



Språk i arkiva:
Ny forskning om eldre talemål frå LIA-prosjektet

Til minne om
Janne Bondi Johannessen 1960–2020
Gjert Kristoffersen 1949–2021

Kristin Hagen, Gjert Kristoffersen, Øystein A. Vangsnes
og Tor A. Åfarli (red.)

Språk i arkiva

Ny forskning om eldre talemål
frå LIA-prosjektet



NOVUS FORLAG
Oslo 2021

Denne boka er utgitt med støtte av:

Institutt for lingvistiske og nordiske studier, Universitetet i Oslo
Institutt for lingvistiske, litterære og estetiske studium, Universitetet
i Bergen
Institutt for språk og kultur, UiT Noregs arktiske universitet
Institutt for språk og litteratur, NTNU

© Novus AS 2021.

Lay-out og omslag: Geir Røsset

Omslagsillustrasjon: Edvard Bondi Knowles

Font: Times New Roman 11/13

eISBN 978-82-8390-082-8

Det må ikke kopieres fra denne boka i strid med åndsverksloven eller avtaler om kopiering inngått med KOPINOR, interesseorgan for rettighetshavere til åndsverk.

Innhald

<i>Kristin Hagen, Øystein A. Vangsnes og Tor A. Åfarli</i> Om LIA-prosjektet og artiklane i denne boka	7
---	---

Norsk talemål

<i>Ivar Berg og Leiv Inge Aa</i> Palatalisering av velarer: historie, variasjon og normering	19
--	----

<i>Gjert Kristoffersen og Randi Neteland</i> Ka farsken? Realisering av /r/ som sibilant foran /k/ i nordnorsk	41
--	----

<i>Signe Laake og Lilja Øvrelid</i> Forskjeller mellom talemål og skriftspråk: Hva kan trebanker fortelle oss?	65
--	----

<i>Bente Selback og Terje Svardal</i> LIA-korpuset som ressurs i revisjonen av tre ordbøker	87
--	----

<i>Karine Stjernholm og Ingunn Indrebø Ims</i> Språkending i Vika: En komparativ analyse av data fra to talespråkskorpus	109
--	-----

<i>Eirik Tengesdal og Björn Lundquist</i> Trykkplasseringa i latinske lånord og partikkelverb i tre norske dialektområde	129
--	-----

<i>Øystein A. Vangsnes og Marit Westergaard</i> Ka LIA fortæll? Eit gjensyn med kv-spørsmål i norske dialektar	155
--	-----

Samisk

<i>Lene Antonsen</i> «Lei niogtredve go byggiimet» Om unormerte lån fra norsk i samisk talespråk	179
<i>Kristine Bentzen</i> VO – OV-variasjon i nordsamisk: Hva kan LIA Sápmi fortelle oss?	201

Amerikanorsk

<i>Ragnhild Eik og Anu Laanemets</i> Å være eller å bli – det er spørsmålet. Ei sammenligning av verbene VÆRE og BLI i amerikanorsk og norgenorsk	217
<i>Ida Larsson och Kari Kinn</i> Argumentplacering i norskt arvspråk i Amerika	241
<i>Brita Ramsevik Riksem, Terje Lohndal og Tor A. Åfarli</i> Adjektivkongruens i amerikanorsk	273
<i>Åshild Søfteland og Arnstein Hjelde</i> «Ein må tenke seg om, ser du veit du, kva ord ein skal bruke» Pragmatiske partiklar i amerikanorsk	295



Om LIA-prosjektet og artiklane i denne boka

Kristin Hagen, Øystein A. Vangsnes og Tor A. Åfarli

Denne boka inneheld 13 fagartiklar som med eitt unntak har sine opphav i føredrag frå LIA-prosjektet sitt sluttseminar i Trondheim 26.–27. november 2019. Arbeida tar for seg vidt forskjellige emne på grunnlag av gamle talemålsopptak som no er gjort tilgjengelege gjennom LIA-prosjektet og tre nye og brukarvenlege elektroniske korpus: 1) LIA norsk – korpus av eldre dialektopptak, 2) LIA Sápmi – Sámegeiela hállangiellakorpus, 3) CANS – amerikanordisk tale-språkskorpus v.3.1.

Om LIA-prosjektet

LIA står for Language Infrastructure made Accessible, eit stort infrastrukturprosjekt finansiert gjennom forskingsinfrastrukturprogrammet til Noregs forskingsråd (NFR) i perioden 2014–2019. Det var eit nasjonalt samarbeidsprosjekt mellom Universitetet i Oslo, Universitetet i Bergen, UiT Noregs arktiske universitet, Noregs teknisk-naturvitskaplege universitet, Norsk Ordbok 2014 og Nasjonalbiblioteket. Initiativtakar til prosjektet var professor Janne Bondi Johannessen, og det har heile tida vore forankra ved Tekstlaboratoriet, UiO. Det vart leidd av Janne saman med ei styringsgruppe med representantar for dei ulike institusjonane i prosjektet.

Det overordna formålet med LIA-prosjektet var å redde eldre talemålsopptak med norsk og samisk språk som levde eit bortgøymt

og i stor grad bortgløymt tilvære i arkiv og instituttbibliotek ved dei «gamle» universiteta våre i Oslo, Bergen, Trondheim og Tromsø, og somme også på private hender. Vidare var det eit mål å transkribere og annotere ei stor mengd av dei viktigaste av desse opptaka og legge dei inn i databasar slik at dei kunne revitaliserast som verdifullt forskingsmateriale. Opptaka vart digitaliserte ved Nasjonalbiblioteket i Mo i Rana, og kopiar er langtidslagra der. Prosjekttilsette ved dei fire universiteta har gått gjennom opptaka, katalogisert dei, utstyrt dei med metadata, og ein stor del av materialet har også vorte transkribert.

Dei transkriberte opptaka har vorte samla i dei tre korpusa nemnde ovanfor, som alle er elektronisk tilgjengelege og søkbare med det brukarvenlege grensesnittet Glossa, utvikla ved Tekstlaboratoriet. Meir om korpusa følgjer nedanfor. Gjennom prosjektet vart også den såkalla B-serien frå Talemålsundersøkelsen i Oslo (TAUS) transkribert og lagt inn i ein ny versjon av det eksisterande TAUS-korpuset (v.3). I tillegg har det vorte oppretta eit søkbart fildepot som inneheld alt digitalisert materiale på norsk, også det som ikkje er transkribert. Her er det mogleg å lytte til lydfilene og å be om å få laste ned filer som ikkje har sensitivt innhald.

For språkteknologiske og andre formål er det laga ein nedlastbar versjon av LIA norsk som omfattar 553 transkripsjonar i tekstformat med tilhøyrande lydfile. Det er også mogleg å laste ned LIA-trebanken, ein trebank som inneheld 5 250 talemålssegment med 55 410 ord og skiljeteikn frå LIA norsk, og som er annotert med morfologisk og syntaktisk informasjon. Annoteringa er gjort maskinelt, men alt er gjennomgått og korrigert manuelt.

Med dette arbeidet har LIA-prosjektet tilgjengeleggjort datamateriale frå fleire tiår med dialektologisk innsamlingsarbeid utført av fleire generasjonar målføregranskarar ved universiteta våre og av andre. Som nemnt ovanfor var det overordna formålet med prosjektet å redde dei eldre talemålsopptaka. Dei fleste av desse opptaka var nemleg lagra på magnetband, og det var grunn til å frykte at dei etter kvart ville gå tapt. LIA-prosjektet var derfor ikkje berre eit lingvistisk prosjekt, men også eit viktig kulturhistorisk redningsprosjekt. Det er fortienesta til Janne at ho såg begge desse aspekta og hadde hand-

lekraft til både å initiere LIA-prosjektet og leie det gjennom heile prosjektperioden og fram til ei vellukka avslutning.

LIA-prosjektet har også vore eit vellukka samarbeidsprosjekt med mange spennande diskusjonar og prosjektmøte ved dei fire universiteta, i Mo i Rana og til sist ein prosjekttur til Finnmark avslutta med seminar på Sámi allaskuvla (Samisk høgskole) i Kautokeino i september 2018. Prosjektet har dessutan hatt til saman fleire titals prosjekttilsette ved dei fire samarbeidsuniversiteta, der mange av dei har vore deltidstilsette studentar som har transkribert, korrekturlese og annotert. Vi er svært takksame for kvar og ein som har bidratt til prosjektet.

Om arbeidet med denne boka

Denne boka er den tredje artikkelsamlinga på Novus forlag med arbeid basert på norske og nordiske talespråkskorpus utvikla ved Tekstlaboratoriet. I praksis, om ikkje formelt, dannar dei tre bøkene ein serie. Den første boka med tittelen *Språk i Oslo: Ny forskning omkring talespråk* redigert av Janne Bondi Johannessen og Kristin Hagen kom i 2008 og inneheld 19 artiklar frå eit seminar i 2006 knytt til det då nyleg ferdigstilte NoTa-korpuset med talemålsopptak frå Oslo samla inn på byrjinga av 2000-talet. I 2014 kom *Språk i Norge og nabolanda: Ny forskning om talespråk*, også redigert av Janne Bondi Johannessen og Kristin Hagen, med 13 artiklar frå eit seminar året før knytt til ferdigstillinga av Nordisk dialektkorpus, eit av resultatata frå det store nordiske samarbeidsprosjektet *Nordisk dialekt-syntaks*.

Det var meininga at også denne boka, *Språk i arkiva: Ny forskning om eldre talemål* frå LIA-prosjektet med bidrag frå avslutningsseminaret for LIA-prosjektet i november 2019, skulle redigerast av Janne Bondi Johannessen og Kristin Hagen. Men våren 2020 vart Janne akutt verre av kreftsjukdomen ho hadde kjempa mot i lengre tid, og 15. juni 2020 døydde ho. I hennar stad gjekk då Tor A. Åfarli, Gjert Kristoffersen og Øystein A. Vangsnes inn som redaktørar. Alle tre

hadde vore med i styringsgruppa for prosjektet som representantar for høvesvis NTNU, UiB og UiT.

Våren 2021 ramma eit nytt dødsfall arbeidet med boka då Gjert Kristoffersen mista livet i ei trafikkulukke 29. mai. Gjert hadde fram til då lagt ned ein stor og uvurderleg innsats med boka. I tillegg til å ha hovudansvar for nokre av artiklane, hadde han også formatert alle bidraga i tråd med den malen som skulle nyttast.

I tillegg til at dei tre bøkene alle har sprunge ut av forskings- og infrastrukturprosjekt med Janne Bondi Johannessen som primus motor, har alle omslaga på bøkene vorte teikna av sonen hennar, Edvard Bondi Knowles. I 2008 var han ein teikneglad tenåring, i 2014 student i landskapsarkitektur ved NMBU og no i 2021 utøvande landskapsarkitekt og dessutan aktiv jazzmusikar. Vi er svært glade også for at dette knyter dei tre bøkene saman.

Om dei tre LIA-korpusa som er nytta i artiklane i boka

Det er empirisk materiale frå dei tre LIA-korpusa nedanfor som er nytta i dei 13 fagartiklane som du finn i denne boka. Samla sett demonstrerer dei kor rikt materialet frå LIA-prosjektet er som utgangspunkt og empirisk grunnlag for teoretisk analyse og språkforskinga generelt. Det er likevel viktig å vere klar over at LIA-korpusa ikkje er balanserte korpus med omsyn til utval av informantar. Sidan LIA-prosjektet har arbeidd med eldre talemålsopptak, har vi måtta bruke det som fanst av opptak, og det har ført til ubalanse, både med omsyn til variablar som kjønn, bustad og alder.

Prosjektet har lagt vinn på å finne opptak av best mogleg lyd kvalitet frå flest mogleg stader. Vi har også prioritert opptak med fri tale og ikkje opplesing av tekst eller oppramsingar av bøyingsparadigme eller stadnamn. Slik har korpusa også vorte gullgruver av kulturhistorisk verdi med tema som matlaging, tømmerdrift, draktskikkar osv.

LIA norsk – korpus av eldre dialektopptak

LIA norsk inneheld opptak og transkripsjonar frå fire universitet: NTNU, UiB, UiO og UiT. Korpuset inneheld også eit delmateriale frå Målførearkivet ved UiO som tidlegare var å finne i Nordisk dialektkorpus.

LIA norsk har 1374 informantar frå 222 kommunar og inneheld ca. 3,5 millionar ord. Korpuset er transkribert både talemålsnært og ortografisk til nynorsk og er hausten 2021 morfologisk tagga med ein nyutvikla talemålstaggar for nynorsk. Korpuset er søkbart både med omsyn til ortografisk form, talemålsnær form, lemma og morfologiske opplysningar. Resultata kan filtrerast gjennom metadata som stad, kommune, fylke, kjønn og opptaksår. Det eldste opptaket er frå 1937.

LIA Sápmi – Sámegiela hállangiellakorpus

LIA Sápmi er det første talespråkskorpuset med samiske dialektar. Korpuset har nesten 190 000 ord fordelt på 122 informantar frå 19 stader. Opptaka er frå tidsrommet 1960–1987, og det er opptak frå store delar av det nordsamiske området. Mange opptak stammar frå samlinga til Nils Jernsletten, som var professor i samisk ved UiT.

Opptaka er ortografisk transkriberte, og korpuset har fått automatisk lingvistisk analyse av Giellatekno ved UiT. Det er dermed mogleg å søke på ordklasse og lemma, i tillegg til ordform. Søka kan sjølvstøtt filtrerast gjennom metadata.

CANS – amerikanordisk talespråkskorpus v. 3.1

CANS - amerikanordisk talespråkskorpus v. 3.1 er det einaste LIA-korpuset som også har nyare opptak. Informantane er eldre norsk-amerikanarar som lærte norsk heime, dei fleste på 1920- og 30-talet. Opptaka blei gjort i perioden 2010 – 2016 av Janne Bondi Johannesen og andre. Korpuset har også ein liten del amerikavensk med

opptak frå Ida Larsson og andre (2011 – 2014). CANS inneheld også gamle opptak frå Didrik Arup Seip og Ernst W. Selmer (1931), Einar Haugen (1942) og Arnstein Hjelde (1987, 1990, 1992). Arbeidet med korpuset starta før LIA-prosjektet fekk klarsignal, men blei mange gonger større på grunn av LIA-midlane.

Korpuset inneheld 268 informantar frå USA og Canada, 22 som snakkar svensk og 246 som snakkar norsk, i alt nesten 775 000 ord. Det er både intervju og samtalar, og opptaka er transkriberte både talemålsnært og ortografisk til bokmål. Transkripsjonane er tagga automatisk med NoTa-taggaren, og korpuset er dermed søkbart på same måten som LIA norsk.

Kort om artiklane i boka

Av dei 13 artiklane i boka tek sju utgangspunkt i LIA norsk, to i LIA Sápmi og fire i CANS, og dei dannar såleis tre tematiske hovudbolkar. Nedanfor gir vi eit kort oversyn over artiklane, inndelte alfabetisk etter namnet til førsteforfattaren innanfor dei tre hovudbolkanane. Vi byrjar med dei sju artiklane som er grupperte under temaet norsk talemål.

Den første har tittelen «Palatalisering av velarer: historie, variasjon og normering» og er skriven av **Ivar Berg og Leiv Inge Aa**. Forfattarane undersøker variasjonen mellom velar plosiv /g, k/ og palatal frikativ /j, ç/ i innlyd i norske dialektar (jf. ord som *skog(j)en*, *tenk(j)e*). Dei peikar på at variasjonen lett blir utjamna der palatalane inngår i morfofonologisk veksling med ein velar, og at desse palatalane er på retur i dialektane. Dei undersøker om LIA-korpuset og Nordisk dialektkorpus kan kaste lys over denne problematikken, og dei ser også på om svekkinga av palatalane i talemålet kan ha samanheng med j-bortfall i nynorsk skrift.

I artikkelen «Ka farsken? Realisering av /r/ som sibilant foran /k/ i nordnorsk» tar **Gjert Kristoffersen og Randi Neteland** for seg bruken av sibilanten [ʂ] som allofon av /r/ i rk-sekvensar i nordnorske dialektar (jf. uttalen av ordet *farsken* i tittelen) med utgangspunkt i LIA-korpuset. Dei finn at sibilantuttale er utbreidd i heile Nord-

Noreg, men at det er store individuelle ulikskapar når det gjeld bruksfrekvens, også blant dei som har mange belegg på variabelen.

Signe Laake og Lilja Øvrelid presenterer i artikkelen «Forskjeller mellom talemål og skriftspråk: Hva kan trebanker fortelle oss?» ein samanliknande studie av norsk talemål og skriftspråk. Dei tar for seg fleire syntaktiske ulikskapar mellom skrift og tale som har blitt undersøkte i tidlegare studiar, som t.d. sideordning vs. underordning og utelating av syntaktiske ledd. I sin studie brukar dei den nyleg utvikla LIA-trebanken og den skriftspråklege Norsk Dependenstrebank og viser korleis desse gjer det mogleg å utføre presise syntaktiske søk som legg til rette for at dei ulike syntaktiske fenomenane kan undersøkjast på ein effektiv måte både i tale og skrift.

I artikkelen «LIA-korpuset som ressurs i revisjonen av tre ordbøker» ser **Bente Selback og Terje Svardal** på kva ressurs eit talemålskorpus som LIA norsk kan vere i arbeidet med revisjonen av tre ordbøker: standardordbøkene Bokmålsordboka og Nynorskordboka og dokumentasjonsordboka Norsk Ordbok. I ein ordboksrevisjon arbeider ein både med å finne kva nye ord (lemma) som bør takast inn i ordbøkene, og ein vurderer dei eksisterande ordboksartiklane: Er eit ord så sentralt at det bør vere med vidare? Er definisjonane i tråd med gjengs språkbruk? Kva bruksdøme fungerer best? I eit slikt revisjonsarbeid er det viktig å ha gode kjelder, og i artikkelen vurderer forfattarane kva ein kan og eventuelt ikkje kan bruke LIA-korpuset til i dette arbeidet.

I artikkelen «Språkending i Vika. En komparativ analyse av data fra to talespråkskorpus» presenterer **Karine Stjernholm og Ingunn Indrebø Ims** ein pilotstudie der dei brukar materiale frå LIA norsk og Nordisk dialektkorpus til å teste ein hypotese om ulik språkutvikling i det vikkværske dialektområdet i perioden 1960–2009. Nærmare bestemt undersøker dei overgang frå *ær*-suffiks til *er*-suffiks i kate-goriane hankjønn, fleirtal, ubestemt, samt nomen agentis og presens av svake verb. Dei finn indikasjonar på at det er ulik bruk av desse suffiksa på tvers av dialektområdet, og dei diskuterer korleis sosiolingvistisk teori kan bida med nyttige perspektiv for å forstå språkutviklinga i området.

I artikkelen «Trykkplasseringa i latinske lånord og partikkelverb i tre norske dialektområde» ser **Eirik Tengesdal og Björn Lundquist** på sambandet mellom trykkplassering i latinske lånord (som *butikk*) og partikkelverb (som *gå ut*), der trykkplasseringa på fyrstestavinga eller sistestavinga varierer. Dei reknar med at partikkelverb viser ein del variasjon som skriv seg frå språkinterne strukturelle faktorar, og at desse faktorane ikkje påverkar latinske ord i like stor grad, og vidare reknar dei med at ei gransking av ikkje-syntaktisk variasjon i latinske ord kan bidra til ei betre forståing av trykkvariasjonen i partikkelverb. Med det som utgangspunkt undersøker dei talarar frå Trøndelag, Hedmark og Finnmark og gir ei grundig skildring og analyse av den relevante trykkvariasjonen. Det empiriske materialet hentar dei frå LIA-korpuset og Nordisk dialektkorpus, og dei finn at LIA og NDK er veileigna for prosodisk gransking.

Til slutt i bolken om norsk talemål presenterer **Øystein A. Vangnes og Marit Westergaard** i artikkelen «Ka LIA fortæll? Eit gjensyn med kv-spørsmål i norske dialektar» ei undersøking som replikerer ei tidlegare korpusbasert undersøking av kv-spørsmål utan V2 i norske dialektar. Den første undersøkinga tok for seg temaet med utgangspunkt i Nordisk dialektkorpus, som har opptak frå perioden 2005–2010 frå 111 stader i landet. Den siste undersøkinga, som er gjennomført for denne artikkelen, tar for seg temaet med utgangspunkt i det norske LIA-korpuset, som inneheld eldre opptak frå 222 ulike stader, men der opptaka er spreidde over ein lengre tidsperiode både med tanke på opptaksår og fødeåret til informantane. Forfattarane finn at data frå LIA-korpuset langt på veg stadfestar det generelle inntrykket frå NDK når det gjeld kv-spørsmål utan V2 i norske dialektar, men ein mogleg skilnad mellom dei to korpusa når det gjeld det vestnorske tilfanget, blir drøfta.

Bolk nummer to er dei to artiklane som tar for seg samisk. Den første, av **Lene Antonsen**, har tittelen «?Lei niogtredve go byggiimet.? Om unormerte lån fra norsk i samisk talespråk». Denne artikkelen undersøker materiale annotert som framandspråkleg og sitatlån i nordsamisk i talespråkkorpuset LIA Sápmi, og vurderer dei mest frekvente orda i høve til ordbøker og tekstkorpus. Analysen viser at

ein del av dette materialet er etablerte lånord i munnleg språk, og mange av dei finst også i ordbøker, men orda er fråverande i skriftleg språk. Vidare viser analysen at størstedelen av materialet er spontanlån der det innlånte leksikalske elementet blir tilpassa samisk morfologi og syntaks.

Det andre bidraget i samisk-bolken har tittelen «VO – OV-variasjon i nordsamisk: Hva kan LIA Sápmi fortelle oss?» og er forfatta av **Kristine Bentzen**. Artikkelen vart opphavleg publisert i *Bauta: Janne Bondi Johannessen in memoriam (Oslo Studies in Language 11: 2, 2020)*, men er tatt med også i denne boka sidan undersøkinga er basert på materiale frå LIA Sápmi. I undersøkinga ser Bentzen på vekslinga mellom VO- og OV-ordstilling i nordsamisk talemål, og ho finn at VO er det mest frekvente mønsteret overordna sett. Men i materialet finst det også mange tilfelle av OV-leddstilling i samband med bruk av hjelpeverb, og i tillegg førekjem det nokre tilfelle der objektet står framfor både hjelpeverb og hovudverb (utan at objektet er tematisert). Forfattaren føreslår ein detaljert formell syntaktisk analyse av dei ulike leddstillingsmønstera.

Tredje og siste bolken omfattar fire bidrag som på ulike vis tar for seg problemstillingar som kan grupperast under temaet amerikansk. Den første er forfatta av **Ragnhild Eik og Anu Laanemets** og har tittelen «Å være eller å bli – det er spørsmålet. Ei sammenligning av verbene VÆRE og BLI i amerikansk og noregorsk». Artikkelen baserer seg på data frå korpusa CANS (Corpus of American Nordic Speech) og LIA norsk, og undersøker verba VÆRE og BLI brukte saman med predikata *født*, *konfirmert* og *gift* i norsk talt i Amerika samanlikna med norsk talt i Noreg. Undersøkinga tar utgangspunkt i tidlegare studiar av amerikansk som viser at VÆRE i nokon grad har tatt over for BLI saman med desse predikata, truleg på grunn av kontakten med engelsk. Forfattarane finn noko overraskande at eit tilsvarende resultat ikkje gjeld for amerikansk, og dei diskuterer korleis dette resultatet skal forståast.

Neste artikkel i denne bolken er «Argumentplacering i norskt arvspråk i Amerika» av **Ida Larsson og Kari Kinn**. Denne artikkelen undersøker distribusjonen til subjekt, objekt og partiklar i eldre og yngre amerikansk, basert på CANS-korpuset, og i talespråket i

Noreg omkring tida for masseutvandringa til Amerika, basert på LIA-korpuset. Eit hovudfunn er at leddstillinga varierer i begge data-materiala, og at det delvis er meir variasjon i distribusjonen i amerikanorsk enn i moderne norsk talemål. Vidare finn forfattarane blant anna at variasjonen held seg relativt stabil over fleire generasjonar i amerikanorsk, men likevel slik at subjektskifte og objektskifte blir mindre frekvent over tid, noko dei føreslår å forklare ved å vise til økonomiprinsipp som favoriserer 'uskifta' posisjonar, men dei drøftar også andre faktorar som kan spele inn.

I artikkelen «Adjektivkongruens i amerikanorsk» ser **Brita Ramsevik Riksem, Terje Lohndal og Tor A. Åfarli** på attributive engelske adjektiv i amerikanorske nominalfrasar som har norsk struktur og norsk substantiv, og dei undersøker om slike engelske adjektiv syner norsk bøyning, basert på data frå det nyare CANS-korpuset. Resultata viser at i dei få tilfella i dette datamaterialet der eit engelsk attributiv adjektiv står saman med eit norsk substantiv, er det ikkje norsk bøyning, noko som er uventa, sidan engelske substantiv og verb som blir brukte i ein norsk kontekst som hovudregel får norsk bøyning. Artikkelen diskuterer kvifor adjektiva oppfører seg annleis enn substantiv og verb og føreslår ein formell syntaktisk analyse av dette.

Den siste artikkelen i den amerikanorske bolken er «'Ein må tenke seg om, ser du veit du, kva ord ein skal bruke.' Pragmatiske partiklar i amerikanorsk» av **Åshild Søfteland og Arnstein Hjelde**. Denne artikkelen tar for seg bruken av dei pragmatiske partiklane *ser du, veit du og trur eg* i CANS. Desse blir samanlikna med bruk av engelsk *you see, you know, I think* og norsk *du veit, du ser, eg trur* i det same datamaterialet. Artikkelen greier ut om frekvens, syntaktisk posisjon og overordna pragmatisk analyse, og forfattarane drøftar også metodologiske problemstillingar. Hovudfunnet er at talarane av amerikanorsk brukar desse pragmatiske partiklane som støtte for informasjonsflyten i samtalesituasjonen, og at dei kan utnytte kombinasjonar av norske og engelske bruksmønster som ein ressurs i språkbruken.

LIA-lenker:

CANS - amerikanordisk talespråkskorpus v.3.1:

<https://tekstlab.uio.no/glossa2/cans3>

Fildepot for LIA:

<https://lia.tekstlab.sigma2.no/>

Glossa:

<https://www.hf.uio.no/iln/om/organisasjon/tekstlab/tjenester/glossa/>

LIA norsk - korpus av eldre dialektopptak:

https://tekstlab.uio.no/glossa2/lia_norsk

LIA-prosjektet:

<https://tekstlab.uio.no/LIA/>

LIA-trebanken:

<https://tekstlab.uio.no/LIA/trebank.html>

LIA Sápmi - Sámegeiela hállangiellakorpus:

<https://tekstlab.uio.no/glossa2/saami>

TAUS v.3:

<https://tekstlab.uio.no/glossa2/taus3>



Palatalisering av velarer

Historie, variasjon og normering

Ivar Berg og Leiv Inge Aa

Artikkelen tar for seg variasjonen mellom velar plosiv /g, k/ og palatal frikativ /j, ç/ eller affrikat i innlyd i norske dialekter. Særlig der palatalene inngår i morfofonologisk veksling med velar, blir variasjonen lett utjamna, og generelt er disse palatalene på retur i dialektene. Vi undersøker hva LIA-korpuset og Nordisk dialektkorpus kan fortelle om utbreiinga av palataler og overgang til velar. Korpusfunna gjør det tydelig at palatalene mange steder har stått svakt lenge. I siste del undersøker vi om palatalavviklinga i målføra også vises igjen med *j*-avvikling i nynorsk skrift. I rettskrivinga blei enkelte *j*-lause former tillatt fra 1938 og noen flere fra 1959. Vi viser at det først og fremst er avishusa som holder *j*-formene i hevd, mens det er mer variasjon og jamt færre *j*-former hos dem som ikke er underlagt husnormer.

Stikkord: Palatalisering av velar, språkhistorie, talemål, nynorsk skriftspråkutvikling

1 Innleiing¹

Vi undersøker den geografiske og historiske variasjonen mellom velar plosiv /g, k/ og palatal frikativ /j, ç/ i innlyd i ord som *skog(j)en*, *ligg(j)e*, *tenk(j)e* og *ikk(j)e*. (Vi forenkler til /j, ç/ her, men ofte er det tale om affrikatene [ʃj] og [çç]; se nedafør.) Historisk dreier det seg om målmerket som i norsk dialektologi er kalt *palatalisering av velar*, noe som skjedde før /j/ og fremre vokaler. I mange tilfelle førte dette til morfofonologisk veksling mellom velar plosiv og palatal frikativ i bøyingsparadigma til enkeltord (se avsnitt 2.1).

Et tradisjonelt poeng i målførelitteraturen er at den morfofonologiske vekslinga ofte blir utjamna og heilt allment er på vikende front i dialektene (t.d. Sandøy 1987, 233 f.; Skjekkeland 1997, 98–100; Mæhlum og Røyneland 2012, 94). Derfor er det interessant å se hvor godt den holder seg – og hvor godt den kan dokumenteres – i de historiske talemålsopptaka i LIA-korpuset. Er det systematisk geografisk variasjon i bevaring eller utjamning av slik variasjon, eller er palatalene bare et generelt konservativt drag i dialekter ulike steder?

Tradisjonelt er palatal uttale markert med ⟨j⟩ i nynorsk skriftspråk, men ikke i alle posisjoner. Morfofonologisk veksling i substantiv blir ikke markert i skrift, og heller ikke palataliseringa i intetkjønnsord som *ønske* og *stykke*, jamvel om den ofte er leksikalisert i former som /ynʃe/ ~ /ønʃe/ og /styçe/. Et ortografisk poeng er at 2012-reforma for nynorsk gjennomførte konsekvent valgfridom i slike tilfelle, med unntak av *ikkje* og *mykje* (og noen ord der konservativt valg av rotvokal medfører obligatorisk ⟨j⟩). Avslutningsvis spør vi åssen dette blir gjennomført i nynorsk skriftspråk, og om avviklinga av palataliseringa i talemåla går hand i hand med tilsvarende *j*-lause skriftformer.

Nedafør gjør vi veldig kort greie for den historiske bakgrunnen (bolk 2) før vi ser på variasjonen i norske talemål slik den er dokumentert i LIA-korpuset, til dels jamført med Nordisk dialektkorpus

¹ Vi takker to anonyme fagfeller for gode forslag til forbedringer, og særlig redaktør Gjert Kristoffersen, som gav gode råd i heile prosessen med denne artikkelen, men som dessverre gikk bort rett før ferdigstillinga.

(NDK; Johannessen mfl. 2009) (bolk 3). LIA-korpuset inneholder eldre talemålsopptak fra ulike samlinger; brorparten er fra 1970- og 1980-åra, men en god del av materialet vi har sett nærmere på, er tatt opp i 1950-åra. Opptaka til den norske delen av NDK blei gjort i perioden 2006–2012. Sjøl om gamle talere i NDK overlapper med unge i LIA, gjengir korpusa grovt sett norsk talemål med omtrent en generasjons mellomrom, slik at de kan vise endring over tid. De to korpusa er tilgjengelige i samme søkegrensesnitt (Glossa) og er dermed lette å sammenlikne. Her gjør vi overordna, landsomfattende korpus-søk før vi tar stikkprøver på utvalgte frekvente infinitivsformer i Vestland fylke. Målet med denne delen er å vise åssen LIA-korpuset og NDK kan brukes til å studere en fonologisk variabel på makronivå, samstundes som de er innganger til opptak og opplysninger på individnivå. Skriftspråkundersøkingene gjør vi i Nynorsk-korpuset og Normkorpuset (bolk 4). Totalt gir denne vesle undersøkinga døme på åssen en kan utnytte moderne korpusdata til å kaste nytt lys over gamle dialektologiske spørsmål.

2 Historisk opphav

2.1 Palatalisering som historisk prosess

Palatal frikativ /ç/ har flere opphav. Norrønt /tj/ og /θj/ falt sammen og har stort sett blitt til /ç/ (f.eks. *tjørn* ‘tjern’ og *þjófr* ‘tjuv’), men her konsentrerer vi oss om norrøn /k/ framfor fremre vokal (t.d. *kinn* > /çin/) og norrøn /kj/ (t.d. *kjot* > /çøt/). Historisk kan en trulig regne med ei stegvis utvikling som skissert i figur 1, der det skifter om dialektene har kommet til steg 2 eller 3 (jf. Kristoffersen og Torp 2016, 189–191). Mellom /k/ og fremre vokaler kom det inn en halvokal, og nyutvikla /kj/ følger opprinnelig /kj/ videre. I framlyd er dette allment i norsk (ord som *kinn* og *gift*), variasjonen gjelder i innlyd.

<i>Tidlig norrønt</i>	<i>Steg 1</i>	<i>Steg 2</i>	<i>Steg 3</i>
<i>k_fremre vokal</i>	<i>kj</i>	/ç̥ç/	/ç/
<i>kj</i>	<i>kj</i>	/ç̥ç/	/ç/

Figur 1: Stegvis palatalisering av velar /k/

Utviklinga av /g/ var tilsvarende, med det unntaket at den frikative allofonen [ɣ] av /g/ trulig blei til /j/ uten å gå via affrikat (jf. Sandøy 1987, 182). Derfor kan det i dialektene være slik at lang /g:/ veksler med /j/ eller /j/ og kort /g/ med /j/ i ord som /ryg/ – /ryjjen/ ‘rygg – ryggen’ og /sku:g/ – /sku:jen/ ‘skog – skogen’.

