



Fra gammelt seddelarkiv til nye forskningsresultater: Nytt lys på gammelnorsk

Av Juliane Tiemann

Denne artikkelen tar for seg livssyklusen til datamateriale med utgangspunkt i det fysiske materialet i Språksamlingene ved Universitetet i Bergen. Materialet er gamle ordbokssedler over norrønt/gammelnorsk som senere ble overført til digitale løsninger. Det blir skissert hvordan dette materialet har blitt brukt om igjen for å bygge opp og supplere et nytt digitalt korpus (Ko-NoKs = Korpus over den norske *Konungs skuggsjá*) utformet for å besvare nye forskningsspørsmål. Som eksempel på en ny analyse basert på dette gamle forskningsmaterialet, rapporterer artikkelen om noen resultater fra en korpusstudie av syntaktisk variasjon i gammelnorsk.

1 Introduksjon

Studier av eldre språk er i sin natur korpusbaserte. Analysearbeid krever først og fremst at materialet er tilgjengelig, helst allerede i digital form, og at korpus inneholder opplysninger som trengs for analysen. I dag kan vi bygge videre på allerede eksisterende korpus for språkmaterialet til eldre norsk og norrønt i både fysisk og digital form. Språksamlingene ved Universitetet i Bergen inneholder blant annet et seddelarkiv (opprettet under arbeidet med det tidligere Gammalnorsk ordboksverk¹) med om lag 650 000 ordbokssedler over norrønt/gammelnorsk. En god del av denne ressursen ble overført til digitale arkiver under Dokumentasjonsprosjektet ved Universitetet i Oslo på 1990-tallet. Men opplysningene for flere tusen ord er fremdeles kun lagret på sedler som ligger i arkivskuffene i Språksamlingenes depot.

1. Del av Dokumentasjonsprosjektet (delprosjekt gammelnorsk) som var et samarbeid mellom de humanistiske miljøene ved Universitetet i Bergen, Universitetet i Oslo, Universitetet i Trondheim og Universitetet i Tromsø.

Denne artikkelen beskriver de forskjellige stegene fra det fysiske materialet og merking av sedlene i seddelarkivet til oppretting av digitale ressurser og analyse av ordstillingsvariasjon i gammelnorsk, basert på et nytt XML-basert korpus. Ved å bruke det fysiske materialet i Språksamlingene bygges det videre på det språkdokumentariske arbeidet som ble satt i gang for nesten 100 år siden. Rammen for arbeidet er PhD-prosjektet «Syntactic Variation and Information Structure in Old Norwegian: An Investigation of *Konungs skuggsjá* in AM 243 bā fol.» ved Universitetet i Bergen. Dette prosjektet fokuserer særlig på samspill mellom syntaks og informasjonsstruktur. Informasjonsstruktur er et komplekst fenomen som refererer til grensesnittet mellom struktur og betydning av språklige ytringer, og som opererer både på setnings- og kontekstnivå. For en analyse av informasjonsstrukturelle virkninger på den endelige setningsstrukturen kreves det derfor en detaljert analyse av språkmaterialet med merking på forskjellige lingvistiske nivåer. Dette er et stort og tidkrevende arbeid, der analysen som allerede har blitt gjort på ordbokssedlene, er en ressurs som forenkler denne prosessen, og som i tillegg forhindrer at det samme arbeidet må gjøres flere ganger. På grunnlag av eksisterende seddelmateriale ble det nye korpuset (KoNoKs = Korpus over den norske *Konungs skuggsjá*) utvidet med annotasjon for flere språklige detaljer spesifikt knyttet til analyse spørsmålene i studien (se kap. 3). I artikkelen vil et forskningseksempel brukes for å vise fram både arbeidet med det gammelnorske materialet og hvilke typer resultater som kan komme ut av et slikt forskningsprosjekt (se kap. 4). Før beskrivelsen av arbeidet med seddelarkivet vil den teoretiske rammen for dette arbeidet skisseres i kap. 2. Det vil også legges vekt på betydningen av analyser av gammelnorsk språkmateriale separat fra gammelislandsk språkmateriale.

2 Syntaktisk variasjon: Nytt prosjekt om gammelnorsk

Variasjon i og endring av ordstilling over tid er godt dokumentert for eldre germanske språk. I løpet av de siste tiårene har fokuset i språkstudiene skiftet mot analyser og beskrivelser av de mekanismene som styrer og utløser variasjon, der forskere særlig har sett på informasjonsstruktur og prosodisk vekt (jf. Taylor og Pintzuk 2012 a,b; Elenbaas og van Kemenade 2014; Walkden 2014). Selv om dette er et aktuelt forskningstema

som er hyppig diskutert for flere språk, mellom annet gammelengelsk, gammelhøytysk og gammelislandsk, har syntaktisk variasjon i gammelnorsk ikke fått mye oppmerksomhet i studier innen vestnordisk. Beskrivelser av syntaktisk variasjon i den vestnordiske språkgreinen skiller ikke mellom gammelislandsk (ca. 1100–1540) og gammelnorsk (ca. 1050–1350), og det blir antatt at det som gjelder for gammelislandsk også gjelder for gammelnorsk. Dette skjer selv om studier innen syntaks hovedsakelig har blitt utført på grunnlag av gammelislandsk materiale (jf. t.d. Thorbjörg Hróarsdóttir 2000, 2009). Praksisen med å ikke skille mellom disse to språkene på syntaktisk nivå blir gjentatt i anerkjent oversikt litteratur over norrønt språk, der t.d. Faarlund (2004: 2) skriver at “[t]here are a few minor phonological and morphological differences between Old Icelandic and Old Norwegian, but no known syntactic differences”. Også Mørck (2016: 320f.) viderefører dette i den nye norske språkhistorien, *Norsk språkhistorie I: Mønster*, der han skriver at «[s]iden det ser ut til å være ganske små syntaktiske forskjeller innanfor det norrøne området i denne perioden, kan vi tillate oss å gå ut fra at beskrivelsene av norrønt med sitt i stor grad islandske grunnlag i det store og heile er gyldige også for gammelnorsk.»²

Hovedgrunnen til at kildene for forskjellige typer analyser primært er islandske og ikke norske, er at omfanget på kilder og opparbeidet/annotert korpusmateriale er større for gammelislandsk (jf. The Icelandic Parsed Historical Corpus/IcePaHC)³ enn for gammelnorsk. Både synkrone og diakrone analyser av gammelnorsk syntaks gir dermed et ufullstendig bilde, mens ordstillingsvariasjon i gammelislandsk har blitt analysert fra flere forskjellige vinkler og med materiale som strekker seg over flere år-

2. Slike påstander er basert på ordstillingsvariasjon som kan observeres i begge språk, men tar ikke inn detaljerte komparative analyser. Slike analyser kan imidlertid vise tendenser for forskjellige ordstillingsmønstre, og med dette også syntaktiske forskjeller mellom disse språkene. Ulike tendenser i språkene adskiller disse over lengre tid som synliggjør forskjeller mellom islandsk og norsk (og islandsk og de andre fastlandsskandinaviske språkene), t.d. er rekkefølgen av det finitte verbet og setningsadverbialer i leddsetninger forskjellig. I moderne islandsk står verbet i andre posisjon, både i hoved- og leddsetninger (jf. Franco 2010: 26f.; Holmberg og Platzack 2008: 437f.), mens i moderne norsk følger verbet setningsadverbialet. Også rekkefølgen av PP-er relativt til NP-er viser forskjeller mellom insulære og fastlandsskandinaviske språk (jf. Thorbjörg Hróarsdóttir 2000: 191).
3. http://www.linguist.is/icelandic_treebank/. Versjon 0.9 (2011). Prosjektet der korpuset ble bygget opp, ble finansiert av Icelandic Research Fund, U.S. National Science Foundation og Rannsóknasjóður Háskóla Íslands. Se Wallenberg m.fl. (2011).

