The (Non-) Effect of Input Frequency on the Acquisition of Word Order in Norwegian Embedded Clauses

Marit Westergaard and Kristine Bentzen
University of Tromsø, CASTL

1. Introduction

In this paper we investigate the acquisition of word order in Norwegian embedded clauses. More specifically, we look at how Norwegian children acquire verb placement in embedded *wh*-questions and all types of embedded clauses containing negation or an adverb. We also consider some data of child-directed speech, as we believe that it is important for studies in first language acquisition take into account the role of input in language development. Whether or not one assumes that children are innately endowed with something like Universal Grammar (UG), it is obvious that certain parts of language, such as vocabulary and phonetic inventory, have to be learned from the primary linguistic data (PLD). Lately, the effect of input on the acquisition process has received considerable attention. In much recent work on language acquisition within the constructivist framework (e.g. Tomasello 2003, Theakston et al. 2004), it is argued that input frequency is vital to understanding both the order of acquisition of particular constructions and children's non-target-consistent production. In fact, it is often argued that children's early multi-word utterances are not the result of rule-governed behavior at all, but simply the result of a functionally-based distributional analysis of the input. Thus, in children's early production, there is little or no syntactic structure underlying their utterances. This stands in stark contrast to the generative approach to language acquisition where it is commonly assumed that UG provides the child with the necessary functional structure and constraints, and that all the child needs to do is learn lexical items and the setting of certain language-specific parameters.

Here we argue that input frequency plays a role in the acquisition of word order, but only in combination with other factors. Thus, our approach is in line with several of the contributions to the present volume, e.g. Roeper, Kupisch, and Bohnacker. The children in this study are acquiring a Northern dialect of Norwegian spoken in the city of Tromsø. Two constructions with similar input frequencies are investigated: embedded questions and (all) embedded clauses containing negation or an adverb. Both constructions are very infrequent in the input. It is shown that children make mistakes in embedded clauses with negation or an adverb, overgeneralizing the word order from main clauses (producing structures with verb movement across negation or an adverb). On the other hand, they do not overgeneralize main question word order into embedded questions (producing structures with verb movement across the subject). This is accounted for within a Split-CP model of clause structure and a weak continuity structure-building approach to language acquisition, where input and economy principles interact in the development of word order. Thus we argue that the lack of input cues for the target-consistent word order in itself is not the reason for children's non-target-consistent production. However, low input frequency may be one of the contributing factors causing the target word order in embedded clauses with negation or an adverb to be acquired relatively late. While children have to rely on input to acquire the word order in lower domains of the clause, UG provides them with the information that embedded questions are different from main clause

questions with respect to illocutionary force. Consequently children do not project the same functional architecture for the two constructions, and overgeneralization of features from main to embedded questions should therefore be impossible.

The paper is organized as follows. In the next section we outline the relevant word order facts of Norwegian. In section 3 we present the acquisition data from the children in this study, while section 4 contains an investigation of some of the adult data in the acquisition corpus. Then, in section 5 we analyse the child data within an economy-based account of language acquisition. Here we also discuss the role of input frequency in the acquisition process. Section 6 is a summary with concluding remarks.

2. The word order of Norwegian

Norwegian is a VO language with a rule of verb second (V2), which means that the finite verb has to appear in second position in all main clauses. This is standardly analysed as verb movement to the topmost head position of the clause, C (see e.g. Vikner 1995). This can be seen in both subject-initial and non-subject-initial clauses, as illustrated in in (1) and (2), respectively. Norwegian also shows V2 effects in main *wh*-questions, as in (3).

(1) John **liker ikke** tog.
John likes not trains
'John does not like trains.'

(2) Ifjor **dro John** til Peru to ganger.
last-year went John to Peru two times
'Last year John went to Peru twice.'

(3) Hvorfor **liker John** tog?
why likes John trains
'Why does John like trains?'

In embedded clauses, the finite verb remains within the VP. This is illustrated in (4) where the verb has to follow negation. As we see in example (5), most embedded clauses (such as embedded *wh*-questions) do not allow V2, as the verb must also follow the subject.

(4) Jeg kjenner en mann [som **{*liker} ikke {liker}** tog].
I know a man who likes not likes trains
'I know a man who doesn't like trains.'

(5) Har du hørt [hvorfor **{*liker} John {liker}** tog]?
have you heard why likes John likes trains
'Have you heard why John likes trains?'

There are some exceptions to the generalization that the verb does not move in embedded contexts. First of all, Norwegian in general optionally allows verb movement in *that*-clauses that are complements of so-called bridge verbs (*say, know, tell, believe,* etc.). In the subject-initial embedded clause in (6) verb movement past negation is optional (cf. the subject-initial main clause in (1)). In the non-subject-initial embedded clause in (7) V2 is obligatory (cf. the non-subject-initial main clause in (2)).

(6) Hun sier [at John **{liker} ikke {liker}** tog lenger].
she says that John likes not likes trains longer
'She said that John didn't like trains any longer.'

(7) John sa at [ifjor **dro han** til Peru to ganger].
John said that last-year went he to Peru two times
'John said that he was in Peru twice last year.'

Although verb movement past negation is accepted in sentence (6), the preferred option in Norwegian is generally to leave the verb in the VP, according to Garbacz (2005).[1] However, as shown in (7), these constructions allow embedded topicalization, and then subject-verb inversion is obligatory. Thus, embedded clauses under bridge-verbs, like those in (6) and (7), are arguably contexts where embedded V2 is available.

Secondly, Bentzen (2005, to appear) has shown that several Northern Norwegian dialects also allow verb movement past certain adverbs in non-V2 contexts, such as embedded *wh*-questions, relative clauses, and adverbial embedded clauses. The Tromsø dialect, which is the target dialect of the children in the current study, allows finite auxiliaries preceding certain adverbs such as *ofte* 'often' and *allerede* 'already,' as illustrated in (8) below. Crucially, however, verb movement past negation is never possible in these contexts, as we see in (9). Again, verb movement is not the preferred option, and we may thus assume that the word order in (8) is relatively infrequent in the input.

(8) Vi begynte å bli spente nå
we began to become excited now
... ettersom vi **ville allerede** kunne vite resultatet på fredag.
... as we would already could know result.the on Friday
'We started to get excited now as we would be able to know the result already on Friday.'

(9) *... ettersom vi **ville ikke** kunne vite resultatet før på fredag.
... as we would not could know result.the before on Friday
'... as we would not be able to know the result until Friday.'

There are also some exceptions to the V2 requirement in main clauses. In several Norwegian dialects verb movement is optional in main *wh*-questions (cf. Westergaard and Vangsnes 2005, Vangsnes 2005). The Tromsø dialect has optional V2 in main *wh*-questions with the monosyllabic *wh*-words *kem* 'who,' *ka* 'what,' and *kor* 'where' (cf. Westergaard 2003), as we see in (10)-(12). The example in (13) shows that in the non-V2 cases the verb also has to follow negation, in line with the restriction on moving the verb past negation in non-V2 contexts (cf. example (9) above):

[1] Garbacz (2005) searched the Big Brother corpus of spoken Norwegian (*Big Brother-korpuset*), and according to his findings, the order S-Neg-V_{fin} is by far the most frequent word order in *that*-clauses, constituting 64% of embedded clauses with negation. The second most common word order is S-V_{fin}-Neg, occurring 29% of the time, whereas Neg-S-V_{fin} is the most infrequent pattern (7% of the time). He also ran a similar search in the Oslo corpus of Standard (written) Norwegian, *Bokmål* (*Oslo-korpuset av taggede norske tekster*). In this corpus, the preference for the S-Neg-V_{fin} order was even more significant. As much as 96% of the sentences had this order, whereas the other two orders were each used only about 2% of the time.

(10) Kem {**like} han John {like}** best?
who likes he John likes best
'Who does John like the best?'

(11) Ka {**er} favorittlandet ditt {er}**?
what is favourite-country.the yours is
'What is your favourite country?'

(12) Kor {**parkerte} han {parkerte}** bilen henne?
where parked he parked car.the LOC
'Where did he park the car?'

(13) Kem han **{*lånte} ikke {lånte}** penga til?
who he lent not lent money to
'Who didn't he lend money to?'

For *wh*-questions with long *wh*-phrases and the disyllabic question words *koffer* 'why,' *korsn* 'how,' and *katti* 'when', V2 is obligatory in main *wh*-questions in this dialect, as illustrated in (14)-(16):

(14) Koffer **gikk han** hjem så tidlig?
why went he home so early
'Why did he go home so early?'

(15) Korsn **visste du** kem det var?
how knew you who it was
'How did you know who it was?'