Det er sannsynlig at /g/ og /k/ hadde palatale allofoner før fremre vokaler i norrønt, og allerede på 1200-tallet blir dette så smått markert i norske tekster, oftest etter /g/. Det er flere døme som pres. *giængr* (norr. *gengr*) og *skiæmta* (norr. *skemta*) i *Legendariske Olavs-saga* fra ca. 1250 (Hægstad 1922, XXX f.; jf. Seip 1955, 182 f.). Slike skrivemåter blir vanligere etter hvert, men aldri heilt dominerende i den norske skrivetradisjonen (Pettersen 1975, 291–293). Vi kan regne med at palataliseringa er eldre enn framskytinga av kort /o/ og /a/ i austlandsmåla, for det heter alltid /kø:ɾ/ ‘kol’ (norr. *kol*) og /kæɾv/ ‘kalv’ (norr. *kalfǫr*) i motsetning til /çø:ɾe/ ‘kjøle, gjøre kald’ (norr. *kóla*) og /çæ:ɾ/ ‘kjær’ (norr. *káerr*). Palataliseringa er også eldre enn endringer i bøyingssuffikset i bestemt form entall av sterke hunnkjønnsord, f.eks. norr. *bókin* > /bu:ça/ med palatalisering /k/ > /ç/ før bøyingssuffikset *-in* blei til /a/.

Palataliseringa er ikke lenger en synkron fonologisk prosess framfor trykkting vokal, men leksikalsk bestemt; Kristoffersen og Torp (2016, 191) meiner at «palataliseringa ikkje har vore noko levande prinsipp på mange hundre år». Nye lånord som *kimono* og *gitar* viser dette. Likevel er vi vant til at ⟨ki⟩ i skrift skal uttales [çi] el. [ççi], og ⟨gi⟩ [ji]. Dette kan overføres til nye leksem; det relativt nye lånordet *kino* har f.eks. alltid palatal. Vi trur derfor at døme på slik «nyere palatalisering» er basert på skriftbildet jamført med kjente ord, ei analogisk overføring til nye leksem, og nettopp leksikalsk bestemt.

Det var mange former som hadde velar + *j* eller velar før fremre vokal, omgivelsene der velarene blei palatalisert. I tillegg til noen enkeltord, særlig de frekvente adverb *ekki* (> *ikkje*) og *mikit* (> *mykje*), gjaldt det følgende grammatiske kategorier:

- Mange verb hadde i norrønt /j/ i infinitiv (f.eks. *byggja*, *leggja*), noen også i andre former (f.eks. *eggja*).

- Mange svake substantiv hadde like ens /j/ før endingsvokalen (f.eks. *sleggja*, *bylgja*).
- Nøytrale *ija*-stammer hadde stammeutlyd på /i/ og fikk palatalisering i alle bøyingsformene (f.eks. *merki*, *stykki*).
- Den bestemte artikkelen *-inn/-in/-it* osv. begynte i de fleste formene på /i/, så substantiv med stammeutlyd på velar plosiv fikk palatalisering i bestemt form (f.eks. *bekinn* (akk.), *bókin*, *líkit*).
- Norrøne maskuline *i*-stammer fikk dessuten palatalisering i heile flertall (f.eks. *bekkir*, *leggir*).

I de fleste av disse tilfella var det ikke grunnlag for palatalisering i alle bøyingsformene (de nøytrale *ija*-stammene er et viktig unntak), slik at denne lydendringa førte til omfattende morfofonologisk veksling mellom velar plosiv og palatal frikativ. Denne morfologiske vekslinga er ennå levende for mange norsktalende og kan til og med overføres til nye lånord, men kan heller ikke regnes som en synkron fonologisk regel i norsk (jf. ovafor om stillinga trykktingt). Sia palatalen i dag må regnes som leksikalsk bestemt, følger det også at overgang til velar plosiv skjer form for form (jf. Sandøy 1987, 233).

2.2 Morfofonologisk variasjon

Palataliseringa førte til omfattende morfofonologisk variasjon mellom ubestemt og bestemt form. Dette kan gi paradigmer som i (1) i deler av Trøndelag og Nordland:²

(1)	a.	/bæk/ – /beççen/	‘bekk – bekken’
	b.	/røg/ – /ryjjen/	‘rygg – ryggen’

Fordi palataliseringa blokkerte senking av korte, fremre vokaler, får nordafjelske mål også variasjon i rotvokalen, som vi ser i (1). Det viser samstundes at senkinga er yngre enn palataliseringa, og sammen med det som var nevnt ovafor om andre endringer som er sei-

² Dette og følgende døme viser variasjonen, og liknende system fins flere steder, *mutatis mutandis*.

nere enn palataliseringa, tyder det på at den må ha vært gjennomført i seinmellomalderen.

Også i verba vil den palatale frikativten inngå i morfofonologisk veksling med velar plosiv, f.eks. i bøyingsmønster som i (2), i Nordfjord:

(2)	a.	<i>liggje</i>	<i>ligg</i>	<i>låg</i>	<i>lege</i>
		/liʝje/	/lige/	/lo:g/	/leje/
	b.	<i>tenkje</i>	<i>tenkjer</i>	<i>tenkte</i>	<i>tenkt</i>
		/teŋçe/	/teŋçe/	/teŋkte/	/teŋkt/

Der det er morfofonologisk veksling, er det alltid mulig med nye analogiske former, altså morfologisk utjamning. Mange har regna utjamning innafor et paradigme som en eigen type analogisk endring, men det kan òg forklaras som tradisjonell proporsjonal analogi bygd på bøyinga til substantiv i samme bøyingsklasse som *ikke* har stammeutlyd på velar (jf. Hill 2007 om ulike typer morfologisk endring). Utjamning av ei bøyning som /bu:k/ 'bok' – /bu:çe/ i Nord-Gudbrandsdalen (med *-e* i bf.sg. av sterke hunkjønnsord) kan til dømes settes opp i en proporsjonal analogi som vist i (3):³

(3)	<i>sol</i> – <i>sole</i>	
	<i>bok</i> – X	X = <i>bok</i> e (Sel, Nord-Gudbrandsdalen)

Tilsvarende likninger kan lett settes opp for endringer i andre bøyingsformer. Det er faktisk bare adverb *ikk(j)e* og *my(kj)e* og de nøytrale *ija*-stammene som ikke viser morfofonologisk veksling. Det er altså gode indrespråklige grunner til at denne vekslinga kunne utjammes, og i et tidlig arbeid pekte Thorson (1973 [1949], 343) på at det er «analogiverknaden som driv den palatale uttalen attende». Uten å gjøre poenget eksplisitt observerer Mæhlum og Røyneland (2012, 94) at *ikkje* og *mykje* er de orda som «har halde aller best på den pa-

³ *Bok*e er oppgitt av Per Esben Myren-Svelstad (personlig kommunikasjon). Legg merke til at selværene har beholdt den tradisjonelle endingsvokalen, men utjamna den morfofonologiske vekslinga i stammen. Det er altså ikke snakk om direkte overgang til *boka*.

latale uttalen», og det er ord som ikke inngår i et bøyingsmønster med former med /k/. Samstundes spiller nok også påvirkning fra talemål uten palatalisering inn, særlig i nyere tid, og ved de formene som *ikke* varierer, er slik ytre påvirkning den mest naturlige forklaringa på overgang fra /ç/ til /k/ i f.eks. *ikkje* > *ikke*.

3 Utbreiing

Allerede Larsen (1897, 28) kommenterte at det er mindre av slik morfofonologisk veksling i austnorsk enn i vestnorsk, sjøl om vi trulig kan regne med at palataliseringa opphavlig gjaldt allment. I nyere tid har *ija*-stammer som *stykke* tradisjonelt hatt palatal i all norsk, mens palatalisering i bestemt form entall av sterke hankjønnsord mangler på Austlandet (jf. Sandøy 1987, 183 f.). Søkemulighetene i LIA og NDK har gjort det mulig å prøve å skaffe noen store tall for å gi et overordna bilde. Deretter ser vi nøyer på noen former som peker fram mot diskusjonen av tilhøvet til språknormering i neste bolc.

3.1 Overordna tendenser

Generelle søk i korpuset på ord som i rettskrivinga har *-ke-* i inneller utlyd, men i den talemålsnære transkripsjonen er notert med <kj/tj/sj> (de naturlige måtene å notere palatal konsonant på etter det norske bokstaveringsssystemet, bekrefta av stikkprøver), fanger opp en god del relevante former som ikke blir skrevet med <j> i den normerte transkripsjonen, f.eks. infinitiv og bestemt form entall.⁴ Søket gir treff over heile landet, se figur 2a. Dette inkluderer søraust, f.eks. /styçe/ i Eidsberg (Østfold) og /løyçi/ 'Løken' i Fet (Akershus). (Noen sammensatte ord blir med i søket, og treffet i Aremark heilt i søraust er «spisesji» 'spiseskei'. Ellers var det få irrelevante treff.) Hoff (1946, 217–222) nevner en god del spor etter palataliseringa i Indre Østfold – tenk på gards- og bynavnet *Askim* av norrønt *Askheimr* – men også at /styçe/ blir avløst av /styke/ (op.cit., 218).

⁴ Søketrykk: [word=".+ke." %c & phon=".+(tj|kj|sj)+."].

Samme søkestreng i NDK (med unntak av at ⟨sj⟩ i talemålsnær skrivemåte blei ekskludert fordi den gav mange irrelevante treff) gir framleis mange treff (figur 2b), men de er likevel mer avgrensa på to viktige måter: Treffa er i større grad geografisk avgrensa til vestnorsk og midlandsmål, og de er leksikalsk avgrensa: De fire nørdeste treffa er alle substantivet *stykke*, mens de aller fleste treffa fra austlandsk er verbet *tenke*.



Figur 2a (LIA t.v.) og 2b (NDK t.h.): notert palatal som svarer til ⟨ke⟩ i skrift.

Vi har prøvd å gjøre et tilsvarende breitt søk for å fange opp ord som har morfofonologisk veksling, ved å søke etter substantiv med stammeutlyd på /g/ eller /k/ i bestemt form entall. Det er totalt 17 964 slike former i LIA. Av disse er 14,3 % (2574) notert med ⟨j⟩ før ei bøyingsending i talemålsnær transkripsjon.⁵ Tilsvarende tall for NDK

⁵ Søkeuttrykk: [lemma=".*(g|k)" %c & ((pos="subst" & num="eint" & defn="bu")) & phon=".*(a|e|i|en)"]; totalen er samme søk uten "phon". Dette inkluderer *ikke* svake substantiv, for da blir det vanskelig å skille ut de som tradisjonelt ikke har hatt palatal i noe talemål, f.eks. svake hunkjønnsord som *uke/veke*.



Figur 3: Palatal i adverbet *mye/mykje* (LIA).

er 6027 former totalt og 452 notert med ⟨j⟩, dvs. 7,5 %.⁶ Det kan være noen feil her, men som en overordna indikasjon kan det være bra nok, og forskjellen mellom korpusa er tydelig.

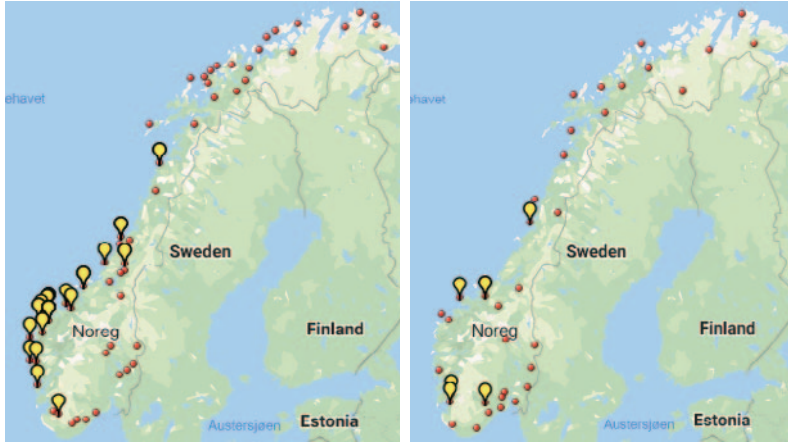
I verbet *tenke* kan vi vente palatal i infinitiv og presens (og stundom analogisk i imperativ, men det er ingen døme på det i LIA). I LIA er det totalt 1561 forekomster + 6 ført under lemmaet *tenkje*.⁷ Av disse er 344 + 3 transkribert med ⟨tj|kj⟩, dvs. 22,2 % palataler. I NDK er det 303 belegg der 86 viser palatal, dvs. 28,4 %. I dette leksemet viser altså NDK større andel palatal konsonant enn LIA, noe som viser at palatalene kan holde seg godt i frekvente leksem.

Det ubøyde adverbet *mye/mykje* gir mange treff med palatal også i Nord-Norge (figur 3). Her er palatal frikativ markert med gult (treffet i Østfold (Råde) er feiltranskribert), mens de andre formene i hovedsak er /my:/ og /my:e/. Generelt har nordnorsk oftest spor etter palataliseringa i leksema *ikke*, *mye* og *stykke* (Bull 1990, 163).

Skal en jamføre morfofonologisk veksling i spesifikke leksem i de to korpusa, må en finne noen som er frekvente, ellers blir talla så små at en eventuell forskjell gjerne kan være tilfeldig. Leksemet

⁶ Søkeuttrykk: [lemma="+(g|k)" %c & ((pos="noun" & num="sg" & defn="def")) & phon="*j(a|e|i|en)"].

⁷ Søkeuttrykk: [lemma="tenke" %c & ((pos="verb" & temp="pres|inf")) & phon="*(tj|kj).*"].



Figur 4a (LIA t.v.) og 4b (NDK t.h.): Palatalisering (gult) i bf.sg. av *fisk*.

stykke gir totalt 1001 treff i LIA-korpuset, 35,6 % av disse med palatal. Tilsvarende tall i NDK er 460 treff med 32,2 % palatal. Det er altså liten forskjell. Leksemet *merke* gir 795 treff i LIA-korpuset og 484 i NDK, med 18,6 % palatal i LIA og 13,0 % i NDK. Forskjellen mellom korpusera er heller ikke her overveldende, men forskjellen mellom leksema er derimot tydelig.

Det frekvente leksemet *fisk* har tradisjonelt hatt palatalisering i bestemt form entall, men ikke ellers i paradigmet. Søk på bestemt form entall i korpusera gir 285 treff i LIA med 33,0 % palatal (figur 4a), mens NDK gir 66 treff med 13,6 % palatal (figur 4b). I dette ordet med morfofonologisk veksling er forskjellen mellom korpusera altså tydelig.

3.2 Noen stikkprøver fra Vestland fylke og tilhøvet til nynorsknorma

Vi har tatt noen regionale stikkprøver i LIA og NDK for å sammenlikne med de landsomfattende undersøkningene ovafor og skriftundersøkningene i bolk 4. Der ser vi på infinitivformene til sju «viktige» verb (se tabell 4) som blei råka av *j*-avviklinga i Læreboknormalen fra 1959 (s. 58). For å se om talemålskorpusera avslører ei avvikling av palatal uttale i kjerneområda for nynorskten, viser tabell 1–3 ned-

<i>Sted/Korpus</i>	<i>LIA</i>	<i>NDK (ca. 2010)⁸</i>
<i>Førde</i>	1978: UK+, UM-	
<i>Selje</i>	1958: GK+	
<i>Austevoll</i>	1981: UM+	
<i>Bømlo</i>	1978: GM2+, UM2±	GM+, UM-, UK-
<i>Kvinnherad</i>	1951: GM2+	UM-, GM?
<i>Stord</i>	1980: UM-, UK+	
<i>Eidfjord</i>		UM-, UK-
<i>Fusa</i>		UM?, UK-
<i>Lindås</i>		GK+
<i>Voss</i>		GK+

Tabell 1: *Bygg(j)a/-e* i gamle Hordaland og Sogn og Fjordane. Årstalla angir opptaksår, UK = ung kvinne, GM = gammel mann, GM2 = to gamle menn osv. Plusstegn = palatal uttale, minustegn = ikke-palatal uttale, ± = begge deler. Spørsmålstegn = uavklart (til dømes utydelig tale).

<i>Sted/korpus</i>	<i>LIA</i>	<i>NDK (ca. 2010)</i>
<i>Aurland</i>	1955: GK+	
<i>Eid</i>	1978: GM+, GK-	
<i>Eikefjord</i>	1956: GM?	
<i>Førde</i>	1978: UK-	
<i>Hyllestad</i>	1957: GK-	GK-
<i>Jostedal</i>	1959: GM2+	
<i>Kinn</i>	1957: GM-, GK-	
<i>Selje</i>	1958: GK2+	
<i>Kalvåg</i>		GK+
<i>Stryn</i>		UM-

Tabell 2: *Legg(j)a/-e* i Sogn og Fjordane. (Samme koder som i tabell 1.)

slag i LIA og NDK for de to første av disse verba; tabell 1 tar for seg *bygg(j)a/-e*, tabell 2–3 *legg(j)a/-e*. Den overordna tendensen er at palataliseringa lever godt hos eldre talere i begge korpusa, og at avviklinga hovedsakelig er hos yngre, med et par unntak.

Tabellene viser, som nevnt ovafor, at avviklinga av palataliseringa ikke ser ut til å være dramatisk, men jamn. Det er over 50 år mellom

⁸ Innsamlinga og transkripsjonen av NDK-materialet hadde pr. 2014 «pågått i over åtte år» (Johannessen og Hagen 2014: 15), så vi setter 2010 som et medialt årstall.

<i>Sted/korpus</i>	<i>LIA</i>	<i>NDK (ca. 2010)</i>
<i>Bømlo</i>	1978: GM3+, UM2 +, GK+, UK+	
<i>Etnes</i>	1974: GM+	
<i>Kvam</i>	1969: GM+	
<i>Lindås</i>	1957: GK+, 1981: GK+	UM-
<i>Eidfjord</i>	1979: GM2+, GK+	GK+, UM-
<i>Fusa</i>		GM-
<i>Stord</i>	1980: UM-	

Tabell 3: *legg(j)a/-e* i Hordaland. (Samme koder som i tabell 1.)

de eldste døma i LIA og opptaka i NDK, og resultatene er klart i tabell 1 (*bygg(j)a/-e*), der det er atskillig flere NDK-treff enn i tabell 2 og 3. Sjøl om vi må ta forbehold om ujamnt datagrunnlag, synes utviklinga å være bra tydelig. Som det går fram av tabell 2, har LIA døme på eldre informanter i 1957 i Hyllestad og Kinn kommune som ikke palataliserer, samt ei eldre dame i Eid kommune i 1978. Det er også liknende funn for *ligg(j)a/-e* hos en eldre mann i Jølster i 1959, i tillegg til at det er noe intraindividuell variasjon med *tenk(j)a/-e* hos to informanter fra Førde og Arna rundt 1980. Hovedtendensen er likevel at de frekvente verba *har* palatalisering i LIA-korpuset, særlig i Hordaland. Det er stort sett unge i begge korpusa, litt flere i NDK, som ikke har palatal uttale.

3.3 Oppsummering av talemålssituasjonen

Korpuset har gjort det mulig å skaffe noen overordna tall, samstundes som vi har gjort noen stikkprøver på enkelte leksem. Disse stikkprøvene i korpuset tyder på følgende (definitive resultat kan ikke denne vesle undersøkingsgjøring gjøre krav på):

- Palatal uttale holder seg best i vestnorsk og midlandsmål.
- Palatal uttale holder seg best der det ikke er morfofonologisk veksling, som vist i *my(kj)e* og *stykke* i motsetning til bestemte former som *fisken*.
- Palatal uttale er leksikalsk bestemt og kan derfor skifte mellom ulike leksem, uavhengig av grammatisk kategori.

- d) På makronivå er det tydelig forskjell mellom korpusa, med mest palatalisering i LIA.
- e) Både i LIA og NDK har eldre talere oftest palatal uttale, men det er også informanter i LIA som ikke har det; ingen unge talere i NDK har gjennomført palatalisering.

4 Nyere skriftspråkutvikling og normering

4.1 J-avviklinga i nynorsknorma

Landsmålet tok fra starten inn noen skrivemåter med ⟨gj⟩ og ⟨kj⟩ som svarer til palatale varianter i talemålet. Én grunn til at skrivemåte med ⟨j⟩ i mange tilfelle blei unngått i landsmålet, var at palataliseringa blei regna som automatisk. Derfor kunne Aasmund O. Vinje skrive «ikki» og meine at det skulle uttales med /ç/.

Sjøl ikke i samnorsktida kom dette inn i bokmål. Bare to leksem i bokmål har en avvikende skrivemåte, nemlig *bikkje* og *krykkje* (fuglen), og de framstår nærmest som fremmedord innafor dette systemet med den særnorske forma si (Flydal 1989, 378). Disse orda er henta inn fra norske talemål og fins ikke i dansk (anna enn som norske lånord).

De nøytrale *ija*-stammene har aldri blitt systematisk normert med ⟨j⟩ i nynorsk, trass i uttalen og den sporadiske forekomsten av slike skrivemåter. At ⟨j⟩ normalt ikke blei skrevet i norrønt, kan ha spilt inn. Aasen (1873) har «Ynskje (el. Ynske)», men «Stykke (el. Stykkje)» og «Merke (el. Merkje)», og denne normeringa blir videreført av Skard (1901), uten parentesformene hos Aasen.

I 1938 begynte avviklinga av *j*-skrivemåte i andre kategorier (Kirke- og undervisningsdepartementet 1938, 13f). Da blei det obligatorisk å droppe ⟨j⟩ i verbalsubstantiv på *-ing*, f.eks. *tenking* og *bygging*, og svake hunkjønnsord og verb på *-ggje* fikk sideformer uten ⟨j⟩, f.eks. *brygge* (verb og substantiv). Dette blei ført vidare i læreboknormalen i 1959, som i stor grad åpna for å stryke ⟨j⟩ i sambanda *-kkj-*, *-ngj-*, *-nkj-*, *-rkj-*, *-skj-* og vokal + *-kj-*. Detaljnivået var slik at utgreininga regna opp «[d]ei viktigaste orda det gjeld» (Kirke-

og undervisningsdepartementet 1959, 58), inkludert unntak som *bikkje*, *ikkje* og *krykkje*.

I rettskrivingsreforma i 2012 fikk verb og svake hunkjønnsord med stamme på *-g* eller *-k* valgfri ⟨j⟩ som prinsipp. Eneste unntak i verb og substantiv var noen tradisjonsformer som *fylgje* og *syrgje*, mens de alternative formene med ⟨ø⟩ fikk valgfri ⟨j⟩, altså *følg(j)e* og *sørg(j)e* (Språkrådet 2012). Valgfridommen på dette punktet i 2012-reforma er hevda å være basert på usus i skriftlige kjelder; reforma er den første som er fundert i større undersøkinger i skriftspråskorpus, sjøl om undersøkingene hadde metodiske mangler (Holm og Munkvold 2013).

Etter 2012 skjer normeringa av nynorsk uten at tilnærming til bokmål er et mål. Nynorsk skal primært normeres på grunnlag av eksisterende nynorsk skrift, og henvisninga til talemål i de gjeldende retningslinjene for normering (2015–) er sekundær.⁹ Det fins likevel stadig et talemålsprinsipp for nynorsk (ulikt for bokmål), men som det går fram av det siste avsnittet gjengitt under, har det etter måten låg prioritet:

Når ein strammar inn norma på punkt der det er talemålsvariasjon, bør forma som har best talemålsgrunnlag, bli ståande. Dersom ein utvidar norma, bør den nye forma ha svært godt talemålsgrunnlag. Med talemålsgrunnlag er det tenkt både på frekvens, generell utbreiing og utbreiing i kjerneområda for nynorsk.

Bruksprinsippet går framfor talemålsprinsippet: Unormerte former som er lite brukte i skrift, skal ikkje takast inn i norma sjølv om dei har godt talemålsgrunnlag. (Språkrådet 2015, 13).

Sjølv om talemålsprinsippet kan sies å være noe marginalt, er det relevant å jamføre utviklinga av ulike talemålstrekk i kjerneområda

⁹ Det rår et visst hierarki blant normeringsprinsippa, der stabilitet i norma og sjølstendighet fra bokmål, samt nevnte ususprinsipp, vil trumfe talemålsprinsippet i de fleste sammenhenger. I skrivende stund er retningslinjene under revisjon i Språkrådet, der dette hierarkiet blir enda tydeligere framheva. Noen konsekvenser av prinsippa slik de ser ut i 2015-utgava, er òg kort diskutert i Hovdenak (2015).

nettopp for å vurdere om det er eller seinere kan bli utslagsgivende for nynorsknorma. Språkrådet definerer ikke «kjerneområda», men vi vil i denne sammenheng avgrense oss til Vestland fylke, der nynorsk òg er administrasjonsspråk.¹⁰ Her skal vi derfor se på åssen den skriftlige praksisen for *j*-skrivemåter eventuelt har endra seg, og jamføre med bolken ovafor om palatalisering av velarer i de nynorske kjerneområda fra 1950-tallet og utetter, altså fra og med opptaka i LIA. Der så vi at det store bildet ikke var drastisk endra fra LIA til NDK, avviklinga er saktegående.

4.2 Skrift (*usus*)

Gitt ei massiv bokmåleksponeering òg i de nynorske kjerneområda, i tillegg til svekkinga av palatalisering i dialektene, er det kanskje ventende at *j*-lause former vinner terreng i nynorsk, og at dette – gitt bruksprinsippet og trass i sjølstendighetsprinsippet i Språkrådets (2015) normeringsretningslinjer – kan komme til å påvirke norma ytterligere på sikt. Helset (2017, 109 f.) finner likevel ikke klare spor av dette i undersøkinga si, der det er ei klar overvekt av *j*-former. Han tar for seg verbet *bygg(j)a/-e* og substantivet *rekk(j)e* i tekster fra *Aviskorpuset* (*Bergens Tidende*, *Dag og Tid*, *Hallingdølen*, *Klassekampen*, *Nationen* og *Sogn Avis*; se referanselista) i perioden ca. 2000–2010. Materialet kan altså trulig jamføres direkte med det som lå til grunn for 2012-norma. Funna hans kan oppsummeres slik:

- *bygga/-e* 22 % (n=ca. 900) *byggja/-e* 78 % (n=ca. 3100)
- *rekke* 21 % (n=ca. 600) *rekkje* 79 % (n=ca. 2200)

Nær 80 % tilslag med *j*-former utgjør aleine kanskje et grunnlag for å forfekte det som eneform. Men talla aleine avslører ikke praksisen til den enkelte skribenten, og hva skribentene sjøl meiner om de brukte formene. Om verbformene *bygge* vs. *byggje* kommenterer Helset (2017, 140):

¹⁰ I LIA-korpuset kan en søke på de gamle fylka fra før 2018/2020, mens dialektkorpuset opererer med fem regioner, slik at en sjøl må avgrense på kommunenivå.

[H]eile 30 % av dei i alt 197 skribentane som i det heile brukte det aktuelle leksemet i korpuset, vekslar mellom å bruke dei to formene, medan 53,3 % held seg konsekvent til forma *byggje* og 16,7 % held seg konsekvent til forma *bygge* (Helset 2013), noko som samsvarar bra med det faktum at berre 25 % av informantane mine reknar *bygge* for å vere ei markert radikal form.

Utgangspunktet er altså at *j*-former dominerer, men at skribenter vakler og heller ikke nødvendigvis regner *j*-lause former for å vere radikale. I det følgende bidrar vi med noen stikkprøvebaserte undersøkinger, der vi ser vi på infinitivsformene av noen av de «viktige» verba nevnt i læreboknormalen (1959, 58). Dersom en søker på de fire mulige infinitivsformene for hvert verb i nynorske bøker i Nasjonalbibliotekets N-gram-tjeneste,¹¹ finner en at *e*-infinitiv med *j*-innskott dominerer i fire av de sju verba (*byggje*, *leggje*, *tenkje*, *søkje*) de siste åra, men at *j*-lause *e*-infinitivsformer er vanligere enn før – òg i tilfella der de framleis er i mindretall. *A*-infinitiv uten *j*-innskott er marginalt (og *a*-infinitiv generelt på retur).

Vi avgrensar undersøkinga nedafor til å inkludere to lokalaviser (*Sogn Avis* 2007–2010 og *Hallingdølen* 2002–2009) og to forfattere, Kjartan Fløgstad og Edvard Hoem, begge med korpustilfang fra 1970-tallet fram til kring tusenårsskiftet.¹² Vi har også søkt på de samme verba i nynorskdelen av Normkorpuset,¹³ der tekstene er skrevet av elever på 3., 4., 6., og 7. trinn i grunnskolen. I undersøkinga har vi inkludert sju infinitivsformer nevnt i Læreboknormalen, der *legg(j)a/-e*, *ligg(j)a/-e*, *bygg(j)a/-e* og *tenk(j)a/-e* er de mest frekvente. Prosentandel og antall treff med *j*-innskott er oppgitt i tabell 4.

Med *Sogn Avis* har vi inkludert ei avis sentralt i kjerneområdet til nynorsken, mens *Hallingdølen* i denne sammenheng er ei randso-neavis. Umiddelbart ser en at prosentene for former med ⟨j⟩ er klart høgest i *Sogn Avis*, der alle verba så nær som det minst frekvente

¹¹ https://www.nb.no/sp_tjenester/beta/ngram_1/.

¹² De omtalte kjeldene nedafor er alle spesifisert i *Nynorskorpuset* (se referanse-lista).

¹³ <https://www.hf.uio.no/iln/om/organisasjon/tekstlab/prosjekter/norm/>.

Verb/Kjeldre	Sogn Avis (2007–10)	Halling- dølen (2002– 09)	Kjartan Fløgstad (1972–96)	Edvard Hoem (1978–2004)	Nyn.tekster i Norm- korpuset
<i>Byggja/-e</i>	88,7 % (n=2159)	57,1 % (n=1319)	50,0 % (n=40)	0 % (n=47)	36,6 % (n=41)
<i>Leggja/-e</i>	87,0 % (n=2647)	57,0 % (n=1284)	52,9 % (n=121)	0 % (n=191)	33,3 % (n=39)
<i>Liggja/-e</i>	84,9 % (n=690)	41,9 % (n=458)	51,3 % (n=78)	0 % (n=53)	11,8 % (n=17)
<i>Stengja/-e</i>	81,1 % (n=180)	36,1 % (n=122)	20,0 % (n=5)	0 % (n=7)	100 % (n=1)
<i>Tenkja/-e</i>	89,5 % (n=1376)	67,8 % (n=559)	55,9 % (n=152)	0,6 % (n=155)	30,0 % (n=40)
<i>Steikja/-a</i>	27,8 % (n=18)	20 % (n=5)	0 % (n=2)	0 % (n=4)	0 % (n=4)
<i>Søkja/-e</i>	85,8 % (n=607)	74,2 % (n=299)	17,6 % (n=17)	4,0 % (n=50)	33,3 % (n=3)

Tabell 4. Prosent for *j*-innskott i infinitivsformer av (frekvente) verb. Totalt antall treff er oppgitt i parentes.

(*steik(j)a/-e*) ligger på 80–90 %. Talla fra *Sogn Avis* er gjenkjennbare fra Helset (2017) si undersøking, mens *j*-prosenten i *Hallingdølen* varierer fra 36,1 til 74,2 (om vi ser vekk fra *steik(j)a/-e*, som bare har fem treff totalt). *J*-prosenten er altså langt lågere i *Hallingdølen* enn i de samla talla fra Aviskorpuset hos Helset (2017).

De to forfatterne skriver mer radikalt enn avisene; Kjartan Fløgstads fordeling er sirka 50/50 blant de mest frekvente verba. Går en bøkene hans mer etter i sømmene, finner en at han har brukt mye *j* i storverka fra rundt 1980, mens *j*-innskotta mangler i større grad både i tidlige og seinere bøker. *Ligg(j)a* fins både med og uten *j*-innskott i *Dalen Portland* (1977) og *Fimbul* (1994). Dermed kan en gjette at forfatteren blander ubevisst, og at forlaget i ulik grad har grepet inn. Det kan være tilfellet hos Edvard Hoem òg, som bare er tre tilfeller (1 × *tenkje* og 2 × *søkje*) unna å ha (fått) renska unna alle *j*-formene.

At avisene skriver mer konservativt enn Fløgstad og Hoem, overrasker ikke, da vi kan gå ut fra at journalistene ikke velger form sjøl, men er underlagt ei husnorm (jf. Ims 2007). Det er likevel interessant å se at praksisen er såpass ulik hos de to avisene. Helset si undersøking inkluderer også regionale og nasjonale aviser, som altså ligger nærmere «kjerneområdenynorsken» enn *Hallingdølen*. Blant elevene i Normkorpuset ser en at andelen av *j*-skrivende ligger på ca. 1/3 ved

de tre mest frekvente verba. At den er såpass låg, overrasker kanskje heller ikke, gitt talemålsutviklinga (jf. at unge vestlendinger i NDK ikke har palataler) og at elevene trulig lesar mer bokmål enn nynorsk på fritida. Bokmålspåvirkninga i nynorske elevtekster er ellers kjent blant anna gjennom studien til Bjørhusdal og Juuhl (2017).

Ei undersøking som dette viser bare noen tendenser, men det synes klart nok at avishusa ivaretar *j*-formene i større grad enn enkeltpersoner som skriver, og avisene representerer sjølsagt òg mye tekst. De er mye leste tekster og påvirker derfor folks normoppfatninger eller internaliserte normer. Ims (2007) er riktignok skeptisk til at høgfrekvente former automatisk sammenfaller med internaliserte former og er de mest levedyktige: «Slike opptellingar er barometre på normativ språkbruk hos store tekstprodusentar, men de sier ikke noe om hvilke ordformer og varianter som er internalisert eller ikke hos den jevne språkbruker» (s. 43). Med dette perspektivet er ikke bruksprinsippet entydig, for de brukte formene er ikke nødvendigvis skribentenes valg. Hva som skal eller bør utgjøre tekst- og kunnskapsgrunnlaget i normeringsarbeid, blir diskutert av blant andre Helset (2018) og blir også noe justert i Språkrådets normeringsretningslinjer, som planlegges oppdatert i 2021. Å legge vekt på aviser i normgrunnlaget er i praksis å ivareta en viss stabilitet i norma (jf. stabilitetsprinsippet i normeringsretningslinjene). Men gitt et aukende digitalt tilfang og at digitale avistekster gjennomgår mindre korrektur og redigering enn papirtrykte tekster, er det ikke gitt at aviser i framtida vil representere den samme stabiliteten.