hundrer (jf. bl.a. Halldor Sigurðsson 1985, 1988, 1994; Thorsteinn Indriðason 1987; Eiríkur Rögnvaldsson 1994/1995, 1996; Thorbjörg Hróarsdóttir 1996, 2000, 2009; Haugan 2000).⁴ Fraværet av gammelnorsk i slike studier er en god grunn til å fokusere mer på dette språklige materialet. Behovet for syntaktisk annoterte gammelnorske tekster ble først fanget opp i Menotec-prosjektet⁵ som er publisert i trebanken til PROIEL.⁶ I dette prosjektet valgte man en dependensanalyse som syntaktisk modell. Bortsett fra KoNoKs-korpuset finnes det så langt ingen gammelnorske korpus som er bygget opp etter frase-strukturelle analyser. KoNoKs ble utviklet i tilknytning til forskningsprosjektet «Syntactic Variation and Information Structure in Old Norwegian» som gir en analyse av ordstillingsvariasjon i gammelnorsk fra et pragmatisk perspektiv med søkelys på påvirkningen av informasjonsstruktur og av prosodisk vekt som avgjørende faktorer for ulik ordstilling. Fremgangsmåten for dette var en næranalyse av én gammelnorsk tekst (se kap. 3) med fokus på blant annet variasjon i objektsposisjon, adjektivposisjon og possessivkonstruksjoner. Dataene fra analysene ble sammenlignet med resultater fra lignende analyser for gammelislandsk (særlig Thorbjörg Hróarsdóttir 2000, 2009; Pfaff 2019). Et slikt komparativt perspektiv med gammelislandsk danner et bedre grunnlag for enten å bekrefte at beskrivelser for gammelislandsk syntaks også er gyldige for gammelnorsk, eller for å vise at det faktisk finnes forskjeller mellom de to språkene (og dermed eksisterende forskningshull). Korpuset KoNoKs er dermed et nytt, nyttig og viktig grunnlag for videre arbeid med og analyser av syntaktisk variasjon i gammelnorsk og for videre komparative studier, både når det gjelder syntaks og informasjonsstruktur. Noen av resultatene for syntaktisk variasjon med fokus på objektsposisjonen, dvs. opposisjonen i ordstillingene OV og VO og hvordan informasjonsstruktur og prosodisk vekt styrer denne variasjonen, vil kort skisseres i denne artikkelen.

4. Selv om vi har en rekke publikasjoner som fokuserer på gammelnorske kilder fra forskjellige perspektiver (jf. Flom 1921–1923; Holm-Olsen 1952; Schnell og Simek 2000; Christoffersen 1994, 2002; Haugen og Ommundsen 2010), så er disse studiene ofte mer begrensede og det finnes en mindre mengde lingvistiske analyser enn for gammelislandsk.
5. <https://www.menota.org/menotec.xml>. Dette er et infrastrukturprosjekt finansiert av Norges forskningsråd (2010-2012).
6. <https://proiel.github.io/>. Den syntaktiske analysen i dette prosjektet er basert på et dependens-grammatisk rammeverk og skiller seg dermed fra analysen gjort i PhD-prosjektet som er presentert her.

2.1 *Det teoretiske rammeverket: Kort om informasjonsstruktur*

Ifølge Lambrecht (1994: xv) kan informasjonsstruktur bli forstått som en grammatisk komponent som styrer «the relationship between speaker assumptions and the formal sentence structure». Han (1994: xiii) skriver videre at

[i]n the information structure component of language, propositions as conceptual representations of states of affairs undergo pragmatic structuring according to the utterance contexts in which these states of affairs are to be communicated. Such pragmatically structured propositions are then expressed as formal objects with morphosyntactic and prosodic structure.

Hinterhölzl og van Kemenade (2012) antar at ordstilling er en del av setningsgrammatikken, mens informasjonsstruktur er en del av pragmatikken. På denne måten styrer syntaksen hvordan en setning kan se ut (her refereres det til alle mulige overflatestrukturer samt restriksjoner i et språk), mens pragmatikken bestemmer den spesifikke formen av en setning som passer i en gitt kontekst. Hinterhölzl og van Kemenade (2012: 819) argumenterer videre for at «syntax, [information structure], and prosody all three influence the choice for a particular word order pattern» som viser at «[information structure] interacts with the entire grammatical system, rather than with just one component.» Informasjonsstruktur er altså et komplekst fenomen og definisjon og tilhørende kategorier er hyppig diskutert i litteraturen (jf. Levinson 1983; Kruijff og Duchier 2003; Kruijff-Korbayová og Steedman 2003). I denne artikkelen spiller denne diskusjonen og bruken av forskjellig terminologi (tema-remå vs. topik-fokus, osv.)⁷ en mindre rolle. Her vil det informasjonsstrukturelle nivået, noe forenklet, defineres ved bruk av den følgende passasjen gitt i Matic (2015: 95):

Utterances transmit both the information contained in the message and the implicit or explicit instructions on how this information is to be processed and integrated into the hearer's knowledge stock. ... The information comes about by relating the 'new' elements (i.e., what the hearer is assumed not to be aware of) to these 'given' elements.

7. Dikotomien tema-remå (internasjonalt bruker man heller *topic-comment* som begreper der *comment* inneholder et framhevet fokuselement) beskriver ulike deler av en setning som viser til hva setningen dreier seg om (tema) og hva som er ny informasjon relatert til setningens tema (remå).

Dette beskriver altså hvordan informasjon er presentert i den aktuelle kommunikative utvekslingen (form, uttrykk og strukturering), der kontekstinformasjon/en diskurs fortløpende oppdateres med ny informasjon. Dette er et rekursivt fenomen der informasjonen deles opp i forståelige diskursenheter (jf. Chafe 1976 «information packaging») relatert til det som allerede er kjent i en kontekst (jf. Roberts 2012 «questions under discussion»). Som nevnt ovenfor opererer informasjonsstruktur både på setningsnivå og diskursnivå.

Det teoretiske rammeverket for analyse av informasjonsstruktur ble opprinnelig utviklet for analyse av moderne språk med sterkt fokus på intonasjonsmønstre. Noen av de mest fremtredende informasjonsstrukturelle kjennetegnene slik som prosodisk prominens (trykk) er ikke direkte tilgjengelige i historisk materiale. For en informasjonsstrukturell analyse av språklig materiale uten tilgang til et aktivt intonasjonssystem er det derfor viktig å kunne bygge på veldefinerte annotasjonskategorier på flere språklige nivåer og utprøvde analysemetoder. Flere arbeider de siste tiårene setter søkelys på nettopp dette, og inkluderer store mengder korpusmateriale for forskjellige eldre språk som latin, hettittisk, gammelholtyrisk og gammelengelsk.⁸

8. Flere store prosjekter som tar for seg informasjonsstrukturelle analyser for både moderne og eldre språk er finansiert av Deutsche Forschungsgemeinschaft: a) SFB-prosjektet «Information structure: the linguistic means for structuring utterances, sentences and texts» (<http://www.sfb632.uni-potsdam.de/>, 2003-2015), b) «Informationsstruktur in älteren indogermanischen Sprachen» (<https://gepris.dfg.de/gepris/projekt/109055449?context=projekt&task=showDetail&id=109055449&>, 2009–2016) og c) «Informationsstruktur in komplexen Sätzen – synchron und diachron: Erarbeitung eines diachronen Kontrastkorpus» (<https://gepris.dfg.de/gepris/projekt/199843560>, 2011–2017). I Norge ble det gjennomført et prosjekt som fokuserer på germanske og romanske språk, finansiert av Norges forskningsråd og Institutt for litteratur, områdestudier og europeiske språk ved Universitetet i Oslo: «Information Structure and Word Order Change in Germanic and Romance Languages» (<https://www.hf.uio.no/ilos/english/research/projects/iswoc/index.html>, 2011–2014); jf. Bech og Eide 2014). Dette prosjektet inkluderte også en analyse av gammelnorsk, men selv om norsk var en del av dette prosjektet, ble kun deler av *Óláfs saga ins helga* (Upps DG 8 II) analysert under arbeidet, og så langt finnes det ingen publikasjoner som tar for seg resultatene for det gammelnorske materialet. I tillegg er det viktig å poengtere at annotasjonen i ISWOC følger en dependens-grammatisk analyse (i samme tråd som analysene presentert i PROIEL) som viser til hierarkisk ordning av konstituenten i motsetning til en frase-strukturell lineær analyse.

