

Object shift feeds variable subject and verb placement: implications for linearization and production planning.

What intraspeaker variation can tell us about the nature of grammar.

Björn Lundquist (UiT) and Maud Westendorp (UiT)

October 16, 2020

Many thanks to Ida Larsson, Gillian Ramchand, Craig Sailor and Gustavo Guajardo for inspiring discussions.

Standard word order in North Germanic

Subject initial main clause, with aux. and adverbs

Pre	C/V2	Mid	Verbs	Obj/Part	Lang
Johan John	borde should	förmodligen inte probably not	ha gett have given	Maria boken Mary the book	ALL

- ▶ 60-80% Subject initial clauses
- ▶ Highest verb in V2 position (main verb *in situ* 50-75%)
- ▶ Adverbs in left-right order in midfield.
- ▶ Remaining verbs clustered up in left-right order at the top of VP.
- ▶ Arguments in left-right order in VP (IO – DO – PP)

Holmberg and Platzack (1995), Vikner (1995), Thráinsson (2008), Thráinsson et al (2003), NALS special issue on the ScanDiaSyn project (2014).

Variation within and between languages when arguments and verbs end up in the midfield/IP/TP

Inverted NP subject, V2 main verb, light object pron.

Pre	C/V2	Mid	Verbs	Obj/Part	Lang
Idag	välkomnade	värden mig inte		till festen.	Swe: 6 ord
Today	welcomed	the host me not		to the party.	

- ▶ Norwegian: Strict Subject-Object order, variable Subject-Neg order
- ▶ Danish, Faroese (Icelandic): Strict Subject-Object-Neg order. (Svenonius 2001)

Verb placement embedded clauses/subject placement

Pre	C	Mid	Verbs	Obj/Part	Lang
X	at	Páll lurti altíð		eftir útvapinum	Far
X	that	Paul listened always		to the radio	

Embedded V2 found in MSC languages; V-to-I possibly restricted to Icelandic (and Faroese) (see Heycock et al. 2012)

- ▶ Common pattern: Strict categorical word order in language A, variable word order in language B. (Object shift, Subject shift, Long object shift, V3 in questions), but the variation is usually still highly conditioned by linguistic factors.

Nordic Word order Database (Larsson et al. 2019)

- ▶ Database (available online <https://tekstlab.uio.no/nwd/>) containing elicited sentences from all the North Germanic languages, covering 10 core syntactic variables.

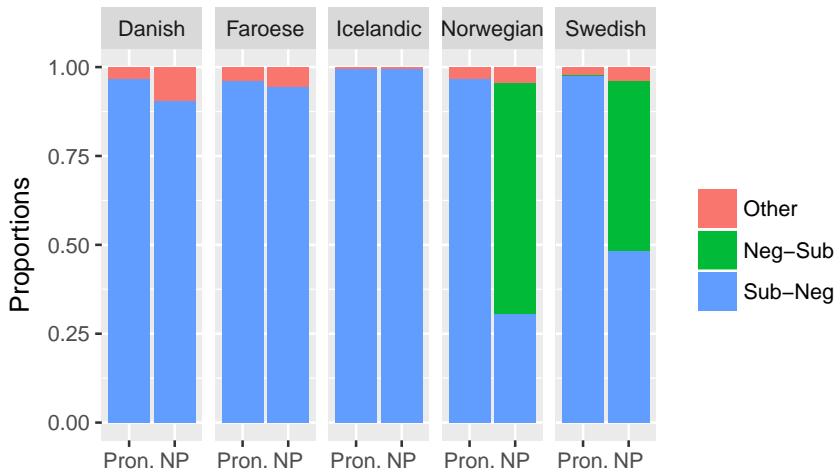
- (1)
 - a. **Subject shift:** Subject (Pro/DP) – Negation/Adverb
 - b. **Long object shift:** Subject (DP) – light pronominal object
 - c. **Reflexive Long object shift:** Subject (DP) – Refl. object
 - d. **“Long” particle shift:** Subject (Pro/DP) – Verb Part.
 - e. **Object shift:** Object (Pro/DP) – Neg/Adv
 - f. **Particle Shift:** Object (Pro/DP) – Verb Part.
 - g. **Embedded V2:** Fin. verb – Adv. in non-factive emb.clause
 - h. **V-to-I movement:** Fin. verb – Adv. in embedded quest.
 - i. **V3 in question:** Fin. verb – Subject (Pro/DP). in questions.
 - j. **V3 after preverbal adverbs:** Fin. verb – Adverb in Main clauses@
 - k. (+ comp, in embed, questions, articles, gender, prosody)

