

NORDIC WORD ORDER DATABASE: MOTIVATIONS, METHODS, MATERIAL AND INFRASTRUCTURE

BJÖRN LUNDQUIST, IDA LARSSON, MAUD WESTENDORP,
EIRIK TENGESDAL AND ANDERS NØKLESTAD

ABSTRACT

In this article, we present the Nordic Word Order Database (NWD), with a focus on the rationale behind it, the methods used in data elicitation, data analysis and the empirical scope of the database. NWD is an online database with a user-friendly search interface, hosted by The Text Laboratory at the University of Oslo, launched in April 2019 (<https://tekstlab.uio.no/nwd>). It contains elicited production data from speakers of all of the North Germanic languages, including several different dialects. So far, 7 fieldtrips have been conducted, and data from altogether around 250 participants (age 16–60) have been collected (approx. 55 000 sentences in total). The data elicitation is carried out through a carefully controlled production experiment that targets core syntactic phenomena that are known to show variation within and/or between the North Germanic languages, e.g., subject placement, object placement, particle placement and verb placement. In this article, we present the motivations and research questions behind the database, as well as a description of the experiment, the data collection procedure, and the structure of the database.¹

[1] INTRODUCTION

The North Germanic languages have undergone massive changes over the last 800 years, visible for instance in the change from OV to VO word order. The Mainland North Germanic (MNG) varieties (Swedish, Norwegian and Danish) and their dialects have followed fairly similar patterns in their development, e.g.,

[1] The Nordic Word order Database was funded by an infrastructure grant from University of Oslo (project leader: Ida Larsson) and the RCN project *Variation and Change in the Scandinavian Verb Phrase* (project number: 250755, PI: Ida Larsson). The infrastructure grant covered costs related to the technical work related to the database and annotation of the collected material. The RCN grant covered expenses associated with field work and project development. Additional funding was provided by the LAVA Lab at UiT The Arctic University of Norway for data collection (Tromsø) and data analysis. Westendorp's contribution was funded by the PhD project *Rethinking Verb Second in Scandinavian: a synchronic and diachronic analysis* (UiT).

with regard to the loss of case, verb agreement and V-to-I movement, whereas the Insular North Germanic (ING) varieties (Icelandic and Faroese) have kept these features (although, more recently, Faroese has moved towards the morphosyntactic patterns of the MNG varieties) (Holmberg & Platzack 1995, Vikner 1995). In many cases, all MNG varieties have moved in the same general direction. However, at a more detailed level, we find diverging paths, resulting in a fairly large amount of word order variation within the language area, for example variation in subject placement, object placement and verb placement in embedded clauses. Some of the variation can be described as between-language, but, in fact, most of the variation occurs *within* the different varieties. We do not yet know if certain syntactic phenomena are restricted to specific dialects or reflect optionality/intraspeaker variation. This uncertainty is especially clear when considering today's language situation, where speakers might get linguistic input from a wide variety of dialects and registers, as well as from mass media and written material. Given this, one question is: are these (relatively infrequent) word order variants stably associated with dialect groupings, in similar ways to omnipresent linguistic variables such as intonation, segmental phonology and highly frequent inflectional material?

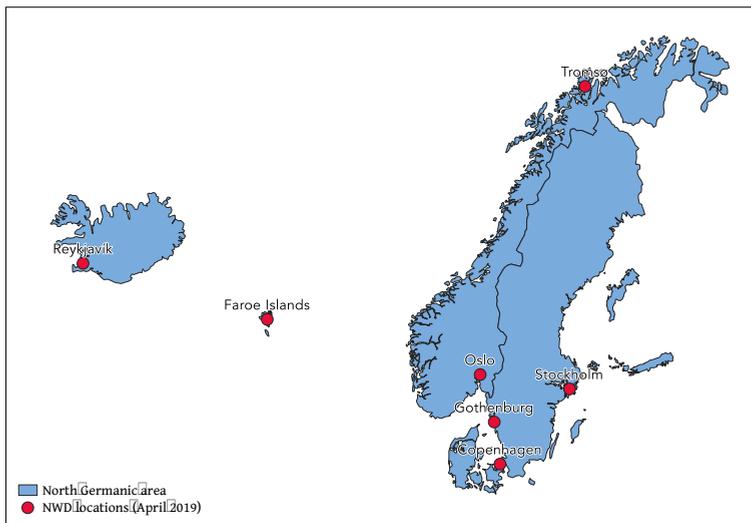


FIGURE 1: Overview of fieldwork locations (up to April 2019): Reykjavík, Faroe Islands (3 locations), Oslo, Tromsø, Copenhagen, Gothenburg, and Stockholm.

In this article, we present the Nordic Word Order Database (NWD), the rationale behind it and the methods used in data elicitation and data analysis. NWD is an online database with a user-friendly search interface, hosted by The Text Laboratory at the University of Oslo launched in spring 2019. It contains elicited

production data from speakers of all the North Germanic languages, including several different dialects. So far (April 2019), fieldwork has been conducted in seven different locations (see Figure 1), and data from altogether around 250 participants (age 16–60) have been collected (approx. 55 000 sentences in total). Given its size, and the systematic methods used in the data collected, several core questions about linguistic variation in the modern speech communities can be answered, e.g., is syntactic variation conditioned mainly by “region” or “register”, or is the variation mainly within or between individuals? Furthermore, it has been claimed that some of the variation is conditioned by higher ranked prosodic or intonational constraints (see e.g., Erteschik-Shir and Josefsson 2017 on object shift). By including sound files of all the elicited sentences, we are now able to seriously evaluate these claims.

The aim of the project was to collect data from at least 20 participants with a similar background and age from each location. By including at least 5 items of each syntactic phenomenon, we get at least 100 instances of each phenomenon per location/dialect, which we take to be good enough for a serious estimation of the amount of variation within each dialect.

Many of the investigated phenomena are highly infrequent in spontaneous production, and systematic elicitation is therefore the only means to investigate both inter- and intra-speaker variation. For example, in a corpus-study of object shift in Mainland North Germanic, Bentzen et al. (2013) found 116 instances of object shift in a corpus of 210 000 utterances (at most 1 shift per 10 000 sentences). Given this relative infrequency, it is hard or impossible to conclude from corpus studies whether the syntactic variation, if found at all, is within or between individuals, or between dialects.

The structure of this article is as follows. In Section 2, we present the syntactic phenomena included in the project and how they can be searched for in the online database. Section 3 describes the method used to elicit the variation, and Section 4 describes the experimental set-up and the coding of the results. Section 5 gives a very short overview of the literature on the syntactic variables tested, and here we also directly address how the data in the NWD can provide new insights into long-standing theoretical debates. Section 6 concludes the paper.

[2] THE SYNTACTIC PHENOMENA COVERED IN THE NWD, AND THE PARAMETERS IN THE SEARCH INTERFACE

The core objective of the NWD is to investigate variation in the order of different constituents in the clause. The investigated phenomena have all been much discussed in the literature on North Germanic, and results from previous studies have informed the experimental set-up and choice of our experimental material,

e.g., the type of embedded clause, sentence adverb or type of pronoun used. We refer the reader to Section 5 for more details about the different phenomena, with relevant references.