2.2 Annotasjonsarbeid: Kort om KoNoKs

For å kunne ekstrahere informasjonsstrukturelle entiteter (fokus og topik/tema) og for å kompensere for manglende intonasjonale faktorer i det eldre språklige materialet på en effektiv måte, må analysen bygge på en kumulativ oppsummering av tydelige kategorier i forskjellige strukturelle domener. En multi-modal korpustilnærming⁹ er dermed nødvendig (jf. også Donhauser 2007; Petrova m.fl. 2009). Denne annotasjonen tar utgangspunkt i en omfattende grammatisk merking inklusive lemmatisering av materialet. For KoNoKs er den første delen av annotasjonen således det som ble gjort på ordbokssedlene; dette er grunnlaget for alt videre arbeid. Den videre annotasjonen følger koherente annotasjonsretningslinjer for eldre korpusspråk som har blitt skissert i nasjonale og internasjonale prosjekter (se fn. 8). Annotasjonen av informasjonsstrukturelle enheter bygger på en kombinasjon av særlig semantiske, pragmatiske og diskursrelaterte annotasjoner, slik som “saliency” (animathet), “givenness”, “definiteness” eller “anaphoric/cataphoric relations”. På den måten kan et element som er annotert for å være kjent, bestemt og har anaforisk referanse, analyseres som en sterk topik-kan-

	179	180	181	182	183	184	185	186	187	188
[text]	cn	þcir	isar	liggia	mæirr	i	lanndnorðr	cða	til	norðs
[lem]	but	this	ice	lie	more	in	north.land	or	to	north
[trans]	en	sá	íss	liggja	mæir	i	landnorðr	cða	til	norðr
[glos-morph]		NOM.M.PL	NOM(M).PL	PRS.IND.ACT.3PL	comp		ACC(N).SG			GEN(N).SG
[pos]	conj	prdem	noun	v.fin	adv	prep	noun	conj	prep	noun
[v-sem]										
[saliency]		pr3.dem	abstract	pr3			abstract			abstract
[givenness]		access-sit	giv				new			new
[definitness]		def					indef			indef
[gramm-def]			indef							
[context]		demon/deict.ref								
[frame]		schemeice								
[l-particle]										
[shift]		continue								
[Topic]		Con-T								
[position T]		initial								
[Comment]	cm			cm						

Fig. 1: Skjermdump av annotasjonen for en topik-kandidat (10r, kol.a:26–27).

- Et multi-modalt korpus er en datakolleksjon som brukes for å studere hvordan to eller mer modaliteter (forskjellige nivåer) knyttes sammen og påvirker hverandre.

	180	181	182	183	184	185	186	187	188	189	190	191	192	193	194	195	196
[text]	þeir	isar	liggja	meiðr i		larndnorðr	eða til	norðrs	firi	larndino	hæll	en	til	suðrs	eða	utsuðrs	
[trans]	this	ice	lie	more	in	north.land	or to	north	country	rather	but	to	south	or	southwest		
[lem]	sá	iss	liggja	meiðr i		landnorðr	eða til	norðr	fyri	land	heldr	en	til	suðr	eða	útsuðr	
[glos-morph]	NOM.M.PL	NOM(M).PL	PRS.IND.ACT.3PL	comp		ACC(N).SG		GEN(N).SG	DAT(N).SG	comp				GEN(N).SG		GEN(N).SG	
[pos]	prdem	noun	v.fin	adv	prep	noun	conj	prep	noun	prep	noun	adv	adv	prep	noun	conj	noun.prop
[v-sem]																	
[salience]	pr3.dem	abstract	pr3			abstract		abstract	abstract					abstract		abstract	
[givenness]	access-sit	giv						new	giv					new		new	
[definitness]	def					indef		indef	def					indef		indef	
[gramm-def]		indef															
[context]	demon/dict.ref																
[frame]	scheme:ice																
[l-particle]										scheme:greenland							
[shift]	continue																
[Topic]	Con-T																
[position-T]	initial																
[Comment]			cm														
[NFocus]																	
[CFocus]						cf									cf		
[position-F]						middle									final		

Fig. 2: Skjermdump av annotasjonen for en fokus-kandidat (10r, kol.a:26–29).

didat og et element som er annotert som nytt, ubestemt og kontrastivt som en sterk fokus-kandidat. Den kumulative analysen som viser forskjellige og overlappende trekk, er illustrert i figur 1 og 2 (over).¹⁰

For slikt analysearbeid er det som nevnt nødvendig at språkmaterialet er opparbeidet med bred grammatisk informasjon (for en detaljert beskrivelse av annotasjonen se Tiemann unpubl. manuskript: Kappa).

3 Datamateriale: Fra sedler til datafil

Utarbeidede forskningsressurser for det gammelnorske språket eksisterer i dag både i fysisk og digital form. En stor del av de digitale ressursene med morfologisk merking har utgangspunkt i seddelarkivet i Språksamlingene. De gamle ordbokssedlene er oppbevart i skuffer og lagret i Språksamlingenes depot (ett av arkivene/magasinene til Språksamlingene i Bergen) som vist i figur 3.

Analysearbeidet i PhD-prosjektet som er presentert her, tar hovedsakelig for seg én større og sammenhengende gammelnorsk tekst, *Konungs skuggsjá* i det norske hovedhåndskriftet, AM 243 bα fol. Verket er en pedagogisk-filosofisk tekst i form av en dialog mellom far og sønn,

10. Referansen gitt i parenteser leses på følgende måte: manuskriptside (recto og verso), kolonne (a og b), og linjenummer for eksempelet.



Fig. 3: Språksamlingenes depot – seddelarkivet.

skrevet på midten av 1200-tallet. I omfang er *Konungs skuggsjá* ingen liten tekst, og selv om det er noen lakuner i AM 243 bā fol., inneholder dette håndskriftet 63 910 ord. Teksten er del av Språksamlingene og har her fått full grammatisk oppmerking på ordbokssedler (se kap. 3.1). Disse sedlene er lagret i spesiallagede syrefrie bokser og er sortert alfabetisk, som vist i figur 4.



Fig. 4: Gammelnorske ordbokssedler – Konungs skuggsjá.

3.1 Digitalisering

Hver ordboksseddel i samlingen har seddelformat ca. 8 x 18 cm og inneholder 6–10 linjer med tekst, som vist i figur 5. Dette sikrer at hvert opp-

slagsord (leksem) er gitt i en best mulig kontekst. Alle oppslagsordene i hvert tekststykke har hver sin seddel.