(16) Katti **lande flyet ditt** i Lima?
when lands plane.the yours in Lima
'When does your plane arrive in Lima?'

Summing up, Norwegian in general has obligatory verb movement to the second position in all kinds of main clauses, but no verb movement in embedded clauses. However, embedded V2 is possible in *that*-clauses embedded under bridge verbs. Furthermore, the Tromsø dialect discussed in the current study optionally allows verb movement past certain adverbs, but not past negation in non-V2 context, such as embedded *wh*-questions, relative clauses, and embedded adverbial clauses. In addition, this dialect has optional V2 in a subgroup of *wh*-questions (those introduced by *kem* 'who,' *ka* 'what,' and *kor* 'where').

In this study we focus on how Norwegian children growing up in Tromsø acquire verb placement in embedded clauses. More specifically, we investigate how they acquire verb placement with respect to the subject in embedded *wh*-questions, as in (7), on the one hand, and how they acquire verb placement with respect to negation and adverbs in (all) embedded clauses and non-V2 main *wh*-questions, as in (6) and (15), on the other. We consider data from three very young Norwegian children below the age of 3, as well as data from two older children, up to the age of 8. In order to consider the potential effect of input frequency, we also investigate a sample of some adult data in an acquisition corpus.

3. Child data

3.1 Previous studies on the acquisition of embedded clauses

Previous studies on the acquisition of word order in V2 languages suggest that verb placement in main clauses is in place from very early on. Westergaard (2005) shows that this is also the case for Norwegian children. As soon as multi-word utterances

appear in the child data, verb movement generally applies in non-subject-initial clauses, questions, and subject-initial clauses with negation or adverbs. Such early acquisition of V2 in main clauses is also attested in Swedish (Santelmann 1995, Platzack 1996), Dutch (Jordens 1990), German (Poeppel and Wexler 1993, Müller 1996), and Lucernese Swiss German (Schönenberger 2001).

Findings concerning the acquisition of word order in embedded clauses are more varied. Clahsen and Smolka (1986) find that German-speaking children correctly place the verb clause-finally in embedded clauses already in the very first instances of such clauses. Penner (1996) reports on data from a Bernese Swiss-German child which indicate that there is correct verb placement in embedded contexts (clause-finally) until about the age of 3;2, but this stage is followed by a period of a few months when the child produces embedded clauses both with and without movement. Occasional incorrect verb movement in German embedded clauses is also reported for monolingual German children by Gawlitzek-Maiwald, Tracy, and Fritzenschaft (1992), and for bilingual German-English children by Döpke (1998). Håkansson and Dooley Collberg (1994) have shown that Swedish-speaking children seem to incorrectly move some modal auxiliaries in Swedish embedded clauses. Finally, Schönenberger (2001) found that her two Lucernese Swiss German subjects consistently moved the finite verb in a non-target-like manner in embedded clauses. This pattern occurred for an extended period, until the age of 4;11, before the verb-final pattern took over.

Previous findings are thus inconclusive, some studies suggesting that verb placement in embedded clauses is unproblematic, whereas others report this to be an area where children make mistakes for an extended period of time.

In the following sections we present data from Norwegian-speaking children indicating that there is evidence of overgeneralization of verb movement past negation or an adverb into constructions that do not allow verb movement in the target language. However, non-target-like verb movement past subjects is not attested in the children's production.

3.2 Young Children

In this section we provide some evidence from three very young Norwegian children. These data come from a relatively large corpus consisting of altogether 70 one-hour recordings of three children between the age of approximately 1;9 and 3;3.[2] Given the young age of these children, there are not many instances of embedded clauses in the data. However, there are a few relevant examples in the later files, altogether 108 embedded questions, 28 embedded clauses with negation or an adverb and one non-V2 main *wh*-question with negation.

Let us start with the 28 embedded clauses with negation or an adverb (all declaratives). As many as 15 of these had to be excluded from our discussion because they are unclear with respect to the question of verb movement. In three of these examples, although they display the target non-V2 word order, the verb involved seems to be non-finite, and judging from the context, there is a modal missing in the

[2] Apart from ten files that have been recorded and transcribed by the first author, the corpus has been collected by Merete Anderssen. See Anderssen (2005) and Westergaard (2005) for details about the corpus.

structure.[3] The remaining 12 of the excluded examples display the most common word order in the embedded clauses, viz. the order where negation *ikke/ikkje* 'not' occurs immediately following the complementizer (if present), i.e. above the verb as well as the subject, as illustrated in (17). This Neg-S-V_{fin} word order is also possible in the adult grammar, although as mentioned in section 2, it is much less frequent than S-Neg-V_{fin} (see footnote 1). It is also unclear exactly what position negation is attached to in such sentences.[4] Nevertheless, we must conclude that these sentences cannot reveal anything about possible verb movement.

(17) nei ho skal passe på mæ **ikkje reven komme** å ta mæ. (Ina.18, 2;8.12)
no she shall watch on me not fox.the comes to take me
'She is to watch out so the fox doesn't come and take me.'
Target form: Ho skal passe på mæ så ikkje reven kommer og tar mæ.

The 13 remaining embedded clauses with negation were relevant to the current study. Four of these 13 sentences can be said to be true examples of target-consistent word order without verb movement in embedded contexts. Two of these are illustrated in (18) and (19), where negation appears between the subject and the verb, indicating that no verb movement has taken place.

(18) ikke da [//] at det da **ikke blir** stramt. (Ole.18, 2;9.15)
not then that it then not becomes tight
'... that it doesn't get (too) tight.'

(19) bare når dem **ikke hold** på da dette dem xxx. (Ina.27, 3;3.18)
only when they not hold on then fall they xxx
'Only when they are not holding on, then they fall.'
Target form: Bare når dem ikke hold(er) (fast?), da dætt dem xxx.

But nine of the examples in the data do in fact display verb movement in embedded contexts. Five of these are complements to bridge verbs, one of which is illustrated in (20), which means that these are grammatical (but not preferred) with verb movement in the target language. The four remaining embedded clauses in the corpus exhibit verb movement in non-V2 contexts where it is clearly ungrammatical in the target-language. An example is given in (21).

(20) han sa han **ville ikke** spise <han> [?]. (Ann.17, 2;8.4)
he says he would not eat him

(21) det er ho mamma som **har også** tegna. (Ina.26, 3;2.5)
it is she mommie who has also drawn

Another construction where the target language does not show verb movement is non-V2 main *wh*-questions, as we saw in (13). In the data from the younger

[3] In the dialect the children in this study are acquiring, the *–er* ending of the present tense has been reduced to *–e*, which means that the infinitive and the present tense verb forms are identical for the majority of verbs (the two classes of regular verbs, which make up approximately 96% of all verbs in the language, according to Endresen and Simonsen 2001).

[4] A relatively standard assumption is that negation and adverbs may be adjoined to VP as well as to TP, but not AgrSP. It is, however, possible that certain light adverbs and negation may be adjoined even higher, as suggested in Holmberg (1993). In sentences such as (17) in child language, it could either be the case that negation is adjoined to this higher position, or alternatively, that the child has failed to move the subject (see Westergaard 2005).

children there is one such question containing negation, and in this example the verb has indeed moved across negation. The word order in (22) indicates that there is overgeneralization of verb movement in these cases, not to a position in the clause structure above the subject, but presumably to a lower functional head.

(24) kem som **vil ikkje** være ilag med han? (Ina.25, 3;1.8)
who that will not be together with him
'Who doesn't want to be with him?'
Target form: Kem som ikkje vil være i lag med han?

Thus, 10 out of 14 relevant examples show that all three children seem to prefer verb movement over a word order without movement, both in cases where it is completely ungrammatical in the target grammar and cases where it is only dispreferred (in complements of bridge verbs). This is interesting, as it goes against a minimalist account, where movement is always considered to be a more costly operation than no movement. It could of course be the case that these examples are restarts - i.e. that they are biclausal structures with two main clauses. Such an explanation is also supported by the fact that none of these clauses are introduced by complementizers. However, within a minimalist account it is generally assumed that children start out with the least costly approach to word order, viz. a structure with no movement (cf. Platzack 1996, Roberts 1999), and that they only produce movement if there are strong and consistent cues for this in the input. Given that the option not to move the verb should be available to the children (as they do produce some embedded clauses with target-consistent word order), it is surprising from the point of view of a minimalist account that they seem to prefer the least economical structure in these cases. This indicates that economy presumably interacts with other factors in the acquisition of word order.

We next consider embedded questions, of which there are a total of 108 examples in the corpus. What is striking about these clauses is that virtually all of them occur with target-consistent non-V2 word order; that is, with no verb movement across the subject. Examples of these embedded questions are provided in (23)-(25), both from relatively early and relatively late files in the corpus.