The data elicitation mainly targets the relative order of two constituents in the clause, and the variables are therefore defined as pairs of clause elements. The pairs that are targeted are listed below, under the rubrics of classic North Germanic syntactic phenomena:

1. Subject shift (SS): does a non-initial subject (NP or pronoun) precede or follow negation or other sentence adverbs? (Targeted pair: Subject – Negation/sentence adverb)
2. Object shift (OS): does the direct object (NP, pronoun or reflexive) precede or follow negation/sentence adverb when the main verb has moved to second position? (Targeted pair: Negation/adverb – Object)
3. Particle shift: Does the direct object (NP or pronoun) precede or follow a verb particle? (Targeted pair: Object – Particle)
4. Long object shift (LOS): Does the subject (NP or pronoun) precede or follow a light pronominal object (reflexive or first person)? (Targeted pair: Subject – Object)
5. “Long” particle shift: Does the subject (NP, pronoun or reflexive) precede or follow a verb particle? (Targeted pair: Subject – Particle)
6. Embedded V2: Does the finite verb precede or follow a sentence adverb in an embedded assertive clause? (Targeted pair: Verb – Sentence adverb)
7. Embedded V-to-I: Does the finite verb precede or follow a sentence adverb in an embedded question? (Targeted pair: Verb – Sentence adverb)
8. V3 in main clause non-subject *wh*-questions: Does the finite verb follow or precede the subject in a non-subject question? (Targeted pair: Verb – Subject)
9. V3 in main clauses with sentence adverbs: Does the finite verb precede or follow a modifier/sentence adverb in a main clause? (Targeted pair: Verb – Sentence adverb)

In addition, a couple of morphosyntactic features that are not easily stated as “Pairs” are targeted in the data collection. One is the presence of the complementizer *som* in embedded subject and object *wh*-questions (e.g., “He wondered who *som* came to the party.”, “He wondered what *som* he bought at the store.”). Another one is the presence of *som* in main clause subject questions (e.g., “Who *som* came to the party?”), a phenomenon with an isogloss similar to the V3 non-subject question (see (8) above). The different phenomena and relevant literature will be discussed in Section 5.

In the Nordic Word Order Database search interface, there are five linguistic search parameters that can be specified, presented as drop-down menus, in addition to various filters for metadata (see the appendix for details). The five parameters are the following:

- (i) Pair: The pairs listed above, e.g., Subject – Object.
- (ii) Type Element 1: The form of the first member of the pair, e.g., NP, pronoun or negation.
- (iii) Type Element 2: The form of the second member of the pair, e.g., NP, pronoun or negation.
- (iv) Exact Category: The exact syntactic phenomena targeted, e.g., long object shift (LOS), particle shift with non-spatial particles.
- (v) Produced Word Order: The order that the participant produced in the relevant condition, e.g., Subject-Object (SO) if the subject preceded the object in the elicited sentence.

To give an example, if you are interested in the phenomenon Long object shift, i.e. cases where a pronominal object precedes the inverted subject within the midfield, select the Pair-category “Subject – Object”. If you are then specifically interested in sentences that have a noun phrase subject, select NP in the drop-down menu “Type Element 1”. The form of the second element can further be specified in the same way, so as to restrict the search for instance to sentences with either reflexive objects, or first person pronominal objects. In the Produced Word Order column, the search result can further be filtered to show only produced orders of e.g. “OS” (i.e., objects preceding subjects, as in for example “Yesterday helped me the teacher.”). Figure 2 shows the search interface and the first

hits for the parameter values stated above.

| No. | Informant | Read | Realization | Produced order | Year of birth |
|-----|-----------|---------------------------------|--------------------------------|----------------|---------------|
| 111 | FI02 | Läraren gav mig en ny bok igår. | Igår gav mig läraren en ny bok | OS | 1970 |
| 111 | GB04 | Läraren gav mig en ny bok igår. | Igår gav mig läraren en ny bok | OS | NA |
| 111 | GB13 | Läraren gav mig en ny bok igår. | Igår gav mig läraren en ny bok | OS | 1986 |
| 111 | GB14 | Läraren gav mig en ny bok igår. | Igår gav mig läraren en ny bok | OS | NA |
| 111 | STH03 | Läraren gav mig en ny bok igår. | Igår gav mig läraren en ny bok | OS | NA |

FIGURE 2: The search interface and result presentation. The results do not include the background sentences, since the “Include ‘Read’” box has not been ticked.

Alternatively, the exact syntactic phenomena can be selected in the Exact Category drop-down menu (see appendix for a list of phenomena, and their abbreviations in the database). If verb movement in embedded clauses is the object of study, select the pair “Verb – Adverb”, and then further specify the verb type (here, e.g., Verb_Emb_Q), and if desired, the exact type of adverb in the Type Element 2 column, and Produced Word Order (AV or VA for the orders adverb-verb and verb-adverb, respectively).

As for the metadata filter, the most relevant category will generally be Language (Danish, Faroese, Icelandic, Norwegian or Swedish) and Recording Place. In displaying the results, it is possible to choose whether the so-called background sentence (see Section 3) should be included or not, by ticking the “Include ‘Read’” box. If this box is left empty, only the sentences with possible word-order variation are displayed. These are labelled “Produce” in the Type-column in the result (as opposed to “Read”).

[3] METHODS

The data are elicited by a very simple experiment set up in the open source experiment software OpenSesame (Mathôt et al. 2012). In the experiment, a participant is asked to first read a sentence on a computer screen and is subsequently prompted to perform a transformation of the sentence. We will first go through the different transformations used in the experiment, and then describe exactly how the experiment is implemented and carried out.

[3.1] *The material used in the elicitation*

The data elicitation is carried out with the help of a spoken language elicitation experiment. In the recording sessions, the participant is presented with a written sentence on a computer screen which s/he reads out loud. We will refer to this sentence as the *background sentence* below. Hereafter, the beginning of a new sentence appears on the screen, which we will refer to as the *trigger* below. The participant completes the new sentence, based on the instructions and the background sentence. We refer to the produced sentences as the *target sentence*. The complete test battery involves two “macro-experiments”, one with three parts which target the placement of arguments and adverbs/particles, and one with two parts which target the placement of verbs and adverbs. Each part builds on a special transformation, i.e., a specific relation between the background sentence and the sentence produced after the appearance of the trigger. Each of the parts, and the specific task used within, will be explained in the subsections below. In Table 1, we give an overview of the two macro-experiments and their parts.

| Experiment | Part/Transformation | Target Phenomena |
|---|--|---|
| 1. Placement of arguments w.r.t. adverbs, particles and other arguments | 1.1: Subject-Verb inversion | Subject Shift, Long Object Shift, Long Particle Shift |
| | 1.2: Complex to simple tense, Subject-Verb inversion | Object Shift, Long Object Shift, Long Particle Shift, Subject Shift |
| | 1.3: Passive to active | Particle Shift, (NP) Object Shift |
| 2. Placement of finite verbs in main and embedded clauses | 2.1: Main clause to embedded clause | V-to-T, V-to-C |
| | 2.2: Embedded clause to main clause | Non-V2 in questions and declaratives |

TABLE 1: Overview of the experiments and their subparts.

To the extent possible, we use direct translations of the same sentences

across languages but for certain phenomena, we have had to modify the sentences slightly or in some cases add special sentence structures. There is also some variation as to which adverbs are used to investigate subject shift and object shift. In general, the experiment tests the order between subject and negation, but the Norwegian version of the test also includes sentences with the adverbs *alltid* ‘always’ and *ofte* ‘often’, and for Icelandic, the adverb *líklega* ‘likely’ was used.

In this section, we will refer to the *standard test* as the set of sentences used in all locations. The experiments were piloted in Sweden, Norway and on the Faroe Islands. This led to some changes in phrasing of overly complicated sentences. For Faroese and Icelandic, we also changed the use of (pro)nouns, so that a feminine and masculine pronoun always referred to animates.

Since each part targets more than one syntactic construction (see below), there was little need for true filler items. Each part contains between four and six main conditions (here, constituent pairs), where each main condition usually has two subconditions (for example the form of the first or second member of pair). True filler items were only added in the two final parts, where we tested verb placement, in order to reduce the proportion of sentences with sentence adverbs. Overall, no part had fewer than four conditions, and no condition appeared in more than 25% of the trials within a given part. Part 1B had the lowest number of conditions (and items), while the other parts had at least 6 (sub)conditions. The items were pseudo-randomized, and the same condition never appeared twice in a row. Each part started with 2–4 practice items. Below, we present the the five parts in detail.

Subject-Verb inversion (Experiment 1, part 1)

In the first part (experiment 1, part 1), the participant reads a sentence (the background sentence) in the simple past tense with a sentence-initial subject, like the following one:²

- (1) Johan köpte en ny bil igår.
 Johan buy.PST a new car yesterday
 ‘John bought a new car yesterday.’