5 Konklusjon

Som nevnt innleiingvis er det ingen ny innsikt at målføretrekket «palatalisering av velar» er i ferd med å svekkes i dialektene. En kan spørre seg hvor langt tilbake en må gå for å finne den paradisiske urtilstanden her, for også i et korpus som LIA med tradisjonelle dialekttalere er det mange som *ikke* har palataler i de aktuelle orda og posisjonene. Vi kan slå fast at palatalene står svakt og har gjort det lenge. Utviklinga fortsetter, for det er ei tydelig utvikling fra LIA til

NDK (med forbehold om ujamnt datagrunnlag), om en ikke kan kalle endringa dramatisk.

I det vi kan kalle ei palatal avvikling spiller trulig indre og ytre årsaksfaktorer sammen. I ord med morfofonologisk veksling har det alltid vært grunnlag for analogisk utjamning, og palatalene har alltid vært fremmede for talt riksmål og til dels bymål generelt, som i dag øver stor påvirkning på omliggende dialekter. I den grad palatal uttale er bevart, er den veldig leksemavhengig. Vi vil derfor vente at palatalene holder seg lengst i høgfrekvente leksem. Palatalene er dessuten et allment konservativt trekk i norske talemål, det fins ingen ryddig isogloss for dette fenomenet

Normeringa (nynorsk) og usus har fulgt etter og foreskriver i mindre grad enn før (j). *J*-formene lever godt i avishusnormene og er såleis frekvente i tekstkorpus, men de er i mindre grad brukt av andre skrivere. Når palatalene er mindre brukt i talemåla enn før, ikke fins i bokmål og ikke er enerådende i nynorsk tekst heller, er det ikke anna å vente enn at *j*-formene har dårlige kår blant elever med nynorsk som hovedmål.

Referanser

- Aviskorpuset* (nynorsk). 2016. Utvikla av prosjektet *Norsk aviskorpus*, distribuert av Clarino Bergen Centre. <https://clarino.uib.no/korpuskel/corpus-list>.
- Bjørhusdal, Eli og Gudrun Kløve Juuhl. 2017. Bokmålsavvik frå nynorsknorma i sjetteklassetekstar. *Maal og Minne* 109(1), 93–121.
- Bull, Tove. 1990. Målet i Troms og Finnmark. I *Den store dialektboka*, red. av Ernst Håkon Jahr, 157–177. Oslo: Novus.
- Flydal, Leiv. 1989. La arbeidsfolks språk bli einaste riksmål i landet – så blir det norsk! I *Språket – struktur og samfunn. Uvalgte skrifter II*, red. av Arne Halvorsen og Geirr Wiggen, 355–379. Oslo: Novus.
- Helset, Stig Jarle. 2017. *Norm og røyndom – ein statistisk studie av operative normer i det nynorske skriftspråket*. Ph.d.-avhandling, Universitetet i Bergen.

- Helset, Stig Jarle. 2018. Om kunnskapsgrunnlaget for normeringa av det nynorske skriftspråket. *Norsk lingvistisk tidsskrift* 36(1), 69–93.
- Hill, Eugen. 2007. Proportionale Analogie, paradigmatischer Ausgleich und Formerweiterung. Ein Beitrag zur Typologie des morphologischen Wandels. *Diachronica* 24(1), 81–118.
- Hoff, Ingeborg. 1946. *Skjetvemålet*. Skrifter utgitt av Det Norske Videnskaps-Akademi i Oslo, II. Hist.-Filos. Klasse. 1946. No. 1. Oslo: Dybwad.
- Holm, Thorgeir og Sergej Alexander Munkvold. 2013. Den nye nynorskrettskrivingi. *Norsk Lingvistisk Tidsskrift* 31(2), 165–220.
- Hovdenak, Marit. 2015. Kor går norsk rettskriving? *Språknytt* 2/2015, 18–21.
- Hægstad, Marius. 1922. Um maalet i Ólafs saga hins helga. I Oscar Albert Johnsen (utg.), *Ólafs saga hins helga*, XXVIII–LVII. Utgjeve av Den Norske Historiske Kildeskriftkommission. Kristiania: Dybwad.
- Ims, Ingunn Indrebø. 2007. «Det er vi som bestemmer åssen folk snakker og skriver»: En undersøkelse av norske avisers språknormative prinsipp og praksis. Hovedoppgave, Universitetet i Oslo.
- Johannessen, Janne Bondi og Kristin Hagen. 2014. Om artiklene i denne boka og *Nordisk dialektkorpus*. I *Språk i Norge og nabolanda. Ny forskning om talespråk*, red. av Janne Bondi Johannessen og Kristin Hagen, 7–25. Oslo: Novus.
- Johannessen, Janne Bondi, Joel Priestley, Kristin Hagen, Tor Anders Åfarli og Øystein Alexander Vangsnes. 2009. The Nordic Dialect Corpus – an Advanced Research Tool. I *Proceedings of the 17th Nordic Conference if Computational Linguistics, NODALIDA 2009*, red. av Kristiina Jokinen og Eckhard Brick, 73–80. NEALT Proceedings Series Vol. 4. Tarttu: DSpace. <http://tekstlab.uio.no/nota/scandiasyn/>.
- Kirke- og undervisningsdepartementet. 1938. *Ny rettskrivning 1938*. Oslo: Olaf Nordlis Forlag.
- Kirke- og undervisningsdepartementet. 1959. *Ny læreboknormal*. Oslo: Olaf Nordlis Forlag.

- Kristoffersen, Gjert og Arne Torp. 2016. Fonologi. I *Mønster*, red. av Helge Sandøy, 101–211. Norsk språkhistorie, hovedred. Helge Sandøy og Agnete Nesse, band I. Oslo: Novus.
- Larsen, Amund B. 1897. *Oversigt over de norske bygdemål*. Kristiania: Aschehoug.
- LIA norsk – korpus av eldre dialektopptak*. <http://tekstlab.uio.no/LIA/korpus.html>.
- Mæhlum, Brit og Unn Røynealand. *Det norske dialektlandskapet*. Oslo: Cappelen Damm.
- Normkorpuset*. <http://www.hf.uio.no/iln/om/organisasjon/tekstlab/prosjekter/norm/>.
- Nynorskkorpuset*. http://no2014.uib.no/korpuset/r-webconc_avansertsok.html.
- Pettersen, Egil. 1975. *Språkbrytning i Vest-Norge 1450–1550*. Bergen: Universitetsforlaget.
- Sandøy, Helge. 1987. *Norsk dialektkunnskap*. 2. utg. Oslo: Novus.
- Seip, Didrik Arup. 1955. *Norsk språkhistorie til omkring 1370*. 2. utg. Oslo: Aschehoug.
- Skard, Matias (1901). *Landsmaals-ordlista med rettleiding um skrivemaaten*. Kristiania: Aschehoug.
- Skjekkeland, Martin. 1997. *Dei norske dialektane*. Kristiansand: Høyskoleforlaget.
- Språkrådet. 2012. *Ny nynorskrettskriving frå 2012*. <https://www.sprakradet.no/globalassets/vi-og-vart/publikasjoner/ny-nynorskrettskriving.pdf>.
- Språkrådet. 2015. *Retningslinjer for normering*. <https://www.sprakradet.no/Spraka-vare/Norsk/normering/retningslinjer-for-normering/>.
- Thorson, Per. 1973 [1949]. Om regressiv palatalisering av *g* og *k* i germansk, serleg norsk. I *Frå norsk målføregranskning*, red. av Olav T. Beito og Ingeborg Hoff, 329–355. Oslo: Universitetsforlaget.

English summary

The article discusses variation between the velar plosives /g, k/ and the palatal fricatives /j, ç/ or affricates in Norwegian dialects. Variation between a velar and a palatal consonant in the inflectional paradigm of a single word is often levelled, and generally the palatals are being replaced by velars in many dialects. We investigate how the LIA corpus and Nordic Dialect Corpus can be used to study this phenomenon. The findings confirm that the palatals have been vanishing for some time. Finally, we investigate whether the loss of palatal consonants is also reflected in *Nynorsk* writing by adopting non-*j*-forms, which were allowed in the standards from 1938 and 1959. We demonstrate that the *j*-forms are first and foremost preserved by newspapers, whereas independent writers and pupils in elementary school have a much lower rate of *j*-forms, following the situation in spoken Norwegian.

Ivar Berg

Institutt for språk og litteratur
Norges teknisk-naturvitenskapelige universitet NTNU
ivar.berg@ntnu.no

Leiv Inge Aa

Institutt for lærerutdanning
Norges teknisk-naturvitenskapelige universitet NTNU
leiv.inge.aa@ntnu.no



Ka farsken?

Realisering av /r/ som sibilant foran /k/ i nordnorsk

Gjert Kristoffersen† og Randi Neteland

Artikkelen redegjør for bruken av [ʂ] som allofon av /r/ i *rk*-sekvenser i Nord-Norge med utgangspunkt i LIA-korpuset. Vi analyserer geografisk utbredelse for hele landsdelen og i de enkelte kommunene, samt hvor frekvent varianten er blant dem som bruker den. Resultatene viser at sibilantuttale er utbredt i hele landsdelen, og vi finner ingen klare isoglosser. Det er imidlertid store individuelle forskjeller i hvor ofte varianten blir brukt, også blant dem som har mange belegg på variabelen. Vi har ikke kunnet påvise sammenhenger med bakgrunnsvariabler som kjønn, sosial bakgrunn og alder, men metadataene i LIA er for sparsomme til at vi kan konkludere med at slike sammenhenger ikke finnes. Vi har imidlertid funnet to eksempler på at varianten kan være i mindre bruk i lokalsamfunn som relativt nylig har gjennomgått språkskifte fra kvensk eller samisk til norsk.

Stikkord: Nordnorsk, fonologi, *r*-fonologi, sibilanter, etnolekt

1 Innledning¹

I nordnorsk er realisering av /r/ som sibilant foran /k/ og /p/ et godt kjent og mye omtalt fenomen, men det fins ingen systematisk oversikt over hvor vanlig sibilantuttalen er, over dens geografiske utbredelse og dens ev. samvariasjon med sosiale bakgrunnsvariabler. Den nordnorske delen av LIA-korpuset består av transkriberte eldre lyd-båndopptak av 380 informanter fra 41 ulike steder i Nord-Norge, og bør dermed i utgangspunktet være godt egnet til å framskaffe en oversikt over fenomenet.²

I nordnorsk uttales /r/ normalt som en tapp [r] eller en frikativ/ap-proksimant [r̥] (Foldvik 1977, 111). At stemte sonoranter som /r/ blir assimilert til ustemt foran ustemte obstruenter, er vanlig. I vårt materiale veksler uttalen av /r/ foran /p/ og /k/ først og fremst mellom en ustemt, ikke-sibilantisk frikativ [ɹ̥] og en sibilant som er svært lik [ʁ̥] (Endresen 1988, 50).³ Målet med artikkelen er en auditiv analyse av forekomsten av sibilantuttale. En nærmere fonetisk analyse av de ikke-sibilantiske realisasjonene i denne omgivelsen ligger utenfor målet med artikkelen.

1.1 Tidligere forskning

Ifølge kildene som omtaler sibilantuttalen, fins den i Salten (Brekke 2000, 36, 63; Skånlund 1933, 68; se også Fiva 1996, 135 om Bodø), på Helgeland (Olssen 1958, 56), i Tromsø-området og Vesterålen

¹ Vi takker Institutt for lingvistiske, litterære og estetiske studier ved UiB for økonomisk støtte til kontroll av belegg der forfatterne ikke var samstemte, og Guro Wågan for å ha utført denne kontrollen på en utmerket måte. Anders Nøklestad skylder vi takk for å ha formulert CPQ-søkestrengene vi har brukt for å identifisere beleggene i korpuset. Takk også til Pia Lane (UiO) og Carola Kleemann (UiT) for hjelp med kvensk og samisk fonotaks, og til Hilde Sollid for å ha hentet fram fra arkivet ved UiT og kopiert Hoel (1984) for oss. Vi takker også to fagfeller for nyttige råd og kommentarer.

² Uttalen fins også lenger sør i landet, f.eks. i Trøndelag, men analysen her begrenser seg til de tre nordligste fylkene.

³ Nordlendingene kan altså si [faʁken] eller [faʁʁken], og når det gjelder akkurat dette ordet, fins det henholdsvis 5 og 3 eksempler på disse variantene i LIA-materialet.

(Bull 1996, 177; Elstad 1979, 303; 1982a, 71) og i Finnmark (Hatlebrekke 1996, 220). Ut fra dette kan det se ut til at bruken av sibilantuttale er spredt utover hele landsdelen. Men det er også områder som ifølge kildene skal mangle sibilantuttale: Senja (Iversen 1913, 8), Harstad (Elstad 1979, 303; 1982b, 78) og Lofoten (Christiansen 1933; Elstad 1982a, 71). Disse befinner seg innimellom stedene med sibilantuttale. Det er altså ikke mulig å se noen isogloss eller et kjerneområde for uttalen på grunnlag av den tidligere forskningen.

Et tilsvarende fravær av mønster framgår av målføresynopsisen.⁴ Her finner vi ordene *skarp* og *mark* notert under paragraf 6. I de fleste tilfeller er det notert *r* eller ustemt *r*, men noen steder er det notert sibilantisk uttale. Tabell 1 viser frekvensen av transkribert sibilantuttale.

	<i>skarp</i>	<i>mark</i>
Finnmark	0 % (n=12)	18 % (n=11)
Troms	12 % (n=25)	11 % (n=27)
Nordland	25 % (n=64)	28 % (n=69)

Tabell 1: Forekomst av transkribert sibilantuttale i Målføresynopsisen (n = totalt antall steder der uttalen av *skarp/mark* er angitt)

Som en kan se, er det Nordland som er grundigst dokumentert, mens det er lite data fra de to nordligste fylkene, særlig fra Finnmark. Det er også i Nordland vi finner flest steder med notert sibilantuttale. Totalt er det 19 steder som har fått notert sibilantuttale i ordet *skarp*, og 24 steder som har fått notert sibilantuttale i ordet *mark*.⁵

⁴ Målføresynopsisen er en tabellarisk oversikt over fonologien og morfologien i norske kommuner, basert på *Storms Kortere Ordliste* (1884). Målføresynopsisen er oppbevart hos *Språksamlingane* og tilgjengelig på nett: <http://inger.uib.no/perl/search/search.cgi?appid=145&tabid=2165>.

⁵ Dette gjelder Hasvik (bare i *mark*), Gamvik (bare i *mark*), Balsfjord, Tromsøysund (bare i *skarp*), Dyrøy (bare i *mark*), Medby i Salangen, Værøy (bare i *mark*), Buksnes (bare i *mark*), Sortland, Guvåg i Bø, Steigen, Kjøpsvik i Tysfjord, Korsnes i Tysfjord, Vestbygd i Lødingen, Ballangen (bare i *mark*), Meløy (bare i *mark*), Bodin (bare i *skarp*), Fauske, Fauske, Saltdal (bare i *skarp*), Sørfold (bare i *skarp*), Kjerringøy (bare i *skarp*), Brønnøy, Ylvingen i Vega, Sandnessjøen (bare i *mark*), Lånan i Alstahaug, Herøy (bare i *mark*), Vefsn (bare i *mark*) og Grane.

I flere av oversiktsartiklene om nordnorsk fra 1980- og 90-tallet står det mer eller mindre direkte at sibilantuttale er stigmatisert – for eksempel gjengis det hermer etter folk som bruker sibilantuttale (Elstad 1979, 303; 1982b, 78). Toril Fivas beskrivelse av sibilantuttale i bodødialekten inneholder et eksempel:⁶

Et annet trekk ved dialekten er uttalen av skriftbildets r+p/k: ‘park’, ‘sterk’ og ‘harpe’ uttales gjerne *pašk*, *stæšk* og *hašpe*. I stedet for r uttales altså en sje-lyd som i ‘sjø’ og ‘skjell’. Dette regnes heller ikke for særlig dannet, og mange bodøbarn er her blitt korrigert av sine mødre i åras løp. Det er fort gjort å gå for langt også, slik som den mora som sa til dattera si at «Det heter ikke torsk, det heter tork!». Det samme fenomenet finner vi også i tromsødialekten. (Fiva 1990, 212)

I tillegg skriver flere at uttalen er på vei ut av bruk (Bull 1996, 177; Elstad 1982b, 78; Hatlebrekke 1996, 220).

1.2 Forskningsspørsmål

Med grunnlag i den tidligere forskningen er det tydelig at sibilantuttale av /r/ foran /k/ og /p/ er brukt i mange nordnorske dialekter, men at det ikke er klart hvor vanlig uttalen er. Det er heller ikke klart om det fins et kjerneområde (eller flere) for uttalen. Utsagnene om at uttalen er på vei ut av bruk, og stigmatiseringen av den, tyder på at bruken kan variere langs sosiale dimensjoner. Vi satte derfor opp tre forskningsspørsmål:

- 1) Forekomst: Hva er den geografiske utbredelsen av sibilantuttalen?
- 2) Frekvens: Hvor ofte er /r/ realisert som sibilant foran /k/?
- 3) Fordeling: Fins det sosiale føringer på bruken av sibilantuttale?

For å kunne gi sikre svar på disse spørsmålene burde vi ideelt sett ha hatt et stort, systematisk og geografisk og sosialt balansert korpus til disposisjon. LIA-korpuset er ingen av delene.

⁶ En av fagfellene minner oss på at når det gjelder graden av stigma, kan det være forskjell på byer som Bodø, Harstad og Tromsø på den ene siden, og mindre bygdesamfunn på den andre. Det er vi enige i.

Korpuset består av eldre, transkriberte lydopptak fra mange geografiske lokasjoner som er tatt opp på ulike tidspunkt og med varierende antall informanter i ulike aldersgrupper. Det gjør at det ikke er mulig å gjennomføre systematiske sosiolingvistiske undersøkelser av variasjon i tid og rom og etter ulike sosiale dimensjoner på landsdelsnivå. Siden ingen av opptakene er av nyere dato, er det særlig vanskelig å kunne si noe om forholdene i dag. På tross av disse forholdene mener vi at LIA-materialet kan gi en vesentlig bedre oversikt over den geografiske utbredelsen og over frekvensen av uttalen enn vi har i dag (forskningsspørsmål 1 og 2). Og på enkelte av stedene vi undersøker, er det så mange informanter og belegg på variabelen at vi kan gjennomføre en mer sosiolingvistisk vinkla gransking av om sosiale faktorer har innvirkning på bruk av sibilantuttale (forskningsspørsmål 3).

2 Metode

I tråd med avgrensingen ovenfor opererer vi med to varianter av /r/, [r̥] og [s̥], symbolisert i det følgende som hhv. R og S. Selve variabelen kaller vi S-variabelen. For å identifisere alle potensielt relevante belegg, brukte vi CPQ-søkestrengene gjengitt i (1).⁷

(1) word=".*rk.*"%c / word=".*rp.*"%c

Disse returnerte alle forekomster i den ortografiske transkripsjonen av hhv. <rk> og <rp> innenfor det samme ordet. Søket ble videre avgrenset til fylkene Nordland, Troms og Finnmark. Dette ga 3623 treff på <rk> og 611 på <rp>. Gitt det store området vi undersøker, er dette svært få treff når det gjelder <rp>. Derfor konsentrerer vi oss om <rk>-treffene i resten av artikkelen.

⁷ Det fins også tradisjon for sibilantuttale foran /t/ (se f.eks. Elstad 1982a). Her har imidlertid retroflektierungsregelen overtatt og endret omgivelsen slik at betingelsen for sibilantuttale ikke lenger er til stede.

Vi hørte begge gjennom samtlige treff på variabelen, og annoterte dem som enten R eller S i et regneark der vi hadde kopiert inn alle treffene. Skillet mellom R og S var det eneste skillet vi gjorde, vi skilte ikke mellom stemt og ustemt uttale av den ikke-sibilantiske varianten. For at en gitt realisasjon skal telle som belegg, må både /r/ og den påfølgende lukkelyden være uttalt som separate foner. Etter at annoteringen var fullført, ble følgelig alle treff der lukkelyden ikke ble realisert som egen fon, ekskludert. Dette gjaldt omgivelser der lukkelyden normalt ikke uttales, f.eks. i konsonantgruppene i *mørkt* og *skarpt*, der normaluttalen er ['mœt] og ['skaʔ]. Et annet, leksikalsk eksempel er *Stokmarknes*, som i opptakene fra Hadsel så å si helt konsekvent uttales ['støkma, nəs]. Et siste eksempel er ord der /k/ kan palataliseres, som f.eks. i b.f.sg. *marka*. Her vil uttalen ['maʃça] være prinsipielt mulig, men dette blir likevel en annen omgivelse enn /r/ foran /k/. Uansett er denne typen palatalisering svært sjelden i materialet. Ettersom vi ikke fant eksempler på sibilant i noen tilfeller der /r/ og lukkelyden var skilt av en morfemgrense, som f.eks. i *barkrakene* og *forklare*, ble disse beleggene også trukket ut. Etter dette satt vi igjen med 3013 relevante belegg.

Vi lyttet og annoterte uavhengig av hverandre. Kriteriet for å klasifisere et belegg som sibilant var om lyden vi hørte var karakterisert av den høyfrekvente aperiodiske støyen som karakteriserer sibilanter, og som i prototypiske realisasjoner vil framstå for øret som en [ʃ]. I hvert fall to mulige feilkilder kan svekke reliabiliteten i auditiv basert annotering. For det første kan den auditive terskelverdien knyttet til styrken på den aperiodiske støyen variere individuelt. Og dét problemet kan lett forsterkes når et opptak ikke er av optimal kvalitet, noe som på grunn av alderen karakteriserer en del av innspillingene som inngår i LIA.

Avvikene mellom oss var på 496 belegg, dvs. 16,5 % av alle beleggene.⁸ I tråd med vanlig praksis fikk vi en tredje person til å gå

⁸ At auditiv basert annotering av små, fonetiske forskjeller vil variere fra annotator til annotator, er et vanlig fenomen, se f.eks. Hall-Lew og Fix (2012). Om dette avviket er uventet stort, kan vanskelig vurderes i fravær av et sammenlikningsgrunnlag basert på andres annotering av det samme materialet.

gjennom de beleggene der resultatet var forskjellig. Denne gjennomgangen avgjorde annotasjonen for de beleggene der vi to ikke hadde kommet til samme resultat.

Dette reiser også et mer generelt reliabilitetsproblem knyttet til LIA-korpuset og den såkalte lyd nære transkripsjonen. Ifølge transkripsjonsveiledningen for LIA-prosjektet skal sibilantuttale av /r/ transkriberes som *sj* (Hagen et al. 2018, 22). I prinsippet ville det derfor ha vært mulig for oss å avgrense søkene formulert i (1) til de som var transkribert som *(s)sjk* og *(s)sjp* i den lyd nære transkripsjonen, ved hjelp av søkestrengen gjengitt i (2), for så å regne ut en prosent.⁹

(2) word=".*rk.*"%c & phon=".*sjk.*"

Når vi likevel valgte å annotere med utgangspunkt i egen lytting, var det ut fra det faktum at enhver transkripsjon representerer en individuell tolkning av lydsignalet som ikke nødvendigvis vil samsvare med vår egen tolkning. Dette er etter vår mening et problem med den lyd nære transkripsjonen generelt. Hvis en bruker den som primærdata, er det ikke talen en analyserer, men transkribørens tolkning av denne talen, og denne behøver ikke nødvendigvis være i samsvar med vår egen tolkning. Dette innebærer etter vår mening at kvantitativ variasjonsanalyse av fonologiske fenomener basert på den lyd nære transkripsjonen bare kan brukes eksplorerende og i en tidlig fase av et gitt prosjekt. Skal brukeren av korpuset kunne stå inne for resultatene, må hen selv annotere treffene i korpuset basert på egen lytting, og ideelt av to personer med en tredje som kontrollerer de beleggene der de to første har annotert ulikt. Vi finner klar støtte for dette i resultatene av søkene formulert i (2). Mens vi fant at 780 av de 3013 relevante beleggene ble realisert som sibilant, returnerte søket i (2) bare 326 treff. I prosent utgjør dette en forskjell mellom 10,8 og 25,9 %. I tillegg var det ikke alltid vi hadde annotert sibilantuttale der transkribøren hadde.

⁹ Transkripsjonene av sibilant i den lyd nære transkripsjonen varierte mellom *sj* og *ssj*. Det ble derfor gjort to søk, ett med «sjk» i søkestrengen og ett med «ssjk».

Men reliabilitetsproblemet er naturligvis ikke løst med dette. *Våre* resultater er også basert på en tolkning av lydsignalene, og derfor et stykke på vei utsatt for de samme svakhetene som den lyd nære transkripsjonen. Men mens transkribørene måtte forholde seg til alle aspektene ved lydsignalet, kunne vi konsentrere oss om ett. Og i vår undersøkelse er altså minst 2 av 3 enige om tolkingene. Uansett sikrer kravet om at forskeren må annotere med utgangspunkt i egen lytting, bare det grunnleggende prinsippet at en selv må kunne stå inne for og ta ansvar for resultatet.

3 Resultater

3.1 Forekomst

Av de 3013 beleggene av /r/ foran /k/ er det 780 tilfeller, det vil si 25,9 %, som har sibilant. Tabell 2 viser frekvensen av sibilantuttale og totalt antall tilfeller av /r/ foran /k/ fordelt på de gamle fylkene Finnmark, Troms og Nordland, samt sammenlagt for hele landsdelen. Oversikten viser at sibilantuttalen er noenlunde like frekvent i de tre fylkene. Figur 1 og tabell 3 (s. 50) viser den geografiske fordelingen. I figur 1, som er generert ved hjelp av kartfunksjonen i LIA, er alle stedene i tabell 3 plottet inn på kartet. Sirklene representerer steder der vi har belegg på variabelen, men ingen S-varianter, og de mindre, røde prikkene representerer steder der vi har funnet belegg på S-varianten.

Et hovedfunn som framgår av tabellen, er at det er store forskjeller i frekvensen av sibilantuttale fra kommune til kommune. Det ser heller ikke ut til å være noen geografisk systematikk i hvor sibilantuttalen brukes mye eller lite. For eksempel ser nabokommunene Bardu og Målselv ut til å være forskjellige, og det samme gjelder nærliggende Hadsel og Andøy i Nordland samt Hasvik og Alta i Finnmark.

Finnmark	28,9 % (n=1176)
Troms	24,2 % (n=1348)
Nordland	23,3 % (n=489)
Hele landsdelen	25,9 % (N=3013)

Tabell 2: Frekvens av sibilantuttale av /r/ foran /k/



Figur 1: Geografisk fordeling av sibilantuttale foran /k/

Men på grunn av at det er så ulikt antall informanter fra hvert sted, og at opptakene er av så ulik lengde, er de rene frekvensverdiene i denne generelle oversikten potensielt misvisende. Det er en viktig forskjell på informanter med få belegg og ingen sibilantvarianter og dem med få belegg og likevel minst ett tilfelle av sibilantuttale.

Når det er tale om få belegg, kan vi ikke slutte at talerne i den første gruppa mangler sibilantuttale. Kartet viser at kommunene uten belegg med ett unntak befinner seg i Nordland. Dette gjør det fristende å slutte at fenomenet står svakt der. Dette er det motsatte av hva resultatene fra synopsisen skulle tyde på. Men tallene fra disse kommunene uten sibilantbelegg i Nordland er basert på svært få informanter (én i Leirfjord, Bodø og Sørreisa, to i Bindal og tre i Narvik). I Narvik og Leirfjord er det i tillegg svært få belegg. Datagrunnlaget er med andre ord altfor spinkelt til at vi kan konkludere med at sibilantuttale ikke fins i disse kommunene.

Når det gjelder den andre gruppa, de med få belegg men likevel minst én sibilantvariant, kan vi imidlertid være sikre på at S-varianten er i bruk. Her spiller det lave antallet belegg ingen rolle. Vi kan imidlertid ikke si noe om hvor sterkt varianten står i talen til disse infor-

	<i>Inf. totalt</i>	<i>Inf. med S</i>	<i>Belegg</i>	<i>% S</i>
<i>Finnmark</i>				
Alta	20	7	245	4,1 %
Båtsfjord	4	3	44	59,1 %
Berlevåg	5	4	24	45,8 %
Gamvik	6	4	51	29,4 %
Hammerfest	12	4	138	11,6 %
Hasvik	18	18	136	71,3 %
Kvalsund	2	1	15	40,0 %
Lebesby	3	1	25	8,0 %
Loppa	7	6	98	24,5 %
Måsøy	1	1	22	90,9 %
Vadsø	36	19	329	30,1 %
Vardø	8	4	49	28,6 %
<i>Troms</i>				
Bardu	8	5	90	38,9 %
Berg	8	5	74	16,2 %
Hillesøy	14	7	181	29,3 %
Kvæfjord	1	1	12	16,7 %
Kvænangen	3	2	45	17,8 %
Lavangen	24	14	265	23,8 %
Leinvik	2	2	24	45,8 %
Lyngen	13	7	134	35,1 %
Målselv	19	7	171	7,6 %
Storfjord	7	6	141	17,7 %
Skjervøy	2	1	29	62,1 %
Sørreisa	1	0	16	0,0 %
Torsken	4	3	37	32,4 %
Tranøy	8	5	79	15,2 %
Tromsøysund	9	3	50	30,0 %
<i>Nordland</i>				
Andøya	7	1	39	2,6 %
Bindal	2	0	17	0,0 %
Bodø	1	0	18	0,0 %
Flakstad	8	5	72	11,1 %
Gildeskål	3	1	56	1,8 %
Hadsel	16	13	232	35,8 %
Hemnes	4	3	23	43,5 %
Narvik	3	0	8	0,0 %
Leirfjord	1	0	4	0,0 %
Tysfjord	3	3	16	43,8 %
Vega	1	1	4	75,0 %

Tabell 3: Frekvens av sibilantuttale i /rk/-sekvenser fordelt på kommuner og fylker

mantene. Til det ville vi trenge flere belegg. Det kommer vi tilbake til i neste avsnitt.

Konklusjonen blir uansett at sibilantuttale er svært utbredt, slik figur 1 viser. Den fins i hele landsdelen. Men oversikten i tabell 3 tyder på at den er ulikt geografisk fordelt. På den ene siden har vi en kommune som Hasvik, der samtlige syv informanter bruker varianten, på den andre kommuner som Andøya, Målselv og Alta, der det ser ut til å være relativt langt mellom dem som bruker den.

3.2 Frekvens

Hvor frekvent er sibilantuttale hos dem som har den i sitt fonologiske repertoar? Dette kan ikke oversikten i tabell 3 fortelle oss noe om. At en gitt informant har 100 % eller 0 % basert på ett /rk/-belegg, kan ikke si oss noe om dette, og det kan heller ikke generaliseres basert på to, tre eller fire belegg. For å kunne si noe om dette trenger vi med andre ord data fra informanter som har et visst minimum av belegg.

Hvor stort dette minimumet bør være, fins det ikke noe klart svar på, ut over at jo flere belegg, jo sikrere blir resultatet. I vårt tilfelle må vi likevel ta utgangspunkt i materialet vi har til rådighet. Setter vi minimumet for høyt, blir resultatet at vi sitter igjen med svært få informanter, og resultatene blir tilsvarende lite interessante. Vår vurdering har blitt at et minimum på 7 belegg bør gi en rimelig balanse mellom ønsket om et bredest mulig informantgrunnlag og flest mulig belegg per informant. Det vil si at når vi skal forsøke å si noe om hvor frekvent S-varianten er, ekskluderer vi alle informanter med færre enn syv belegg fra analysen.

<i>Fylke</i>	<i>Inf. totalt</i>	<i>Med S</i>	<i>%</i>	<i>Min.</i>	<i>Maks.</i>	<i>Median</i>	<i>Gj.snitt</i>
Finnmark	69	43	62,3 %	3,8 %	100,0 %	44,2 %	46,8 %
Troms	75	53	70,7 %	2,1 %	100,0 %	27,3 %	37,1 %
Nordland	26	18	69,2 %	6,7 %	93,3 %	24,4 %	36,9 %
Totalt	170	114	67,1 %	2,1 %	100,0 %	30,0 %	41,0 %

Tabell 4: Frekvens av S-varianten blant informanter med minst 7 belegg

Tabell 4 viser hvor mange informanter som totalt har syv belegg eller flere, og hvor mange av disse som har minst ett S-belegg. Vi ser at

det for alle tre fylkene dreier seg om ca. to tredeler av informantene. Gjennomsnittlig S-prosent for informantene med sibilantbelegg framgår av kolonnen helt til høyre, «Gj.snitt». Det som imidlertid er sentralt her, er spennet mellom informanten med færrest S-belegg og den med flest. I alle tre fylkene finner vi det samme mønsteret, et spenn fra godt under 10 % til 100 % i Troms og Finnmark og godt over 90 % i Nordland. Gjennomsnittsprosentene både i tabell 3 og tabell 4 skjuler derfor en svært høy grad av variasjon. Det synes trygt å konkludere med at noen språkbrukere har S-varianten i sitt repertoar, men bruker den sjelden, mens andre bruker den helt eller nesten helt gjennomført. De to siste kolonnene sier litt mer om hvordan fordelingen er mellom den høyeste og den laveste verdien. Medianen er det aritmetiske midtpunktet i settet av individuelle frekvenser, mens «Gj.snitt» er gjennomsnittet av de individuelle frekvensene. Når verdiene er tilnærmet like, vil det si at de individuelle verdiene er noenlunde jevnt spredd over spennet fra laveste til høyeste frekvens. Når medianen for Troms og Nordland ligger godt under gjennomsnittet, viser det at mange lave verdier balanseres mot noen få svært høye i gjennomsnittet. I Finnmark er skårene jevnere fordelt, siden median og gjennomsnitt er omtrent like.