(23) se her ka **Ina gjør**. (Ina.04, 1;11.22)
look here what Ina does

(24) Ann vet ikke kor **han er** henne. (Ann.09, 2;2.19)
Ann know not where he is LOC
'Ann doesn't know where he is.'

(25) skal æ vise # korsen **man trøkke** på knappen? (Ole.20, 2;10.15)
shall I show ... how one pushes on button.the
'Do you want me to show (you) how you push the button?'

There is only one potential exception to the lack of overgeneralization of V2 word order in embedded *wh*-questions, and this is illustrated in (26):

(26) du, ser du ka **er det** der sånn der der? (Ina.27, 3;3.18)
you see you what be that there such there there
'You, do you see what that is/do you see: what is that?'

A possible explanation to the word order in (26) is again that the child is producing a biclausal structure, i.e. that there should be a restart between the question word *ka* 'what' and the preceding part of the sentence, as illustrated by the alternative translation. This analysis is not completely implausible, especially given the linguistic context in the file. It is clear that the adult responding to the child's question has interpreted this as a *wh*-question and not as a *yes*/*no*-question, the reply to (26) being a specification of what "that" refers to.[5] In any case, 1 out of 108 examples does not constitute evidence that there is a rule of verb movement across the subject in embedded questions in the child's internalized grammar, and we therefore conclude that in general there is no evidence of overgeneralization of V2 from main into embedded questions.

Another possible type of verb movement in these embedded *wh*-questions would be movement across negation or an adverb, as we saw in the embedded declarative clauses above. Unfortunately, none of the 108 embedded questions in this corpus include negation or an adverb.

This section has investigated the occasional examples of embedded clauses in early child data. To summarize so far, virtually no cases of verb movement across the subject were attested in embedded questions. In embedded clauses with negation or an adverb, on the other hand, the majority of cases show overgeneralization of verb movement. This was the case in four clear examples, illustrated in (21), and five further examples of embedded clauses under bridge verbs, as shown in (20). In comparison, only four examples of target-consistent word order without verb movement were found in the child data, (18)-(19). Furthermore, the only relevant example of a non-V2 main *wh*-question also displayed verb movement across negation, (22). These findings are summarized in Table 1, and as we can see, the children moved the verb across negation in 10 of the (preferred) non-V2 contexts, whereas they left it in the target-like position following negation only 4 times.

Table 1: Overview of word order in embedded clauses with negation or an adverb in the corpus of three Norwegian children, age approximately 1;9 to 3;0.

	(S)-V_{fin}-Neg/Adv	S-Neg/Adv-V_{fin}
Embedded declaratives	**4**	4
Embedded declaratives under bridge verbs	**5**	0
Non-V2 *wh*-question	**1**	0
Total	**10**	4

3.3 Older Children

The investigations of the two older children are based on sporadic recordings and diary notes from two boys, Henning (2;4.4 - 8;0.17) and Iver (1;8.10 - 5;9.15), as well as the results from a small experimental study with the same two children at the age of 8;0.20 and 5;9.18. The patterns reported for the very young children in the above section are confirmed in the data from the older children.

In the recordings and diary notes several embedded clauses with negation are attested, especially from the age of around 4 years. The data show that the children at

[5] The reply from the investigator is provided in (i). It should be noted, however, that pragmatically, also a *yes/no*-question could call for such an answer.

(i) det [/] det kalles for hyena. (INV, File Ina.27)
it it is-called for hyena
'It is called a hyena.'

this stage display both verb movement and V in-situ in a target-like manner in complements of bridge verbs, as illustrated in (27) and (28). Both of these examples are acceptable, but the word order in (27) is dispreferred in the adult language.

(27) æ vet at æ **har ikke** gjort det. (Henning 4;8.13)
I know that I have not done it
'I know that I haven't done it.'

(28) æ sa at han **ikke sku** ... (Henning 4;8.0)
I said that he not should
'I said that he shouldn't...'

In Bentzen (2003) it was shown that these children also move the verb past negation in many other types of embedded non-V2 contexts. Such movement was attested in relative clauses, as illustrated in (29)-(30), in adverbial subordinate clauses, (31), and in embedded *wh*-questions, as illustrated in (32)-(33). There is also one instance in the data of verb movement in a non-V2 main *wh*-question, given in (34). The following examples are all ungrammatical in the adult language:

(29) æ like alt som **er ikke** sterkt og alt som er sterkt.(Henning 4;2.7)
I like everything that is not hot and everything that is hot
'I like everything that isn't hot, and everything that is hot.'

(30) du må få dæ en biffkniv som **er ikke** sånn. (Iver 5;8.16)
you must get you a steak.knife that is not like-that
'You need to get a steak knife that isn't like that.'

(31) æ må ta på ullæsta for at æ **skal ikke** bli så kald. (Iver 4;11.29)
I must take on wool.socks for that I shall not get so cold.
'I need to put on wool socks in order to not get too cold.'

(32) når han Iver **er ikke** her så kan æ ta med (Henning 4;6.27)
when he Iver is not here then can I take with
den store skjeia.
the big spoon
'When Iver isn't here, I can use the big spoon.'

(33) men æ lik'ikke det når det **er ikke** sånn. (Henning 4;7.16)
but I like-not it when it is not like-that
'But I don't like it when it isn't like that.'

(34) kem som **var ikke** helt i form? (Henning 4;5.0)
who that was not completely in shape
'Who wasn't feeling too well?'

Several embedded *wh*-questions were also attested in the recordings and diary notes of the older children. As (35)-(36) show, none of them exhibit verb movement past the subject.

(35) vet du ka **det her er**, tante? (Henning 3;11.12)
know you what this here is aunt
'Do you know what this is, auntie?'

(36) æ vet korsn **dem lage** et hus sånn her. (Iver 4;7.10)
I know how they make a house like-that here
'I know how to make a house in this way.'

Evidence from the sporadic recordings and the diary notes shows that the overgeneralization of main clause word order into embedded declaratives with negation or an adverb found in the very young Norwegian children is also attested for the two older children at around the age of 4-5 years. Furthermore, the lack of such overgeneralization into embedded *wh*-questions found in the young children is replicated in the data from the older children. However, the sporadic recordings and the diary notes do not give any indications as to how frequently these two children display the non-target-like verb movement, nor for how long such patterns persist in the children's grammars. Therefore, a small experiment was conducted with the two children at the age of 5;9.18 and 8;0.20.

The small experimental study was designed to elicit embedded *wh*-questions with negation or an adverb. In the experiment, we introduced the children to the hippo Hårek. The children were told that Hårek was a very peculiar hippo who had three special features: (i) he claimed to have the best memory in the world, (ii) he did not talk to adults, and (iii) importantly, he would not respond to you unless you started your sentences with 'Do you remember...?'. The investigator (the second author) read a story with the children about a four-year-old boy, Karsten, who was ill and had to stay at home rather than going to kindergarten. The children were told that Hårek also knew the story, and that they were now going to test how much he remembered of it, asking questions starting with 'Do you remember...?'. We attempted to elicit altogether 16 embedded questions, 12 of which were supposed to contain negation or an adverb. The remaining four questions where included as fillers. The elicitation setup is illustrated in (37):

(37) INV: *So, Karsten couldn't go to kindergarten today, and that was because he was ill. Therefore he couldn't go to kindergarten. We remember that that was why, but ask Hårek whether he remembers why.*
CHILD: *Do you remember why Karsten didn't go to kindergarten today?*

The older child, Henning, included negation or an adverb in 11 of the 12 designated questions, and in all cases, negation or the adverb preceded the verb, as shown in (38)-(40):

(38) huske du koffer han Karsten **ikke var** i barnehagen? (Henning 8;0.20)
remember you why he Karsten not was in kindergarten.the
'Do you remember why Karsten wasn't in the kindergarten?'

(39) huske du koffer ho **ikke ville** kjøpe den potta?
remember you why she not wanted buy that pot
'Do you remember why she didn't want to buy that pot?'

(40) huske du koffer en mann **ikke fikk** kjøpe Løveungen?
remember you why a man not got buy Lion.baby.the
'Do you remember why a man didn't get to buy the Lion baby?'

The younger child, Iver, included negation or an adverb in only 8 of the 12 designated questions, and in 7 of these 8, he produced the non-target-like word order with the verb preceding negation or the adverb, as illustrated in (41)-(43):

(41) huske du koffer han Karsten **var ikke** i barnehagen? (Iver 5;9.18)
remember you why he Karsten was not in kindergarten.the
'Do you remember why Karsten wasn't in the kindergarten?'

(42) huske du koffer dama **ville** **ikke** kjøpe en nattpotte?
remember you why lady.the wanted not buy a night.pot
'Do you remember why the lady didn't want to buy a chamber pot?'

