why they say that
    ‘Why are they saying that?’

(4)  Kor mange unga *har* du? / *Kor mange unga* du har?
    how many kids have you
    ‘How many kids do you have?’

(5)  Kor *bor* du? / Kor *du* bor?
    where live you / where you live
    ‘Where do you live?’

In an area approximately 150 kilometers north of Tromsø there has been considerable language contact between Norwegian and the Kven language (spoken by Finnish immigrants to this area arriving in the 18th and 19th centuries), which does not have the V2 property. In the dialects of Norwegian spoken there, non-V2 word order has been claimed to be grammatical also in questions introduced by the long wh-elements, i.e. in questions such as (3) and (4), cf. Nilsen (1996) and Sollid (2003). Non-V2 in questions with long wh-elements is also attested in dialects spoken in the Western parts of the country (see e.g. Åfarli 1985, 1986, Vangsnes 2005, 2007), and this has been analyzed diachronically as a further step in the development from V2 to non-V2 (Westergaard 2009a, Westergaard, Vangsnes & Lohndal 2012, 2017). That is, the historical development has started with the short wh-elements and then spread to the long ones (see section 3 below).4

The variation found in questions with short wh-elements such as (5) is not random: Although both word orders are perfectly grammatical in isolation, there are clear speaker preferences in discourse. The examples in (6a, b) have been taken from a corpus of spontaneous speech produced by eight adults and three children, all living in Tromsø and speaking North Norwegian dialects (see

---

4 In Westergaard, Vangsnes & Lohndal (2012, 2017), the diachronic development is argued to be somewhat more complex, starting in subject wh-questions and then, as a second step, spreading to non-subject questions with short wh-elements.
Anderssen 2006 for more information on the corpus). Based on data from this corpus, Westergaard (2003, 2009a) argues that the word order choice in this context is dependent on information structure. More specifically, V2 is chosen when the subject is new and/or focused information (typically a full DP), while non-V2 is preferred when the subject is informationally given (often a pronoun). This is illustrated by the V2 word order in (6a), which introduces a new referent (skoan hannes ‘his shoes’). In the follow-up question the subject is given information and may thus be referred to by a pronoun. Consequently, non-V2 word order is used in (6b).

(6)  
  a. kor er skoan hannes henne? (INV, file Ole.17) 
     where are shoe.DEF/PL his LOC 
     ‘Where are his shoes?’
  b. kor dem er henne? 
     where they are LOC 
     ‘Where are they?’

This pattern is similar to what has been found in declaratives in Old English (OE) and to a certain extent also ME (Bech 2001, Westergaard 2009b, van Kemenade & Westergaard 2012). That is, while full DP subjects normally require V2 word order, pronominal subjects are virtually always found preceding the verb. This is illustrated in the following OE examples (from van Kemenade & Westergaard 2012: 92).

(7)  
  On twam þingum hæfte God þæs mannes sawe_gegodod (ÆCHom I, 1.20.1) 
  in two things had God the man’s soul endowed 
  ‘With two things God had endowed man’s soul.’

(8)  
  Be ðæm we magon suiðe swutulę oncnawan ðæt… (CP 26.181.16) 
  by that we may very clearly perceive that…
  ‘By that, we may perceive very clearly that…’

Earlier versions of English are also similar to North Norwegian in that there are certain initial elements that behave differently from others with respect to the word order allowed or required in the sentence. Most notably, this concerns elements belonging to the so-called then-group (then, thus, now), see also Warner (2007) and Eitler & Westergaard (2014). That is, even when the subject was a pronoun, declaratives introduced by these elements virtually always occurred with V2 word order in OE. This is shown in example (9) with one of the most frequent of these elements, þa ‘then’, from Bech (2001: 3).

(9)  
  þa siglde he þonan suðryhte be lande. (Or 14:16). 
  then sailed he from there southwards along coast 
  ‘Then he sailed from there southwards along the coast.’

For this reason, the element þa and its variant þonne are often considered to be operators in OE, on a par with wh-elements and the negative element ne, also requiring V2 word order (e.g. van Kemenade 1987, Pintzuk 1991, Kroc...
Taylor 1997, but see Westergaard 2009b for a split-CP model which makes the operator analysis unnecessary). In the data investigated in Westergaard (2009b), the proportion of V2 word order in declaratives introduced by pa/ponne was 99% (192/194) in early OE and 99.4% (158/159) in late OE. That is, there are only occasional examples not conforming to the rule, and these are often disregarded in historical analyses.

Finally in this section, it should be mentioned that there are exceptions to the V2 rule also in declaratives in present-day Norwegian. This mainly concerns declaratives introduced by the initial adverb kanske ‘maybe’, which may appear with either V2 or non-V2, as illustrated in (10).

(10) Kanskje barna vil leke / Kanskje vil barna leke.
    maybe children.DEF will play
    ‘Perhaps the children want to play.’