Som vi alt har sett, er informant- og beleggtallet for lite i de fleste kommunene for en tilsvarende analyse på kommunenivå. Men for noen er det mulig. Tabell 5 viser kommunene der 7 eller flere informanter har 7 eller flere belegg hver av /rk/-variabelen. Kolonnen «Inf. med $N \geq 7$ » viser hvor mange informanter som har flere enn 6 belegg på variabelen, og «Inf. med S-belegg» viser hvor mange av disse som

Kommune	Fylke	Inf. med $N \geq 7$	Inf. med S-belegg	Min.	Maks.	Median	Gj.snitt
Alta	F	9	3	4,5 %	11,1 %	8,3 %	8,0 %
Hasvik	F	7	7	30,0 %	100,0 %	79,8 %	69,4 %
Vadsø	F	20	12	3,8 %	84,2 %	35,2 %	40,0 %
Hillesøy	T	8	5	2,1 %	87,9 %	55,6 %	52,4 %
Lavangen	T	16	13	7,1 %	100,0 %	28,6 %	39,5 %
Lyngen	T	8	6	5,3 %	75,0 %	45,0 %	42,7 %
Målselv	T	11	5	6,5 %	20,0 %	18,2 %	14,7 %
Hadsel	N	14	13	7,1 %	93,3 %	43,5 %	45,5 %

Tabell 5: Spredning av individuelle frekvenser i kommuner med minst 7 informanter med minst 7 belegg

har belegg på S-varianten. Tabellen viser også variasjonsbredden i frekvenser mellom informantene med belegg på S-varianten, laveste frekvens («Min.»), høyeste («Maks.»), medianen og gjennomsnittet.

Tabellen forteller oss for det første hvor utbredt sibilantuttale er i de enkelte kommunene. I Hasvik har 7 av 7 sibilantuttale, og vi kan derfor anta at fenomenet er svært utbredt der. Dette står i motsetning til Alta og Målselv, der under halvparten av informantene har sibilantuttale. Forekomsten varierer med andre ord i høy grad.

Frekvensmålene tyder på at spredningen er liten i Alta og Målselv. Her kan det se ut som om det er mange som mangler sibilantuttale, og de som har, bruker den relativt sjelden. I de andre kommunene er spredningen mye større, noe som tyder på at sibilantuttale er ujevnt fordelt også i befolkningen generelt.

Konklusjonen på dette avsnittet blir at individuelle frekvenser varierer svært mye når vi baserer analysen på informanter med syv belegg eller mer. En del av disse informantene bruker ikke S-varianten i opptakene, jf. den fjerde kolonnen i tabell 5. Av dem som bruker den, synes frekvensen å variere fra svært sjelden til gjennomført. Denne spredningen finner vi igjen på kommunenivå. Tabell 5 tyder på at det fins kommuner, som Alta og Målselv, der sibilantuttale er lite frekvent både på individ og gruppenivå, mens i andre, særlig i Hasvik i vårt materiale, er den svært utbredt og brukes av de fleste av innbyggerne.

4 Sosiale føringer

I innledningen så vi at bruk av sibilantuttale i dialektologiske skildringer av nordnorsk ofte blir framstilt som stigmatisert og/eller på vei ut av bruk. Dette tyder på at det kan være sosiale føringer eller sosiale faktorer som påvirker bruken av sibilantuttale. I denne delen av artikkelen vil vi utforske i hvilken grad det er mulig å identifisere slike faktorer. Vi har valgt å se nærmere på materialet fra kommunene Vadsø og Lavangen fordi det er fra disse kommunene vi har mest data om sibilantuttalen.

4.1 Vadsø

Fra Vadsø kommune i Finnmark har vi belegg fra totalt 36 informanter. Gjennomsnittsbruken av sibilantuttale er 30,1 % (jf. tabell 3), og dette ligger nært opptil gjennomsnittet i fylket (28,9 %, jf. tabell 2). Blant informantene med minst 7 belegg på variabelen og minst ett belegg med S er imidlertid den gjennomsnittlige frekvensen av sibilantuttale noe høyere (40,0 %, jf. tabell 5). Dette er til tross for at 8 av de 20 med minst 7 belegg på variabelen ikke bruker sibilantuttalen i opptakene (tabell 5).

Informantene fra Vadsø er 19 kvinner og 17 menn, og de er født mellom 1895 og 1959 (for fire informanter er fødselsår ikke oppgitt i metadataene). Samtlige opptak er samtalebaserte intervju som ble gjennomført på 1970-tallet av Hildur Hatlebrekke. I tillegg til alder og kjønn har Hatlebrekke klassifisert informanter på bakgrunn av hvilken bygd i Vadsø kommune de kom fra. Hun satte også som krav til informantene at de skulle ha «norsk som morsmål» (Hatlebrekke 1976, 2). Dette er fordi mange i Vadsø hadde vokst opp med kvensk morsmål (se nedenfor).

Tabell 6 viser informantene med minst 7 belegg på variabelen, hvilken del av kommunen de kommer fra, kjønn, fødselsår og yrke, samt frekvensen av sibilantuttale. Det er vanskelig å se noen klare mønstre i tabellen. Vi finner både kvinner og menn i gruppen som bruker mye sibilantuttale, så vel som blant dem med lave frekvenser. Blant dem som bruker mest sibilantuttale, finner vi en kjøpmann, et par fiskere/småbrukere og husmødre, men vi finner også fiskere, småbrukere og husmødre med svært lave frekvenser. Selv om det kan se ut til å være noe aldersforskjell i bruken av sibilantuttale, der de eldste (født rett før/etter år 1900) bruker mer sibilantuttale enn de andre informantene, finner vi også informanter født i 1900 og 1904/05 som ikke bruker sibilantuttale.

Det vi oppfatter som det mest interessante mønsteret i materialet, er geografisk. Deler vi inn informantene etter de fem oppgitte delene av kommunen, finner vi store geografiske forskjeller: Mens de fleste informanter fra de østlige delene av kommunen, Ekkerøy, Kiby og Krampen, bruker sibilantuttale svært ofte, brukes sibilantuttale sjelden i Vadsø by og Vestre Jakobselv. Blant informantene med 7 eller

<i>Informant</i>	<i>Bygd</i>	<i>Kjønn</i>	<i>F. år</i>	<i>Yrke</i>	<i>N</i>	<i>% S</i>
uit_0501	Kiby	M	1907	Fisker, småbruker, snekker	19	84,2 %
uit_1001	Kiby	F	1897	Husmor	17	82,4 %
uit_1401	Ekkerøy	F	1904	Husmor	20	80,0 %
uit_0802+0804	Ekkerøy	M	1908	Fisker, småbruker	16	62,5 %
uio_0501+601+1201	Kiby	M	1903	Kjøpmann	21	57,1 %
uit_0701	Vadsø	F	-		11	45,5 %
uio_0102	Vadsø	F	1958		12	25,0 %
uit_1101	Kiby	M	1942	Bygningssekker	7	14,3 %
uit_0301	Ekkerøy	M	1912	Småbruker	10	10,0 %
uio_0901	Vadsø	F	-	Offentlig ansatt	12	8,3 %
uio_0701	Vadsø	F	1923	Husmor og kontorist	14	7,1 %
uit_1602	Vestre	M	1922	Pensjonert fisker	26	3,8 %
	Jakobselv					
uio_0101	Vadsø	F	1958		11	0,0 %
uio_0601	Vadsø	M	1900	Havnearbeider	13	0,0 %
uio_0802	Vadsø	F	1935	Husmor	7	0,0 %
uit_0201	Vadsø	M	-		12	0,0 %
uit_1204	-	M	-		19	0,0 %
uit_1302	Vadsø	M	1904	Sjømann, fisker, småbruker	10	0,0 %
uit_1303	Vadsø	M	1933		14	0,0 %
uit_1603	Vestre	F	1918	Husmor	9	0,0 %
	Jakobselv					

Tabell 6: Oversikt over informantene i Vadsø med minst 7 belegg på variabelen

flere belegg på variabelen (tabell 6), er det bare 4 av 10 fra Vadsø sentrum som har sibilantuttale, mens 7 av 8 fra Kiby og Ekkerøy bruker den. Gjennomsnittstallene er 8,6 % i Vadsø by og hele 54,2 % i Kiby og Ekkerøy.

Årsakene til disse forskjellene kan selvsagt være mange, men Hatlebrekke (1981) peker selv på en etnisk forskjell mellom Vadsø sentrum og Kiby og Ekkerøy. Vadsø kalles ofte for «kvenhovedstaden» fordi så mange kvener bosatte seg der. I en periode (1865–1885) var det flere kvener enn nordmenn som bodde i byen. Ifølge Hatlebrekke har dette ført til at Vadsø by skiller seg klart fra bygdene omkring (særlig Kiby og Ekkerøy) med å ha et talemål med flere «elementer fra bokspråket», mens bygdene omkring representerer den tradisjonelle dialekten i området (Hatlebrekke 1976, 14). Ut fra vår undersøkelse og tidligere forskning om sibilantuttalen er det ikke så lett å peke på hva som er den tradisjonelle varianten i nordnorsk. I tillegg

kommer det momentet at selve språkkontaktsituasjonen og språkbruken som kommer som følge av denne, også må regnes som tradisjonell i Nord-Norge (jf. Bull 1991; 1995). Hvis vi nå likevel tolker sibilantuttale som et nordnorsk dialekttrekk, ser vi at vår kartlegging av sibilantuttale i Vadsø samsvarer med Hatlebrekke sine funn av at talemålet i sentrum er mer normalisert, og bygdene omkring bruker flere dialekttrekk. Men dette lokalt forankra sosiale mønsteret vil vanskelig la seg generalisere til andre deler av landsdelen, og kan slik sett ikke forklare distribusjonen av sibilantuttale vi finner på landsdelsnivå i undersøkelsen vår.

4.2 Lavangen

Fra Lavangen kommune i Troms fylke har vi totalt 24 informanter med belegg på /rk/-variabelen. Gjennomsnittsbruken i kommunen er 23,8 % (tabell 3), og dette ligger nært opptil gjennomsnittet på fylkesnivå (24,2 %, tabell 2). 16 av informantene har 7 belegg på variabelen eller mer, og blant disse er det 13 med belegg på S-varianten (jf. tabell 5).

Opptakene fra Lavangen er samlet inn i forbindelse med forskningsprosjektet *Språkforhold blant samer i Troms*, som ble gjennomført på begynnelsen av 1980-tallet ved Universitetet i Tromsø (Hoel 1984). Formålet var å undersøke den norske og den samiske dialekten til innbyggerne i Spansdalen, som ligger ovenfor og øst for Tennevoll, sentrum i Lavangen. Ifølge Hoel (1984, 2) hadde Spansdalen på denne tida så å si bare innbyggere med samiske røtter, og bygda hadde gjennomgått språkskifte fra samisk til norsk i løpet av de siste par generasjonene. I tillegg til opptak med 20 informanter fra Spansdalen ble det gjort intervju med 37 informanter fra andre deler av Lavangen (Tennevoll og omegn) for å ha et sammenlikningsgrunnlag for den norske dialekten. Dessverre er ikke alle intervjuene fra Lavangen inkludert i LIA-materialet, og bare 2 personer fra Spansdalen har kommet med i korpuset.

Tabell 7 viser informantene med 7 eller flere belegg på variabelen, hvilken bygd/del av kommunen de kommer fra, kjønn, fødselsår og yrke, samt frekvensen av sibilantuttale.

<i>Informant</i>	<i>Bygd</i>	<i>Kjønn</i>	<i>F. år</i>	<i>Yrke</i>	<i>N</i>	<i>% S</i>
uit_0902	Tennevoll	F	1955	Husmor	7	100 %
uit_1301	Tennevoll	F	1943	Husmor	9	100 %
uit_1001	Tennevoll	M	1915	Maler	13	84,6 %
uit_1401	Tennevoll	F	1947	Legesekretær	17	47,1 %
uit_0901	Tennevoll	M	1910	Lærer, klokker	15	46,7 %
uit_0101	Tennevoll	F	1961	-	10	30,0 %
uit_0102	Tennevoll	M	1950	-	14	28,6 %
uit_1303	Tennevoll	M	1932	Bussjåfør	24	20,8 %
uit_0502	Spansdalen	F	1941	-	11	18,2 %
uit_1103	Hesjevika	M	1931	Lærer	21	14,3 %
uit_0302	Tennevoll	F	1933	Husmor	9	11,1 %
uit_0501	Spansdalen	M	1968	-	14	7,1 %
uit_1201	Lotternes	M	1910	-	19	5,3 %
uit_0903	Tennevoll	M	1959	Lærer	14	0,0 %
uit_1003	Røkenes	M	1904	Småbruker, fisker	20	0,0 %
uit_1501	Lotternes	F	1897	Husmor	21	0,0 %

Tabell 7: Oversikt over informantene i Lavangen med minst 7 belegg

Heller ikke i dette delmateriale fra Lavangen finner vi noen klare kjønnsforskjeller, tydelig aldersfordeling, eller yrkeskategorier som kjennetegnes av at de bruker mye/lite sibilantuttale. Derimot finner vi, i likhet med i Vadsø, forholdsvis klare lokalgeografiske forskjeller. Mens gjennomsnittstallet for bruk av sibilant i Tennevoll (sentrum av Lavangen) er 41,7 % for informantene med 7 eller flere belegg, er sibilanten bare brukt i 4,9 % lenger ute i fjorden (Røkenes, Hesjevika, Lotternes) og 12 % i Spansdalen. Vi finner altså en klar geografisk dimensjon i den språklige variasjonen også i Lavangen, men denne sentrum–periferi-dimensjonen følger det motsatte mønsteret av Vadsø. Dette tolker vi som et tydelig tegn på at det ikke er sentrum–periferi-dimensjonen i seg selv, men befolkningsmønstrene og språkkontaktsituasjonene som er utslagsgivende. Spansdalen er som nevnt et område som nylig har gjennomgått et språkskifte fra samisk til norsk. I tillegg var deler av bosettinga langs fjorden utenfor Tennevoll sjøsamisk eller nordmenn som kunne samisk (jf. Friis 1861, se også Hansen 2003, 337–339), og det kan bety at også de andre småstedene langs fjorden har gjennomgått et språkskifte for noen generasjoner siden.

4.3 Oppsummering og drøfting av sosiale faktorer

Sammenlikningen av dem som bruker mye/lite sibilantuttale på de to stedene, viser ingen andre sosiale fellestrekk enn geografi: Hvor informantene bor i kommunen, ser altså ut til å være mer utslagsgivende for bruken av sibilant enn alder, kjønn og yrke. Begge kommunene har områder der dialekten er påvirket av språkkontakt, og områder med en mer enhetlig norsk dialektbakgrunn. Sibilantuttale ser ut til å være langt mer brukt i de sistnevnte områdene. At språkkontakten mellom samisk, norsk og kvensk har ført til mange språklige endringer opp igjennom språkhistorien, er det ingen tvil om (Bull et al. 2018, 429–457). Språklige forskjeller mellom de «nye» nordnorske dialektene (eller etnolektene) som snakkes i språkkontaktområder, og nordnorsken som snakkes i mer enspråklige områder, er også godt dokumentert tidligere, for eksempel gjennom Tove Bulls undersøkelser av talemålet på Furufleten i Nord-Troms (se f.eks. Bull 1991; 1992a; 1992b; 1995). Men lavfrekvent sibilantuttale i /rk/-sekvenser er etter det vi vet, ikke trukket fram i tidligere forskning som et etnolektisk trekk.

Faktorene som har ført til denne uttalen med /r/ foran /k/, er etter vår mening både lingvistiske og sosiale. Verken i samisk eller kvensk kan /r/ uttales [ʂ] foran /k/, og [ʂ] brukes heller ikke i kvensk ellers (bare i innlånte stedsnavn) (Nickel og Sammallahti 2011; Söderholm 2017, 39). Denne typen interferens i språkskiftesituasjoner kan overføres til senere generasjoner og slik bli del av en etnolekt (jf. Bull 1991; 1995; 2005, 36; Sankoff 2013; Sollid 2009).

I tillegg er det sosiale faktorer som har fremmet /rk/-uttalen foran sibilantuttalen. Språkskiftesituasjonen i Nord-Norge var kjennetegnet av at samene og kvenene først og fremst ble fornorsket gjennom skoleverket og internatskolene (Bull et al. 2018, 441), med lærebøker på riksmål/bokmål og lærere med et skriftmålsnært talemål (Bull 1992a). Dette medførte at deres måte å snakke norsk på har ligget nærmere bokmål enn mange andre nordnorske dialekter og bymål (Martinussen 1996; Bull 1992a). I innledningen så vi at tidligere forskning karakteriserer sibilantuttale som stigmatisert, og da er det

naturlig å tro at lærerne unngikk den i undervisningssituasjonen, og at de muligens også instruerte elevene til å unngå den.¹⁰

Tove Bull (2015) beskriver hvordan språktrekk som i utgangspunktet kan se ut til å være bymålsnære eller bokmålsnære, ofte oppfattes som typiske for talere med samisk eller kvensk førstespråk i det nordlige Nord-Norge. Det er fornorskingsprosessene fra midten av 1800-tallet som kan forklare at bymålsnære og bokmålsnære språktrekk har fått en slik lokal betydning (Bull 1992a; 2005, 35). Ut fra våre resultater kan det se ut til at sibilantuttale av /rk/-sekvensen føyer seg inn i et slikt mønster der lavfrekvent bruk av sibilant ikke nødvendigvis bare signaliserer bymål eller bokmålsnær tale, men også kan signalisere samisk/kvensk substrat. Siden en stor del av informantene i Vadsø sentrum/vest og i periferien i Lavangen har gjennomgått et språkskifte for ikke mange generasjoner siden, er det derfor ikke urimelig å tolke frekvent bruk av /r/ i *rk*-lydsekvenser her som et etnolektisk substrattrekk som har etablert seg som del av det lokale talemålet. I dette tilfellet trekker de lingvistiske forutsetningene og de sosiale faktorene i språkskiftesituasjonen i samme retning, og fremmer bruk av /r/ og ikke [ʂ] foran /k/.

Det ville ha vært interessant å undersøke nærmere om informantene med lite sibilantuttale også har andre kjente etnolektiske trekk i talemålet sitt, for eksempel lite retroflektering, førsteleddstrykk i lånord eller redusert genusmarkering (Bull 1992a; 1992b), men det har vi dessverre ikke rom for i denne artikkelen. LIA-korpuset er heller ikke rikt nok til at vi kan gå i dybden på variasjonen på steder i Nord-Norge som er mindre prega av språkkontakt. Hvorvidt variasjonsmønstrene som fins i f.eks. Bodø, henger sammen med hvorvidt ulike alders-, kjønns- og yrkesgrupper normaliserer talemålet sitt, eller om det er andre faktorer som styrer variasjonen på disse stedene, forblir derfor et åpent spørsmål.

¹⁰ Bull (1992b) antyder en liknende forklaring når det gjelder manglende postalveolar uttale av /rn/, /rd/, /rt/ og /rl/.

5 Konklusjon

Vi innledet artikkelen med å stille tre forskningsspørsmål om sibilantuttale av /rk/-sekvenser som vi ville bruke den nordnorske delen av LIA-korpuset til å undersøke:

- Forekomst: Hva er den geografiske utbredelsen?
- Frekvens: Hvor ofte er /r/ realisert som sibilant foran /k/?
- Fordeling: Fins det sosiale føringer på bruken av sibilantuttale?

Tabell 2 og figur 1 viser for det første at /r/ kan bli realisert som sibilant foran /k/ i hele landsdelen. For det andre viser undersøkelsen at frekvensen av sibilantuttale er svært varierende – vi finner store forskjeller i frekvenser mellom undersøkelsesstedene og høy grad av interindividuell variasjon. For mange av stedene er frekvensen basert på data fra få informanter og/eller få belegg, men også stedene med forholdsvis mange informanter og mange belegg på variabelen viser mye interindividuell variasjon. For det tredje viser undersøkelsen vår fra Vadsø og Lavangen ingen spesielle sosiale kjennetegn ved informantene som bruker mye/lite sibilantuttale, utenom at det er geografiske skillelinjer innad i kommunene. Her ser det ut til at sibilantuttale brukes mindre i områder der talemålet kan være preget av et tidligere språkskifte fra samisk eller kvensk til norsk.

Gjert Kristoffersen 1949–2021

Gjert døde brått 29. mai 2021. Da var denne artikkelen ferdig skrevet og formatert. Vi samarbeidet om alt fra valg av tema til dataanalyse og skrivning. Tapet er stort, men jeg tar med meg videre minnet om en svært hyggelig kollega og dyktig fagperson og om lærerike faglige samtaler.

Referanser

- Brekke, Olga. 2000. *Saltendialekten. En grammatikk*. Fauske: Salten og Skjerstad bygdeboknemnd.
- Bull, Tove. 1991. Språklig identitet i ei nordnorsk fjordsamebygd. *Danske folkemål* 33, 23–35.
- Bull, Tove 1992a. Dialect and standard in a language contact area in northern Norway. I *Dialect and Standard Language in the English, Dutch, German and Norwegian Language Areas*, redigert av J. A. Leuvensteijn og J. B. Berns, 365–378. Amsterdam: Royal Netherlands Academy of Arts and Sciences.
- Bull, Tove 1992b. A contact feature in the phonology of a northern Norwegian dialect. I *Language Contact. Theoretical and Empirical Studies*, redigert av Ernst Håkon Jahr, 17–36. Berlin og New York: Mouton de Gruyter.
- Bull, Tove. 1995. Language maintenance and loss in an originally trilingual area in North Norway. *International Journal of the Sociology of Language* 115, 125–134.
- Bull, Tove. 1996. Tromsø bymål. I *Nordnorske dialektar*, redigert av Ernst Håkon Jahr og Olav Skare, 175–179. Oslo: Novus.
- Bull, Tove. 2005. The concept of ethnolect: Ethnolectification and de-ethnolectification of varieties of Norwegian vernaculars in language contact areas. I *Från översättning till etik. En festskrift till Irma Sorvali på hennes 60-årsdag den 15 oktober 2005*, redigert av Paula Rossi, 33–40. Oulu: University of Oulu.
- Bull, Tove. 2015. Nordnorsk i 200 år. I *Talemål etter 1800*, redigert av Helge Sandøy, 243–266. Oslo: Novus.
- Bull, Tove, Espen Karlsen, Eli Raanes og Rolf Theil. 2018. Andre språk i Noreg. I *Norsk språkhistorie II: Praksis*, redigert av Brit Mæhlum, 418–532. Oslo: Novus.
- Christiansen, Hallfrid. 1933. *Gimsøy-målet*. Skrifter utgitt av Det norske videnskaps-akademi i Oslo. II Historisk-filosofisk klasse 1932. Oslo: Jacob Dybwad.
- Elstad, Kåre. 1979. Målet i Troms. I *Troms. Bygd og by i Norge*, redigert av Ivan Kristoffersen, 297–316. Oslo: Gyldendal norsk forlag.

- Elstad, Kåre. 1982a. *Borgfjerdingsmål* 1–2. Oslo: Novus.
- Elstad, Kåre. 1982b. Nordnorske dialektar. I *Nordnorsk: språkav og språkforhold i Nord-Noreg*, redigert av Tove Bull og Kjellaug Jetne, 9–100. Oslo: Det Norske Samlaget.
- Endresen, Rolf Theil. 1988. *Fonetikk. Ei elementær innføring*. Oslo: Universitetsforlaget.
- Fiva, Toril. 1990. Bodødialekten. I *Den store dialektboka*, redigert av Ernst Håkon Jahr, 211–214. Oslo: Novus.
- Fiva, Toril. 1996. Bodødialekten. I *Nordnorske dialektar*, redigert av Ernst Håkon Jahr og Olav Skare, 134–137. Oslo: Novus.
- Friis, Jens Andreas. 1861. *Ethnografisk Kart over Finmarken* 1–5. 10 Bl. Tilgjengelig fra: <https://www.dokpro.uio.no/friiskartene/1861/1861oversikt.html>, sist besøkt 11.3.2021.
- Foldvik, Arne Kjell. 1977. Realisasjon av r i norsk. *Svenska landsmål och svenskt folkliv* 100, 110–118.
- Hall-Lew, Lauren og Sonya Fix. 2012. Perceptual coding reliability of (L)-vocalization in casual speech data. *Lingua* 122, 794–809.
- Hagen, Kristin, Live Håberg, Eirik Olsen og Åshild Søfteland. 2018. *Transkripsjonsrettleiing for LIA*. Tilgjengelig fra: http://tekstlab.uio.no/LIA/pdf/transkripsjonsrettleiing_lia.pdf, sist besøkt 7.5.2020.
- Hansen, Lars Ivar. 2003. *Astafjord bygdebok. Historie 2*. Lavangen: Lavangen kommune.
- Hatlebrekke, Hildur. 1976. *Språkforhold i Vadsø by: med særlig vekt på sterke verb*. Hovedoppgave, Universitetet i Trondheim.
- Hatlebrekke, Hildur. 1981. *Vadsø-dialekten, med glimt fra dialektene i Kiby og Ekkerøy*. Vadsø: Varanger mållag.
- Hatlebrekke, Hildur. 1996. Snakker finnmarkingene bokmål? I *Nordnorske dialektar*, redigert av Ernst Håkon Jahr og Olav Skare, 218–222. Oslo: Novus.
- Hoel, Thomas. 1984. *Orientering om prosjektet «Språkforhold blant samer i Troms»*. Upublisert rapport. Tromsø: Nordnorsk talemålsarkiv, Universitetet i Tromsø.
- Iversen, Ragnvald. 1913. *Senjen-maalet. Lydverket i hoveddrag*. Skrifter Historisk-Filosofisk klasse. Kristiania: Videnskabs-Selskabet i Christiania.

- Martinussen, Bente. 1996. Mot én dialekt i Finnmark? I *Nordnorske dialektar*, redigert av Ernst Håkon Jahr og Olav Skare, 238–248. Oslo: Novus.
- Måføresynopsisen*. Språksamlingene, Universitetet i Bergen. Tilgjengelig fra: <http://inger.uib.no/perl/search/search.cgi?appid=145&tabid=2165>, sist besøkt 11.3.2021.
- Nickel, Klaus Peter og Pekka Sammallahti. 2011. *Nordsamisk grammatikk*. 2. hápmi/utgave. Karasjok: Davvi girji.
- Olszen, Olaf. 1958. *Brønnøymålet. Lydverket*. Skrifter frå Norsk mål-føreakiv, Vol. XIII. Oslo: A/S Bokcentralen.
- Sankoff, Gillian. 2013. Linguistic outcomes of bilingualism. I *The Handbook of Language Variation and Change*, redigert av J. K. Chambers og Natalie Schilling. Chichester: Wiley.
- Skånland, Einar B. 1933. *Saltamålet. Kort oversikt over lydverket*. Oslo: H. Aschehoug & Co.
- Sollid, Hilde. 2009. Utforsking av nordnorske etnolekter i et faghistorisk lys. *Maal og Minne* 2009 (2), 147–169.
- Söderholm, Eira. 2017. *Kvensk grammatikk*. Oslo: Cappelen Damm Akademisk.

English summary

This article analyzes the use of [ʂ] as an allophone of /t/ before /k/ in Northern Norway by means of the LIA-corpus. We look at the geographical distribution of the variant, both in the region as a whole and in the local communities covered by LIA. Our survey shows that the variant is in use all over the region. But frequency of use is subject to considerable variation, also among those where we have found a reasonable number of relevant tokens. No correlations between background variables such as gender, social background or age seem to be present, but the LIA metadata are unfortunately too sparse for any reliable conclusions to be made. On the other hand, we have discovered evidence pointing towards the variant being less frequently used in two local communities which have fairly recently undergone shift from Fenno-Ugric Kven and Sámi into Norwegian.

Randi Neteland
Institutt for lingvistiske, litterære og estetiske studier
Universitetet i Bergen
randi.neteland@uib.no



Forskjeller mellom talemål og skriftspråk

Hva kan trebanker fortelle oss?

Signe Laake og Lilja Øvrelid

Denne artikkelen presenterer en sammenlignende studie av norsk talemål og skriftspråk som tar for seg en rekke syntaktiske forskjeller mellom skrift og tale som har blitt foreslått i tidligere studier. Den nylig utviklede LIA-trebanken inneholder transkribert talespråk og er annotert for syntaks i overensstemmelse med den skriftspråklige Norsk Dependenstrebant, med visse utvidelser for å dekke talespråksspesifikke fenomener. Den syntaktiske analysen i disse trebankene tillater presise syntaktiske søk som gjør at vi kan undersøke en rekke syntaktiske fenomener, som sideordning vs. underordning, utelating av syntaktiske ledd og sammenligning av leddstilling, og sammenligne disse fenomenene i skrift og tale i norsk.

Stikkord: LIA-trebanken, Norsk Dependenstrebant (NDT), syntaktisk annotasjon, dependensgrammatikk, skriftspråk, talespråk

1 Introduksjon

Tilgjengeligheten av trebanker muliggjør lingvistiske studier ved søk i manuelt annotert syntaktisk struktur. Den nylig utviklede LIA-trebanken inneholder transkribert talespråk og er annotert i overensstemmelse med den skriftspråklige Norsk Dependenstrebant, med visse utvidelser for å dekke talespråksspesifikke fenomener. Dette muliggjør kvantitativ sammenligning basert på dependensrepresentasjoner.

I denne artikkelen presenterer vi en sammenlignende studie av norsk talemål og skriftspråk som tar for seg en rekke syntaktiske forskjeller mellom skrift og tale som har blitt diskutert i tidligere studier. Vi viser her at den syntaktiske analysen i disse trebankene tillater presise syntaktiske søk som gjør at vi kan fange opp en rekke syntaktiske fenomener, som sideordning vs. underordning, utelating av syntaktiske ledd og sammenligning av leddstilling.

Artikkelen er strukturert som følger: I seksjon 2 introduserer vi den syntaktiske formalismen som danner grunnlag for trebankene, nemlig dependensgrammatikk, og i seksjon 3 presenterer vi de to trebankene (NDT og LIA) som danner datagrunnlaget for denne studien. Seksjon 4 beskriver den empiriske studien, der vi begynner med en generell sammenligning av de to trebankene, før vi ser nærmere på spesifikke syntaktiske mønstre i disse. Avslutningsvis konkluderer vi artikkelen i seksjon 5.

2 Dependensgrammatikk og dependenstrebant

Syntaktiske representasjoner i form av såkalte dependensstrukturer, uttrykt ved bileksikale relasjoner mellom ord, har mottatt økt oppmerksomhet innen språkteknologisk forskning i løpet av de siste ti til femten årene. Utvikling av algoritmer som effektivt kan behandle disse strukturene, utgjør en viktig grunn til at dependensgrammatikk nå er blitt den dominerende syntaktiske formalismen innen språkteknologi.

Dependensgrammatikk har en lang historie innen teoretisk lingvistikk (Tesnière 1959; Mel'čuk 1988), men kan på ingen måte hev-

des å utgjøre en enhetlig syntaktisk formalisme. Et sentralt kjennetegn ved ulike tilnærminger til dependensgrammatikk er allikevel at syntaktisk struktur ikke beskrives ved konstituentstruktur, men snarere gjennom binære relasjoner mellom ord, såkalte bileksikale relasjoner. Formelt sett er disse representasjonene *trær*; det vil si en graf der alle noder kan nås fra en *rot*. Roten er toppnoden i grafen og den eneste noden som ikke har noen innkommende kanter. Kantene i dependensgrafene er merket med *dependensrelasjoner*, der relasjonene som regel uttrykker syntaktiske funksjoner (subjekt, objekt osv.), men kan også være basert på eksempelvis semantiske roller.

2.1 Trebanker

Trebanker er tekstkorpus som er manuelt annotert med syntaktisk struktur. Slike tekstsamlinger står helt sentralt i dagens språkteknologi siden de muliggjør maskinlæringsbasert utvikling av automatiske verktøy, som ordklassetaggere og syntaktiske parsere. Trebanker kan også være svært nyttige for lingvistiske studier og lingvistisk hypotesetesting fordi de muliggjør søk direkte i manuelt oppmerket syntaktisk struktur.

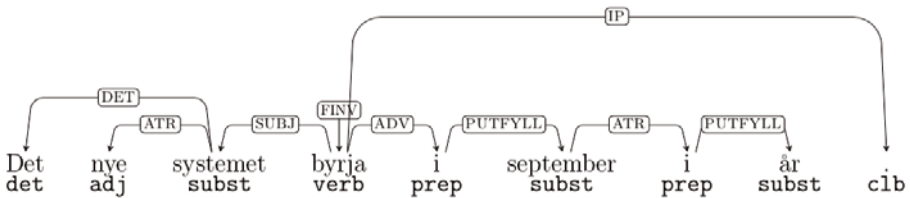
Trebanker har derfor blitt utviklet for en rekke ulike språk, og et stort antall trebanker har i løpet av senere tid blitt annotert med dependensgrammatikk. Der trebanker med konstituentbasert analyse, som den engelske Penn Treebank (Marcus et al., 1993), lenge var dominerende, har dependenstrebanker i senere år blitt klart i flertall. En utfordring har imidlertid vært at ulike dependensgrammatiske tradisjoner har dannet grunnlag for de ulike trebanksprosjektene, og at analysene derfor har vært langt fra enhetlige. Universal Dependencies-initiativet (Nivre et al. 2016) har vært et dugnadsdrevet initiativ for å samle de ulike trebankene for ulike språk under en felles annotasjonsstandard på tvers av språk og tilgjengeliggjøre disse digitalt.