(43) huske du koffer Løveungen **var** **ikke** til salgs?
remember you why Lion.baby.the was not to sale
'Do you remember why the Lion baby wasn't for sale?'

Neither of the children ever moved the verb past the subject in these embedded *wh*-questions. Thus, the sporadic recordings, the diary notes, and the small experiment with older children constitute evidence that children up to the age of (at least) around 6, optionally move the verb past negation and adverbs in non-V2 contexts. Furthermore, the small experiment shows that they do not move verbs past subjects in embedded questions.

One possible explanation for the children's word order patterns in embedded clauses could be that there is a word order change taking place in the language. This, however, seems unlikely, given that most of the examples of verbs preceding negation or adverbs in the diary notes and sporadic recordings are from the older child Henning at the age of approximately 4-5. The fact that at the age of 8 he hardly uses this word order anymore suggests that this is a feature of a certain developmental stage in the acquisition process, rather than e.g. an indication of a syntactic change taking place in the dialect.

Summing up the investigation of both very young and somewhat older Norwegian children, we found substantial evidence for overgeneralization of verb movement past negation and adverbs from main clauses into embedded clauses, at least up to the age of 6. Such verb movement is generally not accepted in the target language. In the few contexts where it is possible in the target language, viz. past certain adverbs in non-V2 contexts, and past both negation and adverbs in *that*-clauses embedded under bridge verbs which allow embedded V2, verb movement is the dispreferred option. Within a minimalist account of language acquisition it is at first sight unexpected that children should prefer verb movement where it is not allowed or dispreferred in the target language. Assuming that economy principles play an important role in language acquisition, one would expect children to avoid costly operations such as verb movement, unless there is strong and consistent evidence for such movement in the input.

Furthermore, it does not seem to be the case that the children in the study are simply applying main clause V2 word order in embedded clauses in general. This is evident from the fact that they do not move the verb past the subject in embedded *wh*-questions in parallel with main *wh*-questions. Thus, what needs to be explained is why children overgeneralize verb movement past negation and adverbs but not past subjects. In the following sections we discuss possible reasons for this asymmetry in the acquisition of verb placement. An important question is whether this is a result of asymmetries in the frequency of the relevant constructions in the input, or whether other factors may play a role here.

4. Input Frequencies

As mentioned in the introduction, there has recently been an increased interest in the role of the input, within functional as well as formal approaches to language acquisition, and especially within the constructivist framework, where input is often argued to be the sole explanation for acquisition orders and children's error patterns.

An example of a constructivist approach relevant to the constructions at hand is Tomasello (2003), who argues that children's early production of embedded clauses provide no evidence for a hierarchical structure in children's linguistic systems. Investigating examples from German child language, he finds that early embedded clauses always appear with the same matrix verbs, normally only two or three different ones. Therefore these are better analyzed as linear constructions, he argues, where the matrix verb is simply stuck onto the beginning of the clause, which remains a main clause structure. Applying this line of reasoning to the Norwegian child data in the previous section, it could be argued that the embedded clauses with negation are not really embedded constructions, but rather main clauses with an initial chunk which looks like a matrix clause. Thus, we find main clause word order in these constructions with the verb preceding negation or an adverb, as was illustrated in e.g. (20) and (21). Presumably the V-Neg/Adv combination in the main clause would on this approach not be the result of verb movement, but simply a linguistic chunk which is reproduced from main to embedded clauses (or rather, structures which look like embedded clauses).

If children learn syntactic structure from input only, we would then expect to find the following frequencies of the relevant constructions in the input that the children in this study are exposed to: Embedded clauses with negation or an adverb should be infrequent in the input, since this is the clause type where children make word order mistakes for an extended period of time. On the other hand, main clauses with negation or an adverb should be relatively frequent, since this is where the V-Neg/Adv pattern that the children are overgeneralizing is found. Furthermore, embedded questions should also be quite frequent, as the children were found not to overgeneralize the V-Subject word order found in main *wh*-questions.

Obviously, it is possible to argue that children's early utterances have more syntactic structure than what is normally assumed within a constructivist approach and still argue for a frequency effect. On such an approach within a generative framework, the word order of main clauses would be the result of verb movement, and because of a high frequency of main clauses with V2, this type of movement would then be overgeneralized to embedded clauses. That means that frequency would override economy in this case, since, as discussed above, a common idea within the minimalist framework is that syntactic movement is always more costly than the lack of movement. For a frequency effect to play a role here, one would expect to find exactly the same input frequencies as was sketched above for the constructivist approach: Embedded clauses with negation or an adverb should be considerably less frequent than the corresponding main clauses, while embedded questions are expected to be quite frequent.

In order get an indication of what child-directed speech may consist of in terms of frequency of syntactic constructions, some samples of the adult material from the Tromsø corpus were investigated in detail. First of all, one file (corresponding to approximately one hour of spontaneous speech) was hand-searched and all complete sentences of the investigator (INV) were counted. In this file, the investigator produced a total of 793 utterances, out of which there were 668 complete clauses, 554 matrix and 114 embedded clauses. There are altogether 123 subject-initial main clause declaratives in the sample, 43 of which contain negation or an adverb, see Table 2. This means that the evidence for verb movement across negation or adverbs in main clauses makes up 6.4% of the total input. Furthermore, there are as many as 337 examples of questions and non-subject-initial declaratives (50.4%), providing the child with evidence for verb movement across the subject. This means that there is

ample evidence in the input that Norwegian is a V2 language in main clauses, see Westergaard (forthcoming) for a more detailed analysis. Similar findings have been attested for much larger samples of Swedish input data in Josefsson (2004), altogether 14,033 adult utterances, where V2 constructions such as *yes*/*no*-questions are attested in 22-28% of all utterances, and non-subject-initial declaratives 12-27%.

Table 2: Overview of evidence for V2 and non-V2 in a sample of child-directed speech, the investigator in the file Ole.14 (age of child 2;6.21), with percentages calculated relative to the total number of complete (matrix and embedded) clauses (N=668).

Evidence for V2		**Evidence for non-V2**	
Subject-initial decl. with Neg/Adv	6.4% (43)	Embedded clauses with Neg/Adv Non-V2 *wh*-questions with Neg/Adv	0.9% (6) 0.1% (1)
Non-subject-initial decl. and questions	50.4% (337)	Embedded questions	1.6% (11)

The evidence that the verb does not move past negation or adverbs in non-V2 contexts should be expressed in all embedded clauses as well as in non-V2 main *wh*-questions with negation or adverbs. These constructions are indeed very infrequent in the input. As illustrated in the right-hand column of Table 2, the investigator produces only six embedded clauses with negation or an adverb in the file, corresponding to 0.9% of the input in the sample. One of these sentences is illustrated in (44) and could be compared to the non-target-consistent child utterances in (20)-(21) and (29)-(33) above. Furthermore, there is only one example of a non-V2 main *wh*-question with negation, which increases the evidence for the lack of verb movement across negation, but only by 0.1%. This example is given in (45) and should be compared with the non-target-consistent child utterances in (22) and (34) above.

(44) pass på at den **ikkje faller** over. (INV, file Ole.14)
watch on that it not falls over
'Watch out so it doesn't fall over.'

(45) kem som **ikkje får** kjøre? (INV, file Ole.14)
who that not gets drive
'Who doesn't get to drive?'

This means that the total evidence for the lack of verb movement across negation or an adverb in Norwegian embedded contexts and non-V2 main *wh*-questions is attested only 1.0% in the input sample. Moreover, there is also an example of an embedded clause under a bridge verb with the word order V-Adv in this file, which is illustrated in (46). As discussed in section 2 above, these are also grammatical in the target language, further complicating the structures that have to be acquired by the child.

(46) æ trur han **må bare** sitte der. (INV, file Ole.14)
I think he must only sit there
'I think he just has to sit there.'

So far our predictions with respect to frequency seem to be borne out: The evidence for Neg/Adv-V word order in embedded contexts and non-V2 main *wh*-questions is extremely infrequent in the input (1.0%), and compared to the 6.4%

evidence for the opposite word order in subject-initial main clauses, it could be argued that the more frequent word order is overgeneralized to the less frequent one. The 50.4% evidence for V2 in non-subject-initial declaratives and questions, i.e. a word order where the verb precedes the subject, could be added to this, as these utterances provide the child with general evidence for verb movement in the language.

But what about the embedded questions, which were also predicted to be frequent in the input? As illustrated in Table 2, it turns out that the evidence that the verb does not move across the subject in embedded (non-subject) *wh*-questions is also very infrequent in the sample of adult data. The investigator produced only eleven such examples, making up as little as 1.6% of the input sample. One of these is provided in (47).