Once the participant has read the sentence aloud, the beginning of a new sentence appears on the screen, while the background sentence is still visible, now with the temporal adverbial in the first position. This is the trigger, and the task for the participant is to complete the second sentence using the content given in the background sentence. In the following we will put the possible completions,

[2] For convenience, examples are given in Swedish, unless otherwise indicated.

i.e. what the participant might produce, in parentheses:

- (2) Igår... (köpte Johan en ny bil.)
 Yesterday buy.PST Johan a new car
 ‘Yesterday, Johan bought a new car.’

In this example, we do not expect any variation, in the traditional North Germanic dialects: when the subject inverts with the verb, it ends up directly adjacent to the verb, preceding the phrasal direct object. However, as soon as the background sentence contains an adverb, a light object pronoun or a particle directly following the finite verb, the inverted subject can be placed either before or after the post-verbal element, depending on language and dialect. Below, we show how subject–verb inversion is used to test subject shift, long object shift and long particle shift. For consistency, we always treat the subject as the element with the variable placement, but this should not be interpreted as an analysis of the word order alternation:

Subject vs. Negation (subject shift), background (3) and target (4):

- (3) Bagaren kom inte till jobbet igår.
 Baker.DEF come.PST not to work.DEF yesterday
 ‘The baker did not come to work yesterday.’
- (4) Igår... (kom {bagaren} inte {bagaren})
 Yesterday come.PST baker.DEF not baker.DEF
 till jobbet.)
 to work.DEF
 ‘Yesterday, the baker did not come to work.’

Subject shift is tested with both pronominal and NP subjects. In the standard version of the experiment, 5 pronominal and 5 NP items are included.

Subject vs. Pronominal first person light object (long object shift), background (5) and target (6):

- (5) Läraren hjälpte mig med läxan igår.
 Teacher.DEF help.PST me with homework.DEF yesterday
 ‘The teacher helped me with the homework yesterday.’

- (6) Igår... (hjälpte {läraren} mig {läraren}
 Yesterday help.PST teacher.DE me teacher.DEF
 F
 med läxan.)
 with homework.DEF
 ‘Yesterday, the teacher helped me with the homework.’

For this pair we used only NP subjects, 5 items, as we are not aware of any North Germanic dialect that would allow an object pronoun shifting over a subject pronoun.³

Subject vs. Simple reflexive object (reflexive long object shift), background (7) and target (8):

- (7) Läraren rakade sig innan jobbet igår.
 Teacher.DEF shave.PST REFL before work.DEF yesterday
 ‘The teacher shaved himself before work yesterday.’
- (8) Igår... (rakade {läraren} sig {läraren}
 Yesterday shave.PST teacher.DEF REFL teacher.DEF
 innan jobbet.)
 before work.DEF
 ‘Yesterday, the teacher shaved himself before work.’

Reflexive long object shift is also tested with both pronominal and NP subjects, 5 of each.

Subject vs. Verb particle (“long” particle shift), background (9) and target (10):

- (9) Löparen gav upp under sista rundan
 Runner.DEF give.PST up during last lap.DEF
 igår.
 yesterday
 ‘The runner gave up during the final lap yesterday.’

[3] As pointed out by one of the reviewers this seem to hold for all Germanic varieties, with the exclusion of some German dialects discussed in Weiß (2015).

- (10) Igår... (gav {löparen} upp {löparen}
 Yesterday give.PST runner.DEF up runner.DEF
 under sista rundan.)
 during last lap.DEF
 ‘Yesterday, the runner gave up during the last lap.’

We include a total of 11 items for this pair: 6 with NP subjects and 5 with pronominal subjects. As we will discuss in Section 5.3, we do not find a lot of syntactic variation with respect to subject–particle order, but we do however find interesting prosodic variation. Note that we only use intransitive verbs here, since a transitive verb would allow for two different word orders of the background sentence (NP – Particle or Particle – NP). It is important that only one word order is allowed in the background, as we do not want to prime the participant to produce a certain word order.

The first part of experiment 1 contains 36 items in total, split over four main conditions (subject shift, long object shift, reflexive long object shift and long particle shift), where three of the four conditions are split into two sub-conditions (pronominal or NP subject). In most cases, the initial element in the trigger in this part is the simple adverb *igår* ‘yesterday’, but some variation occurs (e.g., *last year*). At some locations, we added other items such as for instance other sentence adverbs in addition to negation. This will be detailed in future individual articles.

Change of tense and subject verb inversion (Experiment 1, part 2)

The first part, as presented above, makes use of a very simple paradigm to check the placement of the subject with respect to other elements. It can however not be used for testing the position of the object with respect to negation and other sentence adverbs (i.e., object shift).

This is solved in the second part, where we manipulate the placement of the main verb in addition to the placement of the subject. In part 2, the background sentence has a complex verbal expression denoting future, which in all North Germanic languages can involve an auxiliary in the second position, and the main verb left *in situ* within the verb phrase, as in (11) below. This means that object shift is not an option in the background sentence.

- (11) Hon kommer inte att hjälpa mig med läxan.
 She will not INF help me with homework.DEF
 ‘She will not help me with the homework.’

After reading the background sentence aloud, the trigger appears, which in

this part consists of an initial adverb, but also the main verb in the simple past tense. In other words, the participants do not have to come up with the past tense form of the verb themselves, but only have to linearize the post-verbal elements.

Subject vs. Verb particle (“long” particle shift), background (12), target (13):

(12) Löparen kommer att ge upp under sista rundan.
 Runner.DEF will INF give up during last lap.DEF
 ‘The runner will give up during the final lap.’

(13) Igår gav... ({löparen} upp {löparen}
 Yesterday give.PST runner.DEF up runner.DEF
 under sista rundan.)
 during last lap.DEF
 ‘Yesterday, the runner gave up during the last lap.’

In this part of the experiment, we include 10 particle verbs (as in 12–13), 5 with pronominal subjects and 5 with NP subjects. Again, this is mainly for investigating the order subject–particle and the prosodic properties of the particle verbs (and to provide fillers for the object shift sentences).

We also include two conditions of object shift, one with NP subjects and one with pronominal subjects as the word order possibilities for these are known to be different:

NP subject vs. Negation vs. Reflexive pronoun (long reflexive object shift, and subject shift), background (14):

(14) Läraren kommer inte att raka sig efter jobbet.
 Teacher.DEF will not INF shave REFL after work.DEF
 ‘The teacher will not shave himself after work.’

In this case, there are three elements that need to be ordered with respect to each other.⁴ In Swedish, all six logically possible orderings are available, as given below:

[4] The future construction used in the test is complex in all the languages in the database except Icelandic (here, *mun* + infinitive). In Faroese and Swedish, the infinitive is preceded by an infinitive marker, and in Norwegian and Danish, it is preceded by both an infinitive marker and the preposition *til* ‘to’. Although there is some variation in the placement of negation w.r.t. infinitive markers in North Germanic, the negation has to surface to the left of the infinitive marker and the preposition in these future constructions.

- (15) a. Igår rakade läraren inte sig efter jobbet.
 b. Igår rakade läraren sig inte efter jobbet.
 c. Igår rakade inte läraren sig efter jobbet.
 d. Igår rakade inte sig läraren efter jobbet.
 e. Igår rakade sig inte läraren efter jobbet.
 f. Igår rakade sig läraren inte efter jobbet.
 ‘Yesterday, the teacher did not shave after work.’

5 items of this type are included in the test, all with NP subjects, where the NP consists of a non-modified definite noun (usually denoting a person of a profession, e.g., teacher, baker, policeman).

Pronominal subject vs. Negation vs. Pronominal 1st person object ((long) object shift, and subject shift), background (16):

- (16) Hon kommer inte att hjälpa mig med läxan.
 She will not INF help me with homework.DEF
 ‘She will not help me with the homework.’