2.1.1 Norsk dependenstrebank

I 2014 ble Norsk dependenstrebank (NDT; Solberg et al. 2014) gjort tilgjengelig. Den ble utviklet ved Språkbanken (Nasjonalbiblioteket) i samarbeid med Tekstlaboratoriet, UiO og Språkteknologigruppen

ved Institutt for Informatikk, UiO. NDT inneholder totalt 600 000 ord fordelt ca. 50–50 på bokmål og nynorsk. Kildene er i hovedsak nyhetstekst (ca. 85%), men også noe tekst hentet fra offentlige rapporter og bloggdata. Tekstene ble annotert av to lingvister etter pre-prosesserings med en automatisk parser (videre beskrevet i Solberg et al, 2014). I forbindelse med trebankarbeidet ble også et sett med retningslinjer utformet (Kinn et al. 2014) som gir detaljert beskrivelse av annotasjonen samt kriterier for syntaktiske valg som ble gjort under annotasjonen. Dependensgrafen i figur 1 viser en analyse fra trebanken. Vi ser at det finite verbet *byrja* fungerer som rot i setningen og denne har en subjektsnode (SUBJ) *systemet* som dependent og har også en adverbial dependent (ADV) der preposisjonen *i* er hode.

NDT har etter at den ble utgitt, dannet grunnlaget for trening av maskinlæringsbaserte tagger og parsere for norsk (Hohle et al. 2017; Vellidal et al. 2017). Den er også blitt konvertert til UD-standarden (Øvrelid og Hohle 2016) og via denne blitt brukt i flerspråklige verktøy som spaCy (<https://spacy.io/>) og Stanza (<https://stanfordnlp.github.io/stanza/>).



Figur 1: Dependensgraf fra NDT

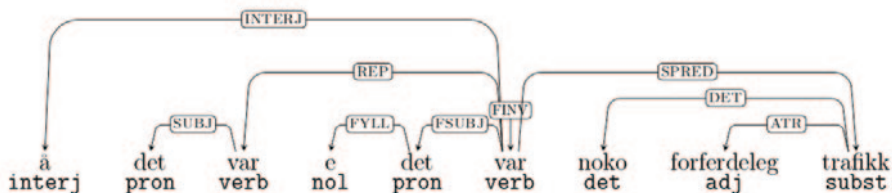
2.1.2 LIA-trebanken

I forbindelse med LIA-prosjektet har en delmengde av LIA-korpuset blitt manuelt annotert med syntaktisk analyse i form av dependensstrukturer. LIA-trebanken (Øvrelid et al. 2018) inneholder 3 651 ytringer og 48 715 løpende ord. Dataene består av transkribert tale (nynorsk), og annotasjonen følger NDTs retningslinjer med noen utvidelser for å gi en syntaktisk analyse av talespråksspesifikke feno-

mener som pauser (som # i (1)), nølelyder (som *e* i (2) og (3)) og reparasjoner (som *det var* i (2)).¹

- (1) ja # og kørde mjølka ut i byen igjen
- (2) å det var e det var noko forferdeleg trafikk
- (3) så det var mykje g- e mykje greier

Figur 2 viser dependensgrafen for ytringen i (2). Der ser vi syntaktiske funksjoner som SUBJ (subjekt) og SPRED (subjektspredikat) og også talespråksspesifikke funksjoner som REP (for reparasjoner) og FYLL (for nølelyder).



Figur 2: Dependensgraf fra LIA

3 Vår studie

I studien vår sammenligner to ulike språkssystemer: skrift og tale. Det er grunnleggende forskjeller mellom disse to språkssystemene som vi viser kommer til syne i de syntaktiske forskjellene vi finner. De to systemene brukes i forskjellige kommunikasjonssituasjoner. Talespråk brukes typisk til kommunikasjon mellom to eller flere som er til stede på samme tid, mens skriftspråk er kommunikasjon mellom sender og mottaker som typisk er adskilt både i tid og rom. Skriftspråk er også preget av monolog i større grad enn talespråk, som gjerne er dialogisk (i hvert fall er materialet i LIA av denne art).

¹ Johannessen og Jørgensen (2006) og Rosén (2008) diskuterer hvor hensiktsmessig det er å analysere slike fenomener på lik linje som syntaktiske forbindelser. I denne artikkelen forholder vi oss likevel til den syntaktiske analysen som foreligger for NDT.

Andre grunnleggende forskjeller er at tale formidles av lydsvingninger, og talestrømmen deles opp av pauser, rytme og tonegang. Ord og lyder glir over i hverandre, og det vil være avbrudd og gjentakelser og korrigeringer. Skriftspråk er grafiske tegn, og vi har tydelige skiller som markerer de ulike delene av teksten slik som markering av ord og setninger. I tillegg er det forskjell mellom de to ulike språk-systemene når det gjelder planlegging og redigering. Vi har mindre mulighet til å planlegge hva vi sier, og det går heller ikke an å gå tilbake for å forandre. I skrift derimot kan en planlegge i mye større grad og også gå tilbake å redigere og gjøre forandringer underveis.

I vår studie sammenligner vi nynorsk skriftspråk slik det framstår i NDT med det talemålet som finnes i LIA. Det er noen forbehold som må tas, spesielt med tanke på talemålsmaterialet i LIA. I LIA finnes det data fra mange ulike dialekter. Dialektopptakene er også gjort på ulike tidspunkt, så korpuset inneholder derfor dialekter fra ulike tidsperioder. Vi har ikke tatt hensyn til dette i vår studie. Ulike dialekter og tidsperioder har også blitt slått sammen.

I den kvantitative delen av studien har vi gjort flere ulike sammenligninger mellom de to trebankene. Fokuset i studien vår er syntaktiske fenomener, og vi benytter oss derfor i hovedsak av den syntaktiske annotasjonen i korpusene for å utføre sammenligningen. Men siden korpusene også er annotert for ordklasse, benytter vi oss av denne informasjonen siden det ofte er relevant i undersøkelsen av syntaktiske fenomener. Innledningsvis har vi sammenlignet ulike egenskaper ved de annoterte dataene, eksempelvis fordeling av dependensrelasjoner og ordklasser i trebankene, samt setningslengde. Dette diskuteres nærmere i henholdsvis seksjon 3.1 og 3.2. Vi har deretter sett på spesifikke syntaktiske konstruksjoner og gjort søk i trebanken som fanger opp disse. For å søke i trebankene benytter vi oss av skript skrevet i programmeringsspråket Python som leser trebankene inn i en intern datastruktur som gjør det lett å navigere grafene. Merk at de fleste søkene nok også ville kunne gjennomføres i eksisterende søkegrensesnitt som eksempelvis Iness. Når det gjelder de spesifikke syntaktiske fenomenene vi har valgt å fokusere på, har vi i stor grad tatt utgangspunkt i hypoteser og observasjoner presen-

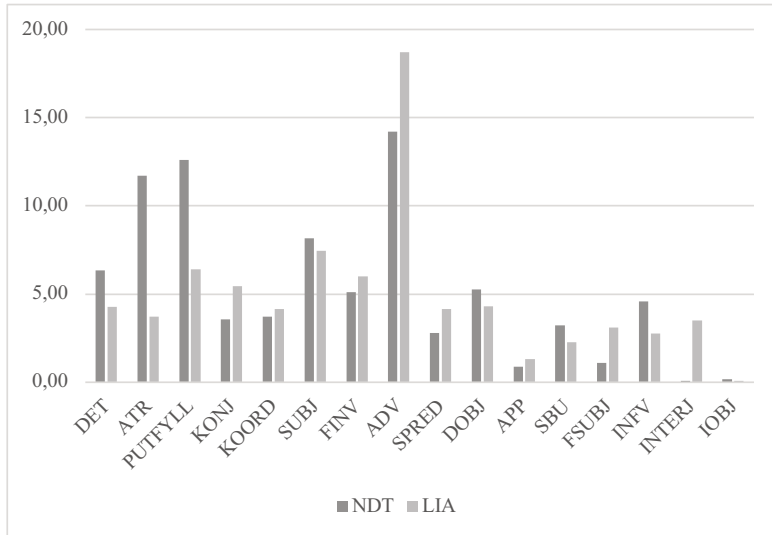
tert i tidligere forskning. Vi begrenser oss videre til nynorsk-delen av NDT fordi LIA-materialet også er transkribert til nynorsk skrift.

3.1 Oversikt: dependensrelasjoner og ordklasser

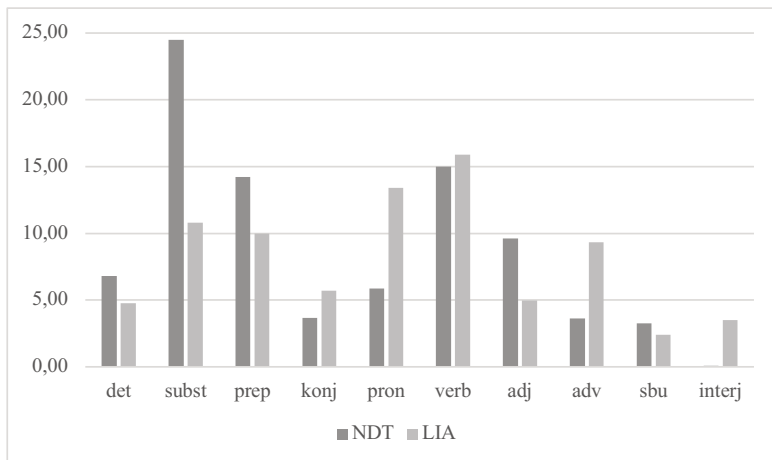
LIA-trebanken inneholder informasjon både om syntaktisk struktur, samt ordklasser, som vist i Figur 1 over. Hvert ord vil derfor være annotert med informasjon om ordklasse og dependensrelasjon, noe som gjør det enkelt å sammenligne distribusjonen av disse i NDT og LIA. Tabell 1 er en oversikt over de ulike dependensrelasjonene i korpusene. Figur 3 og 4 (neste side) viser den prosentvise fordelingen av ulike dependensrelasjoner og ordklasser i de to trebankene.

<i>Dependensrelasjon</i>	<i>Fullt navn</i>	<i>Beskrivelse</i>
ADV	adverbial	Adverbielle dependenter som modifierer verb, tidvis også adjektiver, determinativer, adverb osv., både obligatoriske og ikke-obligatoriske adverbialer.
ATR	attributt	Modifiserer substantiver, ofte adjektiver men også preposisjoner.
APP	apposisjon	Appositive dependenter med substantiver og egenavn.
DET	determinativ	Artikler, men også genitiver og kvantitetssubstantiver (<i>en del</i>).
DOBJ	direkte objekt	Komplementer til verb, både nominale, adjektiviske og verbale ledd (infinitiver, infinitte og finitte leddsetninger).
FINV	finitt verb	Finite hovedverb, utgjør som regel rot i dependensgrafene.
FSUBJ	formelt subjekt	Ikke-referensielle subjekter i subjeksposisjon i presenteringskonstruksjoner, upersonlig passiv og utbrytninger.
INJV	infinitt verb	Infinitte verb som opptrer med finitt hjelpeverb, infinitiver og partisipper.
INTERJ	interjeksjon	Kan stå som rot uten finitt verb (- <i>Ja, innimellom</i>).
IOBJ	indirekte objekt	Brukes kun for nominale elementer i setninger med direkte objekt.
KONJ	konjunksjon	Brukes for konjunksjoner (<i>og, eller</i>).
KOORD	koordinasjon	Relasjon for konjunker utover første konjunkt (som gis det koordinerte uttrykkets syntaktiske funksjon).
PUTFYLL	preposisjonsutfylling	Preposisjonsobjekt, nominal dependent på preposisjon.
SBU	subjunksjon	Brukes for subjunksjoner (<i>at, når</i>).
SPRED	subjektspredikativ	Predikativelement i kopula-konstruksjoner.
SUBJ	subjekt	Referensielle subjekter i subjeksposisjon.

Tabell 1: Oversikt over dependensrelasjonene i Figur 3 med beskrivelse. For ytterligere detaljer se Kinn et al. (2014).



Figur 3: Dependensrelasjoner. Prosentvis fordeling av dependensrelasjoner over totalt antall ord i trebanken, viser de 16 mest frekvente relasjonene.



Figur 4: Ordklasser. Prosentvis fordeling av ordklasser over totalt antall ord i trebanken, viser de 10 mest frekvente ordklassene

Som det kommer frem av figur 3 og 4, er det noen dependensrelasjoner og ordklasser hvor det er tydelige forskjeller mellom NDT og LIA, mens andre relasjoner og ordklasser har en lik fordeling i NDT og LIA som for eksempel fordelingen av subjekt og finitt verbal. Videre kommer vi til å legge vekt på de dependensrelasjonene og ordklassene der forskjellen i frekvens mellom dem forhåpentligvis kan si oss noe om generelle forskjeller mellom tale og skrift.

Figur 3 viser at den mest frekvente relasjonen i begge korpusene er ADV(erbial). Men det er en tydelig forskjell i frekvens i de to korpusene: den er mye mer frekvent i LIA enn i NDT. En kan se den samme tendensen når vi ser på frekvensen av ordklassen adverb i figur 4. Adverb har en høyere frekvens i LIA enn i NDT.

En annen tydelig forskjell mellom NDT og LIA finner vi i dependensrelasjonene PUTFYLL og ATR. PUTFYLL er komplementet i en preposisjonsfrase, mens ATR brukes på beskrivende dependenter på substantiver, gjerne da adjektiv. Begge disse dependensrelasjonene er mindre frekvente i LIA enn i NDT. Siden både PUTFYLL og ATR er komplementer i fraser, kan dette tyde på at talemålet har færre komplekse fraser enn skriftspråket. Papazian (1984, 142) og Kulbrandstad og Kinn (2016, 50) har foreslått at det er enklere syntaktisk konstruksjoner i talemålet, og våre data støtter dette.

Dependensrelasjonen ATR kan sees i sammenheng med ordklassen adjektiv, hvor vi ser at det er lavere frekvens av adjektiv i LIA enn i NDT. Dette kan tyde på at det er færre utfyllende beskrivelser i talemål. Det kan være en effekt av kommunikasjonssituasjonen der en ikke trenger å beskrive noe grundig hvis en får bekreftelse at samtalepartneren vet hva en snakker om.

I NDT har dependensrelasjonen INTERJ(eksjon) en svært lav frekvens (den relative frekvensen for interjeksjoner er 0,07 %), mens det er en hel del av disse i LIA (den relative frekvensen i LIA er 3,5 %). Dette er heller ikke overraskende siden talemålssituasjonen innbyr til bruk av interjeksjoner i større grad enn i skriftspråksituasjoner. Papazian (1984, 184) påpeker at tale gjerne er mer livlig, impulsiv og emosjonell enn skrift og derfor mer preget av interjeksjoner, slik våre data også tyder på.

Noe overraskende er det kanskje at figur 3 viser at det er flere forekomster av SUBJ(ekter) enn av FINV (finitte verb). En skulle tro at forekomsten av disse to skulle være tilnærmet like siden norsk har et subjektskrav (se seksjon 3.4 for en diskusjon om utelatelse av setningsledd). Men forholdet mellom antall SUBJ og FINV er primært knyttet til de valgene som har blitt tatt i hvordan korpusene blir annotert. Det finite verbet får relasjonen FINV i helsetninger, mens det finite verbet i finite leddsetninger angir relasjonen leddsetningen har til verbet i oversetningen, som for eksempel DOBJ eller ADV. Subjekter derimot blir annotert både i leddsetninger og helsetninger som SUBJ. Dette fører til at alle subjekter i korpuset er SUBJ, mens ikke alle finite verb er FINV.

Vi kan også finne forskjeller mellom tale og skrift ved å se på ulike ordklasserfrekvenser i NDT og LIA i figur 4. Den største ordklassen, substantiv, viser en klar forskjell i de to korpusene hvor substantiv er mer frekvent i NDT (24 %) enn i LIA (10 %). Denne ordklassen burde sees i sammenheng med en annen ordklasse, nemlig pronomen. Det er mer bruk av pronomen i LIA (13 %) enn i NDT (5 %). Kulbrandstad og Kinn (2016, 50) påpeker at det er mer bruk av deiktiske uttrykk, for eksempel pronomen, i talemål enn i skriftspråk. Dette er ikke overraskende med tanke på at talesituasjonen er preget av en her og nå-situasjon i større grad enn skrivesituasjonen, og dette legger til rette for utstrakt bruk av deiktiske uttrykk i talemålet.

Søkene i korpuset gir oss også mulighet til å kombinere informasjonen vi har om dependensrelasjoner og ordklasser. Dermed kan vi få vite mer om hva slags form setningsleddene har. Vi kan blant annet finne ut hva slags type subjekter vi finner i de to ulike korpusene. Ved å se på typen subjekt kan vi undersøke den tendensen at et subjekt gjerne vil være kjent informasjon. Et typisk subjekt som har kjent informasjon vil ha formen pronomen. Som vi kan se av tabell 2, er det en sterk tendens til at subjektet er pronominalt i LIA der bare ca. 10 % av subjektene har en annen ordklasse enn pronomen. Disse funnene blir støttet av Søfteland (2013, 142), som har undersøkt ulike typer subjekt i spontantale fra fire samtaler i Nordisk dialektkorpus og fant at 83 % subjektene var pronomielle. I NDT derimot finner

vi ikke tilsvarende forhold mellom subjekt og pronomen. I skriftspråkskorpuset er det en overvekt av substantiver som subjekt.

	<i>LIA</i> (N=3622)	<i>NDT</i> (N=24714)
<i>Pronomen</i>	88,9 %	39,7 %
<i>Substantiver</i>	9,7 %	56,5 %
<i>Adjektiver</i>	0,4 %	2,0 %
<i>Determinativer</i>	0,4 %	1,0 %

Tabell 2: Fordelingen av subjekter

3.2 Setningslengde

Både Papazian (1984, 142) og Kulbrandstad og Kinn (2016, 50) hevder at talemålet typisk har kortere ytringer enn skriftspråket. Hverken Papazian (1984) eller Kulbrandstad og Kinn (2016) har gjort kvantitative undersøkelser av setningslengde i skriftspråket sammenlignet med talemålet. Kvantitative undersøkelser ble derimot gjort i Talemålsundersøkelsen i Oslo (TAUS). I TAUS-materialet ble setningslengden i opptak av uformelle intervjuer med folk fra Oslo fra 1971–73, undersøkt. Der fant Hanssen et al. (1978, 138) at den gjennomsnittlige setningslengden var 10 ord.

For å undersøke setningslengden i de to korpusene har vi måttet bruke ulike definisjoner av hva en setning er. I *NDT* har vi definert en setning som en enhet som blir markert med stor bokstav og punktum selv om dette avviker fra den grammatiske definisjonen av setning i Faarlund et al. (1997, 39) der en setning blir definert som en forbindelse av et nominal som er setningens subjekt, og en verbfrase som må inneholde et finitt verb. En konsekvens av vår definisjon av setning er at ikke alt som er markert med stor bokstav og punktum i skriftlige tekster oppfyller den grammatiske definisjonen av en setning.

I talemål er det notorisk vanskelig å definere enheten “setning”. For å komme frem til en definisjon av setning har vi basert oss på transkripsjonen som er gjort av lydfilene. Her er talestrømmen delt

opp i segmenter. I transkripsjonsveiledning til LIA står det at segmentene skal være meningsfulle enheter som skal ligne mest mulig på setninger (Hagen et al. 2018, 14). Derfor har vår definisjon av en setning i LIA vært at et segment er lik en setning. Et segment vil derfor kunne inneholde avbrudd, nølelyder og reparasjoner. Videre i artikkelen bruker vi *setning* om denne enheten i talemål.

Vi finner at den gjennomsnittlige setningslengden i NDT er 17,24 ord, mens den i LIA er 13,34 ord. Med dette kan vi bekrefte antakelsen i Papazian (1984) og Kulbrandstad og Kinn (2016) om at det er kortere setninger i talemålet enn i skriftspråket. Likevel er våre resultater fra NDT og LIA ganske usikre siden vi har sammenlignet to enheter som ikke nødvendigvis er like siden vi bruker forskjellige definisjoner på setning i de to korpusene.

3.3 Sideordning og underordning

Papazian (1984, 142) hevder at det er mer lineær sideordning i talemålet sammenlignet med skriftspråket, og at det er mer underordning i skriftspråket enn i talemålet blant annet ved mer utstrakt bruk av leddsetninger. I følge Papazian fører mer sideordning i talespråket til blant annet en enklere syntaks og at denne enklere syntaksen også henger sammen med de kortere setningene som vi allerede har diskutert i forrige del.

For å undersøke sideordning og underordning på setningsnivå sammenligner vi frekvensen av dependensrelasjonene KONJ(unksjon) og SBU(subjunksjon). Forventningen om mer bruk av leddsetninger i skriftspråket ble bekreftet ved at vi finner en høyere frekvens av relasjonen SBU i NDT enn i LIA. Tilsvarende blir antakelsen om at det er mer lineær sideordning i skrift enn i tale, bekreftet siden vi finner at det er høyere frekvens av relasjonen KONJ i LIA enn i NDT.

3.4 Utelating av syntaktiske ledd

Utelating av syntaktiske ledd som for eksempel subjekt og objekt er typisk for talemål (Papazian 1984, 143; Stjernholm 2008; Nygård 2018). Som eksemplene under viser, kan en i norsk talemål utelate

ulike syntaktiske ledd. I (4) har et referensielt subjekt blitt utelatt, mens i (5) har et formelt subjekt blitt utelatt.

- (4) Gjekk ein tur til noko trytetjørner som var litt ifrå for å fiske (LIA)
- (5) Var langt t- å gå av garde i kolmørke om morgonen da (LIA)

For å undersøke dette begrenser vi oss til å undersøke setninger der subjektet er utelatt. I norsk er dette sammen med det finitte verbalet det eneste andre obligatoriske setningsleddet (med visse forbehold). Korpuset er ikke tagget for valens eller lignende. Derfor er det ingen mulighet til å for eksempel søke etter setning som skulle ha hatt et objekt, men hvor objektet er utelatt.

Ved å søke etter mønster der roten er et finitt verb uten subjekts-dependent og verbet ikke er i imperativ, fant vi at det var forskjell mellom de to korpusene. I NDT var frekvensen av setninger uten et subjekt 1,9 %, mens det i LIA er 6,7 %. Tidligere studier slik som Stjernholm (2008) og Nygård (2018) har ikke undersøkt frekvensen av utelatte subjekt i talemål selv om de begge kommenterer at dette er vanlig. En kunne kanskje ha forventet en enda høyere frekvens av utelatte subjekt i LIA. Uansett støtter funnene våre tidligere forskning som viser at utelattelse av subjekt er mer typisk for talemål enn for skriftspråk.

Til å være et skriftspråkskorpus er det overraskende at NDT har såpass mange setninger uten subjekt, siden en setning i skriftspråket blir definert som en forbindelse av et nominal som er setningens subjekt og en verbfrase som må inneholde et finitt verb (Faarlund et al. 1997, 39). Men her spiller tekstutvalget en rolle. Siden den største delen av tekstene i NDT er avistekster, vil avisoverskrifter stå for en del av setningene uten subjekt. Eksempel (6) viser dette:

- (6) Veks forbi Danmark (NDT)

I tillegg inneholder NDT bloggtekster. Disse tekstene er gjerne mer uformelle og talemålspreget, og vi finner setninger uten subjekter her slik som i (7). I tillegg kan vi ikke utelukke setninger av typen som

vises i (8), som kanskje heller viser bruk av stor bokstav og punktum selv om det ikke er en helsetning enn muligheter for utelatelse av subjekt i skriftspråket.

- (7) Men ok, skal ikkje gå i detalj på dette her no. (NDT)
- (8) Og vart tekne imot så hjarteleg at det sit i kroppen enno. (NDT)

3.5 Subjektdublering

Det er ikke bare mulig i norsk å utelate visse setningsledd, det er også mulig med dublering av setningsledd. Som oftest er det subjektet som er dubleret, men det kan også være andre ledd. Det er vanligvis slik at siste leddet (gjerne kalt kopien eller det høyredislokerte leddet) er et pronomen slik som eksempel (9) viser.

- (9) **Han** hadde sjølve livet i hendene **han** (LIA)

Borthen (2018, 404) hevder at pronominal høyredislokering (kopien er et pronomen), er et vanlig fenomen i muntlig tale i norsk, men hun har selv bare eksempler fra skriftspråk. Undersøkelsen av LIA og NDT kan bekrefte Borthens antakelse. Ved å søke etter mønster der SUBJ(jektet) har en pronominal dependent med relasjonen APP(osisjon), fant vi at frekvensen av subjektdublering av typen pronominal høyredislokering i LIA i alle setninger er 2,7 % ($N=102$), mens i NDT er frekvensen 0,1 % ($N=35$).

- (10) Men **eg** var jo glad til **eg** for da fekk eg reise ein tur til Gjøvik (LIA)
- (11) **Det** var mykje pengar **det** (LIA)
- (12) Men **vi** var nøgde med den førre finansministeren, **vi** (NDT)
- (13) For å forklåre motivasjonen for arbeidet, peikar han mellom anna på det at **den gamle typen bibellesarar** er borte; **dei som drøfta tolkingar, usemjer og uklåre tekststader nøye**. (NDT)

Hvis vi sammenligner eksemplene fra LIA og NDT, ser vi at de to eksemplene fra LIA følger den prototypiske subjekt-dubleringen som Borthen (2018) undersøker. Slike eksempler finnes også i NDT slik som (12) viser, men det er typen som vi ser i (13) som er den vanligste i NDT, altså hvor det høyredislokerte leddet er en spesifisering av subjektet mer enn en dublering eller kopiering. Det er nok en annen type dublering enn den som er frekvent i talemålet.

3.6 Indirekte objekt

I norsk har vi mulighet for å uttrykke de semantiske rollene benefaktiv og mottaker (blant annet) på to ulike måter enten med et nominalt indirekte objekt som i (14) eller som en preposisjonsfrase som i (15):

(14) Hun ga **hesten** en gulrot.

(15) Hun ga en gulrot **til hesten**.

Allerede Western (1921, 140–141) påpekte at det er en forskjell mellom talespråk og skriftspråk i dette henseende: Nominalt indirekte objekt finnes ofte i «det litterære sprog hvor talesproget foretrekker et komplement». Med komplement mener Western her en preposisjonsfrase slik som i (15).

Ved å se på den relative frekvensen for dependensrelasjonen indirekte objekt finner vi at NDT har en frekvens på 1,7 % ($N=529$), mens LIA har en frekvens på 0,6 % ($N=34$). Det er viktig å påpeke at i annotasjonen av korpusene har bare nominale ledd blitt annotert som indirekte objekt. Preposisjonsfraser slik som i (15) får dependensrelasjonen ADV(erbial).

Faarlund et al. (1997, 727) påpeker at valget mellom nominalt indirekte objekt eller preposisjonsfrase er styrt av pragmatiske og referensielle forhold. Et nominalt indirekte objekt vil gjerne ha unik referanse og representere kjent informasjon. Dette kan vi operasjonalisere for å enkelt undersøke om dette stemmer i NDT og LIA. Vi tar utgangspunkt i at pronomen er både kjent informasjon og også har unik referanse, mens substantiv ikke nødvendigvis vil være kjent informasjon i like stor grad. I LIA finner vi at nesten alle nominale

indirekte objekt er pronomen, (94,1 %), mens resten er substantiver. Det må påpekes at det er få nominale indirekte objekt i LIA, bare 35 tilfeller. I NDT derimot har vi langt flere nominale indirekte objekt ($N=529$), og det ser ikke ut som tendensen til at nominale indirekte objekt typisk er pronomen her er like sterk som i talespråkskorpuset. 57,8 % av de nominale indirekte objektene er pronomen, mens 38 % er substantiv. Forskjellen mellom LIA og NDT i frekvensen av pronomen og substantiver som vi finner i de nominale indirekte objektene, kan sannsynligvis knyttes til de generelle kjennetrekken ved talemål, som gjør at vi også typisk finner flere pronomielle subjekt i LIA enn i NDT, som vi diskuterte i 3.1.

3.7 Rekkefølgen finitt verb–negasjon i leddsetninger

I norsk kan man finne variasjon i rekkefølge på finitt verb og negasjon i leddsetninger. Eksemplet i (16) viser den vanlige rekkefølgen hvor negasjonen *ikke* kommer før det finitte verbet, mens eksempelet i (17) viser motsatt rekkefølge: Negasjonen kommer etter det finitte verbet.

(16) Da mente han [at han ikke kunne være gift] (Neg-V)

(17) Da mente han [at han kunne ikke være gift]
(V-Neg)

Denne variasjonen i rekkefølgen på verb og negasjon i leddsetninger har blitt grundig diskutert blant annet i Julien (2007, 2009) og Wiklund et al. (2009).

For å undersøke om det er noen forskjell mellom disse to leddstillingsmønstrene i skrift og tale søkte vi etter mønstre i trebankene der finitt verb ikke er en rotnode, men likevel har både en subjunksjon og *ikke* som dependenter.

Tabell 3 viser at rekkefølgen hvor *ikke* kommer før det finitte verbet, er helt klart mest frekvent i både NDT og LIA. Dette er ikke overraskende fordi dette er default ordstilling i leddsetninger, mens rekkefølgen hvor negasjonen kommer etter det finitte verbet, har visse restriksjoner. Som vi kan se i tabellen, forekommer denne rekkefølgen nesten ikke i NDT. Overraskende nok har en fjerdedel av

leddsetningene ordstillingen der *ikke* kommer etter det finitte verbet i LIA. Dessverre er ikke LIA-korpuset stort nok til å finne veldig mange eksempler, men vi konkluderer likevel med at det er en klar tendens til at talemålet har flere tilfeller av rekkefølgen V-neg enn i skrift. Våre resultater støttes av funnene i Ringstad (2019). Hun har undersøkt andre talemålskorpus som Scandiasyn, Nota og Big Brother og finner den samme fordelingen mellom rekkefølgen Neg-V (66 %) og V-neg (34 %) som i LIA.

	NDT (N=457)	LIA (N=43)
<i>Neg-V</i>	98,9 %	72,1 %
<i>V-neg</i>	1,1 %	27,9 %

Tabell 3: Rekkefølgen av det finitte verbet og ikke i leddsetninger

3.8 Så-konstruksjon

I norsk kan en ha setninger som (18):

- (18) Da vi hadde gått i fleire timar, **så** kom vi fram til ei hytte (Faarlund et al. 1997, 817)

Faarlund et al. (1997, 817) beskriver denne konstruksjonen med at *så* står i forfeltet etter et ledd (*Da vi hadde gått i fleire timar*) i ekstraposisjon i et løst forfelt. Det vanligste er at det ekstraponerete leddet er alle typer fritt adverbial, ikke bare tidsadverbial slik som i eksempelet. Faarlund et al. (1997, 817) påpeker også at *så* i skrift er sjelden når adverbialet foran bare er en frase, og er mer vanlig i bruk i skrift når det er etter en leddsetning. Salvesen (2017) møtte mye motstand hos informantene som skulle vurdere skriftlige setninger med denne bruken av *så*. Informantene kommenterte ofte at setningene var bedre uten *så* og at setningen ble barnslige med *så*. Dette kan tyde på at en vurderer konstruksjonen som noe som ikke hører hjemme i skrift, men kanskje heller tilhører talemålet.

For å finne denne konstruksjonen i korpusene søkte vi etter mønstre der *så* er dependent på rot verbet og har en apposisjons-dependent. Vi finner at av alle setninger i LIA har 2,5 % av dem *så*-konstruksjo-

nen, mens bare 0,2 % av alle setningene i NDT har denne konstruksjonen.

Våre resultater støtter påstanden om at *så*-konstruksjonen forbindes med talemålet, og at den er mindre frekvent i skrift enn i tale. Vi finner at *så* kan stå etter flere ulike typer ekstraponerte ledd i LIA enn det gjør i NDT. I eksemplene (19)–(22) er det ekstraponerte leddet en adverbial leddsetning, en adverbial preposisjonsfrase og et enkelt adverb. Likevel fant vi *så* etter fraser i NDT som Faarlund et al. (1997, 817) påpeker er sjeldent i skrift slik som i (22) der det ekstraponerte leddet er en preposisjonsfrase.

- (19) Når vi hadde fått krøttera oppover til til ved Auenhaugen **så** kunne vi leike på Auenhagen nesten heile tida (LIA)
- (20) I rettig gammel tid **så** var det nå byggmjøl (LIA)
- (21) Seinare **så** trudde vel kanskje ikkje så mykje på det. (LIA)
- (22) For oss sunnmøringar **så** liknar det mest ein formastelse å servere ball som noko søtt (NDT)

4 Konklusjon

Denne studien har vi presentert en sammenlignende studie av et utvalg av syntaktiske fenomener i talemål og skriftspråk. Vi har vist at søk i syntaktiske strukturer lar oss gjøre denne studien på en presis måte gjennom kvantitative sammenligninger. Den syntaktiske annotasjonen kan fortelle oss en god del om den grammatiske strukturen i språket vårt, og dette gir gode muligheter for videre studier av både skriftspråket og talemålet. Annotasjonen av korpus gir ikke bare muligheter til å gjøre kvantitative studier slik som denne, men en kan også gjøre mer kvalitative studier ved at det gjennom annotasjonen er mulig å finne sjeldne konstruksjoner i språket. Denne typen annotasjon lar oss også bekrefte eller avkrefte tidligere antakelser som ofte ikke var basert på faktisk språkbruk.