- ▶ Containing appr. 60.000 sentences, tagged for word order, with linked sound files (and soon text-speech aligned text grid files for studying word order- prosody correlations, and speech onset latencies).
- ▶ Goal: at least 10 items per variable and speaker, and at least 20 speakers per dialect, to establish if variation is within/between speakers/dialect/languages. (<https://journals.uio.no/NALS/article/view/7529>)

Subject placement in NWD

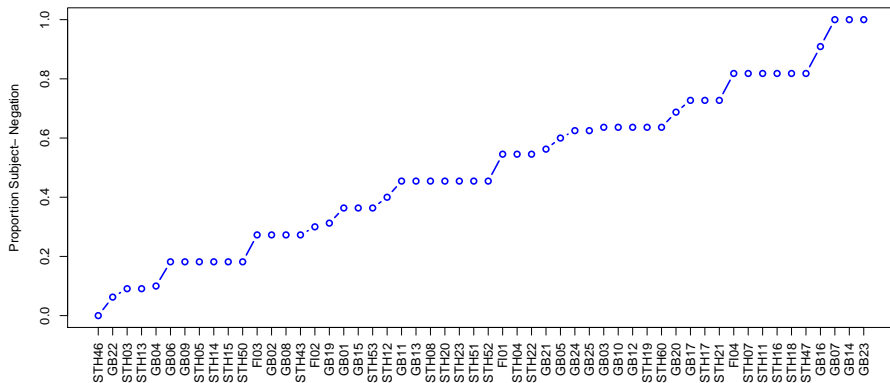
Background: The teacher/he came not to work yesterday.

Target: Yesterday (came (**not**) the teacher/he (**not**) to work yesterday)

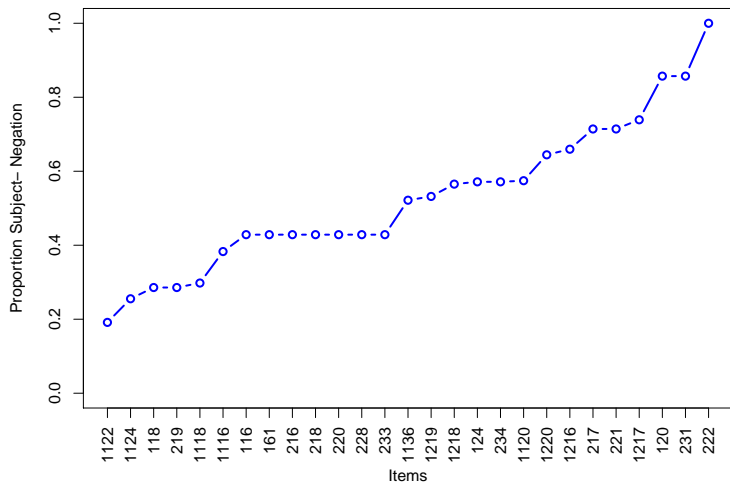


4914 observations (Dan: 555, Far: 1148, Ice: 485, Nor: 1137, Swe: 981)

Swedish, within item variation, NP-subject – Neg



Swedish, within speaker variation, NP-subject – Neg



Three questions/choice points in the study of variable patterns

1. Does the choice of word order have an effect on interpretation?
2. If not, does the two word orders correspond to two different syntactic structures? (e.g., derived by probabilistic ordering of syntactic operations, operations (probabilistically) targeting different types of constituents (Nilsen 2003), or by assuming multiple landing sites associated in an elaborate *fseq.*)
3. If not (i.e., if we accept that syntax may produce LCA-incompatible structures), what is determining the linearization? Is linearization calculated on a left-to-right basis?

We will look at cases where bi-gram frequencies affect word order choices (without affecting meaning).

Observation 1

- ▶ Variable placement of inverted DP subjects in Norwegian:

(2) I går kom ikke bakeren for sent til jobb. (86%)

I går kom bakeren ikke for sent til jobb. (14%)

‘Yesterday, the baker was not late for work’

- ▶ The presence of a light object increases the likelihood of the subject appearing before the adverb ($p < 0.01$, log. glmer):

(3) I går barberte ikke bakeren seg før jobb. (60%)

I går barberte bakeren seg ikke før jobb. (40%)

“Yesterday the baker didn’t shave before work.”

Effects on interpretation? None, as far as we can tell – definite noun phrases and pronouns do not scope-wise interact with negation.

Holmberg's Generalization

(A) A light pronominal object surfaces to the left of a sentence adverb, while a strong pronoun or DP surfaces to the right...

(4) Jeg så ham ikke – Jeg så ikke Erik.
I saw him not – I saw not Erik.
'I didn't see him' – 'I didn't see Erik'

(B) ...as long as the the main verb surfaces to the left of the adverb (no OV!)