The six logically possible orders in (15) are in principle possible here too, but the order of non-contrastive pronominal subjects and negation is quite fixed in all North Germanic varieties (Subject – Negation). The same holds for pronominal subjects and pronominal objects. The main variation we find here is therefore between the pronominal object and the negation (i.e., object shift). The two core orders in the target are the following:

- (17) Igår hjälpte... (hon {mig} inte {mig}
 Yesterday help.PST she me not me
 med läxan.)
 with homework.DEF
 ‘Yesterday, she did not help me with the homework.’

The second part contains in total 20 items, split over three macro-conditions (long particle shift, reflexive (long) object shift, and regular (long) object shift), with the particle sentences split up into two sub-categories (pronoun and NP).

Change of voice from passive to active (Experiment 1, part 3)

In the final part of experiment 1, the main syntactic phenomenon is the placement of direct objects with respect to particles. Note that this phenomenon can-

not be tested in the inversion paradigm (part 1.1) or the tense-changing paradigm (part 1.2), as the order of particle and direct object is not directly affected by the placement of subject or the verb. To target particle placement, we instead manipulate the voice of the verb: the background sentence is a passive sentence, as in (18) (examples given in Norwegian below).

- (18) (Norwegian)
- | | | | | |
|--------------|---------|------------|-----|----|
| Studenten | ble | kastet | ut | av |
| Student.DEF | AUX.PST | throw.PTCP | out | by |
| vaktene | | i går. | | |
| guard.PL.DEF | | yesterday | | |
- ‘The student was thrown out by the guards yesterday.’

The trigger is the subject of the corresponding active clause and the main verb in the simple past (active voice). When the clause contains a particle, as in (18), the object of the verb could then in principle appear either before, or after the particle as in the target in (19).

Object vs. Particle (particle shift), target (19):

- (19) (Norwegian)
- | | | | |
|--------------|-----------|-------------|-----|
| Vaktene | kastet... | {studenten} | ut |
| Guard.PL.DEF | throw.PST | student.DEF | out |
| {studenten} | i går.) | | |
| student.DEF | yesterday | | |
- ‘The guards threw the student out yesterday.’

The third part contains in total 23–25 items that test particle shift, 12 with pronominal subjects and 13 with NP subjects. The items are also split between spatial and non-spatial uses of particles, e.g., *hang up a painting* (spatial) and *buy up a company* (non-spatial, metaphorical).

Furthermore, the spatial particles come in two versions: one in which the particle is followed by a spatial PP (*throw {the student} out {the student} of the pub*) and one where the particle is followed by a temporal adverb (*throw {the student} out {the student} yesterday*).

In addition to particle shift, object shift is tested in this part as well, both with pronominal and NP subjects, as exemplified below for Icelandic:

Object vs. Negation or sentence adverb (object shift), background (20) and target (21):

- (20) (Icelandic)
- | | | | | |
|-------|-------------------|------|------------------|----|
| Kalli | var | ekki | yfirheyrður | af |
| Kalli | AUX.PST | not | interrogate.PTCP | by |
| unga | lögreglumanninum. | | | |
| young | policeman.DEF | | | |
- ‘Kalli was not interrogated by the young policeman.’
- (21) (Icelandic)
- | | | |
|-----------|------------------|-----------------|
| Ungi | lögreglumaðurinn | yfirheyrði... |
| young | policeman.DEF | interrogate.PST |
| {Kalla} | ekki | {Kalla}.) |
| Kalli.ACC | not | Kalli.ACC |
- ‘The young policeman did not interrogate Kalli.’

The exact length of this part differs a little bit across language versions, mainly due to the fact that some of the particle constructions or passives were not available in all of the languages. Most of the items are still identical across the different locations (to be detailed in individual field work articles).

From main clause to embedded clause (Experiment 2, part 1)

The first experiment (parts 1–3) focuses on the order of the arguments, either with respect to each other (long object shift) or with respect to negation, a sentence adverb or a verb particle (subject shift, object shift, (long) particle shift).

Another area of variation within the North Germanic languages is the placement of the finite verb, both in main clauses and embedded clauses. The three paradigms used in the first experiment all relied on main clause V2, and could not be used to test deviations from V2 or verb placement in embedded clauses. In the second experiment, we target verb placement by altering the embedding of a sentence: either from embedded to main clause, or from main clause to embedded clause.

In the first part of this experiment, the participants see a background sentence preceded by a name in parenthesis (and a different coloured font), and they are instructed to read the background sentence without the name. We illustrate this with the following practice item (in Swedish) consisting of a simple subject initial main clause:

- (22) (Erik:) Jag har glömt mina nycklar hemma.
 I have forget.SUP my.PL key.PL home
 'I have forgotten my keys at home.'

The target sentence then starts with the name as a subject followed by an embedding verb, e.g.:

- (23) Erik säger att... (han har glömt sina
 Erik say.PRS that he has forget.SUP his
 nycklar hemma.)
 key.PL home
 'Erik says that he has forgotten his keys at home.'

In the example above, the participant only needs to change the person value of the subject and the possessor (from first to third), but the word order remains the same. The crucial test items however either contain a sentence adverb or consist of embedded questions.

Verb vs. Adverb in embedded clause (embedded verb movement), background (24) and target (25):

- (24) Jag cyklar alltid till jobbet.
 I bike.PRS always to work.DEF
 'I always bike to work.'

- (25) Erik säger att... (han {cyklar} alltid
 Erik say.PRS that he {bikes} always
 {cyklar} till jobbet.)
 {bikes} to work.DEF
 'Erik says that he always bikes to work.'

Here we test the relative ordering of the finite verb and sentence adverbs in three different embedded contexts: in the complement of an assertive/non-factive verb (*say*), in the complement of a factive predicate (*be proud of*), and in embedded *yes/no* questions.

The original experiment contains 12 items with assertive verbs, 12 items with factive verbs and 12 embedded *yes/no* questions. Half of the items contain a reflexive verb, which also provided information about possible interactions between verb movement and object shift. The reports from the individual field

trips contain information about the exact number of items in the different conditions.

The main-to-embedded task also contains items with main/embedded questions. These items partly function as fillers, but they also test for (a) the distribution of complementizers/relative markers in different types of embedded subject and non-subject *wh*-questions, as well as (b) the absence of subject-verb inversion in embedded non-subject *wh*-questions.

Som-insertion in embedded subject question, background (26) and target (27):

(26) Vilka band spelade på festivalen?
Which band.PL play.PST at festival.DEF
'Which bands played at the festival?'

(27) Anna ville veta... (vilka band {som}
Anna want.PST know Which band.PL {that}
spelade på festivalen.)
play.PST at festival.DEF
'Anna wanted to know which bands played at the festival.'

Som-insertion in embedded non-subject question, background (28) and target (29):

(28) Vilken film såg eleverna igår?
Which film see.PST students.DEF.PL yesterday
'Which film did the students see yesterday?'

(29) Anna ville veta... (vilken film {som}
Anna want.PST know which film {that}
eleverna såg igår.)
student.DEF.PL see.PST yesterday
'Anna wanted to know which film the students saw yesterday.'

In the object *wh*-question, we can also see if the participants sometimes switch over to main clause word order, with the verb preceding the subject. This word order is in principle not available but might turn up if the participants use a "quote" strategy, i.e., just repeat the main clause.

From main clause to embedded clause (Experiment 2, part 2)

In the final part, we reverse the strategy from the previous part (2.1). Here, the participant first reads a sentence containing an embedded clause, either a declarative or a question, and is subsequently asked to produce the embedded clause as a main clause. We use similar cues as in the previous experiment, but in this case, it is the target item that is preceded by a name, e.g.: Background: *Erik said that he usually bikes to work.* – Target: *(Erik:) [...] (I usually bike to work).*

As in the previous part, we alternate between declaratives and questions (subject and non-subject *wh*-questions). In the declarative test sentences, we also alternate between two types of adverbs: regular sentence adverbs (*usually, unfortunately*) and adverbs that may be sandwiched between the initial element and the finite verb (*literally, simply*), thus creating main clauses with the verb in the third position. In the embedded clause, the adverb will always precede the verb, independent of which class it belongs to, but in main clauses, both orders are available only for the second group of adverbs:

Verb vs. Adverb, main clause, background (30) and target (31):

(30) Peter sa att Tyskland bokstavligen
 Peter say.PST that Germany literally
 krossade Sverige i finalen.
 crush.PST Sweden in final.DEF
 ‘Peter said that Germany literally crushed Sweden in the final.’