Det er visse utfordringer ved å bruke syntaktisk annotasjon. Det er foreløpig en teknisk barriere som det kan være vanskelig å komme over. Et ønske for fremtiden vil være at det kommer på plass en tek-

nisk løsning som gjør det mer brukervennlig å gjøre søk i de syntaktiske annotasjonene. Dessuten er også materialet foreløpig ganske lite. Likevel er de to trebankene, og spesielt LIA-trebanken med det unike dialektmaterialet, en verdifull ressurs for forskningen på norsk språk.

Referanser

- Borthen, Kaja. 2018. Pronominal høyredislokering i norsk, det er et interessant fenomen, det. *Norsk lingvistisk tidsskrift* 36, 403–450.
- Faarlund, Jan Terje, Svein Lie og Kjell Ivar Vannebo. 1997. *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Hagen, Kristin, Live Håberg, Eirik Olsen og Åshild Søfteland. 2018. Transkripsjonsretteiing for LIA. http://www.tekstlab.uio.no/LIA/pdf/transkripsjonsretteiing_lia.pdf.
- Hanssen, Eskil, Thomas Hoel, Ernst Håkon Jahr, Olaug Rekdal og Geirr Wiggen. 1978. *Oslomål. Prosjektbeskrivelse og syntaktisk analyse av oslomål med henblikk på sosiale skilnader*. Oslo: Novus.
- Johannessen, Janne Bondi og Fredrik Jørgensen. 2006. Annotating and parsing spoken language. I *Treebanking for Discourse and Speech: Proceedings from NODALIDA 2005 Special Session on Treebanks for Spoken language and Discourse*, redigert av Peter Juel Henriksen og Peter Rossen Skadhauge, 83–104. København: Samfundslitteratur.
- Julien, Marit. 2007. Embedded V2 in Norwegian and Swedish. *Working Papers in Scandinavian Syntax* 80, 103–161.
- Julien, Marit. 2009. The force of the argument. *Working Papers in Scandinavian Syntax* 84, 225–232.
- Kinn, Kari, Per Erik Solberg og Pål Kristian Eriksen. 2014. *NDT Guidelines for Morphological Annotation*. Oslo: Språkbanken, Nasjonalbiblioteket.
- Kulbrandstad, Lars Anders og Torodd. Kinn. 2016. *Språkets mønstre*. Oslo: Universitetsforlaget.

- Marcus, Mitchell, Beatrice Santorini og Mary Ann Marcinkiewicz. 1993. Building a large annotated corpus of English: The Penn Treebank. *Computational Linguistics* 19 (2), 313–330
- Mel'čuk, Igor Aleksandrovič. 1988. *Dependency syntax: theory and practice*. Albany, NY: SUNY press.
- Nivre, Joakim, Marie-Catherine de Marneffe, Filip Ginter, Yoav Goldberg, Yoav, Jan Hajič, Jan, Christopher Manning, Ryan McDonald, Ryan, Slav Petrov, Sampo Pyysalo, Natalia Silveira, Reut Tsarfaty og Daniel Zeman.. 2016. Universal dependencies v1: A multilingual treebank collection. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)* Portoroz, Slovenia: European Language Resources Association (ELRA).
- Nygård, Mari. 2018. *Norwegian Discourse Ellipsis. Clausal architecture and licensing conditions*. Amsterdam: John Benjamins.
- Papazian, Eric. 1984. *Talemål og skriftspråk. Forskjellen og samsvaret*. Skriftserie nr. 16, serie B. Bergen: Institutt for fonetikk og lingvistikk, Universitetet i Bergen.
- Salvesen, Christine Meklenborg. 2017. I norsk så bruker vi så. Foredrag, MONS 17, Bergen
- Solberg, Per Erik, Arne Skjærholt, Lilja Øvrelid, Kristin Hagen og Janne Bondi Johannessen. 2014. The Norwegian Dependency Treebank. *Proceedings the Ninth International Conference on Language Resources and Evaluation (LREC'14)*, Reykjavik: European Language Resources Association (ELRA).
- Stjernholm, Karine. 2008. Subjektsellipser: fins pro i norsk talespråk? Masteroppgave, Universitetet i Oslo.
- Ringstad, Tina Louise. 2019. Distribution and function of embedded V–Neg in Norwegian: A corpus study. *Nordic Journal of Linguistics* 42 (3), 1–35.
- Rosén, Victoria. 2008. Mot en trebank for talespråk. I *Språk i Oslo. Ny forskning omkring talespråk*, redigert av Janne Bondi Johannessen og Kristin Hagen, 214–225. Oslo: Novus
- Teleman, Ulf. 1974. *Manual för grammatisk beskrivning av talad och skriven svenska*. Lundastudier i språkvetenskap. Lund: Studentlitteratur.

- Tesnière, Lucien. 1965. *Éléments de syntaxe structurale*. Paris: Librairie C. Klincksieck.
- Velldal, Erik, Lilja Øvrelid, og Petter Hohle. 2017. Joint UD parsing of Norwegian bokmål and nynorsk. *Proceedings of the Nordic Conference on Computational Linguistics (NoDaLiDa'17)*. Linköping: Linköping University Electronic Press.
- Western, August. 1921. *Norsk riksmåls-grammatikk for studerende og lærere*. Kristiania: H. Aschehoug & Co.
- Wiklund, Anna-Lena, Kristine Bentzen, Gunnar Hrafn Hrafnbjargarson og Thorbjörg Hróarsdóttir. 2009. On the distribution and illocution of V2 in Scandinavian that-clauses. *Lingua* 119 (12), 1914–1938.
- Øvrelid, Lilja og Petter Hohle. 2016. Universal dependencies for Norwegian. *Proceedings of the Tenth International Conference on Language Resources and Evaluation (LREC'16)*. Portoroz, Slovenia: European Language Resources Association (ELRA).
- Øvrelid, Lilja, Andre Kåsen, Kristin Hagen, Anders Nøklestad, Per Erik Solberg og Janne Bondi Johannessen. 2018. The LIA treebank of spoken Norwegian dialects. *Proceedings of the Eleventh International Conference on Language Resources and Evaluation (LREC'18)* Miyazaki, Japan: European Language Resources Association (ELRA).

English summary

This article presents a comparative study of Norwegian spoken and written language that addresses a number of syntactic differences between written and spoken language that have been proposed in previous studies. The newly developed LIA treebank contains transcribed spoken language and is annotated in accordance with the written language Norwegian Dependency treebank, with certain additions to cover spoken language-specific phenomena. The syntactic analysis in these two treebanks facilitate precise syntactic searches that allow us to investigate a number of syntactic phenomena, such as co-ordination vs. subordination, omission of constituents and word

order, and compare these phenomena in writing and speaking in Norwegian.

Signe Laake
Institutt for grunnskole- og faglærerutdanning
Oslomet – storbyuniversitetet
signe.laake@oslomet.no

Lilja Øvrelid
Institutt for informatikk
Universitetet i Oslo
liljao@ifi.uio.no



LIA-korpuset som ressurs i revisjonen av tre ordbøker

Bente Selback og Terje Svardal

I denne artikkelen ser vi nærmare på kva ressurs eit talemålskorpus som LIA norsk kan vere i arbeidet med revisjonen av tre ordbøker: standard-ordbøkene *Bokmålsordboka* og *Nynorskordboka* og dokumentasjonsordboka *Norsk Ordbok*. I revisjonen arbeider ein både med å finne fram til nye lemma som bør takast inn i ordbøkene, og ein vurderer dei ordbokartiklane som er med frå før. Redaktørane må mellom anna ta stilling til om eit ord er så sentralt at det bør vere med vidare, om definisjonane er i tråd med gjengs språkbruk, kva bruksdøme som vil fungere best, og kva som er dei viktigaste kollokasjonane. Det er naudsynt å ha gode kjelder, og i artikkelen vurderer vi kva ein kan og ikkje kan bruke LIA-korpuset til i arbeidet.

Stikkord: leksikografi, ordbokrevisjon, skriftspråk, talemål, korpus

1 Innleiing

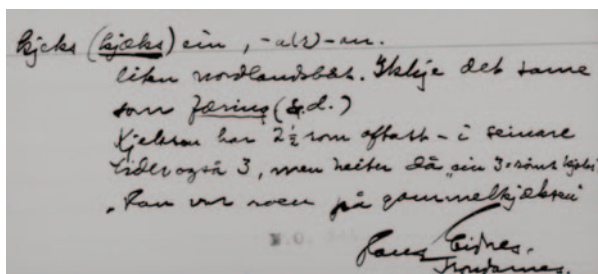
Kvardagen til ein leksikograf består av ei endelaus rekkje med spørsmål: Er dette lemmaet frekvent nok til å få ein plass i ordboka? Korleis kan ein lage ein god definisjon? Kva døme skal ein velje for illustrere bruken av ordet på ein forståeleg og representativ måte? Kva kollokasjonar er dei vanlegaste? Er ordet del av eitt eller fleire faste uttrykk? For å finne svar på desse og mange andre spørsmål treng ein gode og varierte kjelder som kan vise korleis orda blir brukte i både skrift og tale. Kva kjelder som eignar seg best, kjem sjølv sagt an på kva spørsmål ein stiller, og kva ordbok ein arbeider med. I denne artikkelen vil vi ta utgangspunkt i revisjonen som er i gang av standardordbøkene *Bokmålsordboka* og *Nynorskordboka* og av dokumentasjonsordboka *Norsk Ordbok*. Kan eit talemålskorpus som LIA norsk brukast i dette arbeidet, og kva kan det i så fall brukast til? Sidan det er snakk om to ulike ordboktypar, skal vi sjå på standardordbøkene og *Norsk Ordbok* kvar for seg og vurdere kva nytte ein kan ha av LIA i revisjonen av dei eksisterande artiklane og for å finne fram til nytt ordtilfang. Før vi ser nærmare på desse spørsmåla, vil vi kort vise kva overgangen til digitale materialsamlingar har hatt å seie for ordbokarbeidet, og presentere dei tre ordbøkene og revisjonen.

2 Leksikografi før og no – frå setelarkiv til elektroniske korpus

Mykje har endra seg sidan arbeidet med *Norsk Ordbok* og seinare standardordbøkene tok til. Tidlegare måtte leksikografane gå gjennom ordsetlane i arkivet for å finne svar på ulike spørsmål. Var ein til dømes ute etter kva ulike tydingar ordet *kjeks* kunne ha, måtte ein hente fram den aktuelle arkivskuffa, lese gjennom dei ofte handskrivne setlane og sortere dei i bunkar etter tydingane. I dag har ein god hjelp av elektroniske korpus som inneheld store mengder tekst, og der ein har avanserte søkjemoglegheiter som ofte gjer det råd å raskt finne dei opplysningane ein leiter etter.

2.1 Setelarkivet

Setelarkivet til *Norsk Ordbok* (NO-arkivet) omfattar kring 3,4 millionar setlar med kring 500 000 unike lemma (Vikør 2018, 32).¹ Arkivet vart i hovudsak bygd opp frå 1930 og var i jamn vekst til midt på 90-talet. Frå 2002 har det vorte produsert elektroniske setlar. Materialet i NO-arkivet er samla inn dels av ordbokredaksjonen og dels av frivillige medarbeidarar. Dermed er det nokså ueinsarta. Om lag 600 frivillige medarbeidarar landet rundt bidrog, og dei har levert alt frå éin til kring 65 000 setlar. Variasjonen i bakgrunn og interesser har vore relativt stor, noko materialet ber preg av. Det varierer desutan kor mykje tilfang det er frå dei ulike målføra. Eit område som enno er lite dekt, er kyststroka av Agder.



Figur 1: Ein av dei 92 arkivsetlane med ordet *kjeks*

Det finst to hovudgrupper av setlar, dei med litterære ekserpt og dei med bruksdøme frå målføra. Om lag ein halv million av setlane inneheld opplysningar frå målføra om ordformer, bøyning, tydingar og døme på bruk (Vikør 2018, 32). Arkivet finst både som tradisjonelle arkivsetlar, som er skanna og gjort tilgjengelege av Dokumentasjonsprosjektet på 1990-talet (Kristiansen og Ore 1998), og som nyare elektroniske setlar.

¹ Setelarkivet til *Norsk Ordbok*: <http://usd.uib.no/perl/search/search.cgi?tabid=436&appid=8>

2.2 Elektroniske skriftspråkskorpus

Sjølv om ordsetlane framleis er i bruk, nyttar vi i dag primært elektroniske korpus i det leksikografiske arbeidet. Det finst mange ulike skriftspråkskorpus som femner om ulike typar tekst. Særleg korpus som inneheld tekstar som har vore gjennom ein redaksjonell prosess, er interessante for oss. Aller mest bruker vi korpushandteringssystemet Korpuskel (Meurer 2012), som er vidareutvikla for ordbokrevisjonen gjennom Korpuskel-Lex.² Korpuskel-Lex har avanserte søkjemoglegheiter og gjev oss høve til å søkje på tvers av ei samling ulike korpus, mellom anna Norsk aviskorpus og Nynorskkorpuset (Lyse 2020). Til saman inneheld skriftspråkskorpusa i Korpuskel-Lex om lag 2,7 milliardar ord. Det gjev oss betre moglegheiter enn tidlegare til å finne fram til den representative språkbruken. Samstundes kan materialet bli overveldande. Eit trunkert søk på ord som startar med *kjeks*, gjev til dømes 7783 treff, noko som gjer det vanskeleg å få oversikt. Heldigvis er det mogleg å snevre inn søket slik at ein får eit meir handterleg materiale.

nnk	↑ sett fram saft og kjeks, og eg har smakt på	kjeksen	, mest for å vera høflig. – Sjå her! Her er me	2009
nnk	verdas dyraste kjeks gav meg, den påskribla	kjeksen	frå Sir Ernest Shackletons mislykka ekspedis	2008
nnk	rødde istappar, og så midt opp i dette denne	kjeksen	, så komplett absurd, og no kjem alle dei stu	2008
nnk	, fersk kokfisk, meinte eg. Og vesleguten tek	kjeksen	, eller ka veslebaaten no heiter, og dreg utp:	1911
nnk	det er mange ting å finne rom åt i den vesle	kjeksen	eller spissbåten. Båtene må lastast rett, så h	2005
nnk	med kaffikrem, og i valnott- og kokoskaker.	Kjeksen	er produserte av eit japansk selskap, medan	2008

Figur 2: Utdrag frå konkordans i Korpuskel-Lex som viser treff på *kjeks*.* i Nynorskkorpuset

2.3 Elektroniske talemålskorpus

Dei fleste og største korpusa inneheld skriftspråk, men det finst òg fleire tilgjengelege korpus med transkribert talemålsmateriale. Her vil vi berre nemne to av dei vi har brukt mest i revisjonsarbeidet, Nordisk dialektkorpus og Talebanken.³ Dei eldre delane av Nordisk dialektkorpus er flytta over i LIA-korpuset, men den norske delen

² Korpuskel-Lex: <https://clarino.uib.no/lex/>

³ Nordisk dialektkorpus: <https://tekstlab.uio.no/glossa2/ndc2>

Talebanken: <https://clarino.uib.no/korpuskel/corpus-list?collection=Talebanken>

består framleis av opptak frå 111 stader rundt i landet. Opptaka er gjorde i perioden 1998–2015, og dei er transkriberte i både ein tale-målsnær variant og til bokmål. Korpuset er på om lag 2 millionar ord. Talebanken inneheld dei tre korpusa *Dialektendring*, *Industri-stad* og *Talesøk*. Han har både eldre og nyare opptak, primært frå Vestlandet. Opptaka er transkriberte til nynorsk. Talebanken inneheld om lag 7,6 millionar ord.

Sjølv om det alt finst fleire talemålskorpus, har vi ikkje like gode korpusressursar for tale som for skrift. Eit talemålskorpus kan seie oss kva ord og former som er brukte i målføra rundt om i landet, noko som er nyttig for *Norsk Ordbok*. Men det kan òg fortelje oss noko generelt om kva som kjenneteiknar talemål i motsetnad til skriftspråk, og slik vere eit godt supplement til skriftspråkskorpusa. Både LIA-korpuset, Talebanken og Nordisk dialektkorpus inneheld munnleg språkbruk i form av dialogar. Vi kan tenkje oss at det vil vere ei god kjelde til ordtilfang som er typisk for samtale, som preposisjonar, interjeksjonar, pragmatiske partiklar og så vidare, og at det difor særleg kan vere til hjelp i revisjonen av denne typen ord.

3 Ordbøkene og revisjonen

I 2016 tok Universitetet i Bergen over ansvaret for Språksamlingane frå Universitetet i Oslo. Samlingane inneheld store mengder fysisk og elektronisk materiale innanfor fagfelt som namnegranskning, norrønt, dialektologi og leksikografi. Saman med det leksikografiske materialet fekk ein ansvaret for *Norsk Ordbok*, *Bokmålsordboka* og *Nynorskordboka*. Ordbøkene er ikkje ferdige ein gong for alle. Verda og språket endrar seg som kjent, og då er det nødvendig å revidere ordbøkene òg. Difor er det i gang eit revisjonsarbeid for alle dei tre verka.

3.1 Standardordbøkene Bokmålsordboka og Nynorskordboka
Bokmålsordboka og *Nynorskordboka* vart tidlegare kalla *handord-bøkene*. Dette omgrepet viser til at dei var meinte å vere eittbands ordbøker i eit nokolunde hendig format. I dag er ordbøkene elekt-

roniske, og dermed er ikkje plassomsyn like viktig som i ei trykt bok, men dei to ordbøkene skal framleis vere to mellomstore allmennordbøker for dei moderne skriftspråka bokmål og nynorsk. Dette er styrande for omfanget og lemmautvalet.

Standardordbøkene er mykje brukte. I 2020 vart det gjort nærmare 58 millionar søk i dei elektroniske ordbøkene, og fleirtalet av brukarane vel standardsøket som gjev søkjeresultat i begge ordbøkene samstundes (Ordbokstatistikk 2020). Brukarane av ordbøkene inkluderer alt frå profesjonelle skribentar til skuleelevar og personar som prøver å lære seg norsk som andrespråk. Den samansette brukargruppa er noko vi må ha i tankane når vi utformar ordbøkene.

Historia til standardordbøkene tok til i 1966. Då kom innstillinga frå Vogt-komiteen, og eitt av framlegga deira var å opprette eit norsk leksikografisk institutt. Dei skreiv vidare at «Det som i dag trenges kanskje mere enn noe annet, er gode, praktiske ettbinds ordbøker til overkommelig pris for bokmål og nynorsk, eventuelt for begge språkformer under ett» (Innstilling 1966, 13).

Norsk leksikografisk institutt vart oppretta i 1972 og lagt til Universitetet i Oslo. Arbeidet med dei to ordbøkene tok til i 1974 i samarbeid med Norsk språkråd (Vikør og Wetås 2016, 23–24). Arbeidet var organisert slik at det var to redaksjonar, ein for kvar målform, men dei samarbeidde om opplegget og delte alfabetet mellom seg slik at dei starta på kvar sine delar og sidan bytte manus. Denne framgangsmåten skulle sikre at dei to ordbøkene vart mest mogleg like kvarandre, men då førsteutgåvene var klare i 1986, viste det seg at ordbøkene likevel ikkje hadde vorte så like som ein først hadde tenkt. Dette hadde samanheng med ulikt kjeldegrunnlag og ideologiske og historiske ulikskapar mellom målformene. Mellom anna var det viktig for nynorskredaksjonen å få med ein større del talemålstilfang (Worren 1998). *Nynorskordboka* med om lag 90 000 oppslagsord vart difor ein del større enn *Bokmålsordboka* med om lag 65 000 oppslagsord.

Begge ordbøkene er komne i fleire nye utgåver sidan den gongen. I 1994 vart dei tilgjengelege i elektroniske utgåver på nettet (ordbok.uib.no). Nettutgåva av *Nynorskordboka* vart oppdatert i 2012 og *Bokmålsordboka* i 2016. Då fekk dei ny utsjånad og funksjonali-

tet, med ordklasse og leddeling på alle oppslagsord, søkbare fleirordsuttrykk, hyperlenking osv. I 2017 kom dei som appen *Ordbøkene*. Rettskriving og bøyingar er oppdaterte i samband med nyare rettskrivingsvedtak, og ein del nyord er lagde til, men ein har ikkje gått systematisk gjennom lemmautvalet og innhaldet i artiklane sidan førsteutgåvene (Hovdenak 2013).

Arbeidet med revisjonen av begge ordbøkene vart sett i gang i 2018. Prosjektet er eit samarbeid mellom Universitetet i Bergen og Språkrådet, og det har fått økonomisk støtte frå Kulturdepartementet. Arbeidet er organisert litt annleis enn tidlegare. Denne gongen er det berre éin redaksjon som arbeider med begge ordbøkene parallelt. Tre viktige oppgåver er å

- redusere skilnader mellom dei to ordbøkene som ikkje har grunnlag i faktiske målformskilnader
- syte for at definisjonar og bruksdøme i eksisterande artiklar er i tråd med språkbruken i dag
- ta inn nye oppslagsord

3.2 Dokumentasjonsordboka Norsk Ordbok

Norsk Ordbok – Ordbok over det norske folkemålet og det nynorske skriftmålet vart til i tida mellom 1930 og 2015. Dette er eit vitenskapleg ordbokverk som omfattar 12 band og inneheld kring 330 000 artiklar. Formålet med *Norsk Ordbok* (NO) er å gje ei uttømmende vitenskapleg framstilling av ordtilfanget i dei norske dialektane frå 1600 til i dag og i det nynorske skriftmålet (Vikør 2018, 10). I artiklane finn ein ordtydingar, etymologiske opplysningar, bruksdøme frå litteraturen og målføra, dialektformer og grammatiske opplysningar. Definisjonane byggjer på grundig semantisk analyse av eit datamateriale og kan ha kjeldefeste døme frå litteraturen og målføra. Idealet er at verket skal skildre alle aspekt av språket, dekkje alle delar av samfunnet og alle sjangrar. Verket vender seg til vaksne språkbrukarar med norsk som morsmål, vanleg skulegang og sterk interesse for norsk språk. Særlege målgrupper er lærarar, vitkapsfolk og lekfolk med interesse for målføre, nynorsk litteratur og språktradisjon (Vikør 2018, 10).

Verket vart fullført i 2015, men artiklane *a* til *håvøyll* er ikkje digitalt tilgjengelege. Prosjektet NO-AH (Revisjon og digitalisering av *Norsk Ordbok a–h*) tok til i 2019. Det er eit samarbeid mellom Universitetet i Bergen og Høgskulen i Volda. Målet er å gjere heile verket digitalt tilgjengeleg på nettet. Prosjektet omfattar oppdatering av redaksjonelt innhald og tekniske løysingar. Frå første hefte av band 1 vart utgjeve i 1950, har det skjedd endringar i språket og samfunnet. Vi har utvikling til dømes på det tekniske området, der eit felt som data- og kommunikasjonsteknikk har endra seg sterkt og ordtilfanget med dette.

I tillegg til innhaldet er det behov for endring av databaseløysing og redigeringsgrensesnitt. Dei tekniske spørsmåla går vi ikkje inn på i denne artikkelen.

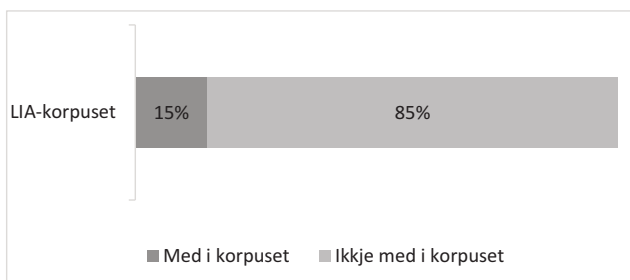
4 LIA-korpuset og revisjonen av standardordbøkene

Ei ordbok er eit utval av eit prinsipielt uendeleg ordtilfang, og kva type ordbok det er snakk om, er med på å avgjere ordutvalet. Som nemnt ovanfor er *Bokmålsordboka* og *Nynorskordboka* middels store allmennordbøker. I teorien er dei normative på rettskriving og bøyning, men ikkje i lemmautval, men i praksis blir lemmautvalet ofte oppfatta normativt av brukarane. Dersom eit ord ikkje står i ordbøkene, vil mange oppfatte det som at ordet dermed er «forbode» å bruke, noko mange av spørsmåla redaksjonen får, viser (Rauset 2019). Vi autoriserer ikkje ordtilfanget, men vel ut dei lemmaa vi meiner det er vesentleg å beskrive i ordbøkene. For at utvalet skal bli så representativt som råd, treng vi gode kjelder, og vi skal sjå nærare på om LIA-korpuset kan vere ei slik kjelde.

4.1 Revisjon av eksisterande artiklar i standardordbøkene

I revisjonen av standardordbøkene er arbeidet organisert på den måten at kvar redaktør får tildelt ein alfabetbolck som han eller ho så reviderer i begge ordbøkene. Ein vanleg arbeidsmåte er å gå systematisk gjennom ordbokartiklane i bolken og søkje opp orda i korpus for å mellom anna sjekke kor utbreidd bruken av dei er, om definisjonen

ser ut til å stemme, og om døma er gode. For å undersøkje om LIA-korpuset eignar seg til denne arbeidsmåten, har vi gjort ein liten test av dei 300 første orda i ein tilfeldig alfabetbolk. Bolken startar med ordet *aviarium* ‘fuglebur’ og sluttar med *avsanne*. Vi har søkt systematisk i korpuset på alle oppslagsorda i bolken for å sjå kor mykje informasjon ein kan finne. Figur 3 viser resultatet av testen.



Figur 3: Test av ein ordbokbolk i LIA-korpuset (N = 300)

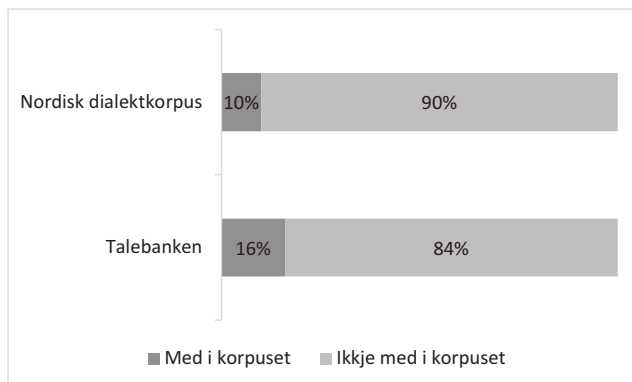
Av dei 300 undersøkte orda var det 45 som gav tilslag i LIA-korpuset, altså berre 15 prosent, og over halvparten av dei orda vi fekk tilslag på, hadde berre eitt eller to belegg. Det betyr at det for dei fleste eksisterande ordbokartiklane ikkje er så mykje informasjon å hente i dette korpuset, med unntak av at vi kan slå fast at dei fleste oppslagsorda ikkje er frekvente i talemål.

Vi har testa den same bokstavbolken i Talebanken og den norske delen av Nordisk dialektkorpus. Heller ikkje i desse talemålskorpusa får vi mange treff (sjå figur 4 neste side). Talebanken gjev om lag like mange treff som LIA-korpuset med treff på 16 prosent av oppslagsorda. I Nordisk dialektkorpus får vi ikkje treff på meir enn kvart tiande oppslagsord.

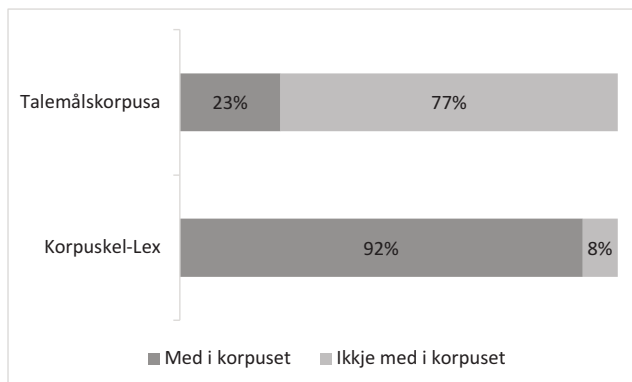
Om vi slår saman treffa for alle dei tre talemålskorpusa, hjelper det litt (figur 5 neste side), men framleis får vi treff på mindre enn kvart fjerde ordbokoppslag. Til samanlikning gjev søk i skriftspråkskorpusa i Korpuskel-Lex ein treffprosent på over 90.

Det er kan hende ikkje så veldig overraskande at vi får få treff når vi testar eit skriftbunde ordtilfang mot ein talemålsressurs på denne

måten. Éi mogleg årsak er storleiken på korpusa. Dei tre undersøkte talemålskorpusa er til saman på om lag 13 millionar ord. Til samanlikning er Nynorskorpuset på over 100 millionar ord og dei samla skriftspråkskorpusa i Korpuskel-Lex på om lag 2,7 milliardar ord. Ei slik skeivfordeling må vi nok berre leve med sidan eit talemålskorpus krev mykje meir arbeid å byggje opp.



Figur 4: Test av ein ordbokbolc i Nordisk dialektkorpus og Talebanken (N = 300)



Figur 5: Test av ein ordbokbolc i tre talemålskorpus (LIA, Talebanken og Nordisk dialektkorpus) og i skriftspråkskorpusa i Korpuskel-Lex (N = 300)

Storleiken på korpusa er likevel ikkje heile forklaringa på dei manglande treffa. Viss vi ser nærmare på kva ord den undersøkte bolken inneheld, er vel 70 prosent substantiv, om lag 20 prosent verb og resten adjektiv og adverb. Dette vil sjølvstundt variere noko alt etter kva alfabetstrekk det er snakk om, og andre strekk vil ha fleire funksjonsord, men det er ingen tvil om at det er innhaldsorda som dominerer ordbøkene.

Ordbøkene har med mange ord som er lite vanlege i daglegspråket. Dei seks første orda i den undersøkte bolken er *aviarium*, *aviatkar*, *aviatikk*, *aviatisk*, *avig* og *aving*, og det må seiast å vere ord dei færreste av oss bruker jamleg, men som det likevel kan vere fint å finne i ordboka dersom ein kjem over dei og lurar på kva dei tyder.

Ordbøkene inneheld dessutan veldig mange samansetjingar, noko som kan vere nyttig, særleg der samansetjingane er lite gjennomslutlege, eller der det kan vere vanskeleg å vite om orda skal ha samansetjingsfuge. Heiter det til dømes *ordbokredaktør* eller *ordboksredaktør*? Ein del av dei samansette orda som er med i dag, er nokså tilfeldige eller utdaterte. Den undersøkte alfabetbolken inneheld ord som *avisklipparkiv*, *avkledningsscene*, *avkomsgransking*, *avlatskremmar*, *avlusingsanstalt*, *avmagringsmiddel* og *avrekningskontor*. Det er relativt lite sannsynleg at vi skal komme over samansetjingar som desse i eit talemålskorpus.

Dei manglande treffa kan òg ha å gjere med kva materiale som er lagt inn i talemålskorpusa. Alle dei undersøkte korpusa er baserte på opptak av intervju med informantar, og slike intervju er ein spesiell sjanger. Det er visse emne ein ofte snakkar om, mens andre emne sjeldan blir nemnde, og det vil sjølvstundt påverke kva ord vi finn i korpuset. Hadde korpusa vore sette saman av radioopptak eller stortingsforhandlingar, ville vi funne eit anna ordtilfang. Dersom vi sorterar orda i alfabetbolken etter domene, vil vi til dømes sjå at mange ord som kan seiast å høyre til livet på garden, er å finne i korpusa (*avkom*, *avl*, *avle*, *avling*, *avlingsmetode*, *avlstdyr*, *avløysar*), mens det derimot ikkje er nokon som snakkar om dei mange språkvitskaplege termene som har fått plass i ordbøkene (*avleiing*, *avleiingsending*, *avleiingsforstaving*, *avleiingsprefiks*, *avleiingssuffiks*, *avlyd*, *avlydshøve*, *avlydssteg*, *avlydsveksel*, *avløysarord*).

Testen viser at ein framgangsmåte med automatiske søk i talemålskorpus på alle eksisterande oppslagsord i ordboka er lite tenleg. LIA-korpuset kan rimeleg nok ikkje hjelpe oss med alle typar ord, men det kan likevel vere nyttig å gjere meir målretta søk på ord der talemålskorpus har kvalitetar skriftspråkskorpusa ikkje har.

Standardordbøkene, og då særleg *Nynorskordboka*, inneheld ein god del tradisjonsord og talemålsord. Ein del av desse har vi få skriftlege belegg på, og det kan vere vanskeleg å vurdere om dei fortener ein plass i ei allmennordbok eller ikkje. Er ordet framleis i bruk i talemåla rundt om i landet, eller er det eit gammalt ord som ikkje er i bruk? Skal det bli ståande berre i *Nynorskordboka*, eller er ordet så vanleg at det bør opprettast ein tilsvarande artikkel for bokmål? Her kan tilgang til talemålskorpus hjelpe oss nærmare eit svar.

Eit døme er artikkelen *flautstokk*, som frå før berre er med i *Nynorskordboka*. Ein flautstokk er ein del av ein slede, og det er eit typisk døme på den typen ord det finst fleire av i *Nynorskordboka* enn i *Bokmålsordboka*. Men er dette eit ord som er vanleg i bruk? Dersom vi søker i Korpuskel-Lex, får vi ingen treff på dette ordet. Eit søk i Bokhylla til Nasjonalbiblioteket gjev ingen treff i aviser og vel 50 boktreff, men dei fleste av dei er treff i diverse ordbøker og ordsamlingar i tillegg til eit par bøker om skogsdrift i gamledagar. Vi kjem altså ikkje så mykje nærmare ei avklaring av om dette er eit ord som høyrer heime i ei moderne allmennordbok. Dersom vi søker opp ordet i LIA-korpuset, får vi derimot heile ti treff fordelt på tre informantar frå Brandbu, Nannestad og Tolga. Det tyder på at dette har vore eit vanleg omgrep i talemålet fleire stader. For to av informantane kjenner vi fødselsåret (1893 og 1903). For den tredje manglar denne informasjonen, men vi veit i alle fall at han var ein vaksen mann i 1968 då opptaket vart gjort, så vi kan tippe at han er fødd om lag på same tida som dei andre. Dei to andre opptaka er frå 1970 og 1973. LIA-korpuset fortel oss altså at ordet har vore brukt i talemålet til eldre folk på byrjinga av 70-talet, men om det framleis er i bruk i dag om lag 50 år seinare, kan ikkje korpuset fortelje oss. Til det er det meste av materialet i eldste laget.