(5) Det var mange som ikke så ham.
*Det var mange som ham ikke så.
'There were many people who didn't see him'

(C) and the subject surfaces to the left of the adverb (exc. Swe).

(6) I går hjalp ikke læreren meg med leksene.
*I går hjalp meg ikke læreren med leksene
lit. Yesterday helped not teacher me with homework

Observation 1

- ▶ Variable placement of inverted DP subjects in Norwegian:

(7) I går kom ikke bakeren for sent til jobb. (appr. 86%)
I går kom bakeren ikke for sent til jobb. (appr. 14%)
'Yesterday, the baker was not late for work'

- ▶ **The presence of a light object increases the likelihood of the subject appearing before the adverb:**

(8) I går barberte ikke bakeren seg føre jobb. (appr. 60%)
I går barberte bakeren seg ikke føre jobb. (appr. 40%)
"Yesterday the baker didn't shave before work."

1. Strict ordering restriction: Subject > Object
2. Strict ordering restriction: Light Object > Neg
3. Variable order, with preference: Neg > Subject

Variable word order in TP/Midfield, without semantic effects

Do we have any reason to suspect that (a-b) have different syntactic representations?

- (9) I går barberte ikke bakeren seg føre jobb.
I går barberte bakeren seg ikke føre jobb.
“Yesterday the baker didn't shave before work.”

- ▶ **Constituency?** No tests (fronting, pronominalization, ellipsis) that indicate that [Neg] [VP] is a constituent with the exclusion of the subject, or [Sub [VP]] is a constituent with the exclusion of the Neg.
- ▶ No indications of [Sub - Obj], [Sub - Neg], or [Obj - Neg] constituency.

- (10) a. *[Inte bakaren] barberte seg...
b. *[Bakeren seg] barberte inte...

Variable word order in TP, w.out semantic effects

Do we have any reason to suspect that (a-b) have different syntactic representations?

- (11) I går barberte ikke bakeren seg før jobb.
I går barberte bakeren seg ikke før jobb.
"Yesterday the baker didn't shave before work."

► **Scope?** We stick to definite noun phrases to avoid scope interactions.

- (12) Förra året gillade inte många studenter showen. (pref. *Neg* > *Many*)
Last year liked not many students the show.

- (13) Förra året gillade inte många studenter showen. (Only *Many* > *Neg*)
Last year liked not many students the show.

Pragmatic reasons for transparent scope marking? *Neg* > *Many* may be constituent negation.

- (14) a. Inte många studenter gillade showen förra året. (Unamb)
Not many students liked the show last year.
b. Många studenter gillade inte showen förra året. (Unamb)
Many students liked not the show last year.

How to linearize a flat TP

1. Strict templates (that are hard to explain derivationally):
Subject > Object (see Fox and Pesetsky 2003, Nilsen 2003, Holmberg 2004):

(15) I går hjalp ikke læreren **meg** med leksene.
*I går hjalp **meg** ikke læreren med leksene
I går hjalp læreren ikke **meg** med leksene.
lit. Yesterday helped not teacher me with homework

2. Linearization governed by prosodic constraints? Possible, but unlikely an important factor in the Scandinavian TP: the variable elements prosodically incorporate into whatever they are next to, and don't affect the prosody more than e.g. past tense marking or definiteness marking.
3. Organize spell-out around frequently co-occurring pairs of words, and/or recently activated (pairs of) words.

- (16) I går kom ikke bakeren for sent til jobb. (appr. 86%)
I går kom bakeren ikke for sent til jobb. (appr. 14%)
'Yesterday, the baker was not late for work'
- (17) I går barberte ikke bakeren seg før jobb. (appr. 60%)
I går barberte bakeren seg ikke før jobb. (appr. 40%)
"Yesterday the baker didn't shave before work."
- (18) XP V2(shave) [TP { [the baker]} , { [SEG] } , { [ikke] } [vP]]
1. Alternative 1: High bigram frequency of [SEG], [ikke], increases likelihood of early ordering. This is followed by a categorical SO-ordering (i.e., subject before object.) – production planning not from left to right!
 2. Alternative 2: High bigram frequency of [SEG], [shave], increases likelihood of early linearization. This is followed by a categorical SO-ordering (i.e., subject before object.) – Left to right planning
 3. In the absence of a reflexive/light object, the bigram frequency of [Verb] – [Neg] will be higher than [Verb] – [NP]. Early linearization of Verb–Neg. Depending on verb, [Verb] – [SEG] and [Verb] — [Neg] frequencies are roughly the same.