(31) ... (Tyskland {bokstavligen} krossade {bokstavligen}
 Germany Literally crush.PST literally
 Sverige i finalen.)
 Sweden in final.DEF
 ‘Germany literally crushed Sweden in the final.’

The standard test contains 8 sentences with V3 adverbs, and 8 sentences with regular (V2) adverbs, i.e., adverbs that are not expected to trigger V3. The standard test also includes subject and non-subject questions, to test whether the speakers deviate from V2 in interrogative contexts, something that one will only find in Norwegian dialects. In the relevant dialects, the relative pronoun/complementizer *som* may appear after a *wh*-subject, which leads to the absence of verb movement to the V2 position. In object questions, the verb may simply stay *in situ* (i.e., inside in the verb phrase), leading to a non-V2 structure (WH – Subject – Adverb – Verb). The relevant word order alternation is thus the placement

of the subject relative to the verb:

Verb vs. Subject, in questions, background (32) and target (33):

(32) (Norwegian)
 Peter spurte hva hun jobbet med.
 Peter ask.PST what she work.PST with
 ‘Peter asked what she was working on.’

(33) (Norwegian)
 ... (Hva {jobber} du {jobber} med?)
 What work.PRS you work.PRS with
 ‘What are you working on?’

The standard test contains 8 subject questions and 8 non-subject questions, with either NP or pronominal subjects.

[4] THE EXPERIMENTAL SET-UP AND THE CODING OF THE RESULTS

As was mentioned above, the data collection was carried out in a highly systematic fashion in a controlled setting. Below, we first go through the technical details of the experiments before we briefly describe the analysis and annotation of the collected data which can be accessed through the NWD. The experiment scripts and material used in the experiments is available in a GitHub repository, where scripts for analysing the results and ELAN templates can be found as well (<https://github.com/BjornLundquist>).

Most of the data collection took place in high schools in the different Nordic countries (details will be provided in future articles). Arrangements were made with contact persons or teachers at the schools in advance so that students were recruited to participate and so that we were given access to rooms for running the language experiments. Each of the macro-experiments described above lasted for about 14–18 minutes, which means that we could collect data from at least 20 participants in the course of a day.⁵ The experiments were run on laptops, and recordings were made through an external recorder, often with lavalier microphones. All participants signed consent forms and filled in a minimal background questionnaire before starting the experiment. Metadata (age, gender, education, place of birth) are available in the database, and can be used to restrict searches.

[5] The individual parts lasted on average 5–8 minutes, including instructions and practice rounds. Instructions were usually only necessary for the first part. For the subsequent parts, the participants generally figured out the procedure themselves based on the background and trigger sentence.

Each part of the experiment started with two to four practice rounds. In the first practice round, we gave the participant very short instructions, e.g.: “Read the first sentence aloud and then produce the second sentence based on the material in the first sentence. Imagine you are at home or with friends, and reading a sentence aloud from a newspaper”. We did not comment on the pragmatic or semantic relationship between the sentences, only that the participant should use “the words in the first sentence”. The participants also seemed to focus exclusively on the syntactic transformations that they had to perform. The experiment was highly intuitive, and most participants understood the paradigm straight away.

Each item in the experiment consists of the following steps:

- (i) The participant sees the background sentence on a screen, white font on black background, for 1000ms.
- (ii) A beep-sound is heard (500ms in experiment 1, 300ms in experiment 2) and the sentence turns from white font to red font at the end of the beep.
- (iii) The participant reads the sentence aloud as soon as it turns red (i.e., 1300–1500ms after the sentence first appeared on the screen).
- (iv) When the participant has read the sentence aloud, the beginning of a new sentence, the *trigger*, appears on the screen in white font below the background sentence (which is still present on the screen).
- (v) 300ms after the appearance of the second line, a beep sound is heard (300–500ms) and the font changes to red.
- (vi) At the offset of the beep, the participant starts reading the new sentence, and finishes it, based on the content of the first sentence.

As laid out in section 2 above, each part consists of 20–36 items, presented in pseudo-randomized order (never two items of the same syntactic phenomenon in succession). The experimenter advances to the next step or item in the experiment with a keystroke or remote click. The result file produced by OpenSesame contains time stamps for the beginning and end of each stage in the experiment (the background sentence and the trigger), as well as the item numbers. This file is later transformed into a format that can be fed straight into ELAN (see GitHub for R-scripts).

We combine the file with the time stamps and the sound file in ELAN, as to

get a pre-segmented ELAN file, with numbered items. In ELAN, we manually add in minimal annotations, indicating which word order was produced, e.g., AV (Adverb – Verb) or VA (Verb – Adverb). The files with the annotations are later exported as .csv-files, and combined with item information (which condition each sentence belongs to), a file with each available word order for each item, and a file with participant metadata (see GitHub repository for R-script).

The sound files and the full datafile are later inserted into the database hosted by The Text Laboratory. The database is a SQLite database with a search interface written in F#. The interface has a back end running on Saturn, a web application framework built on [ASP.NET Core](#), while the front end is based on ReactJS and uses the Fable compiler to transform F# into JavaScript. The sound clips associated with search results are played through the react-sound library.

[5] THE SYNTACTIC PHENOMENA: BACKGROUND LITERATURE AND OPEN QUESTIONS

The syntactic variables targeted in the Nordic Word Order Database are among the most discussed in the North Germanic syntax literature: V2, embedded verb movement, (long) object shift, particle shift and subject positions (see e.g. Holmberg & Platzack 1995, Vikner 1995, Thráinsson 2007). Most of the topics are also covered in NALS online vol. 1, which summarizes some of the patterns found in the Nordic Syntax Database (Lindstad et al. 2009) and the Nordic Dialect Corpus (Johannessen et al. 2009). In this section, we give a short background to the phenomena, with some selected references. We also point to some of the open questions that can be addressed using the data in NWD and mention some preliminary results.

[5.1] *Subject shift: Subject placement with respect to negation and other adverbs*

Subject shift is the topic of a chapter in NALS volume 1 (Bentzen 2014a), including a list of relevant references. Some of the core patterns of the North Germanic subject placement are discussed in Holmberg (1993), Svenonius (2002) and Andréasson (2007).

One clear pattern is found in all the North Germanic languages: unstressed subject pronouns, when not in sentence-initial position, have to precede sentence adverbs, and will surface directly after the finite verb. There are, however, exceptions to this generalization, found in Norwegian dialects where the negation behaves like a clitic (see Østbø Munch 2013). In the NWD data, we have so far not tested any of these dialects, and there are not yet any examples of the order Negation – (subject) Pronoun in the database (compared to over 2000 instances of the order Pronoun – Negation).

When we look at noun phrase subjects on the other hand, we do however observe variation in placement, both within and between the languages. Norwegian and Swedish show a fairly free ordering between an NP subject and negation (or other sentence adverb), while Danish and Insular North Germanic have a fairly strict order, with the subject preceding the negation (see Svenonius 2002 and Nilsen 2003 for discussion). The results from the NWD corroborate this: we find a consistent Subject – Negation order in Danish, Icelandic and Faroese, whereas we find a more balanced distribution in Swedish and Norwegian (with a slight preference for negation preceding NP subjects). Importantly, it seems like most Swedish and Norwegian participants produce both orders in the experiment, without any strong item effect, which suggests that this is random variation not conditioned by semantics or pragmatics. Among other things, it remains to be seen whether the word order variation relates to prosodic factors.