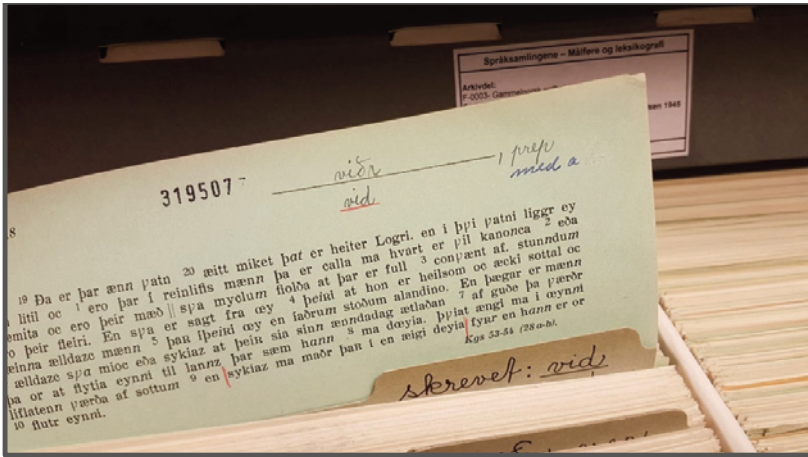


Fig. 5: Ordbokseddel – detalj.

Hvert ord har fått grammatisk bestemmelse (ordklasse og bøyningsform), og er ført opp som oppslagsord i normalisert form øverst på seddelen, ved siden av ordets skriveform i håndskriftet. Alt materialet som finnes på slike sedler, inklusive all grammatisk merking, burde vært overført til digitale databaser. Men selv om en stor del av Språksamlingene allerede er digitalisert, finnes det mellom annet i seddelarkivet fremdeles materiale som ennå ikke er fullstendig overført til digitale arkiver.

Middelaldertekstene fra samlingen blir innlemmet i Menota-arkivet¹¹ der også teksten til *Konungs skuggsjá* er tilgjengelig. Figur 6 viser en del av katalogen i Menota der Språksamlingene er nevnt som hovedredaktør. Som del av Dokumentasjonsprosjektet ble materialet på kartotek kortene fra den eldste fasen av det norske språket innlest, digitalisert og senere overført til menotisk XML,¹² og dermed gjort om til elektroniske tekster. Menotisk XML følger en TEI standard¹³ og den tilføyde annotasjonen er på et morfologisk nivå i tråd med bestemmelse foretatt på ordboks-

11. <https://www.menota.org/forside.xhtml>, (2001 – fortløpende).

12. XML-kodingen som følger retningslinjene oppgitt for Menota. Se Menota Handbook 3.0: https://www.menota.org/HB3_index.xml.

13. <https://tei-c.org/>.

FRA GAMMELT SEDDELARKIV TIL NYE FORSKNINGSRISULTATER

Signature	MS or Work Title	Main Editor(s)	Date	Lang.	Words	Lemma	Form
AM 178 fol (1r.1-17r.17)	Piðriks saga af Bern (part 1 of 22)	Språksamlingane	c. 1600-1700	isl	9216	partly	partly
AM 178 fol (47r.3-49r.9)	Piðriks saga af Bern (part 10 of 22)	Språksamlingane	c. 1600-1700	isl	1375	completely	completely
AM 178 fol (58r.19-58v.6)	Piðriks saga af Bern (part 3 of 22)	Språksamlingane	c. 1600-1700	isl	176	completely	completely
AM 178 fol (60r.16-61r.15)	Piðriks saga af Bern (part 5 of 22)	Språksamlingane	c. 1600-1700	isl	672	completely	completely
AM 243 ba fol	Konungs skuggsjá	Språksamlingane	c. 1275	nor	63910	completely	completely
AM 619 4to	Old Norwegian homily book	Språksamlingane	c. 1200-1225	nor	62237	completely	completely
DG 4-7	Strengleikar	Språksamlingane	c. 1270	nor	38453	completely	completely
Holm perg 17 4to	Thómass saga erkbiskups	Språksamlingane	c. 1300	nor	59884	completely	completely

Fig. 6: Katalog i Menota.

sedlene. Målet med Dokumentasjonsprosjektet var å bygge opp et fullstendig ordarkiv med alle forekomster av ord i norsk språk fra de første skrevne beleggene på norsk fra omkring 1200. Dette skulle gi en samlet og enkel tilgang til norsk middelalderspråk.

Filene i Menota kan lastes ned, og har en omfattende 'header'¹⁴ med utfyllende informasjon om blant annet annotatorer, endringer og lisens.

Figur 7 (neste side) viser hvordan samme tekststykke av *Konungs skuggsjá* ser ut i de forskjellige formatene fra faksimilefotografi til kartotekkort og til slutt i html-visning på Menotas hjemmeside, der teksten er publisert med en diplomatarisk transkripsjon (markeringen av ordet som er tatt opp på kortet, er tilføyd både i fotografi og i den elektroniske transkripsjonen).

Som nevnt ovenfor er en stor del av seddelarkivet allerede digitalisert og overført til Menota. Men omtrent en tredjepart av annotasjonen for *Konungs skuggsjá* hadde uheldigvis falt ut under konverteringen, og den grammatiske bestemmelsen for ordene i denne delen fantes utelukkende i seddelformat. Det var derfor nødvendig å gå tilbake til seddelarkivet for å fullføre annotasjonen som trengs for videre syntaktisk analyse av teksten. Det arkiverte fysiske materialet fikk i denne omgangen igjen en enorm verdi for forskning og kunne brukes som grunnlag til nye korpusanalyser. Uten arbeidet som skaperne av ordbokssedlene allerede hadde gjort, hadde det tatt så lang tid å fullføre all annotasjonen for prosjektet om syntaks og informasjonsstruktur i gammelnorsk at det sann-

14. En XML-header spesifiserer XML-versjonen og gir ytterligere informasjon om kodingen.

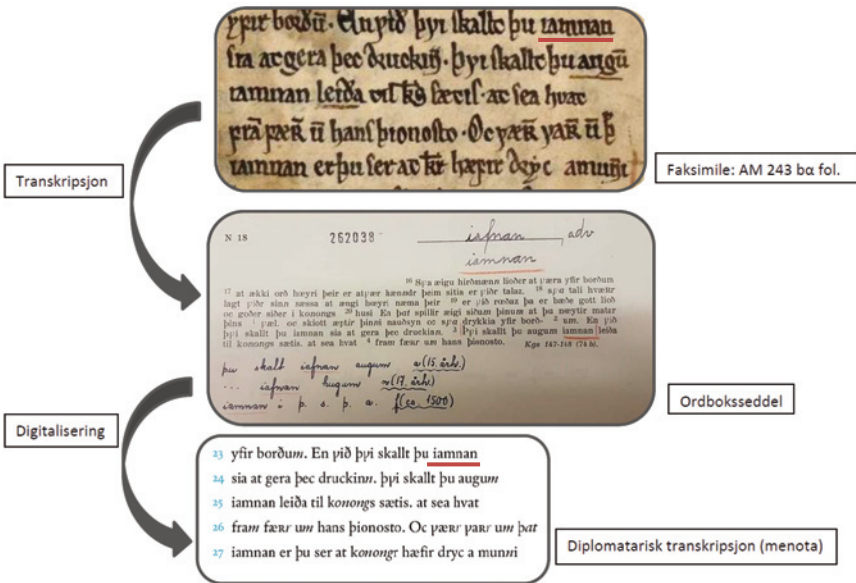


Fig. 7: Eksempel – tekststykke i de forskjellige formatene.

synligvis ville ha gått langt utover prosjektets omfang. Selv om mye av merkingen kan tilføyes nokså raskt i en XML-editor, slik at arbeidet generelt sett går mye forttere enn før det datamaskinelle arbeidet, er det fremdeles tidkrevende. Analysen som er skrevet på sedlene, er da en ressurs som forenkler arbeidsflyten med teksten og annotasjon (jf. også Tiemann 2018). I tillegg vil konsultering av dette materialet forhindre at det samme arbeidet må gjøres flere ganger.