According to an investigation of two corpora (8 and 166 adult speakers respectively), the two word orders appear with very different frequencies in spontaneous speech, non-V2 being attested 95-100% in the production of individual speakers (Westergaard 2008, 2009a). I return to the relevance of such exceptions and the frequency with which they appear in production data in section 7.

3. Word Order in wh-questions in Northern Troms

In a thesis aptly called Koffer dæm sir det? [Why they say that?], Nilsen (1996) discusses the dialect spoken in an area in Northern Troms, Nordreisa, with respect to the option of using non-V2 word order in questions with long wh-elements. Nilsen’s (1996) investigation is based on acceptability judgments and occasional authentic examples. The dialect speakers were supposed to judge altogether 48 sentences in a questionnaire as either acceptable or unacceptable (i.e. a binary choice). Two of her test sentences with long wh-elements and non-V2 word order are illustrated in (11) and (12), from Nilsen 1996: 77. These were judged as acceptable by 28.9% (26/90) and 37.8% (34/90) of the informants.

(11) Koffer ho Ingrid sitt på kjøkkenet og sover? (accepted by 28.9%)
    why DET Ingrid sit on kitchen.DEF and sleep
    ‘Why is Ingrid sitting in the kitchen sleeping?’

(12) Korsn bil han Øyvind har kjøpt? (accepted by 37.8%)
    which/how car DET Øyvind has bought
    ‘What car did Øyvind buy?’

No systematic study has been done on such examples in natural discourse. However, based on data from the Tromsø corpus (Anderssen 2006), Westergaard (2005) studies the spontaneous production of two speakers who are originally from another area in Northern Troms, Manndalen (Kåfjord), a little
to the south of Nordreisa. These two are the parents of one of the three children in this acquisition corpus, and they have both lived in Tromsø most of their adult lives. In the Kåfjord area there has historically been considerable contact with the indigenous language Saami, another non-V2 language. Both of these speakers produce occasional examples of non-V2 in questions with long wh-elements, but only the mother produces enough data to make it possible to see a pattern. Table 1 (from Westergaard 2009a: 61) provides an overview of the wh-questions produced by this speaker in the corpus.

Table 1: The percentage of V2 word order across different wh-questions, female Kåfjord speaker born 1957 (N=863).

<table>
<thead>
<tr>
<th>Wh-element</th>
<th>'ka' 'what'</th>
<th>kor/kem 'where/who'</th>
<th>korsen/koffer/katti 'how, why, when'</th>
<th>Full phrases</th>
</tr>
</thead>
<tbody>
<tr>
<td>% V2</td>
<td>8.7% (46/527)</td>
<td>27.9% (68/244)</td>
<td>80% (36/45)</td>
<td>91.5% (43/47)</td>
</tr>
</tbody>
</table>

There is a clear correlation between the length of the wh-element and the proportion of V2: The shorter the wh-element, the more non-V2 word order is produced. Westergaard (2009a) argues that this pattern is the result of a diachronic development where complexity and frequency play an important role, and where the language contact in Kåfjord is argued to accelerate the process. More specifically, the historical change from V2 to non-V2 is accounted for as a process which starts with the question word 'ka' 'what', which may be argued to be the least complex of all wh-words (see e.g. Bayer 2004). This word order spreads to the other monosyllabic wh-words ('kor' 'where' and 'kem' 'who'), which results in a surprisingly stable distinction between monosyllabic and longer wh-elements across numerous dialects. According to Westergaard (2009a), the reason for this stability is that this distinction corresponds to a difference between heads and phrases (represented by the monosyllabic wh-words and the longer wh-elements respectively). Nevertheless, due to a considerable frequency difference between the short and the long wh-elements in typical everyday speech (the short ones accounting for approximately 90%-95% of the input to children), the development does spread to the disyllabic wh-elements and finally to full wh-phrases in some dialects. This historical development is reflected in the synchronic frequency differences shown in Table 1.

4. Research questions

The sparse previous literature on non-V2 word order in the dialects of Northern Troms leads to a number of new research questions. Since Nilsen (1996) did not collect spontaneous data, and since Westergaard (2005) only considered two speakers who no longer live in this area, we do not know to what extent non-V2 is used in spontaneous production. Perhaps the low percentages found in the production of the Mandalen speakers in the Westergaard (2005) study is due to the fact that they now live in an area where this word order would be considered ungrammatical. Thus, an important question is whether non-V2 in questions with long wh-elements can be attested with the same frequency as in questions
with short wh-words. In Table 1, there also seems to be a difference between disyllabic wh-elements and full wh-phrases with respect to the frequency of V2 (80% vs. 91.5%), but the raw numbers are so small that this difference is uncertain. In a corpus with more examples of non-V2 with the long wh-elements, we might be able to confirm this finding.