(47) vet du ka slags farge **det er**? (INV, file Ole.14)
know you what kind color that is
'Do you know what color that is?'

Since the input sample discussed here is quite small and also produced by only one person, a more focused search of larger samples of the corpus was made, in order to check whether more considerable differences in input frequencies could be attested between the two types of embedded constructions requiring non-V2 word order. More specifically, we searched for negation (no adverbs) and specific question words in the production of the investigator in files Ole.15-22, and in the production of one of the parents in files Ann.01-21 (MOT). As shown in Table 3, the investigator produced altogether 6,351 utterances. Out of these, there were 32 (0.5%) embedded clauses with negation, and no non-V2 main *wh*-question with negation. In addition, the investigator produced 66 (1.04%) embedded (non-subject) *wh*-questions. Ann's mother produced a total of 8,860 utterances, 39 (0.44%) of these were embedded clauses with negation, and 41 (0.46%) were non-V2 main *wh*-questions with negation. Furthermore, she produced 224 (2.5%) embedded (non-subject) *wh*-questions. Thus, the more focused search in the corpus also indicates that the evidence for not moving the verb in non-V2 contexts is relatively infrequent. For both adult speakers the number of embedded questions is somewhat higher than the total number of clauses providing evidence for the lack of verb movement across negation or an adverb, 32 (0.5%) vs. 66 (1.04%) for the investigator, and 80 (0.90%) vs. 224 (2.5%) for Ann's mother.[6]

Table 3: Overview of utterances providing evidence for non-V2 word order in samples of child-directed speech, the investigator (INV) in files Ole.15-22 (N=6351, all utterances) and the mother (MOT) in files Ann.01-21 (N=8860, all utterances).

Evidence for non-V2				
	Emb. clauses w/Neg	Non-V2 *wh*-questions w/Neg	Total	Embedded *wh*-questions
INV	32 (**0.5%**)	0	32 (**0.5%**)	66 (**1.04%**)
MOT	39 (**0.44%**)	41 (**0.46%**)	80 (**0.9%**)	224 (**2.5%**)

[6] It should be noted that the number of embedded questions may be somewhat inflated as an effect of the recordings, especially in the speech of the parents. In an attempt to make the children speak as much as possible, they frequently produce utterances such as the following:

(i) ... har du fortalt ho Merete ka du gjorde i går? (MOT, file Ann.03)
have you told DET Merete what you did yesterday

Thus, in the small hand-searched adult sample as well as in the focused search of the larger corpus samples there is a slightly higher percentage of embedded questions than embedded contexts with negation or adverbs. However, we doubt that a difference between e.g. 0.5% and 1.04% could be the only explanation for children producing a considerable number of non-target-consistent constructions in the former case and displaying a virtually error-free production in the latter. Why would 1.04% be enough input to acquire a certain word order, while 0.5% - or the 0.9% produced by the mother - is not? And even if the children's production were the result of differences in input frequency, one would expect such a small difference to have an effect only for a short period of time. However, as shown in section 3 above, the non-target-consistent word order produced by the children in embedded contexts is quite persistent, possibly lasting well into school age. We thus find it highly unlikely that frequency could be the sole explanation for this, and we therefore reject a purely constructivist approach to the child data. Moreover, we believe that such an approach would also have a problem explaining why the V-Subject word order of all main clauses with V2 does not overgeneralize to embedded questions. After all, in the small input sample investigated in Table 2, the difference in input frequency between main clauses with V2 (V-Subject) and embedded clauses (Subject-V) is as much as 50.4% vs 1.6%, which is much higher than the difference between main and embedded clauses with respect to the position of the verb in relation to negation or an adverb (6.4% vs. 1.0%). Thus, if input frequencies were responsible for overgeneralization from the relatively frequent V-Neg/Adv word order of main clauses to embedded contexts, we see no reason why the extremely frequent V-Subject word order should not overgeneralize in the same way.

We therefore want to argue that the results of our investigation of the input clearly reveal that other aspects of language acquisition such as complexity or economy must be invoked to explain the error patterns described in section 3. This will therefore be the focus of the next section.

5. An Economy-based Account

Having rejected an analysis which explains the children's performance solely by reference to input frequency, we will now turn to an account of the observed facts in terms of economy, complexity and to a certain extent, frequency. The framework we adopt is a Split-CP model of clause structure, and this will be outlined briefly in the next section. In section 5.2 we account for the appearance of the children's non-target-consistent word order in embedded contexts with negation by referring to an economy principle which causes them not to move elements higher up in a clausal structure than there is evidence for in the input. The reason why this does not apply in embedded questions will be related to the syntactic model we adopt, where main and embedded clauses have different clausal architecture, reflecting their different illocutionary force. Finally, in section 5.3 we will discuss some reasons why the children's errors in embedded clauses are so persistent, and here frequency will be argued to play a role.

5.1 The Theoretical Framework

We adopt a Split-CP model of clause structure, which is inspired by Rizzi's (1997) original model and later work on Italian syntax, e.g. Rizzi (2001), Benincà and Poletto (2004), but which is in many ways different from these accounts. The model was originally developed in Westergaard and Vangsnes (2005) and somewhat revised in Westergaard (2005, forthcoming). The most important aspect of the model is that

different clause types have different heads in the CP-domain, reflecting the illocutionary force of the clause type. For example, a *wh*-question is an Int(errogative)P(hrase), a *yes/no*-question a Pol(arity)P, and a declarative a Top(ic)P. The syntactic heads in the CP and IP domains of the clause that are relevant for the present discussion are provided in the bracketed structure in (48):

(48) $_{CP}$[(Int° Top° ...) ... (Wh°) Fin° $_{IP}$[InTop° T°

Another crucial aspect of this model for our present purposes is that embedded clauses have a restricted CP domain. That is, embedded declaratives are assumed to be bare Fin(iteness)Ps, while embedded questions are bare WhPs. This reflects the different illocutionary force of main and embedded clauses. For example, embedded questions are not 'real' questions and lack interrrogative force, and thus there is no Int° head present in the clausal structure.

The model was developed mainly to account for different types of V2 grammars in English as well as different Norwegian dialects, many of which have no strict V2 requirement in *wh*-questions, as mentioned in section 2. The main parametric tool of the model is the presence of a specific EPP *head* feature on individual functional heads in the CP domain, called [X°$_{EPP}$].[7] This feature must be lexicalized, a requirement which may be met by verb movement. Grammars differ with respect to whether a particular head is endowed with this EPP feature, which means that there are several sources for V2 word order. According to this model, Norwegian dialects which have no V2 requirement in *wh*-questions, e.g. the Nordmøre dialect described by Åfarli (1986), will be argued to have no EPP feature on the Int° head, but as they are strictly V2 in declaratives, this feature is present on the Top° head. English, which has subject-auxiliary inversion in all questions but (generally) no inversion in declaratives, has the opposite requirements on these two heads. The Tromsø dialect, which the children in this study are acquiring, is argued to have the EPP head feature on Int° as well as the Top° head. The former is necessary to account for V2 word order in *wh*-questions introduced by long *wh*-phrases (see examples (14)-(16) in section 2), while the EPP feature on the Top° head accounts for verb movement in all declarative sentences, across the subject in non-subject-initial declaratives and across negation or an adverb in subject-initial declaratives.[8] Finally, the distinction between main and embedded clauses with respect to the presence of C-heads accounts for the differences between main vs. embedded clause word order: The heads Fin° and Wh° are not endowed with the EPP feature, and consequently, there is no verb movement to the CP domain in embedded clauses in Norwegian.

For our present analysis we will also adopt a general theoretical framework of language acquisition which could be described as a continuity approach which includes structure-building (see Westergaard 2005). The continuity aspect of this is taken care of by a universal "pool" of possible functional categories, where rules for their relative order (and presumably a number of other rules and constraints) are provided by UG. In the process of language acquisition, children select categories

[7] This abbreviation refers to the Extended Projection Principle (EPP) of earlier versions of generative theory (originally from Chomsky 1982), which ensured that all clauses have a subject, Within the Minimalist model, e.g. Chomsky (1995), an EPP feature on a syntactic head will require that this head projects a specifier in order for the uninterpretable EPP feature to be deleted.