3.2 Videre arbeid med datafilene

For den videre syntaktiske og informasjonsstrukturelle annotasjonen kunne arbeidet ikke lenger bli gjort i TEI XML-kodingen brukt i Menota siden det måtte defineres flere overlappende annotasjonsnivåer. I tillegg var det nødvendig å arbeide videre med teksten i en lineær visning, særlig for å se setningene bedre i kontekst og med mulighet for å ha både en horisontal og vertikal akse. For det videre arbeidet ble det derfor brukt en fritt tilgjengelig editor som også har blitt brukt innen store internasjonale forskningsprosjekter om informasjonsstruktur i eldre språk (jf.

fn. 8): EXMARaLDA-Partitur Editor.¹⁵ Gjennom et XSLT-stilark¹⁶ ble menotisk XML omdannet til EXMARaLDA-XML, slik at all morfologisk annotasjon samt selve teksten kunne vises i vertikale linjer som vist i figur 8. I tillegg ble det tilføyd flere annotasjonsnivåer med spesifisert verdi og noen med tom verdi i dette stilarket. Tomme nivåer kunne fylles inn ved behov under den videre analysen. EXMARaLDA gjør det mulig å definere og å sette annotasjonskategorier helt fritt. Som diskutert ovenfor er informasjonsstruktur et komplekst språklig fenomen som strekker seg over setningsgrenser og der konteksten spiller en viktig rolle. For å kunne peke på informasjonsstrukturelle kategorier er det derfor nødvendig å arbeide med flere overlappende nivåer samtidig.

Som vist i figur 8 er den horisontale akse brukte for verdiene av de forskjellige annotasjonene. Den vertikale akse viser de forskjellige annotasjonene der det er mulig å få en oppsummering av alle trekk i ulike strukturelle domener for å identifisere informasjonsstrukturelle komponenter (jf. fig. 1 og 2). Kategoriene for informasjonsstrukturelle kandidater er analysert helt mot slutten av annotasjonsarbeidet. På grunn av en lett tilgjengelig oppsummering blir en ekstraksjon av disse elementene transparent og enkel å komme tilbake til på et senere tidspunkt.

26	27	28	29	20	21	22	23	24	25	26	27	28
oc	heita	þau	iafnan	siðan	hins	hælgja	Kerinus	æpli	<p>	oc	fara	þau
ok	heita	ei	jafnan	eiðan	tunn	heilagr	Kerinus	æpli		ok	fara	ei
and	call	this	always	since	the	holy	Kerinus	apple		and	go	they
	PRS.IND.ACT.3PL	NOM.N.PL	pos	pos	OEN.M.SG	OEN.M.SG.pos	OEN(M).SG	NOM(N).PL			PRS.IND.ACT.3PL	NOM.N.PL
conj	v.ñn	prdem	adv	adv	prdem	adj	noun.prop	noun		conj	v.ñn	prdem
#												performative
pr3		pr3			pr3.dam		human	concrete			pr3	pr3
		access-rit			access-rit		grv	grv				access-rit
		def			def	def	def	indef				def
hactive		usa.ref			demon/defct.ref						hactive	usa.ref
		schema:apples					schema:Kerinus:apple					schema:apples
		chuff										continue
		S-T										Con-T
		middle										2P
om			om							om		
					rf							
					imal							

Fig. 8: EXMARaLDA-editor visning.

15. <https://exmaralda.org/en/partitur-editor-en/>.

16. Et XSLT-stilark beskriver gjennom et sett med regler en transformasjon fra et XML-dokument til et nytt XML-dokument.

Etter annotasjonen ble filene lastet opp i visualiserings- og søkeverktøyet ANNIS (ANNotation of Information Structure)¹⁷ som inneholder en søkemonitor for å ekstrahere resultatene for den endelige analysen av materialet.

4 Studie om gammelnorsk ordstilling

For å vise noen av resultatene som et slikt videre arbeid med gammelt materiale kan gi, vil siste del av denne artikkelen kort skissere en studie av syntaktisk variasjon og informasjonsstruktur i gammelnorsk som ble gjennomført i forbindelse med PhD-prosjektet «Syntactic Variation and Information Structure in Old Norwegian». Etter gjennomført ekstraksjon av dataer fra ANNIS kunne disse dataene analyseres innenfor en informasjonsstrukturell ramme og i tråd med modeller innen generativ syntaks. Det vil ikke bli gjort detaljert rede for selve analysen eller litteraturen i dette avsnittet. Se Tiemann (2022) for en fullstendig diskusjon.

4.1 Objektsposisjon og informasjonsstatus

Studien av objektsposisjon følger særlig analysearbeider for gammelislandsk, der Thorbjörg Hróarsdóttir (2000, 2009) har undersøkt en stor mengde materiale fra et pragmatisk perspektiv. Tabell 1, hentet fra Thorbjörg Hróarsdóttir (2009), viser samspillet mellom informasjonsstatusen

Texts	New information			Old information		
	OV	VO	% OV	OV	VO	% OV
14th c.	75	204	26.9	156	36	81.3
15th c.	48	136	26.1	110	24	82.1
16th c.	50	132	27.5	77	17	82.0
17th c.	79	314	20.1	215	104	67.4
18th c.	20	98	16.9	39	43	47.6
19th c.	84	1,312	6.0	261	1,072	20.0

Tab. 1: Datsett for OV/VO og informasjonsstatus, Thorbjörg Hróarsdóttir (2009: 264).

17. <https://corpus-tools.org/annis/>. ANNIS ble designet i SFB-prosjektet nevnt ovenfor (se fn. 8).

(kjent – nytt)¹⁸ og ordstilling av objekter (pre- eller postverbal stilling; relativt til hovedverbet) i islandsk.

Tabellen viser at objekter som bærer kjent informasjon, oftere står i en preverbal posisjon. Informasjonsstatusen til et objekt spiller dermed en viktig rolle for dets plassering i setningen. Dette følger den generelle og velkjente tendensen for grammatisk ordning der kjent informasjon kommer før ny informasjon. Siden forskningstradisjonen ikke skiller mellom gammelnorsk og gammelislandsk på et syntaktisk nivå, er nullhypotesen at man finner de samme resultatene for gammelnorsk. Det betyr at fordelingen av objekter¹⁹ på OV/VO i gammelnorsk korrelerer med informasjonsstatusen til objekter på samme måte og i samme grad som i gammelislandsk. Vi forventer altså et relativt homogent sett

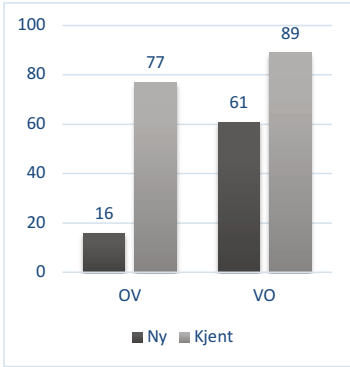


Fig. 9: Plassering av kjente og nye objekter i KoNoKs-datasettet.

av objekter i OV-stilling (kjent; beveger seg til en preverbal stilling på grunn av en informasjonsstrukturell utløser) og et heterogent sett av objekter i VO-stilling (ikke alle objekter vil nødvendigvis bli påvirket av utløseren til bevegelsen; andre faktorer kan forhindre bevegelsen). Datasettet tatt fra KoNoKs består av 243 setninger som har referensielle objekter i leddsetninger med to verb (dvs. hovedverb og hjelpeverb).²⁰ Figur 9 viser plasseringen av kjente og nye objekter i disse setningene (se Tiemann 2022 for gammelnorske eksempler).