A second question concerns the historical development sketched above: If non-V2 word order in questions with long wh-elements is the result of a historical change, then we would expect young people to produce and accept more of the new development. However, there is also the possibility that we could see the opposite situation, with young people actually producing and accepting less non-V2: According to Sollid (2003), the typical contact phenomena of the Nordreisa dialect, one of which is non-V2 with long wh-elements, have become severely stigmatized, and the dialect is currently undergoing a process of leveling to other Norwegian dialects spoken in surrounding areas.

A related question is thus how the dialects that have been in contact with Kven/Finnish and Saami (Nordreisa and Manndalen respectively) compare to other dialects in the area that have been affected by this language contact to a much lesser extent. A comparison should also be made with North Norwegian dialects further away from these areas, e.g. in Southern Troms, where there has been no language contact, as far as we know.

Since previous studies have either used acceptability judgments or corpora, we do not know how the results of these methodologies compare to one another. Will production of a high percentage of non-V2 correspond to a high acceptance rate of this word order in judgment tasks, or is there no relation between production and acceptance? Related to this, it should be mentioned that speakers typically have no awareness of the two different word orders in their production and there is also extreme individual variation (e.g. ranging between 2.5% and 68.4% V2 produced by speakers in the Tromsø corpus, cf. Westergaard 2009a).

The final research question concerns the effect of information structure. In questions with short wh-elements, which are very frequent in spontaneous discourse, it has been shown that the choice between V2 and non-V2 is dependent on the information status of the subject as given vs. new and/or focused information (cf. section 2). But what about questions with the long wh-elements? Given that Nilsen’s (1996) analysis is mainly based on acceptability judgments, and Westergaard’s (2005, 2009a) two subjects produce so few examples, we have simply not had enough spontaneous data so far to make any claim about the role of information structure in questions with the long wh-elements. A further question related to information structure is whether this distinction may also be attested in speakers’ judgments.

Summarizing, our research questions are the following:

1. To what extent do speakers living in Northern Troms produce non-V2 word order – with short and long wh-elements? Furthermore, is there a
frequency difference between non-V2 with disyllabic wh-words and full wh-phrases?
2. Is there a historical change in progress, i.e. do we find differences between young and old speakers?
3. Are there differences between Manndalen (Saami influence), Nordreisa (Kven/Finnish influence) and North Norwegian dialects that have not been affected (or less affected) by language contact?
4. How does spontaneous production correspond to acceptability judgments?
5. Does information structure play a role also with the long wh-elements in discourse? And will the difference between pronominal and DP subjects be detectable also in acceptability judgments?

5. Participants and methodology

Data from Northern Troms were collected in December 2006 during fieldwork carried out by Gunnar Hrafn Hrafnbjargarson, Hilde Sollid, Øystein A. Vangsnes, and the author of the present paper. The locations chosen were Manndalen (with predominantly Saami influence), Nordreisa (with predominantly Kven/Finnish influence), and Djupvik (assumed to have a dialect that is more similar to other northern dialects in that there has been less language contact). The three locations are marked on the map of Troms county in Figure 1, as well as the city of Tromsø and the island of Senja in Southern Troms, where corresponding data were collected (cf. section 6.2).
Both acceptability judgments as well as semi-spontaneous production data were collected from altogether 28 speakers: Twelve speakers from Manndalen, ten from Nordreisa, and six from Djupvik. Table 2 provides an overview of the speakers in terms of sex and year of birth.

### Table 2: Overview of speakers, Northern Troms (N=28).

<table>
<thead>
<tr>
<th>Manndalen</th>
<th>Nordreisa</th>
<th>Djupvik</th>
</tr>
</thead>
<tbody>
<tr>
<td>M5 – M, b. 1935</td>
<td>N5 – M, b. 1948</td>
<td>D5 – M, b. 1945</td>
</tr>
<tr>
<td>M6 – M, b. 1935</td>
<td>N6 – M, b. 1938</td>
<td>D6 – M, b. 1948</td>
</tr>
<tr>
<td>M7 – F, b. 1959</td>
<td>N7 – F, b. 1948</td>
<td></td>
</tr>
<tr>
<td>M8 – F, b. 1963</td>
<td>N8 – F, b. 1927</td>
<td></td>
</tr>
<tr>
<td>M10 – M, b. 1991</td>
<td>N10 – M, b. 1955</td>
<td></td>
</tr>
<tr>
<td>M11 – F, b. 1926</td>
<td></td>
<td></td>
</tr>
<tr>
<td>M12 – F, b. 1920</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

In order to investigate possible historical change in progress in this area, the speakers were divided into three groups based on their age at the time of recording. This is illustrated in Table 3.5

### Table 3: Overview of speakers divided into three age groups, Northern Troms (N=28).

<table>
<thead>
<tr>
<th>Group</th>
<th>Age</th>
<th>Participants</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>58-86</td>
<td>M5, M6*, M11*, M12, N5, N6, N7, N8, D5, D6</td>
</tr>
<tr>
<td>2</td>
<td>26-53</td>
<td>M1, M2, M7, M8, N1, N2, N9, N10, D1, D2, D3, D4</td>
</tr>
<tr>
<td>3</td>
<td>13-16</td>
<td>M3, M4, M9, M10, N3, N4</td>
</tr>
</tbody>
</table>