[8] The optional word order in questions with the monosyllabic question words is accounted for by another C-head, the head of the Foc(us)P (see Westergaard and Vangsnes 2005, Westergaard 2005).

from this universal set, based on principles of UG and cues in the input. Additionally, children need cues to know how the different functional projections are realized syntactically in their particular language, e.g. by verb movement triggered by the EPP feature. We will also argue that in this process children are guided by economy principles. One of these is the principle of structural economy proposed in the Lexical Learning Hypothesis of Clahsen, Eisenbeiss and Penke (1996), originally from Safir (1993). Another economy principle, which is crucial for our analysis of the child data at hand, is a principle of economy of movement (see also Westergaard 2005). These principles will ensure that children do not build more structure or move elements higher in the structure than there is evidence for in the input. This means that movement operations should initially target positions that are as low as possible in the clause structure. This corresponds to what is often found in early child language: To the extent that children produce non-target forms, they normally seem to be due to children producing *less* movement than in the adult language, see e.g. Schaeffer (2000) on the lack of scrambling in Dutch, or Radford (1994) on the lack of inversion in some English-speaking children's *wh*-questions. Superficially, this is of course the opposite of what we see in the acquisition data presented in this paper, and in the next section we will therefore consider this economy principle in more detail in relation to the Norwegian child data.

5.2 Economy of Movement

As mentioned in section 2, V2 word order is in place more or less immediately in those clause types that require it, and Norwegian children must therefore realize very early that their language requires some filled C° head. Still, the question could be asked whether early verb movement indeed targets the same head positions as in the adult grammar. In the syntactic model adopted here it is assumed that the verb in all Norwegian main clauses moves to the highest head, Int° in *wh*-questions, Pol° in *yes/no*-questions, and Top° in all declaratives, subject-initial as well as non-subject-initial clauses. According to the approach to language acquisition we are assuming, UG provides children with the knowledge that all main clauses have a CP domain, and specifies the head that is necessary to produce different clause types, e.g. Int° for *wh*-questions or Pol° for *yes/no*-questions. Thus, it is not unlikely that early verb movement in questions and non-subject-initial declaratives is in fact movement to the appropriate heads. In subject-initial main clauses, on the other hand, this is not immediately obvious.

Since subjects in the world's languages are not universally in the specifier position of the highest C-head (e.g. not in English), UG will not provide Norwegian children with the information that subjects are default topics in this language and move to SpecTopP. This must therefore be learned from input. Likewise, that the verb is attracted by the [$X°_{EPP}$] head feature on Top° and moves to the head of this functional projection must also be learned. Unfortunately, there is no clear evidence in subject-initial declarative main clauses that the verb (and accordingly also the subject) moves all the way to the TopP in Norwegian. Nevertheless, there should be ample evidence in the input that there is verb movement in these sentences, as illustrated by the relatively high frequency of main clauses with negation in the sample of child-directed speech investigated above, obviously displaying the target V-Neg word order (see Table 2). There were 43 examples in the speech sample, making up 6.4% of the total (43 out of 668 clauses). However, if we assume that children only focus on the relevant clause type when searching for cues, as the Split-CP model indicates (see Westergaard 2005, forthcoming), then the evidence for verb movement is much

higher, in fact as much as 35%, as the relevant figure to relate this to is the total number of subject-initial declaratives in the sample, which is 123. Because of this, children apparently realize very early that finite verbs move across negation (and adverbs) in subject-initial main clauses, and they produce target-consistent forms from the onset of the appearance of relevant constructions, as illustrated in (49).

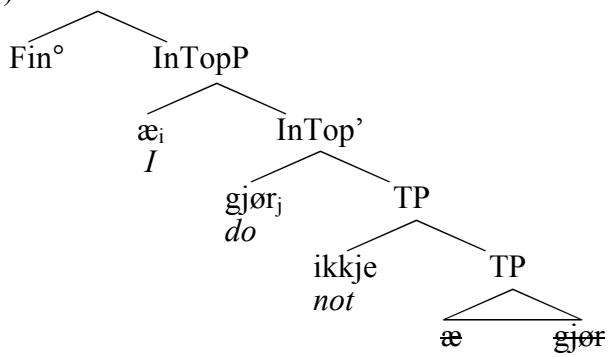
(49) æ gjør ikkje. (Ina.09, 2;2.12)
I do not
'I'm not doing (it).'

However, there are many other head positions than Top° that could serve as the landing site for verb movement in such sentences. Given that the economy principle discussed above ensures that children do not move constituents any farther than they have to, we argue that they will initially pick a lower head as the target for verb movement in these constructions.

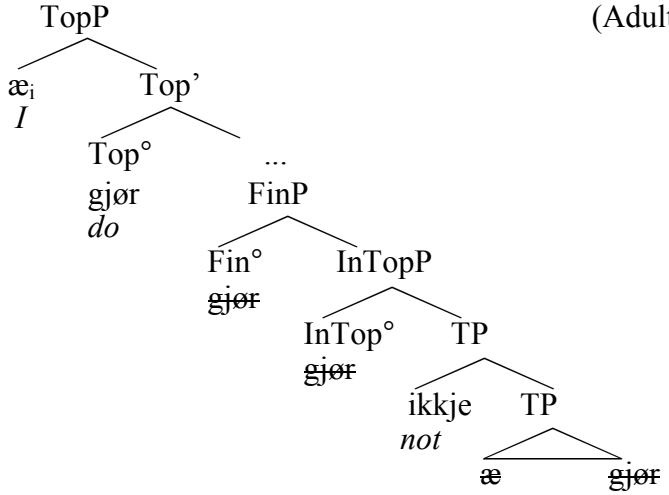
In this model, there are two functional projections in the IP-domain of the clause, the In(ner)Top(ic)P and T(ense)P, as was illustrated in (48) above. Sentence adverbs (including negation) normally occur adjoined to TP in the clause structure, i.e. in a position between the InTopP and TP. Thus, in order for the verb to appear in front of negation in sentence (49), it will minimally have to move to the head of the InTopP, the highest functional projection in the IP domain. Since the InTopP is the lowest possible projection in the clause structure that will ensure that the verb precedes negation, this is also in accordance with the children's shortest move approach. In other words, the children will initially assume that there is verb movement to the head InTop° in Norwegian. Note that this corresponds to V°-to-I° movement in traditional terminology, i.e. the children are in fact missetting a parameter, assuming that there is general verb movement to the IP domain in Norwegian.

Verb movement to the InTop° head will result in the surface word order V-Neg, which corresponds to what is found in the adult language. However, the syntactic representation of a sentence like (49) differs from that of the target grammar, in that the child version of the sentence is a bare FinP with verb movement to InTop°, as illustrated in the partial structure in (50). The corresponding adult structure is a full TopP, as illustrated in (51), where the [X°$_{EPP}$] feature attracts the verb to the head position of this projection.

(50) (Child structure)



(51) (Adult structure)



If this analysis of children's main clause declaratives is correct, we would expect to see verb movement to the InTopP also in embedded contexts. In fact, we expect to see the verb in front of negation and adverbs in all clauses where the verb does not move to the CP domain. This means that children should not only produce the non-target-consistent word order in all kinds of embedded clauses, but also in the non-V2 main *wh*-questions. This is of course exactly what we saw above in the child data presented in section 3, as illustrated in e.g. examples (27) and (34), repeated here for convenience. In all other clause types (questions and non-subject-initial declaratives), the verb moves to a head in the CP domain, and this will mask the V-to-I movement that the children seem to be assuming for their language.

(27') æ vet at æ **har ikke** gjort det. (Henning 4;8.13)
I know that I have not done it
'I know that I haven't done it.'
Target form: Æ vet at æ ikke har gjort det.

(34') kem som **var ikke** helt i form? (Henning 4;5.0)
who that was not completely in shape
'Who wasn't feeling too well?'
Target form: Kem som ikke var helt i form?

An additional example from the Tromsø corpus is illustrated in (52). This sentence is a non-subject-initial declarative, where the verb has failed to move across the subject to Top°. Thus, this sentence displays non-target-consistent word order, as the adult grammar requires V2 (see Westergaard 2004). As the children produce V2 constructions from the onset of multi-word utterances, there are very few such cases attested in the child data, and only one which includes negation. Note that the verb in sentence (52) does occur to the left of negation, suggesting that verb movement has in fact taken place, but not to the position above the subject, which would be expected in the target grammar. The word order in this example indicates that the verb has moved, and in accordance with the argumentation presented here, it has moved to the head of the InTopP:

(52) <ogs+>[/] og så du **kan ikke** tegne mer sånn. (Ann.17, 2;8.4)
and s+…. and so you can not draw more such
'And then you can't draw more like that.'
Target form: Og så kan du ikke tegne mer sånn.

Thus, the children's overgeneralization patterns in embedded clauses (and occasional examples from main clauses) provide some support for the analysis that initially in Norwegian child language there is verb movement to a lower head than Top° in main clause declaratives. This means that the children's choice of an uneconomic word order pattern in embedded clauses (involving verb movement) is actually caused by an economy principle operative in main clauses, viz. the principle of economy of movement.