18. Kjent informasjon er enten allerede nevnt i konteksten eller identifiserbar (dette inkluderer delt kunnskap og kunnskap om verden); ny informasjon er ikke nevnt før, uforutsigbar og ikke konkluderbar.
19. *Objekt* defineres her som komplement som er krevet av hovedverbet. Både nominale objekter og preposisjonalobjekter er tatt med i analysen.
20. Strukturer med et referensielt objekt og to verb sikrer at man har strukturer som viser bevegelse av objektet og ikke bevegelse av verbet. Setninger med negative og kvantifiserende objekter ble ekskludert fra studien. Disse objektene er ikke referensielle og kan vise ulik syntaktisk oppførsel (se Struik og van Kemenade 2018). I tillegg ble eksempler med uklar informasjonsstatus ekskludert (som i tilfeller av generisk referanse).

Resultatene ligner på det som Thorbjörg Hróarsdóttir (2009) har beskrevet for gammelisländsk med hensyn til objekters informasjonsstatus. Dette indikerer at behandlingen av gammelisländsk og gammelnorsk under fellesbegrepet 'norrønt' er legitimt når det gjelder det syntaktiske nivået. Men resultatene viser kun til en generell observasjon og til at man finner de samme ordstillingsmønstrene på overflatestrukturen med en lignende fordeling. En mer detaljert komparativ analyse av materialet og tallene (fra studien presentert her og fra Thorbjörg Hróarsdóttirs studie) tyder derimot på at de to språkene divergerer også på et syntaktisk nivå allerede på den tidsperioden som er undersøkt her. Resultatene av denne sammenligningen er oppsummert i tabell 2. Tallene i denne tabellen er sett i relasjon til VO-stillinger med nye og kjente objekter (tallene for VO-stillingen er ikke oppgitt i denne tabellen). For gammelisländsk er det oppgitt to århundrer (hentet fra Thorbjörg Hróarsdóttir 2009: 264) for å i tillegg vise en interessant utvikling med tanke på hvordan informasjonsstatus korrelerer med ordstilling over tid.

	OV-ny	OV-kjent
Gammelnorsk 13. årh.	20,8 % (16/77)	46,4 % (77/166)
Gammelisländsk 14. årh.	26,9 % (75/279)	81,3 % (156/192)
Gammelisländsk 18. årh.	16,9 % (20/118)	47,6 % (39/82)

Tab. 2: Prosent for OV-stilling i gammelnorsk og gammelisländsk.

Det første som er påfallende, er en nedgang av OV-stilling i gammelisländsk over tid. Også for norsk kan vi stadfeste at OV-stillinger forsvinner som mulige overflatestrukturer på et eller annet tidspunkt.²¹ Men uten komplekse annoterte korpus for norsk for forskjellige århundrer er det vanskelig å diskutere dette i ytterligere detalj. Mer generelt viser tallene at påvirkningen av informasjonsstatusen er svekket allerede ganske tidlig i gammelnorsk sammenlignet med gammelisländsk. Sannsynligheten for at informasjonsstatus er en faktor for OV-stilling i gammelnorsk på 1200-tallet, er dobbelt så høy når objektet er kjent fra konteksten. For gammelisländsk er sannsynligheten tredoblet når objektet er kjent, og denne effekten er stabil over tid, selv om man finner en generell nedgang

21. Når man ser bort ifra topikalisering og 'stylistic fronting' og bare regner med den dominante ordstillingen som klassifiserer dette språket som et VO-språk. Se også 'Object Shift' i moderne isländsk (jf. Thorbjörg Hróarsdóttir 2009: 266, spes. 274ff.).

i OV-stillinger i islandsk. Dette betyr for det første at språksystemet for både islandsk og norsk reagerer sterkt på faktoren 'informasjonsstatus'. For det andre tyder tallene på at det allerede på et relativt tidlig tidspunkt finnes en syntaktisk forskjell mellom de to språkene. Når man da også vurderer språkhistorien til norsk, så kan den svakere effekten av informasjonsstatus på ordstilling forstås som en indikator og en del av en større syntaktisk utvikling, nemlig en økende fiksering av strikte VO-stillinger i dette språket.²² Som et resultat av dette er informasjonsstatusen til objekter mindre reflektert i ordstillingen. Resultatene for gammelnorsk i dette prosjektet er påfallende, men det er viktig å huske at KoNoKs-korpuset kun representerer én tekst fra gammelnorsk tid. Med hensyn til dette kan analysen av materialet i KoNoKs-korpuset bare tilby foreløpige resultater.²³ Likevel antyder en slik detaljert analyse av tallene fra ulike syntaktiske mønstre i språket at en divergerende syntaktisk utvikling i gammelislandsk og gammelnorsk skjedde allerede på 1200-tallet og dermed nokså tidlig i språkhistorien. Denne tilnærmingen til materialet med en sterk komparativ ansats bør oppmuntre til flere detaljerte analyser som tar for seg direkte sammenligning av gammelnorsk og gammelislandsk syntaks.

4.2 *Objektsposisjon: Informasjonsstatus og prosodisk vekt*

I studien av objektsposisjonen i gammelnorsk (se Tiemann 2022) som bare er kort skissert her, undersøkes også prosodisk vekt og dens samspill med informasjonsstatus. For gammelislandsk beskriver Thorbjörg Hróarsdóttir (2009) at prosodisk vekt er den avgjørende faktoren for ordning av setningselementer (jf. også Eiríkur Rögnvaldsson 1996: 70), som også kan overstyre informasjonsstrukturelle effekter.²⁴ Hun viser at fulle

22. Språkhistorisk spiller kاسوبortfall i norsk en rolle i at syntaksen blir mer restriktiv enn i islandsk. Men med et kاسوبystem som fremdeles er aktiv i gammelnorsk blir dette bare en tilleggsfaktor som gjennom tid kan være interessant å se på i samspill med informasjonsstruktur og prosodisk vekt.
23. Siden kun én tekst er representert og annotert her, kan man ikke unngå mulige sjangereffekter. Resultatene må således testes for bredere konklusjoner på grunnlag av analyse av flere gammelnorske tekster fra ulike tidsperioder og ulike sjangere. Det er kjent fra studier av andre eldre språk at tekster kan være nokså ulike med hensyn til ordstilling og informasjonsstruktur.
24. Andre faktorer som ikke er gjort rede for her er rollen av kasus (A-movement; se Chomsky 1981 og Thorbjörg Hróarsdóttir 2000: 115f.). Også Taylor og Pintzuk (2012 a, b) nevner en korrelasjon mellom kasus og ordstilling for gammelengelsk, der objekter i dativ og genitiv oftere står i postverbal posisjon.

(tunge) DP-er oftere står i VO-stilling enn pronomener og lette DP-er, som er foretrukket i en preverbal posisjon.²⁵ Pronomener viser en tverrspråklig tendens til å stå i en høyere posisjon i den syntaktiske strukturen enn fulle NP-er og DP-er. På bakgrunn av dette forventes det for gammelnorsk at tunge objekter står i VO-stilling og lette objekter i OV-stilling. Det forventes videre at prosodisk vekt fungerer som en blokkeringskonisjon for den informasjonsstrukturelle utløseren, noe som resulterer i at kjente, men tunge objekter blir stående i en VO-stilling.