The methodology chosen was developed by Øystein A. Vangsnes in his study of the Sogn dialect in Western Norway (Vangsnes 2007). Thus, the Northern Troms data consist of semi-spontaneous production as well as acceptability judgments. The production data consist of conversations between pairs of dialect speakers, usually matched for sex and age. These conversations typically lasted approximately a half hour. As questions with long wh-elements are typically infrequent in normal discourse (cf. section 4), the speakers were given three different question/answer tasks specifically designed to elicit a high number of such questions. In the first task, speakers played a game where one is to think of a person and the other is to guess who this person is. In this process, only wh-

---

5 Speakers M6 and M11, marked with an asterisk in Table 3, were not included in the calculations of the acceptability judgments, as they did not seem to understand the task.
questions are allowed, e.g. \textit{where does he/she live?}, \textit{what color hair does he/she have?}, etc. In the second task, one person was given a picture with many details (people, houses, cars, animals, etc.) and was able to study it for one minute. Afterwards the picture was given to the other person who would ask the first person what he/she remembered from the picture, e.g. \textit{what is she doing?}, \textit{how many cars are there}, etc. In the third task, the speakers were supposed to interview each other about their lives, and in order to elicit a variety of wh-questions, we gave them a list of wh-items which they were supposed to use at least once. One of the investigators was normally present during these conversations, which typically lasted approximately 30 minutes.

The participants were also asked to provide acceptability judgments of pre-recorded sentences, using a Likert scale from 1 (unacceptable) to 7 (acceptable). These sentences had been recorded by speakers from the relevant dialect areas and were presented aurally to the informants. There were altogether 108 sentences in this judgment task, investigating a variety of syntactic phenomena. This task was not timed, and the investigators discussed individual sentences with the participants, making sure they focused on the relevant aspects when making their judgments and often noting additional comments they had.

There were 26 (randomized) sentences testing word order in wh-questions, 13 with V2 and 13 with non-V2 (identical except for the word order used). Four of these had a monosyllabic wh-word (\textit{ka} 'what' x2, \textit{kem} 'who' x2), five had a disyllabic wh-word (\textit{katti} 'when' x2, \textit{korsn} 'how' x2, \textit{koffer} 'why' x1), and four had a full wh-phrase (\textit{kor mange} x 'how many' x x2, \textit{kor lenge} 'how long' x1, \textit{ka slags} x 'what kind of' x x1). The subject type also varied, such that 8 sentences had pronominal subjects and 5 DPs. In order to test whether information structure would play a role also in speakers' judgments, the test items were always given in context, where the subject of the following question was either already mentioned in the previous utterance, such as \textit{du} 'you' in (13), or was introduced in the question as new information, such as \textit{mor di} 'your mother' in (14).

(13) Det e ei stund sia æ satt på skolebenken, ja!
    \textit{It's really been a while since I sat on school-bench, yes}
    a. Katti \textbf{du gjekk} \textit{ut av ungdomsskolen?}
       \textit{When you went out of youth-school}
    b. Katti \textbf{gjekk} \textit{du ut av ungdomsskolen?}
       \textit{When went you out of youth-school}

(14) Det e fint å høre om gamle daga, syns æ.
    \textit{It is nice to hear about old days, think I}
    a. Katti \textbf{mor di e fødd?}
       \textit{When mother your is born}
    b. Katti \textbf{e mor di fødd?}
when is mother your born
’When is your mother born?’

6. Results

6.1 Acceptability judgments

Table 4 gives the acceptance rate for the 13 wh-questions with the two different word orders, divided according to the length of the wh-element.

Table 4: Average acceptance rate for questions with different wh-elements on a scale from 1 (unacceptable) to 7 (acceptable); three areas in Northern Troms.

<table>
<thead>
<tr>
<th>Wh-element</th>
<th>Manndalen (N=10)</th>
<th>Nordreisa (N=10)</th>
<th>Djupvik (N=6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-V2/V2</td>
<td>Non-V2/V2</td>
<td>Non-V2/V2</td>
</tr>
<tr>
<td>Monosyllabic wh (4 examples)</td>
<td>5.8/6.4</td>
<td>5.8/6.5</td>
<td>6.2/6.7</td>
</tr>
<tr>
<td>Disyllabic wh (5 examples)</td>
<td>5.7/6.7</td>
<td>5.4/6.6</td>
<td>3.9/7</td>
</tr>
<tr>
<td>Wh-phrases (4 examples)</td>
<td>4.7/6.6</td>
<td>4.2/6.8</td>
<td>3.1/6.8</td>
</tr>
</tbody>
</table>

One of the most striking findings is that V2 word order is considered completely acceptable in all cases by all speakers, with a score between 6.4 and 7. This corresponds to findings from the Sogn dialect in Vangsnes (2007). This result means that this word order is clearly part of the dialect grammar, but the reason for the generally high scores may also be related to this being the only word order possible in the standard language (cf. section 2).