[5.2] *Object Shift and Long Object Shift: object placement with respect to adverbs and the subject*

The phenomenon called Object shift, and the generalization that captures the distribution of objects in North Germanic, *Holmberg's Generalization* (Holmberg 1986, 1999), is one of the most discussed topics in the North Germanic syntax literature; see Bentzen's NALS article (2014b) for an overview of the literature and dialect language variation. The database includes object shift data for four different types of objects: light reflexive objects (*sig/seg*), first person singular objects (*mig/meg*), third person pronouns (plural or singular), and full noun phrases.

In North Germanic, objects normally surface inside the VP, but in contexts where the main verb has moved to the V2 position (or possibly to I/T in Icelandic), the object may move to a position preceding sentence adverbs. When the object surfaces in a position preceding a sentence adverb, we say that the object has undergone Object sShift.

The exact properties of this phenomenon differ slightly between the North Germanic languages. First, object shift is restricted to light/non-contrastive pronominal objects in all contemporary North Germanic languages except Icelandic (see Thráinsson 2013 for discussion). Furthermore, shifting of light objects seems to be more or less obligatory (in the relevant contexts) in all the North Germanic languages except Swedish, where object shift appears to be genuinely optional. This cross-language variation is found in the data in NWD as well.

The languages also differ with respect to whether a light object can shift across an NP subject, i.e., whether so-called Long object shift is possible or not. Among the contemporary languages, only Swedish allows long object shift, but

as our results show, long object shift is almost completely restricted to simple reflexive objects.

There has been a lively discussion within the theoretical linguistics literature on how to characterize object shift: is it phonologically conditioned (see Erteschik-Shir 2005 and comments by Hellan 2005 and Svenonius 2005), conditioned by pragmatics/information structure (Andréasson 2007, Bentzen and Anderssen 2019) or more generally a post-syntactic operation? With the help of NWD, we can now test some of these hypotheses, specifically concerning the relationship between object shift and prosody.

[5.3] *Particle shift and long particle shift*

There is a fair amount of variation with respect to particle placement in North Germanic; see Lundquist's NALS article (2014) for an overview, and see e.g. Sandøy (1976), Taraldsen (1983ab), Åfarli (1985) and Svenonius (1996) for discussion.

Overall, the most common pattern in North Germanic is the one where objects precede the verb particle, especially if the object is a light pronoun. In general, light pronominal objects obligatorily precede particles in Danish, Faroese, Icelandic, and Norwegian. There is more variation in the distribution of NP objects. In Danish, all objects precede particles, while the order is more variable in Norwegian. From the NWD results, we can note that Faroese shows a pattern that is similar to Danish, whereas Icelandic, like Norwegian, allows for particles to precede NP objects (but to a lesser extent than Norwegian). Swedish is the odd one out here: all particles obligatorily precede the direct object, irrespective of the form of the object.

In all the North Germanic languages, the particle stays inside the VP, i.e., it never moves along with the verb to the second position (unless lexically prefixed), even though the particle in prosodic terms has a clitic-like behaviour. There is, however, anecdotal evidence that some children and even teenagers produce word orders that suggest a clitic analysis of particles: particles occasionally precede inverted subjects (see Lundquist 2018 fn. 7). As far as we are aware, this is not observed in the adult language of any dialects of Swedish, Norwegian, Danish or Faroese. In our data, we find particles preceding subjects in Icelandic. One possibility is that these subjects are VP-internal, as Icelandic allows VP-internal subjects. If this is the case, the attested Icelandic pattern "XP – V – Part – Subject" can not be taken as evidence for a clitic-like behaviour of the particle. An argument against treating the subjects as VP-internal is, however, that all NP subjects in our results precede sentence adverbials, as is expected for definite NPs in Icelandic.

[5.4] *Verb placement in embedded clauses*

All the North Germanic languages are verb second (V2) in so far that they typically place finite verbs in the second position in declarative main clauses. In embedded clauses, a split between the Insular North Germanic varieties and the Mainland North Germanic varieties is standardly assumed: while ING has so-called independent V-to-I movement across negation and adverbs, MNG does not allow such verb movement (see e.g., Holmberg & Platzack 1995, Vikner 1995). The loss of V-to-I movement (and generalized V2) in embedded clauses in MNG is a well-studied phenomenon and has been claimed to coincide with the loss of rich verbal inflection (see e.g., Platzack & Holmberg 1989, Holmberg & Platzack 1995).

In certain *that*-clauses, MNG still allows embedded V2, but typically only in cases where the complement is (or could be) assertive (e.g., Bentzen et al. 2007, Julien 2007 and Wiklund et al. 2009). The standard analysis is that MNG has lost V-to-I movement, but utilises V-to-C movement in embedded clauses where the finite verb precedes sentence-medial adverbs. It has been argued that contemporary Faroese is in a late stage of losing V-to-I and moving to a system similar to that of the MNG varieties (Jonas 1996, Heycock et al. 2010, 2012). There however continues to be considerable disagreement as to the current status of V-to-I movement in the language. In support of the claim that Faroese has almost lost V-to-I movement, Heycock et al. (2010) show no quantifiable difference in the acceptance of the order Verb – Negation in embedded clauses between Faroese and Danish in a magnitude estimation test.

In our data, we tested verb placement in three types of embedded contexts: complements of assertive/bridge verbs; complements of factive verbs; and embedded *yes/no* questions. The NWD data support previous findings in the literature: embedded verb second in Faroese is significantly more common after a verb that selects for an assertive complement than in a non-assertive complement or an embedded question. The pattern we find thus shows that Faroese has developed a system where verb movement in embedded clauses is constrained by similar semantic and information structural factors as in MNG. Surprisingly, we find embedded verb second only very rarely in material from Norway, Sweden and Denmark and only very few speakers of these varieties even produce this word order. Yet, corpus studies show this word order to be quite common (e.g., Julien 2007 for Norwegian and Swedish; Jensen & Christensen 2013 for Danish).

[5.5] *V3 in main clauses*

There are two exceptions to the standard main clause verb second word order in North Germanic: V3-triggering adverbs and non-V2 *wh*-questions. These phenomena are structurally different, and the latter is regionally constrained whereas the former is not (as far as we know).

In all the North Germanic languages, so-called “preverbal adverbials” can be found in second position between a sentence-initial element and the finite verb. This phenomenon has received surprisingly little attention in the Germanic V2-research, but see e.g. Thráinsson (2007:37–40) for discussion of these adverbs in Icelandic; Brandtler & Håkansson (2017) for Swedish; Julien (2018) for Norwegian; and Nilsen (2003) and Klæsvik-Pettersen (2018) for general discussion. The group of adverbials that can occur in the “V2”-position do not seem to have a unified semantics, and both verb-modifying and sentence-modifying adverbs can occur in preverbal position (Lundquist 2018).

The derivation of the V3 word order in this construction is also debated. Brandtler & Håkansson (2017), as well as Julien (2018), argue that it is a case of V-to-C movement with subsequent merge of the preverbal adverb; in contrast to this Lundquist (2018) argues that the V3-sequence is the result of phrasal movement, where the adverbial and verb form a unit in the verb phrase prior to movement to V2-position.

Not only the semantics and syntax but also the prosodic characteristics of this construction are under discussion in the literature. The data in the NWD will allow us to further test these semantic and syntactic hypotheses and examine the prosodic characteristics of the phenomenon

In varieties of Norwegian, V3 word order is found not only with preverbal adverbs but also in main clause *wh*-questions. Many dialects allow non-V2 word order in these questions alongside regular V2 order. In such non-V2 constructions, the verb remains within the verb phrase: in subject *wh*-questions, the complementizer *som* is inserted in the second position of the clause; in non-subject *wh*-questions, the verb follows the subject. Factors playing a role in the word order possibilities and choice have been extensively discussed in Norwegian dialectology (see e.g. Áfarli 1986, Westergaard & Vangsnes 2005, Westergaard et al. 2017, Westendorp 2018). This construction is thought to occur almost exclusively in Norwegian dialects, but in the NWD we now have a few examples of non-V2 subject *wh*-questions in Faroese.