Prosodisk vekt ble analysert ved hjelp av to separate steg i studien. I en første analyse ble objektene delt inn i tre kategorier: 1) pronomener og demonstrativer (lette objekter), 2) NP-er (nakne substantiver, mellomtunge objekter), og 3) fulle DP-er og PP-er (tunge objekter). I en alternativ analyse ble objektene delt inn i to grupper på basis av stavelsestall: lette objekter med 1–2 stavelser og tunge objekter med 3–6 stavelser. Stavelsestall er målt fra nukleus til nukleus. Objekter som klassifiseres som egen setningsfrase, følger hoved verbet obligatorisk og ble ekskludert fra de to vektanalysene. Tabell 3 og 4 viser resultatene av samspill mellom informasjonsstatus og prosodisk vekt i gammelnorsk (se Tiemann 2022 for gammelnorske eksempler).

	Pronomener og demonstrativer			NP-er			Fulle DP-er og PP-er		
	OV	VO	%OV	OV	VO	%OV	OV	VO	%OV
Kjent	45	18	71,43 %	14	14	50,00 %	18	57	24,00 %
Ny	0	0	0 %	11	28	28,21 %	5	33	13,16 %

Tab. 3: Samspill mellom informasjonsstatus og prosodisk vekt (objekttype)

	Lette objekter (1–2 stavelser)			Tunge objekter (3–6 stavelser)		
	OV	VO	%OV	OV	VO	%OV
Kjent	56	30	65,12 %	21	59	26,25 %
Ny	11	15	42,31 %	5	46	9,80 %

Tab. 4: Samspill mellom informasjonsstatus og prosodisk vekt (stavelsestall)

Resultatene viser at vekt er en signifikant faktor for ordstilling i gammelnorsk. Den samme tendensen kan leses ut av begge de to tabellene,

25. DP betyr ‘determiner phrase’ som har en determinativ som kjerne, ofte med en nominalfrase (NP) som komplement.

nemlig at lette elementer favoriserer en posisjon høyere i den syntaktiske strukturen og dermed oftere står preverbalt. I motsetning til dette står tyngre objekter oftere i VO-stilling. Disse resultatene er også i tråd med en tverrspråklig tendens som viser at lette elementer står foran tunge elementer (jf. Behaghel 1932; se også Hawkins 1994 som argumenterer for at all variasjon i ordstilling er et resultat av syntaktisk vekt). Det er også vanlig at kjente konstituenten i en diskurs er realisert som lette elementer (slik som pronomener), mens nye og fokuserte konstituenten er uttrykt med tunge elementer (slik som substantivfraser). Mer konkret viser resultatene at et kjent objekt som er lett ifølge dets vektanalyse (objekttype/stavelsestill), hovedsakelig står i OV-stilling. Men når et kjent objekt klassifiseres som relativt tungt, står dette hovedsakelig i postverbal stilling. Nye objekter (både tunge og lette) viser derimot en generell tendens for VO-stilling (unntatt pronomener og demonstrativer som vanligvis refererer til kjent informasjon, og derfor ikke forekommer som nye entiteter), se tab. 3. Vekt er altså en signifikant faktor som til en viss grad kan overstyre informasjonsstrukturelle forhold. Dette betyr at selv om OV er foretrukket med kjent informasjon, så forhindrer vektbetingingen en preverbal posisjon hvis objektet er tungt. I tillegg viser tallene at lette objekter i gammelnorsk har en tendens til å være fordelt på VO/OV på basis av informasjonsstatus. Disse følger altså ikke prinsippet om vektfordeling der alle lette objekter skulle stå i en preverbal posisjon. Og her skiller gammelnorsk seg igjen fra gammelisländsk. Thorbjörg Hróarsdóttir (2009: 265f.) skriver nemlig for gammelisländsk at

[L]ight objects with old information (unfocused) are always preverbal and heavy objects with new information (focused) are always postverbal. Light objects with new information are also usually preverbal, while heavy objects with old information are usually postverbal.

Lette objekter i gammelnorsk viser en informasjonsstrukturell effekt og kan stå både pre- og postverbalt, mens tunge objekter oftere dukker opp i postverbal stilling – uansett om de bærer kjent eller ny informasjon. En oppsummering av dette er gitt i tabell 5 og 6 (igjen: tallene er sett i relasjon til VO stilling; asterisken i tabell 5 viser hvor det er forskjell sammenliknet med gammelisländsk).

Objekt	Lett	Tung
Kjent	Preverbal	Postverbal
Ny	*Postverbal	Postverbal

Tab. 5: Oppsummering informasjonsstatus og prosodisk vekt.

%OV	Lett	Tung
Kjent	65,1 %	26,3 %
Ny	42,3 %	9,8 %

Tab. 6: Sannsynlighet for OV-stilling (prosentfordeling).

5 Oppsummering: Gammelt materiale og nye korpusanalyser

Denne artikkelen har hatt som mål å vise hvordan det gamle materialet som er arkivert i Språksamlingene, kan brukes i ny forskning. Artikkelen redegjør for omdanningen fra nedskrevet informasjon på papir til XML-koding og videre taggingsprosesser. På grunnlag av den grammatiske merkingen som ble gjort på ordbokssedlene, var det mulig å gjennomføre en omfattende annotasjon av språklige data innen rammene av et PhD-prosjekt. Digitalisering og tilgjengeliggjøring av slikt materiale bidrar til å utvikle nye korpus og til å analysere og dokumentere språket med nye metoder og tilnærminger.

Seddelarkivet representerer slik en nyttig og viktig ressurs der gjenbruk av gammelt arbeid gir en stor gevinst når det gjelder forskningstid og arbeidskraft. PhD-prosjektet om syntaks og informasjonsstruktur i gammelnorsk viser at det er behov for mer detaljerte studier om norsk språk og språklig utvikling, og fremhever også behovet for komparative næranalyser med gammelislandsk. Her må det selvfølgelig trekkes frem at kun én tekst fra den gammelnorske perioden ble analysert i detalj i løpet av dette prosjektet, og at resultatene for denne teksten ikke nødvendigvis representerer den språklige tilstanden for alt materiale fra gammelnorsk tid. Men de skisserte resultatene tyder på noen interessante forskjeller som kun er tilgjengelige gjennom detaljerte korpusanalyser – noe vi følgelig trenger flere av. Denne artikkelen vil altså oppmuntre forskere til videre undersøkelser av gammelnorsk syntaks, separat fra, men sammenlignet med gammelislandsk.

Referanser

- Bech, Kristin og Kristine, Eide (red.). 2014. *Information Structure and Syntactic Change in Germanic and Romance Languages*. Amsterdam: John Benjamins.
- Behaghel, Otto. 1932. *Deutsche Syntax*. Vol. 4. Heidelberg: Winter.
- Chafe, Wallace. 1976. Givenness, Contrastiveness, Definiteness, Subjects, Topics and Point of View. *Subject and Topic*, red. Charles Li, 27–55. New York: Academic Press.
- Chomsky, Noam. 1981. *Lectures on Government and Binding*. Dordrecht: Foris.
- Christoffersen, Marit. 1994. Sentences with Initial Adverbials in the Law of Magnus Lagabøter with Particular Emphasis on the Position of the Subject. I: *Language Change and Language Structure. Older Germanic Languages in a Comparative Perspective*, red. Toril Swan, Endre Mørck og Olaf Jansen, 75–90. Berlin/New York: Mouton de Gruyter. [Trends in Linguistics. Studies and Monographs 73].
- . 2002. Har leddsetninga i norrønt ‘forfelt’? Et studium av underordna setninger i Magnus Lagabøters landslov. *Norsk Lingvistisk Tidsskrift* (20), 153–176.
- Donhauser, Karin. 2007. Zur Informationsstrukturellen Annotation Sprachhistorischer Texte. *Sprache und Datenverarbeitung* (31), 39–45.
- Eirikur Rögnvaldsson. 1994/1995. Breytileg orðaröð í sagnlið. *Íslenskt mál og allmenn málfræði* (16–17), 27–66.
- . 1996. Word Order Variation in the VP in Old Icelandic. *Working Papers in Scandinavian Syntax* (58), 55–86.
- Elenbaas, Marion og Ans van Kemenade. 2014. Verb particles and OV/VO in the history of English. *Studia Linguistica* (68.1), 140–167.
- Faarlund, Jan Terje. 2004. *The Syntax of Old Norse. With a Survey of the Inflectional Morphology and a Complete Bibliography*. Oxford: Oxford University Press.
- Flom, George. 1921–1923. The Language of the Konungs skuggsja (*Speculum regale*). *University of Illinois Studies in Language and Literature* (7.3), 1–152; (8.4), 155–323.
- Franco, Irene. 2010. Issues in the syntax of Scandinavian embedded clauses. *Working Papers in Scandinavian Syntax* (86), 137–177.
- Halldor Sigurðsson. 1985. *Diachronic Icelandic Syntax: Some Practical and Theoretical Problems*. Ms., University of Stockholm.