Furthermore, we find that there is only a slight distinction between mono- and disyllabic wh-words with respect to the acceptance of non-V2 in the contact dialects in Manndalen and Nordreisa, the scores being 5.8 in both locations for monosyllabic wh-words and 5.7 and 5.4 respectively for disyllabic ones. On the other hand, there is a clear distinction between the two wh-word types in Djupvik, i.e. 6.2 and 3.9 respectively. This supports the claim made in Nilsen (1996) that non-V2 in questions with the long wh-elements is the result of language contact. We also find that there is a difference in acceptability between disyllabic wh-elements and wh-phrases in all three areas, the latter being generally less acceptable with non-V2 (5.7 vs. 4.7 in Manndalen, 5.4 vs. 4.2 in Nordreisa, and 3.9 vs. 3.1 in Djupvik).

A statistical analysis of the judgment data (using the statistical program R) reveals that there is an effect of type of wh-element (p=.005), such that non-V2 with short wh-words is rated significantly higher than either disyllabic or phrasal wh-elements, and disyllabic wh-elements higher than wh-phrases.6 Unfortunately, due to the many different variables and the generally low raw numbers in this study, none of the other results presented in this paper sections display any statistical significance.

---

6 The statistical analysis has been carried out by Martin Corley, University of Edinburgh.
6.2 Comparison with similar data from Southern Troms

In October 2006 (two months before the data collection in Northern Troms) a similar study to the one in Northern Troms has been carried out by Øystein A. Vangsnes and the present author in connection with fieldwork organized by the NORMS project on the island of Senja (Southern Troms), cf. the map in section 5.\(^7\) The methodology was the same as the one described above, but as we had no reason to expect any non-V2 with long wh-elements in this area, there were fewer items with long wh-elements in the acceptability judgments. Thus, the test consisted of six sentences with monosyllabic wh-words, five with disyllabic wh-elements and none with wh-phrases. Most of the sentences with disyllabic wh-words had pronominal subjects, as these are the most likely ones to occur with non-V2 word order. The results are provided in Table 5.

Table 5: Average acceptance rate for questions with mono- and disyllabic wh-elements on a scale from 1 (unacceptable) to 7 (acceptable); Senja, Southern Troms (N=17).

<table>
<thead>
<tr>
<th>Wh-element</th>
<th>Non-V2/V2</th>
</tr>
</thead>
<tbody>
<tr>
<td>Monosyllabic wh (6 examples)</td>
<td>6.5/6.9</td>
</tr>
<tr>
<td>Disyllabic wh (5 examples)</td>
<td>2.1/6.8</td>
</tr>
</tbody>
</table>

The total scores for monosyllabic and disyllabic wh-words show that there is a clear distinction between these two with respect to speakers’ acceptance of non-V2: While this word order is considered perfectly acceptable with the short wh-items (score 6.5), non-V2 is clearly unacceptable in questions with disyllabic wh-elements, the score being only 2.1, even lower than in Djupvik (cf. Table 4). This also means that there is a considerable difference between Southern and Northern Troms with respect to this phenomenon.

6.3 The effect of information structure

In previous studies (e.g. Westergaard 2003, 2009a), a clear difference has been attested between pronominal and full DP subjects in spontaneous production, pronouns appearing predominantly with non-V2 and full DPs typically with V2 word order. Table 6 shows how the subject type affects the scores given to different question types in the acceptability judgment task.

---

\(^7\) NORMS (Nordic Center of Excellence in Microcomparative Syntax) was directed from CASTL, University of Tromsø, by Øystein A. Vangsnes and Peter Svenonius and funded by NOS-HS 2005-2010.
Table 6: Average acceptance rate for non-V2 word order in questions with different wh-elements on a scale from 1 (unacceptable) to 7 (acceptable), with pronominal and full DP subjects.

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Monosyll. wh</td>
<td>6.4/5.3</td>
<td>6.5/5.1</td>
<td>6.8/5.6</td>
<td>6.8/6.2</td>
</tr>
<tr>
<td>Disyll. wh</td>
<td>5.8/5.6</td>
<td>5.3/5.5</td>
<td>3.6/4.4</td>
<td>2.1/2.2</td>
</tr>
<tr>
<td>wh-phrases</td>
<td>4.8/4.4</td>
<td>4.3/3.6</td>
<td>3.2/2.7</td>
<td>-</td>
</tr>
</tbody>
</table>

Considering the questions with monosyllabic wh-items, we see that non-V2 is slightly degraded with DP subjects, the difference between the two subject types being approximately 1.1 across the four locations. This confirms Vangsnes’s (2007) findings from the Sogn dialect and indicates that speakers have a certain sensitivity to the information structure patterns involved in the word order choice, even in experimental situations where the context is quite limited (cf. examples 13-14). While there is also a slight difference between pronominal and DP subjects in the questions with wh-phrases (only 0.4-0.7), there turns out to be no such distinction in questions with disyllabic wh-elements: The difference between the scores for pronouns and DPs in Manndalen is only 0.2, and in the other locations the score for DP subjects is actually somewhat higher than for pronouns.