But why don't we find any overgeneralization of V2 word order in embedded questions? This was illustrated by sentences such as (24) above, repeated here, where the verb correctly appears following the subject, unlike the word order in main clause questions.

(24') Ann vet ikke kor **han er** henne. (Ann.09, 2;2.19)
Ann know not where he is LOC
'Ann doesn't know where he is.'

The reason for this lack of word order overgeneralization is due to the functional architecture of the Split-CP model provided by UG. Recall that main clause questions are IntPs, while embedded questions are bare WhPs, lacking interrogative force. Being endowed with this knowledge, children know that embedded questions are not real questions and consequently do not project an IntP in these cases. Overgeneralizing the [$X°_{EPP}$] head feature on the Int° head to embedded clauses is therefore not possible, simply because that functional head is not present in this context.[9]

On the other hand, according to the account given for children's error patterns above, they will of course be expected to "transfer" verb movement to the InTop° head also in embedded questions. The prediction is that although young children will not move the verb across the subject in an embedded question, they should in fact overgeneralize verb movement across negation. Thus, an ungrammatical sentence such as (53) should be unattested in child data, while non-target forms such as the hypothetical sentence illustrated in (54) are predicted to occur.

(53) *Æ vet ka **vil han** gjøre.
I know what will he do
(54) ?Æ vet ka han **vil ikkje** gjøre.
I know what he will not do

The first part of this prediction is generally borne out in both the Tromsø corpus of younger children, as well as in the diary notes and recordings of the older children. As for the second part of this prediction, the results from the small experiment described in section 3 suggest that children at least up to the age of 6 overgeneralize verb movement to embedded *wh*-questions as well, moving the verb past negation and adverbs, as illustrated in e.g. (41), repeated here.

[9] Languages which do display V2 word order in embedded contexts, e.g. Belfast English, must then be assumed to have verb movement also to the head Wh°.

(41') huske du koffer han Karsten **var ikke** i barnehagen? (Iver 5;9.18)
remember you why he Karsten was not in kindergarten.the
'Do you remember why Karsten wasn't in the kindergarten?'
Target form: Huske du koffer han Karsten ikke var i barnehagen?

5.3 The Way to the Target Grammar

But if children have misset a parameter, how can they reach the target grammar? We argue that in order for children to reset the V-to-I parameter and revise their initial hypothesis, they need to pay attention to the word order in sentences that do not display V2, i.e. all embedded contexts and non-V2 main *wh*-questions. Note that this is different from the degree-0 learnability of e.g. Lightfoot (1999), which argues that children can only detect cues in unembedded contexts. Within the Split-CP model that we are assuming, where main and embedded clauses have different functional architecture, children must pay attention to the word order of relevant clause types separately in order to acquire the status of the EPP head feature with respect to the individual syntactic heads in the CP domain. We also believe that in order for children to be able to distinguish between Norwegian and V2 languages which do display V-to-I movement, e.g. Icelandic, they will have to be sensitive to embedded word order. Icelandic in fact displays exactly the word order in embedded clauses that the children in our study produce in Norwegian, and as far as we can tell, there is no difference between Norwegian and Icelandic *main* clauses that will indicate to children which type of language they are learning. Thus, we argue that the cue that a V2 language also has V-to-I must be found in non-V2 contexts, i.e. generally in embedded clauses.

Embedded clauses are naturally more complex structures than main clauses, and searching for cues in these contexts is arguably more difficult than finding cues in main clauses. This could be one reason why the non-target-consistent word order is so persistent in children's production, possibly lasting beyond the age of six, as indicated by the results of our small experiment. Compared to the extremely early acquisition of word order in general, in Norwegian as well as in other languages, the target-consistent word order in embedded clauses indeed falls into place very late.

Here frequency may also play a role. Recall that it is not sufficient for Norwegian children to pay attention to just any embedded clause; it must also contain negation or an adverb, otherwise the word order will be identical to that of main clauses. And as we saw in Tables 2 and 3 in section 4, these clause types are extremely infrequent in the input, attested only between 0.5% and 1% in the samples of child-directed speech that we investigated. Thus, we would argue that the lack of input frequency does have an effect in this case, viz. the effect that it takes a considerable time for children to revise their initial hypothesis that Norwegian has V-to-I movement. However, their initial hypothesis is not directly *caused* by the lack of frequency, but rather the principle of economy of movement, as we argued above.

One piece of evidence that may support the idea that there is a frequency effect here is the fact that this type of overgeneralization is generally not found in German child language, as we saw in section 3.1. Being an SOV language with verb movement also across objects and adjuncts in main clauses, German will provide considerably more input evidence to children that embedded clauses are different from main clauses. That is, the difference will not only be visible in embedded clauses containing negation or sentence adverbs (which are generally infrequent in child-directed speech), but also in all embedded clauses containing an object or an adjunct. Without having performed a study on German child-directed speech to this end, we think we can safely assume that a German-speaking child will be exposed to the

relevant contexts for non-V2 considerably more often than Norwegian-speaking children.

We may also compare this to another non-target-consistent word order pattern produced by young Norwegian children in so-called 'subject-shift' constructions, where the target language requires pronominal subjects to appear preceding negation in questions and non-subject-initial declaratives. In Westergaard (2005) it is shown that the three children in the study (age approximately 1;9 to 3;0) all initially produce pronominal subjects in a lower position, following negation. In Westergaard (submitted) this is argued to be due to the same economy principle that is discussed in the present article, as well as the general complexity of the construction. However, in this case the children's error pattern is relatively short-lived, as the target-consistent word order falls into place between age 2;6 and 3;0. The difference between the subject-shift constructions and the embedded contexts discussed here may partly be due to different input frequencies. In the same sample of child-directed speech which was investigated in Table 2, evidence for word order in the subject-shift constructions is attested in 4.2% of the total input (28 out of 668) and in 8.3% of all relevant contexts, i.e. questions and non-subject-initial declaratives. This is of course considerably more than the 1.0% evidence for word order in non-V2 contexts, and input frequency may therefore be argued to play a certain role here.

6. Summary and concluding remarks

In this paper we have considered two similar constructions in Norwegian child language, embedded questions and (all) embedded clauses containing negation or an adverb. In the former clause type the children's word order is error-free from the beginning, in that they never overgeneralize verb movement across the subject from main clause questions. In the latter clause type, on the other hand, children produce non-target-consistent word order for a considerable period of time, possibly beyond the age of six. That is, they move the verb across negation or an adverb, and this is a word order which also appears in other non-V2 constructions. An investigation of some samples of child-directed speech revealed that both constructions are extremely infrequent in the input, and that the difference between the two is too small to have such a considerable effect on the children's production. A possible account of the error pattern as a result of input frequencies only was therefore rejected.

Instead, within a weak continuity/structure-building account to language acquisition, we explored an analysis which assumes an economy principle of movement, which generally says that children will not move elements higher in a structure than there is evidence for in the input. More specifically, we argued that Norwegian children's early subject-initial main clauses display verb movement to a lower functional head than in the target grammar, i.e. to a head in the IP domain. This will ensure target-consistent word order in main clauses (V-Neg/Adv), but result in non-target-consistent word order in non-V2 contexts: embedded clauses and non-V2 main *wh*-questions. The reason why there is no overgeneralization of V2 from main to embedded questions is related to the Split-CP model of clause structure that we assume, where different clause types have different functional heads in the CP domain. While main *wh*-questions have an Int° head, embedded questions are bare WhPs, reflecting the fact that they have no interrogative force. Thus, the different functional architecture for the two clause types accounts for the lack of overgeneralization, as a feature value on the Int° head cannot be transferred to a clause type where this head is not present.

However, input frequency is also argued to play a role in this analysis. Together with the general complexity of the relevant constructions, the lack of input frequency may be a reason why the non-target-consistent word order produced by the Norwegian children is so persistent, compared to word order in other constructions.

Thus, we argue that there may certainly be effects of input frequencies in language acquisition, but we doubt that input frequency alone can account for acquisition orders and children's non-target-consistent production. Rather, we believe that explanations must be sought in a variety of areas. In the particular case discussed in the present paper, we have argued that economy as well as complexity interact with frequency to produce the particular error patterns found in the child data.

References

Åfarli, Tor. 1986. 'Some syntactic structures in a dialect of Norwegian.' *Working Papers in Linguistics* 3, pp. 93-111. University of Trondheim.

Anderssen, Merete. 2006. *The Acquisition of Compositional Definiteness in Norwegian*. Ph.D dissertation, University of Tromsø.

Benincà, Paola and Cecilia Poletto. 2004. 'Topic, Focus and V2: defining the CP sublayers.' In Luigi Rizzi (ed.) *The Structure of CP and IP: The Cartography of Syntactic Structures* 2, pp. 52-75. Oxford/New York: Oxford University Press.