Ei anna gruppe ord der eit talemålskorpus kan vere særleg nyttig, er i revisjonen av funksjonsorda. Desse er ein sentral del av ordtil-

fanget sjølv om dei ikkje utgjer ein stor prosent av ordbøkene. Ein del av dei har alt fått etter måten store artiklar, men dei kan justerast og supplerast ved hjelp av talemålsmateriale. Dessverre har ikkje prosjektet ressursar til å gjere djuptgåande analysar av store mengder nytt materiale, til det er det alt for mange artiklar som skal gjennomgåast på kort tid. Her må vi i staden stø oss på andre sitt arbeid. I dei seinare åra er det komme mange studiar som bruker data frå talemålskorpus, til dømes Søfteland og Borthen (2018) som har undersøkt den pragmatiske partikkelen *sjø* i Nordisk dialektkorpus.

4.2 Nytt tilfang i standardordbøkene

I revisjonen skal vi ikkje berre gå gjennom dei artiklane som alt er med i ordbøkene, vi skal òg leggje til nye artiklar. Kan LIA-korpuset hjelpe oss med det? Når det gjeld *Nynorskordboka*, har transkribørane i LIA alt gjort ein del av jobben for oss. Dei talemålsnære transkripsjonane er translittererte til nynorsk, og materialet er lagt til rette slik at alle ord som ikkje er funne i *Nynorskordboka*, er merkte med ein x. Det betyr at det er mogleg for oss å søkje opp alle dei merkte orda og slik få ei liste over moglege kandidatar til ordbøkene. At eit ord er med på x-lista, vil ikkje automatisk seie at det skal takast inn i standardordbøkene, men det er eit godt utgangspunkt for å sjå nærare på ordet.

Dei to mest frekvente lemmaa på x-lista er *derre* (3141 treff) og *herre* (1434 treff). (1) viser to døme frå korpuset på korleis dei er brukte.

- (1) a) ja # tidig korleis dei liksom får dei derre # dei derre dokkene der til å # bevege seg sånn og (aasane_uib_0202)
 b) ned # så køyrer du ned den herre bakken som heiter Storgjerdsbakken (alta_uit_0103)

Ruth Vatvedt Fjeld (2008, 17–18) etterlyste ordet *derre* i *Bokmålsordboka* etter ein analyse av NoTa-materialet. Sidan den gongen har det fått artikkel i *Bokmålsordboka*, men *derre* og *herre* manglar framleis i *Nynorskordboka*. Sjølv om mange av oss nok har lært at dette er eit talemålstrekk som bør lukast ut i skriftlege tekstar, gjev LIA-

korpuset oss eit hint om at vi bør undersøkje om orda likevel er i bruk i skrift. Eit søk i Korpuskel-Lex viser at begge orda er ein god del brukte, særleg i tekstar med eit munnleg eller uhøgtidleg preg. Før dei eventuelt kan komme med i ordboka, må Språkrådet ta stilling til om dei framleis skal vere unormerte i nynorsk (og altså ikkje takast inn i ordboka), om dei skal vere valfrie variantar av *derne* og *herne*, eller om dei skal reknast som separate lemma og få eigne artiklar.

Om vi elles ser på lista over x-orda, er det interessant å registrere at ho inneheld mange av dei orda som tradisjonelt har vore haldne utanfor nynorsken, mellom anna ei rekkje av dei såkalla anbeheitelse-orda (*antal*, *begynnelse*, *bebyggelse* osv.) og ord som *selyfølgelig*, *spise*, *kun* og *forrige*. For ord av tysk og dansk opphav er det Språkrådet som vedtek om dei kan takast inn i *Nynorskordboka* (Språkrådet 2015). Om vi ser vekk frå desse meir problematiske orda, kan x-lista hjelpe oss med å finne lakunar i ordbøkene. Av og til hender det at ord som er heilt vanlege og uproblematiske i språket, manglar. På x-lista finn vi til dømes *fjortis* og *tyggis*, og dei hadde ikkje artiklar i ordbøkene før revisjonen, men fortener nok ein plass i dei reviderte utgåvene.

5 LIA-korpuset og revisjonen av *Norsk Ordbok*

I prosjektet NO-AH skal ein leggje til rette for publisering av ei fullstendig utgåve av *Norsk Ordbok* på nettet. I arbeidet med revisjon og tilrettelegging for nettpublisering ser vi særleg to område der LIA-materialet kan brukast. Artiklar som allereie er trykte og finst i den digitale ordbokdatabasen, kan få tillegg av opplysingar knytte til seg. Dessutan kan vi identifisere lemma som til no ikkje har vore dokumenterte i språksamlingane som ordboka byggjer på.

Synet på kva som er «akseptabelt» folkemål å dokumentere, har kanskje endra seg sidan innsamlinga tok til på 1930-talet. Vi ser til dømes at det folkelege *dass* ‘toalett’ har fått artikkel. Det same har *dasshus*, men elles har ingen samansetjingar fått eigne artiklar. Tid-

legare undersøkingar i talemålskorpus peikar i retning av at til dømes *dassapir* er frekvent nok til å få ein artikkel (Fjeld 2008, 19–20).

5.1 Revisjon av eksisterande artiklar i Norsk Ordbok

For eksisterande artiklar i verket kan LIA-materialet gje oss opplysningar som stadfester det vi veit frå før, eller det kan gje oss tilleggsopplysningar. Som døme fører vi *kjeks* i tydinga ‘båt’ (sjå figur 6).

I kjeks m [Nhl1750 *kiex*, G og Lof1807 *kjæx*, Lyngen-1810-12 *kæx*, C, A₂; målf òg genus f (Sandstad, Lyngen); trul. frå eng. *ketch* 'eit slags tomastra seglbåt'] liten nordlandsbåt med to årepar og to (og eit halvt) rom (mindre enn ein færing) (Nhl1750; G; Lof1807; Lyngen1810-12; C; A₂; Sandstad, NN): *han hadde så nett eit lite segl til tjeksen sin* (EidnesTT 129).

Figur 6: Artikkelen *kjeks* i *Norsk Ordbok*

Basert på funn i LIA-korpuset (2) får vi ei opplysning om tydinga som stadfester definisjonen som alt står i *Norsk Ordbok*. *Kjeks* kan brukast om ein båt.

- (2) så tok eg då å d- tjækksa ## en tjækksbåt å å dønn lå- laga ee laga sånn ee jåa ti mek # sånn tarrejåa (lyngen_uit_0702)

I artikkelen *kjeks* (sjå figur 6) står NN for Nord-Noreg, og i den elektroniske basen har vi heimfestingslista «Sandstad, Nesna, Lurøy, Fauske, Lofoten, Vesterålen, Bø i Vesterålen, Sør-Troms, Trondenes, Senja A², Lyngen, Talvik, Måsøy, Lebesby, Gamvik». A² etter ‘Senja’ refererer til Aasens *Norsk Ordbog*. I LIA-korpuset (3) finn vi eit døme på at ordet er brukt av ein informant frå Andenes. Vi kan dermed leggje til heimfestinga Andenes, som vi ikkje hadde frå før.

- (3) å då sa hann M2 att hann sku ittje få låv å ro meire # åleina på de # tjækks'n di skulle # ta tjækks'n ifrå hann (andenes_uio_0201)

Med belegget frå Lyngen (2) får vi stadfest tydinga, og dessutan gjev dømet ei opplysning om genus femininum som vi kan føre i tillegg til opplysninga frå Sandstad som står frå før i den trykte artikkelen.

5.2 Nytt tilfang i Norsk Ordbok

Dei eldre samlingane av målføre er omfattande, men er prega av at materialet er samla av frivillige medarbeidarar. Sjølv med rettleiing frå ordbokredaksjonen vil det vere ulikt kva den einskilde har sendt inn. Talemålskorpusa vil difor kunne gje oss nytt tilfang å skrive ordbokartiklar om. Nye artiklar kan bli laga om ord som ikkje før har vore beskrivne, eller materialet kan gje opphav til nye artiklar på grunn av endra vurdering av kva ordklasse orda er tilordna.

Her hjelper annoteringa oss i og med at vi som nemnt ovanfor kan søkje på ordformer merkte med x. Desse kan takast ut som liste, og ho kan samanliknast med lemmalista frå *Norsk Ordbok*. Der vil ein kunne fange opp ord som ikkje tidlegare har artikkel i verket.

Ein mindre presis framgangsmåte, men som kan gje resultat, er å søkje med utgangspunkt i redaktøren sin kjennskap til innhaldet i *Norsk Ordbok* og kva tema som i mindre grad er omtalte i dei eldre samlingane. Vi veit at tradisjonelle kvinnedomene er dårlegare dokumenterte enn dei tradisjonelle mannsdomena. Difor gjorde vi eit enkelt søk på alle ord som byrjar med *bakst*. Dette søket gav oss dessverre ikkje nytt tilfang om bakst og baking, men det gav ein uventa bonus i form av ordet *bakstovelem* ‘loft over stove’. Ettersom sistelekken i denne samansetjinga er eit oppslagsord i *Nynorskordboka*, er ho ikkje med på x-lista, jamfør transkripsjonsrettleiinga for LIA. *Bakstovelem* fann vi i eit opptak frå Eikefjord i Sunnfjord (4).

- (4) å i dennj anndre ænnda der va et # romm so dei kallte bakståva # elle bakståeva # åg på dennj va dær åg læmme så dei kallte bakståeulemmen # bakståeulemmen (eikefjord_uib_0102)

Ordet er ikkje registrert i samlingane frå før av, og funnet kan vere grunnlaget for ein ny artikkel i samband med revisjonen av *Norsk Ordbok*.

LIA-korpuset kan dessutan føre til ei ny vurdering av kva ordklasse eit ord høyrer til. Korpuset har døme (5 a–c) på bruk av preposisjonen *ani*:

- (5) a) inngen så turrde å vær ani hann (bergen_uib_0402)
 b) mått ikkje vær ani pærsoner så add lus (bergen_uib_0402)
 c) å så bare sånn så visst enn gutt å enn jænnte i klass'n bare vi kamma ani veranndre vi tullar sånn vi sann (bergen_uib_5102)

II an adv [mlty *an*, sm o s gno *á*, nno *á* 'på'; i tyd 3 frá ty *an*, kan henda ett ital *a* 'til'].
1) lbr, i samb *an i* bort i, innåt: *koma, ná, vera an i*. **2)** i sume faste vend (vanl i talemål, lbr i skr): *gå an vera godtakande; la seg gjera; serl i vend det går (ikkje) an det går (ikkje), ein kan (ikkje) | koma an på stå på, skilja seg: det kjem an på om eg får lov* (òg: det spørst, det er tvilsamt) | *kom an kom og slåst | kom an, din rædde hund* (H.Rytt.Sh IV,95); òg med avveikt tyd, kom (og ver med) | *slá an a* (i samb med *med* el *på*), kvard, (av erotiske grunnar) gjeva seg i snakk (med) og prøva få fylgje (med ein person av anna kjønn), b) mus., lbr, gjeva, slå tone el akkord på eit strenginstrument, serl fyrste tonen el akkorden i eit musikkstykk, c) verta omtykt, fengja. **3)** forretn, (i bokføring) nytta føre namnet til kreditor el (no forald) føre postane på debetsida i konto el på rekning; mots *per* (*pr.*).

Figur 7: Artikkelen *an* i *Norsk Ordbok*, band 1

ani prep i berøring med; borte i (Arendal, Brg); jf
II an 1: ein måtte ikkje vere ani personar som hadde lus (Brg).

Figur 8: Forslag til ny artikkel for *ani* i *Norsk Ordbok*

Denne preposisjonen er registrert i samlingane frå før, men i den trykte utgåva av *Norsk Ordbok* er *ani* omtalt under artikkelen som omhandlar adverbet *an* (sjå figur 7, linje 3–4). Ein kan argumentere for å strukturere om informasjonen på den måten at preposisjonen *ani* får eigen artikkel. Resultatet av vurderinga så langt er eit utkast til artikkel i den digitale ordbokbasen (figur 8).

6 Oppsummering og konklusjon

Eit talemålskorpus som LIA er utan tvil eit nyttig verktøy for ein leksikograf. Kor mykje hjelp ein kan finne, vil likevel avhenge av kva delar av ordtilfanget det er snakk om, og kva type ordbok ein arbeider med. *Bokmålsordboka* og *Nynorskordboka* på den eine sida og *Norsk Ordbok* på den andre representerer to ulike ordboktypar både med omsyn til omfang, innhald og kor godt dokumentert eit lemma må vere for å takast med.

Standardordbøkene skal innehalde det mest sentrale og representative ordtilfanget i bokmål og nynorsk. Dei er dessutan primært skriftspråksordbøker. Artiklar med grunnlag i talespråket som allereie finst i ordbøkene, blir vanlegvis med vidare, men nye artiklar byggjer først og fremst på skriftlege belegg. Som vi har sett over, er talemålskorpusa små samanlikna med skriftspråskorpusa, og det er berre ein liten del av dei eksisterande oppslagsorda som er med i LIA-korpuset. Ei anna utfordring er at revisjonen skal vise fram endringar i språket sidan arbeidet med ordbøkene tok til for over 40 år sidan. Sidan størstedelen av LIA-materialet er eldre opptak, er korpuset mindre eigna. Det kan likevel vere til hjelp i revisjonen av enkeltartiklar, fordi det kan gje oss ein peikepinn om ord som er særleg vanlege i talemål, og ord som blir brukte litt annleis i tale enn i skrift.

Vi meiner LIA-korpuset er mest relevant for det vidare arbeidet med *Norsk Ordbok*. Verket skal dokumentere norsk både munnleg og skriftleg. Eit talemålskorpus er såleis ei nyttig kjelde. LIA-materialet har god geografisk spreining, og det er ein styrke med omsyn til ordbokarbeidet. *Norsk Ordbok* er ei dokumentasjonsordbok og har ikkje like strenge krav som standardordbøkene når det gjeld frekvens

og utbreiing. Også svakare dokumenterte ord frå målføra kan få ein plass der, og LIA-korpuset er difor ein ressurs for revisjonen av *Norsk Ordbok*.

Referansar

- Bokmålsordboka*. 1986–. Språkrådet og Universitetet i Bergen. <https://ordbok.uib.no/>.
- Fjeld, Ruth Vatvedt. 2008. Talespråksforskningens betydning for leksikografien. I *Språk i Oslo: ny forskning omkring talespråk*, redigert av Janne Bondi Johannessen og Kristin Hagen, 15–28. Oslo: Novus.
- Hovdenak, Marit. 2013. Bokmålsordboka og Nynorskordboka gjennom ein generasjon. *Nordiske studier i leksikografi* 12, 229–246. Oslo: Nordisk forening for leksikografi.
- Innstilling om språksaken fra komitéen til å vurdere språksituasjonen m.v. oppnevnt ved kongelig resolusjon 31. januar 1964*. 1966. Oslo: Kirke- og undervisningsdepartementet.
- Kristiansen, Nina og Christian Emil Ore. 1998. *Sluttrapport Dokumentasjonsprosjektet*. Oslo: Dokumentasjonsprosjektet.
- Lyse, Gunn Inger. 2020. Ut med adamsslekt og inn med arveprinsesse? Leksikografiske metodar i revisjonen av Bokmålsordboka og Nynorskordboka. *Nordiske studier i leksikografi* 15, 215–224.
- Meurer, Paul. 2012. Corpuscle – a new corpus management platform for annotated corpora. I *Exploring Newspaper Language: Using the Web to Create and Investigate a Large Corpus of Modern Norwegian*, redigert av Gisle Andersen, 31–50. Philadelphia: John Benjamins.
- Norsk Ordbok. Ordbok over det norske folkemålet og det nynorske skriftmålet*. 1966–2016. Oslo: Det Norske Samlaget. <http://no2014.uib.no/>.
- Nynorskordboka*. 1986–. Språkrådet og Universitetet i Bergen. <https://ordbok.uib.no/>.
- Ordbokstatistikk*. 2020. Universitetet i Bergen. <https://ordbok.uib.no/perl/ordbokstatistikk.cgi>.

- Rauset, Margunn. 2019. Bokmålsordboka og Nynorskordboka – einnegga, toegga eller siamesiske tvillingar? *LexicoNordica* 26, 155–175.
- Språkrådet. 2015. *Retningslinjer for normering av bokmål og nynorsk*. <https://www.sprakradet.no/Spraka-vare/Norsk/normering/retningslinjer-for-normering/>.
- Søfteland, Åshild og Kaja Borthen. 2018. 'Æ e trønder, æ, sjø!' Den pragmatiske partikkelen 'sjø' i midt-norske dialektar. *Norsk Lingvistisk Tidsskrift* 36 (2), 249–280.
- Vikør, Lars S. og Åse Wetås. 2016. Norsk Ordbok: om folket – av folket – for folket. Historia til ordboksverket fram til i dag. I *Livet er æve, og evig er ordet. Festskrift til Norsk Ordbok 1930-2016*, redigert av Lars S. Vikør, Helene Urmland Karlsen og Åse Wetås, 15–39. Oslo: Det Norske Samlaget.
- Vikør, Lars Sigurdson. 2018. *Inn i Norsk Ordbok – Brukarrettleiing og dokumentasjon*. Oslo: Det Norske Samlaget. <http://no2014.uib.no/eNo/tekst/Inn%20i%20Norsk%20Ordbok.pdf>.
- Worren, Dagfinn. 1998. Om å avgrense eit ordtilfang – saga om målføreorda i Norsk handordbok. I *Normer og regler. Festskrift til Dag Gundersen 15. januar 1998*, redigert av Ruth Vatvedt Fjeld og Boye Wangenstein, 59–70. Oslo: Nordisk forening for leksikografi.

English summary

In this article, we take a closer look at how a spoken language corpus like LIA Norwegian can be a resource in the task of revising the three Norwegian dictionaries *Bokmålsordboka*, *Nynorskordboka*, and *Norsk Ordbok*. One of the aims of the revision is finding new lemmas to include in the dictionaries and assessing and revising the dictionary entries which the dictionaries consist of at present. The tasks include deciding whether a word is a central part of modern vocabulary and therefore should be included, whether the definitions are in line with current language use, which usage examples are the most typical, and which are the most important collocations. It is necessary to have

good language sources, and in this article, we consider whether it is helpful or not to use the LIA corpus in the revision of the dictionaries.

Bente Selback
Institutt for lingvistiske, litterære og estetiske studium
Universitetet i Bergen
bente.selback@uib.no

Terje Svardal
Universitetsbiblioteket
Universitetet i Bergen
terje.svardal@uib.no



Språkendring i Vika

En komparativ analyse av data fra to talespråkkorpus

Karine Stjernholm og Ingunn Indrebø Ims

Det vikværske dialektområdet dekker et stort geografisk område på det sentrale Østlandet, og flere beskrivelser av området viser til at vikværingene har ei oppfatning om at de snakker et felles talemål (Endresen 1990; Johnsen 2015). I denne artikkelen presenterer vi en pilotundersøkelse der vi bruker materialet i LIA norsk (heretter LIA) og Nordisk dialektkorpus (Johannessen m.fl. 2009, heretter NDK) til å teste en hypotese om ulik språkutvikling innafor det vikværske dialektområdet i perioden mellom 1960–2009, som er tidsrommet dataene i korpuset dekker. Vi undersøker spesifikt overgang fra *ær*-suffiks til *er*-suffiks i kategoriene hankjønn, flertall, ubestemt, nomen agentis og presens av svake verb (etter Johnsen 2015, 136). Analysen antyder ulik bruk på tvers av dialektområdet, og vi diskuterer hvordan sosiolingvistisk teori (som Eckert 2008 og Woolard 2016) kan bidra med nyttige perspektiver for å forstå språkutviklinga i området.

Stikkord: Korpuslingvistikk, sosiolingvistikk, verditilskrivning, talemålsutvikling på Østlandet, vikværsk

Prolog

Det er med et visst vemod vi skriver denne artikkelen ferdig. Da Janne inviterte oss til å delta på LIA-seminaret i Trondheim i november 2019, hadde vi aldri trudd at dette var siste gang vi skulle være på seminar sammen med henne. Året 2020 har vært tungt på flere måter, med tapet av Janne som det aller vondeste. Samtidig er det nettopp derfor viktig for oss å bidra til denne publikasjonen, og denne studien er forsøkt gjennomført i Jannes ånd ved å se på tradisjonelle spørsmål i norsk språkforskning med nye metoder. Jannes inkluderende væremåte – både faglig og sosialt – vises godt i variasjonen av bidrag i denne boka. Tusen takk, Janne, vi vil minnes deg med takknemlighet for alt vi fikk lære og alt du lot oss være med på.

Dessverre skulle også 2021 bli et år prega av smittefare og tragiske hendelser. At Gjert også skulle gå bort nå, er ufattelig. Vi ønsker å rette ei stor takk til han for arbeidet han la ned i denne artikkelen! Gjert møtte et tidlig og uferdig manusutkast med stort engasjement, og hjalp oss videre med innsiktsfulle kommentarer og høyt verdsatte fonologiske digresjoner. Dessverre fikk ikke Gjert lese den endelige versjonen av artikkelen, og alle feil og mangler tar vi fullt personlig ansvar for.

1 Innledning og forskningsspørsmål

Dialektområdet for vikværsk spenner over et stort geografisk område på begge sider av Oslofjorden (Vika), og består av områder som stedvis er svært tett befolka. Tradisjonelle dialektologiske framstillinger beskriver det vikværske dialektområdet som sammenhengende med flere felles språktrekk som skiller dem fra områda rundt (f.eks. Endresen 1990; Helleland og Papazian 2005; Mæhlum og Røyneland 2012; Skjekkeland 2005). Ifølge Endresen (1990, 89) føler moderne vikværingar at talemåla deres utgjør ei enhet, og folk fra andre landsdeler skal også ha denne oppfatninga og har vanskelig for å skille dialektene i området fra hverandre. På de tretti åra som er gått etter 1990, kan sjølsagt mye ha skjedd, men også Johnsen (2015, 128)

nevner at «folk innanfor dette området sjølve tykkjer dei taler ”same målet” på bae sider av fjorden».

Begge forfatterne av denne artikkelen er bosatt i Vika, på hver si side av Oslofjorden. Ut fra våre erfaringer i dag framstår ikke det vikværske dialektområdet så ensarta som det skildres her. Med tanke på hvor stort geografisk område dialektområdet dekker, at det sentrale Østlandet er det tettest befolka området i landet, og dessuten at det ikke finnes noen åpenbare samlende institusjoner, større industri eller annet som binder øst- og vestsida sammen, lurer vi på om det hovedsakelig er det språkhistoriske opphavet til vikværsk som kan sies å være felles for dette dialektområdet i dag. Det kan innebære at nyere språkutvikling i ulike deler av Vika kan være forskjellig.

«Det fremste kjenneteiknet på vikværsk er den vide bruken av vokalen *æ* i trykklette stavingar» skriver Johnsen (2015, 128), og han argumenterer for at det er nettopp bruken og fordelinga av vokalen *æ* i grammatiske endelser som er noe av grunnen til at vikværingene oppfatter talemålet sitt som ei enhet. Han viser også at flere studier har vist overgang fra *æ/a* til *e* i denne omgivelsen (sst., 130). I denne artikkelen forstår vi denne overgangen som nivellering mot oslomål eller standard østnorsk, jf. nyere talemålsutvikling på Østlandet (se f.eks. Røyneland 2018).

Sjøl om området er prega av nivellering for enkelte språktrekk, skjer dette *samtidig* som flere lokale språklige trekk fortsatt er i bruk, og noen av disse lokale trekka må også anses som nødvendige for å signalisere lokal tilhørighet (se f.eks. en analyse av dette for Østfold i Stjernholm & Søfteland 2019). Røyneland (2005) definerer dialektnivellering som et dynamisk dialektkontaktfenomen som innebærer at den strukturelle forskjellen mellom nærliggende dialekter gradvis blir redusert, og at variasjonen innen en dialekt også gradvis blir redusert. Dersom vår antakelse om at det fortrinnsvis er språkhistoriske forhold som binder dette dialektområdet sammen, er det ikke gitt at de *samme* lokale trekka signaliserer de samme sosiale betydningene i forskjellige deler av Vika i dag.

Eckert (f.eks. 2008 og 2012) påpeker at meningsaspektet ved språklige trekk er viktig for endring og variasjon, og at dette perspektivet i større grad bør inn i studier av språklig variasjon. Dette fordi

den sosiale betydninga ulike variabler kan ha for ulike mennesker, kan påvirke og endre språklig praksis. Woolard (2008) etterspør på si side større fokus på språklig form i studier av språkideologi. Sammen understreker disse perspektiva at fokus både på ideologi og form er viktig for å forstå språklig endring og variasjon. Antakelsen om at vikværing er å snakke samme mål, handler altså om holdningene vikværing har til dialekten sin og ikke språket i bruk. Men følger vi Eckerts resonnement, vil den sosiale betydninga til språklige trekk også kunne endre bruken av dem. På den måten kan forskjellig bruk også kunne tolkes som ulik ideologisk utvikling i ulike deler av Vika, altså som uttrykk for ulik verditilskrivning til historisk sett felles dialekttrekk, det vil si det Agha (2007, 157–58) kaller konkurrerende verditilskrivning (competing valorization). Denne sammenhengen mellom holdninger til språk og endring i språklig praksis gjør at undersøkelser av språkbruk også kan bidra til å belyse den sosiale betydninga til språklige trekk, særlig i endringsprosesser. Og vi vil derfor understreke at observert bruk *over tid* vil kunne indikere verditilskrivning, fordi man vil kunne observere endringer i språklig praksis.

LIA og NDK inneholder data fra de samme målepunkta med et tidsintervall på opptil 50 år (1960–2009), og i denne artikkelen vil vi bruke materialet til å teste en hypotese om ulik språkutvikling innafor det vikværske dialektområdet i denne perioden. Johannessen (2003, 136) framhever bruk av korpus for å kartlegge språklig variasjon. I denne artikkelen velger vi å fokusere spesifikt på ett trekk som faller innafor beskrivelsen av *æ* i trykklette stavinger, nærmere bestemt overgang fra *-ær* til *-er* i kategoriene hankjønn, flertall, ubestemt (m.pl.ub.), nomen agentis (nom.agent.) og presens av svake verb (pres.sv.v.) (Stausland 2015, 136). I analysen undersøker vi bruken av dette trekket i de eldre opptaka fra Vika i LIA og i de nyere opptaka i NDK. Stedene vi har data fra, er Østfold, Vestfold og Telemark (Grenland) som alle ligger i det vikværske dialektområdet.

I analysen har vi ei kvantitativ tilnærming til materialet der vi undersøker endring i frekvensen til de ulike suffiksa fra LIA til NDK. Dessuten går vi mer spesifikt inn på bruken av ei ordform for å diskutere om bruken av denne kan belyse spørsmålet om lokal verditil-

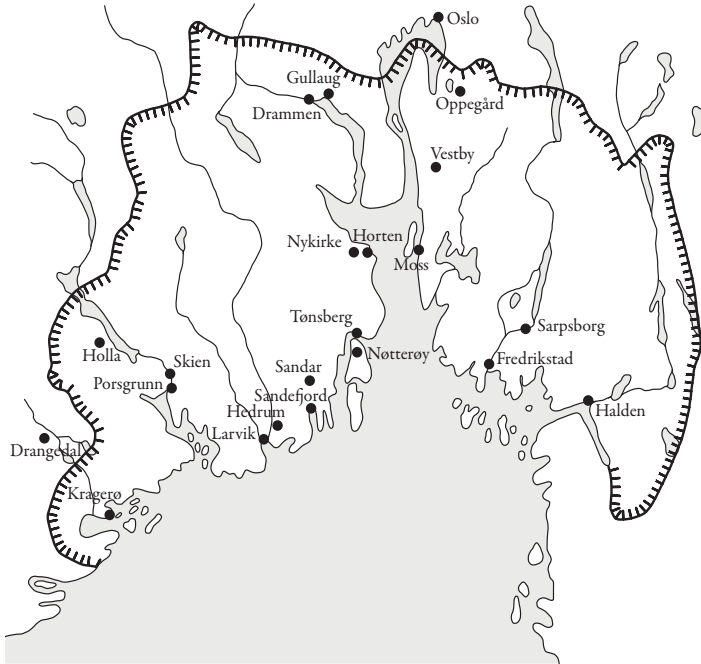
skrivning. Garrett (2011, 515–516) skriver at sameksistensen av flere koder i det samme språksamfunnet sjelden er nøytralt, men heller er uttrykk for sosial forhandling. Cornips, de Rooij og Stengs (2017, 70) skriver: «Differences between dialects are never free of social meaning, nor are the differences between dialects and the standard language or its regional varieties». Stjernholm & Søfteland (2019, 107) tolker spenninga mellom lokale og nivellerte trekk i Østfold i lys av Woolards (2016) beskrivelse av autentisitetssideologi og nøytralitetssideologi. Mens autentisk språkbruk er lokalt forankra til *et sted*, forbindes nøytral språkbruk gjerne med ei rasjonell, ikke-lokal offentlighet som alle i utgangspunktet kan ta del i. I denne artikkelen går vi ikke vidare inn i en teoretisk sosiolingvistisk diskusjon utover å slå fast at ei spenning mellom affiksene *-ær/-er* vil bli tolka som uttrykk for lokal kontra ei mer ikke-lokal tilhørighet, eller mer spesifikt, at *-ær* tolkes som et lokalt uttrykk mens *-er* anses som nøytralt eller ikke stedsspesifikt.

Johannessen (2003, 168) skriver at korpus «kanskje alltid bør brukes i det minste i en første fase, [i] en pilotundersøkelse» av en studie. Hensikten med analysen vi presenterer her, er å teste hypotesen beskrevet over basert på den nevnte variabelen. Materialet vi undersøker, har noen begrensninger som gjør at vi ikke kan trekke sikre konklusjoner om språkutvikling i Vika basert på funna vi presenterer (se mer om dette i del 3 og 5), og analysen bør derfor betraktes som en pilotundersøkelse.

Artikkelen er lagt opp slik at vi i del 2 presenterer det vikkværske dialektområdet og diskuterer valg av variabel for undersøkelse. I del 3 diskuterer vi metode og utvalg, og i del 4 presenterer vi resultatene av analysen. I del 5 diskuterer vi analysen og returnerer til noen av de sosiolingvistiske perspektiva på verditilskrivning og språkending som vi har presentert innledningsvis her.

2 Vika, vikkværsk og valg av variabel for undersøkelse

Det vikkværske dialektområdet omfatter de tidlige fylkene Østfold, Vestfold, deler av nedre Telemark (Grenland) og Akershus, samt



Figur 1 Kart over vikværsk, med de aktuelle målepunkta i analysen (Henta fra Johnsen 2015 med videre referanse til Endresen 1990)

nedre deler av Buskerud. Fordi området omkranser Vika, eller Oslofjorden, har talespråket her også blitt omtalt som vikværsk (jf. Endresen 1990, 89). I avgrensinga av det vikværske området innlemmes deler av Akershus og Buskerud (jf. Endresen 1990; Mæhlum og Røyneland 2012), men de vikværske delene av disse fylkene er ikke med i NDK, så i denne studien ser vi kun på områda Østfold, Vestfold og deler av nedre Telemark (Grenland). Det vikværske dialektområdet går også under betegnelsen sørøstlandsk i flere framstillinger av det norske dialektlandskapet (se f.eks. Mæhlum og Røyneland 2012, 55).

Talespråket i området har, med noen unntak, flere fellestrekk med resten av det østlandske talespråksområdet (som f.eks. tjukk l, jamvekt og lavtone), men det skiller seg også fra de omkringliggende områda med noen særegne språklige trekk, som altså *-ær* i trykklette

stavelser (se også Endresen 1990, 89).¹ Det er likevel ikke *utelukkende* vikkværske dialekter som har *ær*-endelser, men når disse ledsages av pronomenet *jæi* (trykksvakt *jæ*) i 1. person entall, er dette en språklig kombinasjon som har vært knytta spesifikt til dette området (Endresen 1990, 89). Den distinkte æ-en gjør seg gjeldende i blant annet hankjønns substantiva på *-ær* og *-ane* (tidligere *-æne*) i flertallsformene (som i *bilær*, *bilane* og *båtær*, *båtane*), og i tillegg har også svake verb i den såkalte kasta-klassen hatt *-ær* som endelse i presensformene (som i *kastær*, *bråkær*, *huskær*) (jf. Endresen 1990, 89). Johnsen (2015, 128–129) skriver at vikkværske bymål har *æ* i trykklette stavelser bare foran *r*, ellers har de *a* (Johnsen 2015, 135), og han viser til at flere sosiolingvistiske studier har vist seinere overgang fra *æ/a* til *e* i denne omgivelsen, det vil si ei endring som kan tolkes som nivellering. De aktuelle kategoriene som blir undersøkt her, er vist i tabell 1 (basert på Johnsen 2015, 136):²

<i>Kategori</i>	<i>Tradisjonelle former</i>	<i>Nivellerte former</i>	<i>Eksempel</i>
<i>m.pl.ub.</i>	-ær	-er	hestær, hester
<i>nom.agent.</i>	-ær	-er	bakær, baker
<i>pres.sv.v.</i>	-ær	-er	kastær, kaster

Tabell 1: Utvikling av trykklette stavelser *-ær* i vikkværske bymål (basert på Johnsen 2015, 136)

3 Metode og utvalg

I analysen bruker vi en korpuslingvistisk metode som blant annet er brukt i Johannessen (2015) og også i Stjernholm (2013). Sistnevnte brukte metoden til å undersøke mulige endring eller spredning av *a-*

¹ Her kan det legges til at dialektene i de sørligste områda i Østfold, som Lundeby kaller ytre mål eller ytre østfoldsk, faller utafør denne beskrivelsen av vikkværsk på en rekke områder og har også en del særegne trekk, som at enkelte steder har bevart utlydende *-t* i bestemt form entall av intetkjønns substantiv (se f.eks. Lundeby 1995, 15).