6.4 Semi-spontaneous production

Table 7 provides a breakdown of the two word orders used across four different types of wh-elements in the three dialects in Northern Troms: ka ‘what’, the other two monosyllabic wh-words (kor ‘where’ and kem ‘who’), disyllabic wh-words, and wh-phrases.8

Table 7: The proportion of V2 word order in questions with different wh-elements, semi-spontaneous speech, Northern Troms.

<table>
<thead>
<tr>
<th>Wh-element</th>
<th>Manndalen (N=12)</th>
<th>Nordreisa (N=10)</th>
<th>Djupvik (N=6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• ka ‘what’</td>
<td>44.2% (65/147)</td>
<td>25.5% (25/98)</td>
<td>17% (9/53)</td>
</tr>
<tr>
<td>• kor/kem ‘where’/‘who’</td>
<td>47.6% (20/42)</td>
<td>40% (16/40)</td>
<td>32% (8/25)</td>
</tr>
<tr>
<td>Total monosyll. wh</td>
<td>45.0% (85/189)</td>
<td>29.7% (41/138)</td>
<td>19.3% (17/88)</td>
</tr>
<tr>
<td>Disyll. wh</td>
<td>97% (32/33)</td>
<td>100% (50/50)</td>
<td>100% (31/31)</td>
</tr>
<tr>
<td>Wh-phrases</td>
<td>98.6% (285/289)</td>
<td>100% (161/161)</td>
<td>100% (152/152)</td>
</tr>
</tbody>
</table>

The most striking and somewhat disappointing finding is that there are extremely few examples of non-V2 word order with the long wh-elements in the production data. In fact, there are only five clear examples of non-V2 after long

8 The spontaneous data from Senja have not been transcribed, but the investigator who was present during all recordings reported no examples of non-V2 with long wh-elements, as expected.
wh-elements, one with a disyllabic wh-word and four with wh-phrases, accounting for 0.8% (1/114) and 0.6% (4/602) respectively. This corresponds to only 0.7% (5/716) for the total corpus. These examples are produced by only two speakers, both of them from Manndalen, belonging to the young and middle-aged groups respectively. Examples are provided in (15)-(16). It should be noted that the subject in all five cases is du ‘you’, which, unless stressed, must be considered always to be given information.

(15) å korr fin du va på håre – katti du har årrna de? (M4)
    ‘Oh, how nice your hair looks – when did you have it fixed?’

(16) ee ja k a slass kjønn du tænngke på då? (M7)
    ‘Eh, well – what kind sex are you thinking of then?’

Nevertheless, we see that there is somewhat more non-V2 with ka ‘what’ than the other two short wh-elements, especially in Nordreisa and Djupvik. This is similar to findings from previous studies on other dialects (e.g. Westergaard 2009a). Somewhat surprisingly, however, there is more non-V2 with the monosyllabic wh-words in Djupvik than the other two areas. This means that there is in fact an inverse relationship between the acceptance of non-V2 with long wh-elements and the production of non-V2 with the short ones.

6.5 Age differences

As discussed above, the variation attested across the dialects may be analyzed as a historical development in progress. It is therefore important to compare the different age groups. In the following table, I have excluded Djupvik speakers, since they do not seem to really have any productive rule allowing non-V2 with long wh-elements.

Table 8: Average acceptance rate for questions with different wh-elements on a scale from 1 (unacceptable) to 7 (acceptable) across three age groups, Northern Troms (Manndalen and Nordreisa).

<table>
<thead>
<tr>
<th>Wh-element</th>
<th>OLD (N=6)</th>
<th>MIDDLE (N=8)</th>
<th>YOUNG (N=6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Non-V2/V2</td>
<td>Non-V2/V2</td>
<td>Non-V2/V2</td>
</tr>
<tr>
<td>Monosyllabic wh</td>
<td>6.1/6.5</td>
<td>5.8/6.7</td>
<td>5.5/6.2</td>
</tr>
<tr>
<td>(4 examples)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Disyllabic wh</td>
<td>6.1/6.7</td>
<td>5.6/6.7</td>
<td>4.9/6.4</td>
</tr>
<tr>
<td>(5 examples)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Wh-phrases</td>
<td>5.2/6.9</td>
<td>4.8/6.8</td>
<td>3.7/6.5</td>
</tr>
<tr>
<td>(4 examples)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Although differences between the groups are relatively small, the figures in Table 8 indicate that the younger the speakers are, the less non-V2 they accept, especially with the longer wh-elements. There also seems to be a difference between disyllabic wh-words and full wh-phrases in all age groups (6.1 vs. 5.2 – 5.6 vs. 4.8 – 4.9 vs. 3.7), especially among the youngest speakers.
We may also have a look at how this matches production across the three age groups. In Table 9, the production of the speakers in Manndalen and Nordreisa is displayed in terms of age. The lack of word order variation with the long wh-elements makes it impossible to detect any differences between age groups. However, it is clear that it is not the case that older speakers freely produce this word order, while younger speakers do not, as might have been expected given the acceptability scores in Table 8. With respect to the short wh-items, it is also difficult to see a pattern, as young speakers produce somewhat less non-V2 than the middle-aged speakers, but so do the older speakers. It should be noted that the raw numbers are quite low here, which means that individual variation clearly plays a role (as was shown in Westergaard 2009a, cf. section 4 above).