[6] FUTURE PLANS FOR THE DATABASE

The database will continue to grow over the next couple of years. Data collection has been planned for additional dialect areas in Scandinavia, including the Swedish speaking areas in Finland. We will furthermore include material that has been elicited using slightly different methods: currently, we are collecting material that has been elicited solely through spoken dialect material, so as to get away from potential effects of standardisation triggered by the written input. We have also made experiment scripts, analysis scripts and guidelines for annotation available online (<https://github.com/BjornLundquist>) so that in principle anyone can collect data using the same methods and include the data in the database. Hopefully, we will have a database that covers a substantial part of the North Germanic dialect variation within a few years.

The experiment was originally designed to test word order variation and prosodic patterns associated with certain word order patterns, but thanks to the controlled sampling methods, the data can also be used to address questions about processing/production difficulties linked to atypical word orders, and atypical relations between syntax and prosody. A plan for the future is to include text-to-speech alignment in the database, so that production latencies for difference conditions and word order can easily be extracted.

ACKNOWLEDGEMENTS

Many people have been involved in different parts of this project: data collection, pilot testing, annotation and commenting on the elicitation material and the database. We are grateful for comments from three reviewers. In addition, we would like to thank: Rakul Absalonsen, Sigríður Mjöll Björnsdóttir, Gunnvør Brimnes, Elisabet Engdahl, Eline Grønvoll, Isabel Nadine Jensen, Jóhannes Gísli Jónsson, Tor-Willy Knudsen, Filippa Lindahl, Julian Lysvik, Ida Keihl Olsen, Anna Katharina Pilsbacher, Annika Simonsen, Tor Håvard Solhaug, Bror-Magnus Sviland Strand, Anne Caroline Tennøe, Ida Toivonen, as well as all the participants who took part in the experiments, and all the schools and universities that hosted us: Frederiksværk gymnasium og HF, Miðnám á Kambsdali, Tekniski Skúlin í Klaksvík, Kongsbakken VGS, Tromsdalen VGS, NTI Gymnasiet Sollen-tuna, Anna Whitlocks Gymnasium, Fróðskaparsetur Føroya, Háskóli Íslands, Københavns Universitet and Stockholms Universitet.

REFERENCES

- Andréasson, Maia. 2007. Satsadverbial, ledföljd och informationsdynamik i svenskan. *Göteborgsstudier i nordisk språkvetenskap* 7. <https://hdl.handle.net/2077/17006>
- Bentzen, Kristine, Gunnar Hrafn Hrafnbjargarson, Þorbjörg Hróarsdóttir, and Anna-Lena Wiklund. 2007. The Tromsø guide to the Force behind V2. *Working Papers in Scandinavian Syntax* 79, 93–118.
- Bentzen, Kristine. 2014a. Subject placement with respect to negation. *Nordic Atlas of Language Structures (NALS) Journal* 1, 344–352. <https://dx.doi.org/10.5617/nals.5403>
- Bentzen, Kristine. 2014b. Object Shift. *Nordic Atlas of Language Structures (NALS) Journal* 1, 332–343. <https://dx.doi.org/10.5617/nals.5402>
- Bentzen, Kristine and Merete Anderssen. 2019. The form and position of pronominal objects with non-nominal antecedents in Scandinavian and German. *Journal of Comparative Germanic Linguistics* 22(2), 169–188. <https://doi.org/10.1007/s10828-019-09105-w>
- Bentzen, Kristine, Merete Anderssen, and Christian Waldmann. 2013. Object Shift in spoken Mainland Scandinavian: A corpus study of Danish, Norwegian, and Swedish. *Nordic Journal of Linguistics* 36(2), 115–151. <https://doi:10.1017/S0332586513000218>
- Brandtler, Johan and David Håkansson. 2017. V2 eller V3? Om preverbal placering av adverbial i svenskan. *Norsk Lingvistisk Tidsskrift* 35, 11–26. <https://ojs.novus.no/index.php/NLT/article/view/1396>
- Erteschik-Shir, Nomi. 2005. Sound Patterns of Syntax: Object Shift. *Theoretical Linguistics* 31(1–2), 47–93. <https://doi.org/10.1515/thli.2005.31.1-2.47>
- Erteschik-Shir, Nomi and Gunlög Josefsson. 2017. Scandinavian object shift is phonology. In Laura R. Bailey and Michelle Sheehan (eds.), *Order and structure in syntax I: Word order and syntactic structure*, chap. 5, 99–115. Language Science Press. <https://doi.org/10.5281/zenodo.1117686>
- Hellan, Lars. 2005. Comments on Erteschik-Shir's article. *Theoretical Linguistics* 31(1–2), 137–145. <https://doi.org/10.1515/thli.2005.31.1-2.137>
- Heycock, Caroline, Antonella Sorace and Zakaris Svabo Hansen. 2010. V-to-I and V2 in subordinate clauses: an investigation of Faroese in relation to Icelandic and Danish. *Journal of Comparative German Linguistics* 13, 61–97.

<https://doi.org/10.1007/s10828-010-9035-7>

- Heycock, Caroline, Antonella Sorace, Zakaris Svabo Hansen, Frances Wilson and Sten Vikner. 2012. Detecting the late stages of syntactic change: the loss of V-to-T in Faroese. *Language* 88(3), 558–600. <https://doi.org/10.1353/lan.2012.0053>
- Holmberg, Anders. 1986. *Word order and syntactic features in the Scandinavian languages and English*. Doctoral dissertation, Stockholm University.
- Holmberg, Anders. 1993. Two subject positions in Mainland Scandinavian. *Working Papers in Scandinavian Syntax* 52, 29–41.
- Holmberg, Anders. 1999. Remarks on Holmberg’s Generalization. *Studia Linguistica* 53(1), 1–39. <https://doi.org/10.1111/1467-9582.00038>
- Holmberg, Anders and Christer Platzack. 1995. *The Role of Inflection in Scandinavian Syntax*. Oxford University Press.
- Jensen, Torben Juel and Tanya Karoli Christensen. 2013. Promoting the demoted: The distribution and semantics of “main clause word order” in spoken Danish complement clauses. *Lingua* 137, 38–58. <https://doi.org/10.1016/j.lingua.2013.08.005>
- Johannessen, Janne Bondi, Joel Priestley, Kristin Hagen, Tor Anders Áfarli, and Øystein Alexander Vangnes. 2009. The Nordic Dialect Corpus – an advanced research tool. In Kristiina Jokinen and Eckhard Bick (eds.), *Proceedings of the 17th Nordic Conference of Computational Linguistics NODALIDA 2009*, 73–80. <https://tekstlab.uio.no/nota/scandiasyn/>
- Jonas, Dianne Elizabeth. 1996. *Clause structure and verb syntax in Scandinavian and English*. Doctoral dissertation, Harvard University.
- Julien, Marit. 2007. Embedded V2 in Norwegian and Swedish. *Working Papers in Scandinavian Syntax* 80, 103–161. <https://lup.lub.lu.se/record/1032083>
- Julien, Marit. 2018. Om preverbale adverbialer. *Norsk lingvistisk tidskrift* 36, 161–178. <https://ojs.novus.no/index.php/NLT/article/view/1493>
- Klævik-Pettersen, Espen. 2018. *Inversion, V-to-C, and verb-second: An investigation into the syntax and word order of Old French and Late Latin*. Doctoral dissertation, University of Oslo.