- (19) I går kom han ikke.... > I går kom ikke han...
Yesterday came he not ... > Yesterday came not he...
- (20) I går kom bakeren ikke.... < I går kom ikke bakeren...
Yesterday came the baker not ... < Yesterday came not the baker...
- (21) I går rakade bakeren seg ikke \approx I går rakade ikke bakeren seg...
Yesterday shaved the baker REFL not ... > Yesterday shaved not the baker REFL

- ▶ Still, probabilistic patterns that can be violated.

(22) I går kom bakeren ikke ...
Yesterday came the baker not...

- ▶ Intuition (Nilsen 1997, Svenonius 2001, Bentzen 2007): Pre-negation NP more likely if NP is already introduced in context (not necessarily "specific").
- ▶ Further research: priming vs. (n-gram) frequencies (Kinoshita 1995, Dewhurst et al. 1998). Prediction: this effect is most salient if the very word pair has recently been produced.

Predictions, related phenomena

- ▶ If this is n-gram effects on the lexical level (and not the category level), we predict an effect verb-frequency on Neg-sub/Sub-Neg ratio. Very low frequencies should have a negligible effect on speech planning.
- ▶ Ongoing corpus studies: correlation between Neg-sub/Sub-Neg ratio and verb frequency (highly significant, but small effect size).

Predictions, related phenomena, old unsolved problems (Andersen et al.)

Verb placement embedded clauses/subject placement

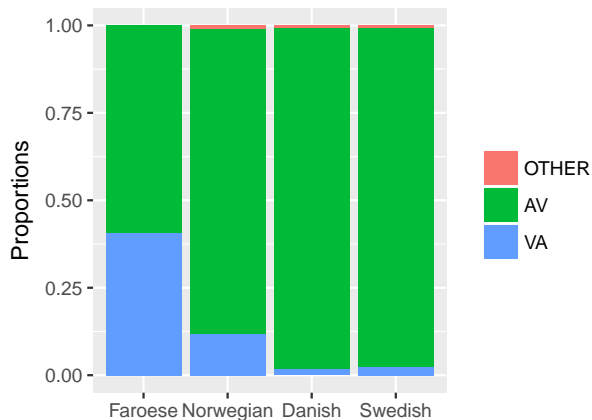
Pre	C	Mid	Verbs	Obj/Part	
Igår	kom	ikke bakeren		till jobbet	
Hon sa	at	bakeren ikke	kom	till jobbet	
Yday/She said	that/came	baker not	came	to work	

- ▶ High [Verb] – [Neg] bigram freq. is a result of default subject initial clauses.
- ▶ [Comp] – [Neg] very low: no subject position preceding C in embedded clauses.
- ▶ If this is effects on the lexical level, we expect relative clause Comp *som* to be more likely to be followed by [Neg – Subject], due to subject-less (subject) relative clauses. (ongoing corpus work).

Observation 2: Verb movement in embedded clauses

Background: I listen always to the radio in the car.

Target: John said (that he (listens) always (listens) to the radio in the car).



The large majority of the Faroese speakers produce both Neg-Sub and Sub-Neg in the experiment.

Observation 2

- ▶ Variable placement of finite verbs in *that*-clauses in Faroese:

(23) Páll segði at han **lurti altíð** eftir útvapinum um morgunin.
Páll segði at han **altíð lurti** eftir útvapinum um morgunin.
'Paul said that he always listens/listens always to the radio in the morning.'

- ▶ The presence of a light object increases the likelihood of the finite verb appearing before the adverb (35% vs 46%, $p < 0.05$)

(24) Páll segði at han **raki sær altíð** um morgunin.
Páll segði at han **altíð raki sær** um morgunin.
Paul said that he always shaves in the morning.

1. Strict ordering restriction: Verb > Object
2. Strict ordering restriction: Light Object > Neg
3. Variable order, with preference: Neg > Verb

The presence of a light object increases the likelihood of the finite verb appearing before the adverb (35% vs 46%, $p < 0.05$)

- (25) Páll segði at han raki sær altíð um morgunin.
Páll segði at han altíð raki sær um morgunin.
Paul said that he always shaves in the morning.

- ▶ Effect much smaller than for Subject placement.
- ▶ Verb movement in Faroese actually has syntactic consequences: creates a syntactic island (except for low adverbs, like *often*). **Frequency should not affect core syntax operations.**
- ▶ Other possible bigram effects: verbs with a “selected subject”, e.g. *weather* verbs.

- (26) Hon sa att det snöar ofta i Tromsø– Hon sa att det ofta snöar i Tromsø. She said that it snows often in Tromsø– She said that it often snows in Tromsø.

- ▶ There are no designated information structure position in the Scandinavian languages: arguments rarely occupy this space,
- ▶ Both adult speakers and children regularise the language.
- ▶ We have good reasons to believe that the internal order of the TP is more or less templatic, where the positions are not associated with semantic entailments.