### Table 9: The percentage of V2 word order in questions with different wh-elements, across three age groups, Northern Troms (Manndalen and Nordreisa).

<table>
<thead>
<tr>
<th>Wh-element</th>
<th>OLD (N=8)</th>
<th>MIDDLE (N=8)</th>
<th>YOUNG (N=6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>• ka 'what'</td>
<td>35.8% (19/53)</td>
<td>27.2% (22/81)</td>
<td>53.3% (49/92)</td>
</tr>
<tr>
<td>• kor/kem 'where, who'</td>
<td>64.3% (9/14)</td>
<td>34.5% (10/29)</td>
<td>48.3% (14/29)</td>
</tr>
<tr>
<td>Total monosyll. wh</td>
<td>41.2% (28/68)</td>
<td>29.1% (32/110)</td>
<td>52.1% (63/121)</td>
</tr>
<tr>
<td>Disyllabic wh</td>
<td>100% (14/14)</td>
<td>100% (47/47)</td>
<td>94.1% (16/17)</td>
</tr>
<tr>
<td>Wh-phrases</td>
<td>100% (89/89)</td>
<td>97.9% (189/193)</td>
<td>100% (135/135)</td>
</tr>
</tbody>
</table>

#### 7. Discussion

We may now return to the research questions formulated in section 4, repeated here for convenience:

1. To what extent do speakers living in Northern Troms produce non-V2 word order – with short and long wh-elements? Furthermore, is there a frequency difference between non-V2 with disyllabic wh-words and full wh-phrases?
2. Is there a historical change in progress, i.e. do we find differences between young and old speakers?
3. Are there differences between Manndalen (Saami influence), Nordreisa (Kven/Finnish influence) and North Norwegian dialects that have not been affected (or less affected) by language contact?
4. How does spontaneous production correspond to acceptability judgments?
5. Does information structure play a role also with the long wh-elements in discourse? And will the difference between pronominal and DP subjects be detectable also in acceptability judgments?

With respect to the first question, we have shown that speakers in all areas investigated produce non-V2 in questions with short wh-words, confirming findings from previous literature on northern dialects (cf. references in section 2). The most striking finding of this study, however, is that the speakers of the contact dialects, contrary to our expectations, generally do not produce non-V2 in
questions with long wh-elements, this phenomenon only being attested 0.7% (cf. Table 7). The reason for this may be sought in Solli's (2003) claim that this phenomenon has become a stigmatized dialect feature that many speakers attempt to avoid, leading to the loss of this particular word order in Manndalen and Nordreisa. That is, the spread of the historical development from V2 to non-V2 to include all wh-questions, which has arguably taken place in these contact dialects, is now being reversed, due to a process of leveling to other Norwegian dialects in the area. This also corresponds to the answer we may provide to question 2, about age differences in the data: The younger speakers accept less non-V2 than the older speakers (cf. Table 8), which also suggests that there is a reversal of the historical development taking place. Returning to the lack of non-V2 in production, it is also possible that aspects of the recording situation may have contributed to this, since one of the investigators was always present during the recordings, and despite our attempts to create a relaxed atmosphere, the speakers may have felt that this was a relatively formal situation. In any case, due to this result, it is unfortunately not possible to answer the question whether there is any frequency difference between the two types of long wh-elements, disyllabic wh-words and wh-phrases.

Moving on to question 4, we find that we are able to say something about speakers’ acceptability judgments in this respect, as there is in fact little correspondence between the production and judgment data. As noted in section 6.1, speakers make clear and statistically significant distinctions between different types of wh-elements in relation to non-V2, accepting less of this word order the more complex the wh-element is. This provides further evidence for the historical account advocated in Westergaard (2009a), which argued that the development progressed according to the complexity of the wh-element. We now have some judgment data confirming the production data of previous studies.