- Lindstad, Arne Martinus, Anders Nøklestad, Janne Bondi Johannessen, and Øystein Alexander Vangsnes. 2009. The Nordic Dialect Database: Mapping Microsyntactic Variation in the Scandinavian Languages. In Kristiina Jokinen and Eckhard Bick (eds.), *Proceedings of the 17th Nordic Conference of Computational Linguistics NODALIDA 2009*, 283–286. <https://tekstlab.uio.no/nsd>
- Lundquist, Björn. 2014. The verb phrase: argument structure and particle placement. *Nordic Atlas of Language Structures (NALS) Journal* 1, 107–109. <https://dx.doi.org/10.5617/nals.5371>
- Lundquist, Björn. 2018. Mer om de preverbala adverbialens syntax, semantik och prosodi. *Norsk Lingvistisk Tidsskrift* 36, 179–195. <https://ojs.novus.no/index.php/NLT/article/view/1494>
- Mathôt, Sebastiaan, Daniel Schreij, and Jan Theeuwes. 2012. OpenSesame: An open-source, graphical experiment builder for the social sciences. *Behavior Research Methods* 44(2), 314–324. <https://doi.org/10.3758/s13428-011-0168-7>
- Nilsen, Øystein. 2003. *Eliminating Positions. Syntax and Semantics of Sentential Modification*. Doctoral dissertation, Utrecht University. <https://dspace.library.uu.nl/handle/1874/624>
- Platzack, Christer and Anders Holmberg. 1989. The role of Agr and finiteness in Germanic VO languages. *Working Papers in Scandinavian Syntax* 43, 51–76.
- Sandøy, Helge. 1976. *Laust samansette verb i vestnordisk: ein samanliknande leddstilingsanalyse for islandsk, færøysk og romsdalsmål*. Mag. phil. thesis, Universitetet i Oslo.
- Svenonius, Peter. 1996. The optionality of particle shift. *Working Papers in Scandinavian Syntax* 57, 47–75.
- Svenonius, Peter. 2002. Subject positions and the placement of adverbials. In Peter Svenonius (ed.), *Subjects, Expletives and the EPP*, 201–242. Oxford University Press.
- Svenonius, Peter. 2005. How Phonological is Object Shift? *Theoretical Linguistics* 31(1–2), 215–227. <https://doi.org/10.1515/thli.2005.31.1-2.215>
- Taraldsen, Knut Tarald. 1983a. Some phrase structure dependent differences between Swedish and Norwegian. *Working Papers in Scandinavian Syntax* 9, 1–45.
- Taraldsen, Knut Tarald. 1983b. *Parametric Variation in Phrase Structure*. Doctoral dissertation, University of Tromsø.

- Thráinsson, Höskuldur. 2007. *The Syntax of Icelandic*. Cambridge University Press.
- Thráinsson, Höskuldur. 2013. Full NP Object Shift: The Old Norse Puzzle and the Faroese Puzzle revisited. *Nordic Journal of Linguistics* 36, 153–186. <https://doi.org/10.1017/S033258651300022X>
- Vikner, Sten. 1995. *Verb movement and expletive subjects in the Germanic languages*. Oxford University Press.
- Weiß, Helmut. 2015. When the subject follows the object. On a curiosity in the syntax of personal pronouns in some German dialects. *Journal of Comparative Germanic Linguistics* 18(1), 65–92. <https://doi.org/10.1007/s10828-015-9071-4>
- Westendorp, Maud. 2018. New methodologies in the Nordic Syntax Database: Word order variation in Norwegian *wh*-questions. *Nordic Atlas of Language Structures (NALS) Journal* 3, 1–18. <http://dx.doi.org/10.5617/nals.6533>
- Westergaard, Marit and Øystein Alexander Vangsnes. 2005. *Wh*-questions, V2, and the Left Periphery of Three Norwegian Dialect types. *Journal of Comparative Germanic Linguistics* 8(1–2), 117–158. <https://doi.org/10.1007/s10828-004-0292-1>
- Westergaard, Marit, Øystein A. Vangsnes, and Terje Lohndal. 2017. Variation and change in Norwegian *wh*-questions: The role of the complementizer *som*. *Linguistic Variation* 17(1), 8–43. <https://doi.org/10.1075/lv.17.1.02wes>
- Wiklund, Anna-Lena, Kristine Bentzen, Gunnar Hrafn Hrafnbjargarson, and Thorbjörg Hróarsdóttir. 2009. On the distribution and illocution of V2 in Scandinavian *that*-clauses. *Lingua* 119(12). 1914–1938. <https://doi.org/10.1016/j.lingua.2009.03.006>
- Østbø Munch, Christine B. 2013. *North Germanic Negation: A Microcomparative Perspective*. Doctoral dissertation, University of Tromsø. <https://hdl.handle.net/10037/5410>
- Áfarli, Tor A. 1985. Norwegian verb particle constructions as causative constructions. *Nordic Journal of Linguistics* 8(1), 75–98. <https://doi.org/10.1017/S0332586500001268>
- Áfarli, Tor. 1986. Some Syntactic Structures in a Dialect of Norwegian. *University of Trondheim Working Papers in Linguistics* 3, 93–111.

CONTACTS

Björn Lundquist
UiT – The Arctic University of Norway
bjorn.lundquist@uit.no

Ida Larsson
University of Oslo
ida.larsson@iln.uio.no

Maud Westendorp
UiT – The Arctic University of Norway
maud.westendorp@uit.no

Eirik Tengesdal
University of Oslo
eirik.tengesdal@iln.uio.no

Anders Nøklestad
University of Oslo
anders.noklestad@iln.uio.no

APPENDIX

The drop-down menus in the search interface of the database contain abbreviations, which in most cases are transparent. Below we provide a list with all abbreviations, organized by the order of the menus.

TypeElement1

Adv: Adverb

NP: Noun phrase

Neg: Negation

Non-reflexive: This refers to verb class used in verb placement test. The verbs were either reflexive (e.g., ‘wash oneself’) or non-reflexive (e.g., ‘walk’).

PRO: Pronoun

Reflexive: See non-reflexive above.

TypeElement2

1st_PRO: First person pronoun (always singular)

3rd_PRO: Third person pronoun (SG or PL, FEM or MASC)

Adv: Adverb

NP: Noun phrase

Neg: Negation

PRO: Pronoun

Particle: Verb particle

Refl_PRO: The reflexive pronoun *sig/seg/sér/sær* (dative in some cases)

Aldrig: The adverb *never*, annotated in verb placement conditions.

Alltid: The adverb *always*, annotated in verb placement conditions.

Ofta: The adverb *often*, annotated in verb placement conditions.

ExactCategory

BridgeV: Bridge verb, category of matrix verb in embedded V2 condition

EmbObQ: Embedded object question (“X wondered *what* Y bought.”)

EmbQ: Embedded *yes/no* question, category of matrix type in embedded V2 condition

EmbSubQ: Embedded subject question (“X wondered *who* bought Y.”)

FillerDecl: Filler item, embedded declarative

LOS: Long object shift, only first person SG movement over inverted NP subject

Non-BridgeV: Non-bridge verb, category of matrix verb in embedded V2 condition

OA: Object shift over adverb

OS: Object shift over negation

Obj-Part: Object-particle, simple directional particle without following PP

Obj-PartGround: Object-particle, object functioning as ground of particle, e.g., ‘clear off the table’

Obj-PartMet: Object-particle, metaphoric/non-directional particle, e.g., ‘buy up the company’

Obj-PartPP: Object-particle, directional particle followed by source/goal PP

Obj-PartPrep: Object-particle, prepositional particle

Part-Sub: Particle and subject, the relative order between these two elements

Refl-LOS: Long object shift, with simple reflexive object (*sig/seg/sér/sær*)

SS: Subject shift, the placement of a subject with respect to a sentence adverb

SS-Refl: Subject shift in the context of a reflexive object

ProducedWordOrder

This menu contains the exact order of elements in the elicited responses. We list abbreviations for each element below:

A: Adverb

N: Negation

O: Object

R: Reflexive

S: Subject

V: Verb

In addition, the two annotations *SOM* and *NON* are used for the category embedded subject question. *SOM* indicates that a complementizer was inserted after the *wh*-word (“X wondered who *that* came to the party.”) and *NON* indicates the absence of such an element. Embedded object questions with an extra complementizer have also been annotated with *SOM*, but in absence of *SOM*, the order between embedded subject and verb has been annotated.