Bentzen, Kristine. 2003. 'V-to-I Movement in the Absence of Morphological Cues: Evidence from Northern Norwegian.' *Nordlyd: Proceedings from the 19th Scandinavian Conference of Linguistics* 31.3, pp. 573-588, University of Tromsø.

Bentzen, Kristine. 2005. 'What's the better move? On verb placement in Standard and Northern Norwegian.' *Nordic Journal of Linguistics* 28.2, pp. 153-188.

Bentzen, Kristine. to appear. 'The degree of verb movement in embedded clauses in three varieties of Norwegian.' *Nordlyd: Scandinavian Dialect Syntax* 33, University of Tromsø.

Big Brother-korpuset. Available at http://www.tekstlab.uio.no/talespraak/bigbrother.

Chomsky, Noam. 1982. *Some Concepts and Consequences of the Theory of Government and Binding.* Cambridge, MA and London: MIT Press.

Chomsky, Noam. 1995. *The Minimalist Program*. Cambridge, MA and London: MIT Press.

Cinque, Guglielmo. 1999. *Adverbs and Functional Heads: A Crosslinguistic Perspective*. New York: Oxford University Press.

Clahsen, Harald and Klaus-Dirk Smolka. 1986. 'Psycholinguistic evidence and the description of V2 phenomena in German.' In Hubert Haider and Martin Prinzhorn (eds.) *Verb Second Phenomena in Germanic Languages*, pp. 137-168. Dordrecht: Foris.

Clahsen, Harald, Sonja Eisenbeiss and Martina Penke. 1996. 'Lexical learning in early syntactic development.' In Harald Clahsen (ed.) *Generative Perspectives on Language Acquisition: Empirical Findings, Theoretical Considerations and Crosslinguistic Comparison* [Language Acquisition and Language Disorders 14], pp. 129-159. Amsterdam and Philadelphia: John Benjamins.

Döpke, Susanne. 1998. 'Competing language structures: The acquisition of verb placement by bilingual German-English children.' *Journal of Child Language* 25, pp. 555-584.

Endresen, Rolf Theil and Hanne Gram Simonsen. 2001. 'The Norwegian verb.' In Hanne Gram Simonsen and Rolf Theil Endresen (eds.) *A Cognitive Approach to the Verb:*

Morphological and Constructional Perspectives, pp. 73-94. Berlin and New York: Mouton de Gruyter.

Garbacz, Piotr. 2005. 'Relationen mellan negationen och de finita verbet i fastlandsskandinaviska bisatser.' Talk presented at *Grammatik i fokus*, Lund.

Gawlitzek-Maiwald, Ira, Rosemary Tracy and Agnes Fritzenschaft. 1992. 'Language acquisition and competing linguistic representations: The child as arbiter.' In Jürgen Meisel (ed.) *The Acquisition of Verb Placement*, pp. 139-179. Dordrecht: Kluwer.

Holmberg, Anders. 1993. 'Two subject positions in Mainland Scandinavian.' *Working Papers in Scandinavian Syntax* 52, pp. 29-41.

Håkansson, Gisela and Sheila Dooley Collberg. 1994. 'The preference for Modal + Neg. An L2 perspective applied to L1 acquisition.' *Second Language Research* 10:2, pp. 95-124.

Jordens, Peter. 1990. 'The acquisition of verb placement in Dutch and German.' *Linguistics* 28, pp. 1407-1448.

Josefsson, Gunlög. 2004. 'Input and Output: Sentence patterns in Child and Adult Swedish.' In Gunlög Josefsson, Christer Platzack and Gisela Håkansson (eds.) *The Acquisition of Swedish Grammar* [Language Acquisition and Language Disorders 33], pp. 95-133. Amsterdam and Philadelphia: John Benjamins.

Lightfoot, David. 1999. *The Development of Language: Acquisition, Change and Evolution.* Malden, MA and Oxford: Blackwell.

Müller, Natascha. 1996. 'V2 in first language acquisition. Early child grammars fall within the range of UG.' *Linguistics* 34, pp. 992-1028.

Oslo-korpuset av taggede norske tekster. Available at http://www.tekstlab.uio.no/norsk/bokmaal/

Penner, Zvi. 1996. 'From empty to doubly-filled complementizers. A case study in the acquisition of subordination in Bernese Swiss German.' Fachgruppe Sprachwissenschaft der Universität Konstanz. *Arbeitspapier* Nr.77.

Platzack, Christer. 1996. 'The initial hypothesis of syntax: A minimalist perspective on language acquisition and attrition.' In Harald Clahsen (ed.) *Generative Perspectives on Language Acquisition: Empirical Findings, Theoretical Considerations and Crosslinguistic Comparison* [Language Acquisition and Language Disorders 14], pp. 369-414. Amsterdam and Philadelphia: John Benjamins.

Poeppel, David and Kenneth Wexler. 1993. 'The full competence hypothesis of clause structure in early German.' *Language* 69, pp. 365-424.

Radford, Andrew 1994. 'The syntax of questions in child English.' *Journal of Child Language* 21, pp. 211-236.

Rizzi, Luigi. 1997. 'The fine structure of the left periphery.' In Liliane Haegeman (ed.) *Elements of Grammar: Handbook of Generative Syntax*, pp. 281-337. Dordrecht: Kluwer.

Rizzi, Luigi. 2001. 'On the position "Int(errogative)" in the left periphery of the clause.' In Guglielmo Cinque and Giampolo Salvi (eds.) *Current Studies in Italian Syntax*, pp. 287-296. Amsterdam: Elsevier.

Roberts, Ian. 1999. 'Verb movement and markedness.' In Michel DeGraff (ed.) *Language Creation and Language Change: Creolization, Diachrony, and Development*, pp. 287-327. Cambridge, MA: MIT Press.

Safir, Ken. 1993. 'Perception, selection, and structural economy.' *Natural Language Semantics* 2, pp. 47-70.

Santelmann, Lynn. 1995. *The Acquisition of Verb Second Grammar in Child Swedish: Continuity of Universal Grammar in* Wh-*questions, Topicalizations and Verb Raising*. Ph.D dissertation, Cornell University.

Schaeffer, Jeanette. 2000. *The Acquisition of Direct Object Scrambling and Clitic Placement: Syntax and Pragmatics*. Amsterdam and Philadelphia: John Benjamins.

Schönenberger, Manuela. 2001. *Embedded V-to-C in Child Grammar: The Acquisition of Verb Placement in Swiss German*. Studies in Theoretical Psycholinguistics. Dordrecht: Kluwer.

Theakston, Anna, Elena Lieven, Julian Pine, and Caroline Rowland, C. 2004. 'Semantic generality, input frequency and the acquisition of syntax.' *Journal of Child Language* 31, pp. 61-99.

Tomasello, Michael. 2003. *Constructing a Language: A Usage-based Theory of Language Acquisition.* Cambridge: Harvard University Press.

Vangsnes, Øystein A. 2005. 'Microparameters for Norwegian *wh*-grammars.' *Linguistic Variation Yearbook* 5, pp. 187-226. Amsterdam and Philadelphia: John Benjamins.

Vikner, Sten. 1995. *Verb Movement and Expletive Subjects in the Germanic Languages*. New York: Oxford University Press.

Westergaard, Marit R. 2003. 'Word Order in *wh*-questions in a North Norwegian Dialect: Some Evidence from an Acquisition Study.' *Nordic Journal of Linguistics* 26.1, pp. 81-109.

Westergaard. Marit R. 2004. 'The Interaction of Input and UG in the Acquisition of Verb Movement in a Dialect of Norwegian.' *Nordlyd: Tromsø Working Papers in Language Acquisition* 32.1, pp. 110-134.

Westergaard, Marit R. 2005. *The Development of Word Order in Norwegian Child Language: The Interaction of Input and Economy Principles in the Acquisition of V2*. Ph.D dissertation, University of Tromsø.

Westergaard, Marit R. Forthcoming. 'Triggering V2: The Amount of Input Needed for Parameter Setting in a Split-CP Model of Word Order.' To appear in *Proceedings of GALA 2005*, Cambridge Scholars Press.

Westergaard, Marit R. (Submitted). 'Verb Movement and Subject Placement in the Acquisition of Word Order: Pragmatics or Structural Economy?' Submitted to *Language Acquisition and Language Disorders*, John Benjamins.

Westergaard, Marit R. and Øystein A. Vangsnes. 2005. '*Wh*-questions, V2, and the left periphery of three Norwegian dialects.' *Journal of Comparative Germanic Linguistics* 8, pp. 117-158.