The combination of the findings from production and judgment data is perhaps the most important one from the perspective of historical linguistics. When studying the diachronic development of a language, one typically only has production data available, but based on this alone, we often make claims about the grammar of the language at a particular stage. The discrepancy found between production and judgments in the present study shows that there is reason to be cautious about this, especially when analyzing phenomena that are infrequent in historical texts. In the present production data, non-V2 in questions with long wh-elements is attested only 0.7%, and such a low percentage in a historical text would presumably lead to an analysis where this phenomenon is considered ungrammatical. A case in point is non-V2 word order in declaratives with the initial elements pa/ponne ‘then’ in the history of English (cf. section 2), which is typically assumed to be ungrammatical and these elements thus argued to have a different categorical status in OE (as operators triggering obligatory V2). But the present data show that many speakers who produce exclusively or predominantly V2 in questions with long wh-elements nevertheless also accept non-V2 in such contexts (cf. Table 4).

This leads us to question 3, about possible distinctions between the two contact dialects (Manndalen and Nordreisa) and the other dialects in Northern and
Southern Troms (Djupvik and Senja). If we only consider production (which admittedly, is relatively sparse), there is no difference to be detected between the four varieties: In addition to the general non-existence of non-V2 with long wh-elements, there is extensive non-V2 in questions with short wh-elements, attested with somewhat varying proportions in all areas. This is presumably due to individual variation, which has also been found in other studies. It is only when we compare the scores for the acceptability judgments of non-V2 that we find differences between the dialects: With disyllabic wh-elements, the average acceptability rate is 5.7 and 5.4 in the two contact areas, while it is 3.9 in Djupvik and only 2.1 in Senja.

Finally, we consider question 5, about the effect of information structure. Previous studies have found that the choice of V2 vs. non-V2 in questions with the short wh-elements is dependent on the status of the subject as given or new and/or focused information. Unfortunately, since there are so few examples of non-V2 with long wh-elements, it is impossible to investigate the effect of information structure, although the simple fact that it is possible to talk about informationally given subjects (which is typically the case in these recordings) without using non-V2 suggests that information structure does not play any major role in questions with the long wh-elements. This is supported by the judgment data (cf. Table 6): These show that, while there is an effect of subject type on word order in questions with the short wh-words, there seems to be no correlation between pronominal subjects and non-V2 with the long wh-elements. This is a somewhat surprising finding, given the fact that, when syntax allows two word orders, information structure typically becomes the factor that distinguishes between the two (e.g. Bresnan & Nikitina 2009).

However, an explanation of this last finding may be sought in another comparison with the history of English, where V2 breaks down at the end of the ME period. While in previous stages of the language, information structure had been the most important factor for the word order choice in declaratives, this effect is argued to give way to syntactic factors in late ME (van Kemenade & Westergaard 2012). The reason for this is that the substantial variation at this time (presumably partly due to dialect contact) no longer gave unambiguous cues for the information structure patterns. Likewise, the situation in Northern Troms, with increased dialect contact in recent decades, may have led to the breakdown of previously existing information structure patterns in questions with long wh-elements. The fact that these patterns are no longer intact indicates that non-V2 in this context is no longer a truly productive part of these contact dialects, despite considerable acceptance of this phenomenon by present-day speakers.

8. Summary / Conclusion

In this paper I have investigated both production and judgment data collected from speakers of several North Norwegian dialects, focusing on the word order variation found in wh-questions, i.e. V2 or non-V2. The findings show that non-V2 word order is widespread in questions with monosyllabic wh-words, thus
confirming previous studies. This word order is also fully accepted in questions with disyllabic and phrasal wh-elements by speakers in two areas where there has been extensive language contact with Saami and Kven/Finnish, both non-V2 languages, as previously argued by Nilsen (1996) and Sollid (2003). Furthermore, there is a statistically significant difference in the acceptability of non-V2 word order corresponding to the complexity of the wh-element, providing support for the historical account of the development from V2 to non-V2 in Westergaard (2009a). The age differences found in the present data suggest that this process is being reversed in the language contact areas, as argued by Sollid (2003).

The most surprising finding in the present study is that the production of non-V2 in questions with phrasal wh-elements is negligible (0.7%). The large discrepancy between judgments and spontaneous production thus shows the importance of collecting different kinds of data in order to get a full picture of a grammatical system at a certain stage. It also indicates that language acquisition proceeds despite extremely low input frequencies, as claimed in Westergaard (2008). Finally, it suggests that one should exhibit some caution when analyzing phenomena that are attested with low frequencies in historical data.

References


Taraldsen, Knut Tarald. 1985. ‘Ka du trur?’ [What do you think?]. In Tove Bull and Anton Fjellstad (eds.), *Heiderskrift til Kåre Elstad* [Festschrift for Kåre Elstad], University of Tromsø, Department of Languages and Literature, 37-47.


Westergaard, Marit, Øystein A. Vangsnes and Terje Lohndal. 2017. ‘Variation and change in Norwegian wh-questions: The role of the complementizer som.’ Linguistic Variation 17:1, 8-43.

Marit Westergaard
UiT – The Arctic University of Norway
Department of Language and Culture/CASTL
9037 Tromsø
Norway
marit.westergaard@uit.no