- . 1988. From OV to VO: Evidence from Old Icelandic. *Working Papers in Scandinavian Syntax* (34), 1–41.
- . 1994. *Um frásagnarumröðun og grundvallarorðaröð í forníslensku*. Reykjavík: Bókaverzlun Ársæls Árnasonar.
- Haugen, Jens. 2000. *Old Norse Word Order and Information Structure*. Doktoravhandling. Norwegian University of Science and Technology.
- Haugen, Odd Einar og Åslaug Ommundsen. 2010. *Vår eldste bok: Skrift, miljø og biletbruk i den norske homilieboka*. Oslo: Novus.
- Hawkins, John. 1994. *A performance theory of order and constituency*. Cambridge: Cambridge University Press.
- Hinterhölzl, Roland og Ans van Kemenade. 2012. The interaction between syntax, information structure and prosody in word order change. *The Oxford Handbook of the History of English*, red. Terttu Nevalainen og Elizabeth Traugott, 803–821. Oxford: Oxford University Press.
- Holmberg, Anders og Christer Platzack. 2008. The Scandinavian languages. I: *The Oxford Handbook of Comparative Syntax*, red. Guglielmo Cinque og Richard S. Kayne, 420–458. Oxford: Oxford University Press.
- Holm-Olsen, Ludvig. 1952. *Håndskriftene av Konungs skuggsjá: En undersøkelse av deres tekstkritiske verdi*. København: Munksgaard. [Bibliotheca Arnamagnæana 13].
- Kruijff, Geert-Jan og Denys Duchier. 2003. Information structure in topological dependency grammar. *Proceedings of the tenth conference on European chapter of the Association for Computational Linguistics* (1), red. Ann Copestake og Jan Hajic, 219–226. New York: ACM Press.
- Kruijff-Korbayová, Ivana og Mark Steedman. 2003. Discourse and Information Structure. *Journal of Logic, Language and Information* (12.3), 249–259.
- Lambrecht, Knut. 1994. *Information Structure and Sentence Form*. Cambridge: Cambridge University Press.
- Levinson, Stephen. 1983. *Pragmatics*. Cambridge: Cambridge University Press.
- Matić, Dejan. 2015. Information Structure in Linguistics. *International Encyclopedia of the Social & Behavioral Sciences* (12), 2. ed., 95–99.
- Mørck, Endre. 2016. Syntaks. I: *Norsk språkhistorie I: Mønster*, red. Helge Sandøy og Agnete Nesse, 317–445. Oslo: Novus.

- Petrova, Svetlana, Michael Solf, Julia Ritz, Christian Chiarchos og Armir Zeldes. 2009. Building and using a richly annotated interlinear diachronic corpus: The case of Old High German Tatian. *Traitement Automatique des Langues* (50.2), 47–71.
- Pfaff, Alexander. 2019. Reunited after 1000 Years. The Development of Definite Articles in Icelandic. *Nordic Journal of Linguistics* (42.2), 165–207.
- Roberts, Craige. 2012. Information Structure: Afterword to Roberts (1996). I: *Semantics and Pragmatics* 2012 (hefte 5), 1–19.
- Schnall, Jens Eike og Rudolf Simek. 2000. *Speculum regale. Der altnorwegische Königsspiegel (Konungs skuggsjá) in der europäischen Tradition*. Wien: Fassbaender. [Studia Medievalia Septentrionalia 5].
- Struik, Tara og Ans van Kemenade. 2018. On the givenness of OV word order: A (re)examination of OV/VO variation in Old English. *English Language & Linguistics* (24.1), 1–22.
- Taylor, Ann og Susan Pintzuk. 2012 a. Rethinking the OV/VO Alternation in Old English: The Effect of Complexity, Grammatical Weight, and Information Status. *The Oxford Handbook of the History of English*, red. Terttu Nevalainen og Elizabeth Traugott, 199–213. Oxford: Oxford University Press.
- . 2012 b. Verb Order, Object Position and Information Status in Old English. *York Papers in Linguistics Series* (2.12a), 29–52.
- Thorbjörg Hróarsdóttir. 1996. The Decline of OV Word Order in the Icelandic VP: A Diachronic Study. *Working Papers in Scandinavian Syntax* (57), 92–141.
- . 2000. *Word Order Change in Icelandic: From OV to VO*. Amsterdam: John Benjamins.
- . 2009. Information Structure and OV order. *Information Structure: Theoretical, Typological and Experimental Perspectives*, red. Malte Zimmermann og Caroline Féry, 258–281. Oxford: Oxford University Press.
- Thorsteinn Indriðason. 1987. *Skýrsla um orðaröð í sagnlið*. Ms., University of Iceland, Reykjavík.
- Tiemann, Juliane. 2018. Nå kan en ny del av det gammelnorske materialet publiseres i digital form. *På Høyden* (jan. 2018). <https://pahoyden.khrono.no/debatt-gamalnorsk-lle/na-kan-en-ny-del-av-det-gammelnorske-materialet-publiseres-i-digital-form/392014>.

- . 2022. The Object Position in Old Norwegian: An Interplay between Syntax, Prosody, and Information Structure. I: *Language Change at the Interface. Intrasentential and intersentential phenomena*, red. Nicholas Catasso, Marco Coniglio og Chiara De Bastiani, 61–94. Amsterdam: John Benjamins.
- . Upublisert manuskript. Syntactic Variation and Information Structure in Old Norwegian: An Investigation of *Konungs skuggsjá* in AM 243 bǫ fol. Upubl. doktoravhandling. University of Bergen.
- Walkden, George. 2014. Object position and Heavy NP Shift in Old Saxon and beyond. *Information structure and syntactic change in Germanic and Romance languages*, 313–340. Amsterdam: John Benjamins.
- Wallenberg, Joel C., Anton Karl Ingason, Einar Freyr Sigurðsson og Eiríkur Rögnvaldsson. 2011. *Icelandic Pared Historical Corpus (IcePaHC)*. Version 0.9. http://www.linguist.is/icelandic_treebank.

Abstract

This article describes the life cycle of data based on the physical material in the Language Collections (Språksamlingene) at the University of Bergen. The material was compiled on index cards over Old Norse/Old Norwegian which was later transferred to digital solutions. It is outlined how this material has been reused to build up and supplement a new digital corpus (KoNoKs = Korpus over den norske *Konungs skuggsjá*) designed to answer new research questions. As an example for a new analysis based on this older research material, this article reports on some results of a corpus study on syntactic variation in Old Norwegian.

Juliane Tiemann
Universitetet i Bergen
Institutt for lingvistiske, litterære og estetiske studier
HF-bygget, Sydneplassen 7
NO-5007 Bergen
juliane.tiemann@uib.no