² Vi bruker grov lydskrift for å framheve den aktuelle variabelen her.

endelse i bestemt form entall av ikke-nøytrale substantiv i Oslo i en trend-studie (se f.eks. Gregersen 2009, 4) som likner den vi bruker her. Der blei data fra talespråkskorpuset TAUS (Talemålsundersøkelsen i Oslo) fra 1970-tallet sammenlikna med data fra NoTa – Oslo (Norsk talespråkskorpus – Oslodelen) fra 2000-tallet. Også Johannessen (2015) bruker de samme to korpusene for å vise språklige endringer i talespråket i Oslo de siste to hundre åra.

Undersøkelsen i Stjernholm (2013) tar utgangspunkt i påstanden om at *a*-endelser i bestemt form entall i oslomål er på frammarsj i hele Oslo (se f.eks. Western 1977, og også flere referanser og diskusjon i Stjernholm 2013, 135). I utgangspunktet kan vi forvente at alle substantiv i oslomål som kan bøyes med *-a* i bestemt form entall, også kan bøyes med *-en*,³ og kanskje også at flere tradisjonelle *en*-ord får ending på *-a* (Lødrup 2011). For å undersøke om *-a* kunne sies å være ekspansivt eller ikke i denne kategorien, var det derfor nødvendig å få en oversikt over *alle* substantiva i korpusene i denne kategorien med suffikset *-a* eller *-en*, og forholdet mellom *alle* ikke-nøytrale substantiv i bestemt form entall blei undersøkt. Sjøl om flere har hevda at *a*-endelser er på frammarsj i Oslo, viste resultatet at det er vanskelig å finne støtte for et slikt utsagn i en slik diakron analyse av et større talemålsmateriale som TAUS og NoTa. Det vil si at det ikke fantes dekning for ei kvantitativ økning av *a*-endelse mellom de to korpusene (Stjernholm 2013, 142).

Ei utfordring med denne metoden, er at resultatene påvirkes av hva informantene snakker om. Dersom man ikke finner belegg for noe i korpuset, betyr ikke det nødvendigvis at det ikke finnes (Johannessen 2003, 147). Stjernholm (2013, 139) eksemplifiserer dette med forholdet mellom *a*- og *en*-endelser i bestemt form entall av ikke-nøytrale substantiv i TAUS og NoTa. Snakker informantene mye om 'jenta', men lite om 'bussen', vil forholdet mellom treff på *-a/-en* påvirkes, og vil potensielt gi store utslag på resultatene. I eksempelet med *jenta* og *bussen* spiller (den språkhistoriske) genusdistinksjonen ei rolle for suffiksfordelingene, noe som ikke vil være relevant for den va-

³ En diskusjon av sammenhengen mellom genus og affiks i oslomål, finnes i Stjernholm (2013; 2019), se også Lødrup (2011) og Busterud m.fl. (2020).

riasjonen vi undersøker i vikkværsk mellom *-ær* og *-er*. Det kan likevel være sosiolingvistiske faktorer som kan påvirke datatilfanget i den analysen vi gjør her, og som av den grunn kan bidra til at resultatene gir et skeivt bilde av den faktiske talemålssituasjonen. Eksempelvis vil variasjonen ‘boka/boken’ i oslomål kunne sies å ha relativt lik sosial verdi med begge de potensielle endelsene, mens det samme ikke vil gjelde for eksempel ‘hytta/hytten’. Dette kan gjelde for spesifikke ordformer med *-ær/-er* i Vika også, nemlig at enkelte ord kan ha en emblematiske funksjon med den ene eller den andre endelsen. Etter å ha diskutert dette med kolleger på Høgskolen i Østfold, har det kommet fram at for eksempel *guttær* kan fungere som et eksempel på ei ordform med sterk lokal verdi.⁴ Dersom for eksempel de spesifikke informantene fra Østfold i LIA snakker mye om *guttær*, vil det kunne påvirke resultatene betydelig og bidra til å gi et skeivt bilde av dette suffiksens mer generelle status og bruk i dette området sammenlikna med de andre områdene vi undersøker.

Målepunkta vi undersøker i LIA og NDK, er presentert i informantmatrisen i tabell 2 på neste side sammen med opptaksår. Det er stor forskjell på LIA og NDK når det gjelder spredning i antall informanter fra sted til sted. Østfold er representert med 12 informanter i LIA og fire i NDK. I Vestfold er det tre informanter i LIA, tatt opp i 1968, mens det er 12 talere fra Vestfold i NDK, tatt opp i 2009. Fra Telemark er det 22 talere fra de områdene som regnes som en del av det vikkværske dialektområdet, det vil si Grenland. Disse er Kragerø og Skien i LIA, mens det kun er 4 informanter fra Grenland (Langesund) i NDK.

Johannessen (2003, 163) legger vekt på at dersom et korpus skal være brukbart for kartlegginger og frekvenstillinger, er det viktig at det er balansert med hensyn til de variablene forskeren er interessert i, eventuelt at det er så balansert det er praktisk mulig å få til. Skeivheten i informantfordelingen som kommer fram i matrisen i tabell 2 (neste side), medfører en ubalanse som bidrar til at resultatene av undersøkelsen vår er usikre. Johannessen (sst.) skriver videre at de fær-

⁴ Den noe anekdotiske informasjonen her har ingen direkte funksjon i analysen, men tas med som et eksempel på ei utfordring med metoden.

LIA (37 talere)		NDK (20 talere)	
Sted	Opptaksår	Sted	Opptaksår
Østfold			
Fredrikstad (3 talere)	1977, 1981	Fredrikstad (4 talere)	2009
Sarpsborg ⁵ (9 talere)	1979		
Totalt 12 talere		Totalt 4 talere	
Vestfold			
Brunlanes (3 talere)	1968	Brunlanes (4 talere)	2009
		Hof (4 talere)	2009
		Lardal (4 talere)	2009
Totalt 3 talere		Totalt 12 talere	
Telemark			
Kragerø (21 talere)	1975	Langesund (4 talere)	2008
Skien (1 taler)			
Totalt 22 talere		Totalt 4 talere	

Tabell 2: Informantmatrise

reste korpus er så balanserte som noen språkforsker kunne ønske seg, og derfor må man ta kvantitative undersøkelser basert på korpus med «mange store klyper salt». Videre skriver hun at ved å først bruke korpus, kan man likevel få en pekepinn på hvorvidt spørsmåla man stiller, er generelle nok, presise nok, om de er flertydige eller på sida av det man ønsker å vite noe om (sst.). Siden vi regner denne undersøkelsen som en pilot, er dette viktig i vår analyse. Vi regner med å finne noen tendenser i materialet som kan brukes som utgangspunkt for videre undersøkelser.

For å undersøke en eventuell ekspansjon av bruken av *-er* på be- kostning av *-ær* i det vikværske dialektområdet, gjorde vi følgende søk. Vi søkte på «ær», [end], [Phonetic form] i begge korpus. Tilsvarende søk blei gjort med «er». Deretter gikk vi inn på *Statistics*, og huka av for *Phonetic form*, *Kjønn*, *Tal*, *Bestemheit* og *Tempus*. Til

slutt lasta vi ned resultatata fra søka som ei Excel-fil. Vi fjerna irrelevante treff, det vil si ord som ender på *-ær*, men som ikke var et grammatisk suffiks (som *militær*, *sekretær*, *besvær*, *bær*, *ær* (*er*)).

Vi lytta ikke til hvert enkelt resultat. Verdien i korpus som dette ligger blant annet i den enkle tilgangen til store data, og med tanke på hvor ressurskrevende det er å lage talespråkkorpus, er det viktig å ha tillit til det arbeidet som er gjort med innsamling, transkribering og tagging, særlig i en pilot som vi presenterer her. Dette medfører likevel et potensielt reliabilitetsproblem som handler om hvor mye vi kan stole på lyttereenskapene til transkribøren (se f.eks. også Kristoffersen og Neteland i denne publikasjonen). Og det betyr heller ikke at det aldri finnes tilfeller der det likevel vil være nyttig å sjekke om transkripsjonen stemmer med lydopptaket. Eksempelvis vil behovet for å lytte antakelig være større i en fonologisk studie enn i en syntaktisk studie. Studien vi presenterer her, er en fonologisk analyse, så dette forbeholdet får betydning for diskusjonen av resultatata.

4 Funn

4.1 Overgang fra *-ær* til *-e* i trykklette stavelser

Søket beskrevet i del 3 over, ga resultatet som kommer fram i tabell 3 for distribusjonen mellom *-ær* og *-er*.⁵ Tabellen viser kun frekvens av *-ær* i prosent med totalt antall belegg i parentes. Tabellen viser resultatata for alle målepunkta totalt, samt for Østfold, Vestfold og Telemark (Grenland).

Totalt		Østfold		Vestfold		Telemark	
<i>LIA</i>	<i>NDK</i>	<i>LIA</i>	<i>NDK</i>	<i>LIA</i>	<i>NDK</i>	<i>LIA</i>	<i>NDK</i>
24 %	30 %	30 %	18 %	34 %	34 %	20 %	28 %
(1951)	(2068)	(665)	(402)	(62)	(1406)	(1226)	(229)

Tabell 3 (alternativ): Frekvens av *-ær*. Talla i parentes er totalt antall belegg.

⁵ Endelse på *-ar* blei også inkludert i analysen, men treffa var så få på denne varianten at vi har valgt å se bort fra dem i analysen.

For alle steder totalt viser tabellen at 24 % av treffa i LIA hadde endelse på *-ær*; som altså tilsvarer 76 % på *-er*. I NDK har andelen treff på *-ær* økt til 30 %, noe som peker i retning av ei økning i bruk av den tradisjonelle varianten.

Ser vi på distribusjonen fylkesvis i LIA, er andelen treff på *-ær* 30 % i Østfold, 34 % i Vestfold, og 20 % i Telemark. Vestfold er altså det fylket i LIA som har høyest frekvens av den tradisjonelle varianten, med Østfold noe bak, mens den tradisjonelle varianten ser ut til å være mindre i bruk i de delene av Telemark som tilhører området. I NDK er fordelinga 18 % i Østfold, 34 % i Vestfold og 28 % i Telemark. Her er altså bruken av den tradisjonelle varianten lavest i Østfold, mens forholdet mellom tradisjonell og nivellert variant i Vestfold er den samme i LIA og NDK. I Telemark viser talla ei økning for den tradisjonelle varianten, altså fra 20 % i LIA til 28 % i NDK. Trenden her er altså nedgang for den tradisjonelle varianten øst for Oslofjorden, mot stabilitet og oppgang på vestsida.

4.2 En nærmere beskrivelse av ei spesifikk ordform

For å komme tettere inn på hvordan denne variasjonen fungerer i praksis, leita vi spesifikt i korpusene etter ordformer der vi kunne observere den variasjonen vi undersøker på tvers av tid (LIA og NDK) og sted (Østfold, Vestfold og Telemark). Dette var utfordrende å finne. Det er svært få ordformer i LIA og NDK som er så frekvente at vi kan undersøke bruk av *samme* ordform på tvers av områda i undersøkelsen. Det aktuelle ordet vi har funnet flest belegg for i hele Vika-området, er verbet 'husker'. Dette ordet er såpass frekvent at vi vil antyde en forsøksvis analyse av bruken. Resultatet fra søk på 'hugsar'/'husker' i presens på de aktuelle målepunkta ga resultatet i tabell 4 med eksempler i tabell 5.

Tabell 4 viser antall treff i prosent med affikset *-er/-ær*. Tabell 5 viser hvilke former som hører innunder *-nivellert* og *+nivellert* i tabell 4. 'n=' i tabell 4 angir hvor mange belegg prosentene er regna ut fra for hvert fylke, tallet i parentes er det totale antall belegg som også omfatter treff som faller utafør analysen, som *husska*, *hussk*.

Talla i LIA fra Østfold viser at andelen ikke-nivellert form, det vil si former på *huss**, dominerer, da særlig ikke-nivellert form som

	LIA		NDK	
	-ær	-er	-ær	-er
<i>Østfold</i>	n=28 (32)		n=11 (12)	
-nivellert	25,0 %	50,0 %	0,0 %	0,0 %
+nivellert	3,6 %	21,4 %	18,2 %	81,8 %
<i>Vestfold</i>	n=0		n=91 (93)	
-nivellert	Ingen treff		71,4 %	12,1 %
+nivellert			2,2 %	14,3 %
<i>Telemark</i>	n=53 (51)		n=13 (13)	
-nivellert	5,7 %	20,8 %	15,4 %	7,7 %
+nivellert	0,0 %	62,3 %	15,4 %	61,5 %

Tabell 4: Distribusjonen av variasjonen for 'husker/hugsar' i LIA og NDK

	-ær	-er
-nivellert	hus(s)æ(r)	hus(s)e(r)
+nivellert	huskæ(r)	husker

Tabell 5: Eksempler til tabell 4. Bokstaver i parentes viser variasjon i transkripsjonen.

ender med affikset *-er* (*husser*) med 50 % treff. Andelen ikke-nivellert med *ær*-suffiks (*hussær*) er på 25 %, andelen nivellert form med *er*-suffiks (*husker*) er på 21,4 %. I NDK er andelen ikke-nivellerte former borte, mens forholdet mellom nivellert form med *ær*-/*er*-suffiks er 18,2/81,8 %.

Fra Vestfold fins det ingen treff i LIA, men talla fra NDK viser dominans av ikke-nivellert form med *ær*-suffiks (*hussær*) med en andel på 71,4 %.

Talla i LIA fra Telemark viser dominans for *er*-suffiks, særlig for den nivellerte forma. Det er også ikke-nivellert form med *er*-suffiks som dominerer i NDK med 61,5 %, men med en høyere andel *ær*-suffiks, 30,8 % fordelt likt mellom nivellerte og ikke-nivellerte former.

Sammenholdt med resultatene fra tabell 3, er inntrykket fra del 4.1 forsterka på den måten at de tradisjonelle formene går mest tilbake i Østfold. I Vestfold ser de ut til å holde stand, mens talla fra Telemark viser ei økning for *-ær*-suffiks.

5 Diskusjon og videre oppfølging av studien

I denne artikkelen har vi sett på overgang fra *-æ* til *-e* i trykklette stavelser i LIA og NDK i det vikværske dialektområdet, nærmere bestemt overgang fra *-ær* til *-er* i kategoriene hankjønn, flertall, ubestemt, nominativ agent og presens av svake verb. Dessuten har vi gått spesifikt inn på bruken av ei ordform (*husker*) for å diskutere om bruken av denne kan belyse spørsmålet om lokal verditilskrivning. I analysen har vi undersøkt bruken av dette trekket i eldre opptak fra Vika i LIA og nyere opptak i NDK, som til sammen representerer et tidsrom fra 1960–2009. Stedene vi har data fra, er Østfold, Vestfold og Telemark (Grenland) som alle ligger i det vikværske dialektområdet.

Analysen vår viser ulik distribusjon for dette trekket på ulike steder i det vikværske området. De tradisjonelle formene, som altså tolkes som lokale uttrykk i analysen, går mest tilbake i Østfold, holder seg stabile i Vestfold og har ei økning i deler av Telemark. Disse dataene styrker antakelsen om ulik språkutvikling innafor det vikværske dialektområdet i denne perioden, iallfall for enkelte sentrale språktrekk, som har blitt regna som distinkte for området.

Det er betydelig færre informanter fra området i NDK (4) enn i LIA (12), og som vi tidligere har nevnt, gjør det at vi må ta forbehold når det gjelder hva analysene faktisk viser. Vestfold er det fylket som har høyest andel av den tradisjonelle varianten i LIA (34 %), men bare tre informanter står bak disse treffa. I NDK er det derimot 12 informanter fra samme område, og andelen er lik for den tradisjonelle varianten (34 %) her. I Telemark viser talla fra NDK ei økning for den tradisjonelle varianten (28 %) sammenlikna med LIA (20 %). Fra dette området er det 22 talere i LIA og fire i NDK. I LIA er det særlig Vestfold som peker seg ut med færre talere (tre stykker), mens det i NDK er færrest talere i Østfold og Telemark (fire på hvert sted). At andelen informanter fra Vestfold i NDK (12 talere) er høyere enn fra de andre stedene, gjør at resultatene herfra framstår som noe mer robust enn fra de andre områda vi ser på.

Diskusjonen om antall informanter og skeivfordeling i utvalget gjør det relevant å trekke fram NORM-prinsippet som har stått sterkt

i norsk dialektologi. Kort skissert dreier dette seg om at dialektforskere har hatt ei forkjærlighet for informanter som representerer et tradisjonelt talemål, også beskrevet som NORM (Non-mobile Old, Rural, Male, se Chambers og Trudgill 1998), og vi kan muligens anta at LIA-informantene er plukka ut etter slike kriterier. Det bør dessuten legges til at også i NDK var det et mål at talerne skulle representere den lokale dialekten på opptaksstedet, men her var kjønnsvariabelen balansert. Dette kan altså farge resultatene fra begge korpus, særlig når antallet informanter er lavt, og dette får også konsekvenser for representativiteten i undersøkelsen. Siden vi kun gjør ei kvantitativ vurdering på gruppenivå her, veit vi heller ikke om enkelte informanter står for en stor andel av bruken av den ene eller andre varianten, noe som gjelder alle stedene vi har undersøkt.

Resultatene bør altså tas med ei klype salt, som Johannessen (2003, 163) påpeker. Like fullt er dette interessante funn, som viser at det er verdt å følge opp med flere og mer utvida studier av det vikværskede dialektområdet. Det LIA og NDK ubestridelig viser gjennom disse analysene, er at det er en sameksistens av flere koder i det vikværskede området, og som Garret (2011, 515–516) og også Cornips, de Rooij og Stengs (2017) skriver, vil variasjon sjelden oppleves nøytralt av brukerne sjøl, men vil nettopp være knyttet til ulike verdier og sosial forhandling. Informantene i LIA og NDK sier ingenting om si oppfatning av eget språk og ulike språklige trekk, men den variasjonen som *de facto* fins i området, er et godt utgangspunkt for videre undersøkelser.

I lys av Woolard (2016) tolker vi bruken av *ær*-suffiks som et lokalt uttrykk (autentisk) mens *er*-suffikset tolkes som nøytralt eller ikke stedsspesifikt (nøytralt). Da er det særlig interessant at informantene, i tillegg til å variere mellom standardnære *husker* og den tradisjonelle *hussær*, også bruker varianter som må sies å ligge mellom disse, som *huskær* og *husser*. Mens autentisk språkbruk altså er lokalt forankra til *et sted*, forbindes nøytral språkbruk gjerne med en rasjonell, ikke-lokal offentlighet som alle i utgangspunktet kan ta del i. Stjernholm & Søfteland (2019) viser hvordan visse lokale trekk i Østfold nesten er nødvendige å bruke lokalt, sjøl for lokale talere som representerer et nokså nivellert talemål. Ulik verditilskrivning

dreier seg om hvilke verdier og normer som kobles til ulike måter å snakke på i et språksamfunn, både av språkbrukerne som sjøl snakker sånn, men også av andre som gjenkjenner måten å snakke på som noe som skiller seg fra andre varieteter (Agha 2007: 154, se også Stjernholm 2013; 2019 og Ims 2013; 2019 om verditilskrivning og språkbruk i Østfold og Oslo). Bruken av det som kan kalles overgangsformer mellom ikke-nivellert og nivellert form, som *huskær* og *husser*, kan tolkes som ei strategisk tilnærming til det å balansere en lokal og en nøytral identitet. I lys av dette mener vi altså at verditilskrivning er en faktor det vil være særlig interessant å fokusere på i videre studier fra det vikværske området. Språkbrukerne kan altså navigere og forholde seg ulikt, eller tilskrive ulik verdi til det samme språklige uttrykket øst og vest for Oslofjorden.

I jakten på faktorer som kan forklare språklig variasjon og endring, og som følgelig vil kreve oppfølgingsstudier når det gjelder forhold vi har antydta i denne artikkelen, vil vi avslutningsvis også komme med ei oppfordring og et ønske for framtidige norske innsamlinger av korpusdata. Dersom informantene i større grad blei spurt om oppfatninger knytta til egen og andres språkbruk i intervju-situasjonen, ville korpusdata også hatt en langt større verdi for sosiolingvistiske og mer kvalitative undersøkelser. I norsk sammenheng kan vi nevne UPUS-korpuset (Utviklingsprosesser i urbane språkmiljø) som et prosjekt som i større grad innlemmer verdi- og holdningsspørsmål knytta til språk i informantintervjuet. Det å få mer kunnskap om språkstrukturelle trekk har ofte vært utgangspunktet for større innsamlinger, også på 2000-tallet, men som Eckert (2008) påpeker, vil man forstå bedre hvilke faktorer som bidrar til språklig endring og variasjon dersom man i større grad innlemmer språkideologiske forhold i språkstrukturelle beskrivelser, men også har et tydeligere fokus på språkstrukturelle trekk i mer språkideologiske analyser (Woolard 2008).

Referanser

- Agha, Asif. 2007. *Language and social relations*. Cambridge: Cambridge University Press.
- Busterud, Guro, Terje Lohndal, Yulia Rodina og Marit Westergaard. 2020. Hvor mange genus er det i Trondheims-dialekten? *Maal og Minne* (2), 1-36. <http://ojs.novus.no/index.php/MOM/article/view/1898>
- Chambers, Jack K. og Peter Trudgill. 1998. *Dialectology*. 2nd ed. Cambridge: Cambridge University Press.
- Cornips, L., de Rooij, V., & Stengs, I. 2017. Carnavalesque language use and the construction of local identities: A plea for language-culture as a field of research. *Jahrbuch für Europäische Ethnologie Dritte Folge 2017*(12), 61–89. https://doi.org/10.30965/9783657788750_006
- Eckert, Penelope. 2008. Variation and the Indexical Field. *Journal of Sociolinguistics* 12 (4), 453–476. DOI: <https://doi.org/10.1111/j.1467-9841.2008.00374.x>.
- Endresen, Rolf Theil. 1990. Vikværsk – målet i Østfold, Vestfold, Grenland og Nedre Buskerud. I *Den store dialektboka*, redigert av Ernst Håkon Jahr, 89–99. Oslo: Novus.
- Garrett, P.B. 2011. Language Socialization and Language Shift. I A. Duranti, E. Ochs & B.B. Schieffelin (red.), *The Handbook of Language Socialization*. Malden, Mass: Wiley-Blackwell, 515–535. <https://doi.org/10.1002/9781444342901.ch22>
- Gregersen, Frans. 2009. The data and design of the LANCHART study. *Acta Linguistica Hafniensia* 41, 3–29. DOI: <https://doi.org/10.1080/03740460903364003>.
- Ims, Ingunn Indrebø. 2013. Språklig registerdanning og verditilskrivning. Betegnelser på nye måter å snakke norsk på i Oslo. *NOA – Norsk som andrespråk* 2013 (2), 37–71. <http://ojs.novus.no/index.php/NOA/article/view/844>
- Ims, Ingunn Indrebø. 2019. A-ender og lokal indeks blant ungdommer i flerspråklige miljøer i Oslo. *Maal og Minne* 2019 (2), 19–67. <http://ojs.novus.no/index.php/MOM/article/view/1740>.

- Johannessen, Janne Bondi. 2003. Innsamling av språklige data: Informanter, introspeksjon og korpus. I *På språkjakt - problemer og utfordringer i språkvitenskapelig datainnsamling*, redigert av Janne Bondi Johannessen, 133–173. Oslo: Unipub forlag.
- Johannessen, Janne Bondi. 2015. Oslo-språket de siste to hundre år. I *Talemål etter 1800. Norsk i jamføring med andre nordiske språk*, redigert av Helge Sandøy, 269–300. Oslo: Novus.
- Johannessen, Janne Bondi, Joel Priestley, Kristin Hagen, Tor Anders Åfarli, og Øystein Alexander Vangsnes. 2009. *The Nordic Dialect Corpus - an Advanced Research Tool*. Volum 4 av *NEALT Proceedings Series: Proceedings of the 17th Nordic Conference of Computational Linguistics NODALIDA 2009*. Redigert av Kristiina Jokinen og Eckhard Bick.
- Johnsen, Sverre Stausland. 2015. Språkendringar langs Oslofjorden. I *Talemål etter 1800. Norsk i jamføring med andre nordiske språk*, redigert av Helge Sandøy, 125–158. Oslo: Novus.
- Lundebj, Einar 1995. *Østfoldmål*. Gressvik: Østfold Mållag.
- Lødrup, Helge. 2011. Hvor mange genus er det i Oslo-dialekten? *Maal og minne* (2), 120–136. <http://ojs.novus.no/index.php/MOM/article/view/330/328>.
- Mæhlum, Brit og Unn Røynealand. 2012. *Det norske dialektlandskapet: Innføring i studiet av dialekter*. Oslo: Cappelen Damm Akademisk.
- Kristoffersen, Gjert og Randi Neteland. 2021. Ka farsken? Realisering av /r/ som sibilant foran /k/ i nordnorsk. I denne boka, s. 41–64.
- Papazian, Eric og Botolv Helleland. 2005. *Norsk talemål. Lokal og sosial variasjon*. Kristiansand: Høyskoleforlaget.
- Røynealand, Unn. 2005. Dialektnivellering, ungdom og identitet. Ein komparativ studie av språkleg variasjon og endring i to tilgrensande dialektområde, Røros og Tynset. Doktoravhandling, Universitetet i Oslo.
- Røynealand, Unn. 2018. Sosial og geografisk variasjon, utjamning og ny variasjon frå ca. 1945 til i dag. I *Norsk språkhistorie. NSH-II Praksis*, redigert av Brit Mæhlum, 240–257. Oslo: Novus.

- Skjekkeland, Martin. 2005. *Dialektar i Noreg. Tradisjon og fornying*. Kristiansand: Høyskoleforlaget AS.
- Stjernholm, Karine. 2013. Stedet velger ikke lenger deg, du velger et sted. Tre artikler om språk i Oslo. Ph.D.-avhandling, Universitetet i Oslo.
- Stjernholm, Karine. 2019. Variasjon som virkemiddel – hiphop, språkvalg og identitetskonstruksjon i Oslo. *Norsk Lingvistisk Tidsskrift* 37, 25–57. <https://hiof.brage.unit.no/hiof-xmlui/handle/11250/2628732>.
- Stjernholm, Karine, & Åshild Søfteland. 2019. Talemål i Sør-Østfold: Ideologi, struktur og praksis. *Målbryting*, 10, 101–131.
- Western, Knut. 1977. *a-endinger i Oslo-mål*. Skrift nr. 5. i *Talemålsundersøkelsen i Oslo*. Oslo: Novus.
- Woolard, Kathryn A. 2008. Why *dat* now?: Linguistic-anthropological contributions to the explanation of sociolinguistic icons and change. *Journal of Sociolinguistics*, 12(4), 432–452. <https://doi.org/10.1111/j.1467-9841.2008.00375.x>
- Woolard, Kathryn A. 2016. *Singular and plural. Ideologies of Linguistic Authority in 21st Century Catalonia*. New York: Oxford University Press. <https://dx.doi.org/10.1093/acprof:oso/9780190258610.001.0001>

Samla oversikt over korpus:

Nordisk dialektkorpus: <http://tekstlab.uio.no/nota/scandiasyn/>.

LIA norsk – korpus av eldre dialektoptak:

<https://www.tekstlab.uio.no/LIA/korpus.html>

Norsk talespråkkorpus – Oslodelen:

<http://www.tekstlab.uio.no/nota/oslo/index.html>

TAUS-korpuset: <http://www.tekstlab.uio.no/nota/taus/>.

UPUS – Utviklingsprosessar i urbane språkmiljø:

<https://www.hf.uio.no/iln/forskning/prosjekter/upus/>

English summary

The dialect area Vika covers a large geographical area in central Eastern Norway, and several descriptions of the dialect indicate that the inhabitants of Vika have a common perception that they speak the same dialect (Endresen 1990; Johnsen 2015). In this article, we present a pilot study in which we use the material in LIA Norwegian and the Nordic dialect corpus (Johannessen et al. 2009) to test a hypothesis of different language development within the dialect area in the period between 1960–2009, which is the amount of time the data in the corpus covers. We examine specifically the transition from *ær*-suffix to *er*-suffix in the categories masculine, plural, indefinite; nomen agentis; and the present tense of weak verbs (after Johnsen 2015, 136). The analysis suggests differences across the dialect area, and we discuss how sociolinguistic theory such as Eckert (2008) and Woolard (2016) can contribute with useful perspectives for understanding language development in the area.

Karine Stjernholm
Institutt for språk, litteratur og kultur
Høgskolen i Østfold
karine.stjernholm@hiof.no

Ingunn Indrebø Ims
Institutt for språk og litteratur
Universitetet i Sørøst-Norge
ingunn.i.ims@usn.no



Trykkplasseringa i latinske lånord og partikkelverb i tre norske dialektområde

Eirik Tengesdal og Björn Lundquist

Denne artikkelen granskar sambandet mellom trykkplassering i latinske lånord (som *butikk*) og partikkelverb (som *gå ut*), der trykkplasseringa på fyrstestavinga eller sistestavinga varierer. Med di partikkelverb har ein del av variasjon som skriv seg frå innoomspråklege strukturelle faktorar, reknar me med at desse faktorane ikkje påverkar latinske ord i like stor grad. Ved å granska ikkje-syntaktisk variasjon i latinske ord, vonar me å forstå trykkvariasjon i partikkelverb betre. Med LIA og NDK granska me talarar frå fylka Trøndelag, Hedmark og Finnmark. Ingen av dialektane hadde ei kategorisk trykkplassering, men Trøndelag skilde seg ut med mest fyrstestavingstrykk både i latinske ord og partikkelverb. Hedmark hadde meir variabel distribusjon, med overvekt av fyrstestavingstrykk. Finnmark hadde berre sistestavingstrykk i dei latinske orda, men det var ganske jamt fordelt mellom fyrste- og sistestavingstrykk i partikkelverba. Fyrstestavingstrykk i partikkelverb, altså på verbet, var den klåre preferansen for alle områda utanom Finnmark, som syntest å ha ein blanda preferanse mellom trykk på verb eller partikkelen. LIA og NDK er velegna for prosodisk gransking.

Nykelord: korpuslingvistikk, latinske lånord, LIA-korpuset, NDK, partikkelverb, prosodi, trykkplassering, variasjon

1 Innleiing^{1 2 3}

I dei nordgermanske språka finn me stor variasjon med omsyn til partikkelverb, både syntaktisk og prosodisk. Syntaktisk ser me ein kategorisk skilnad mellom svensk og dansk, som kjem til synes i at verbpartikkelen alltid er stilt føre eit direkte objekt i svensk (1), og etter eit direkte objekt i dansk (2), utan omsyn til om objektet er ein heil nominalfrase eller eit pronomen:

- (1) Vakterna kastade ut studenten/honom. (Sve)
- (2) Vagterne smed studenten/ham ud. (Dan)

I dei andre nordgermanske språka, framfor alt norsk, er partikkel-plassering med omsyn til objekt friare. I norsk er pronomielle objekt gjerne stilte føre partiklar (3), medan plassering av NP-objekt er meir variabelt. Trass variasjon, er der generelt ein preferanse for partikkel føre NP-objekt i norsk (sjå m.a. Aa 2015, 2), som i (4).

- (3) Vaktene kasta han ut. (Nor, Pron)
- (4) Vaktene kasta ut studenten. (Nor, NP)

¹ Me takkar to anonyme fagfellar for verdifulle tilbakemeldingar på ei tidlegare utgåve av denne artikkelen, og for nyttige innspel frå redaktørane Tor A. Åfarli og Gjert Kristoffersen. Me takkar au for innspela frå deltakarane på sluttseminaret for LIA-prosjektet 26.–27.11.2019. Med denne artikkelen vil me heidra professor Janne Bondi Johannessen, som var leiar for LIA-prosjektet. Me saknar deg, Janne. Tusen takk for alt! ET vart finansiert av NFR-prosjektet 250755, og BL vart finansiert av NFR-prosjektet 302524.

² Definisjonar: LIA: LIA norsk – korpus av eldre dialektopptak: <https://www.tekstlab.uio.no/LIA/korpus.html>

NDK/NDC: Nordisk dialektkorpus/Nordic Dialect Corpus (Johannessen mfl. 2009): <http://www.tekstlab.uio.no/scandiasyn>

³ I denne artikkelen omtalar me ulike storleikar med liknande merkelappar. Ei rekkje fylke er frå 2018–2020 samanslegne, t.d. er Nord-Trøndelag og Sør-Trøndelag no samanslegne til Trøndelag (Regjeringen.no 2019). Trøndsk kan likevel omfatta meir enn mål som ein finn innom eit geografisk avgrensa fylke, m.a. reknar ein måla på Nordmøre som trøndske (Jahr 1990, 10).

Men ein finn variasjon i dette mønsteret: I visse dialektar, framfor alt i Trøndelag og Hedmark, finn ein ofte det «svenske» mønsteret, med både pronomen og NP-objekt etter partiklar (jf. 1), medan NP-objekt tolleg ofte er stilte føre ein partikkel i andre dialektar.

I norsk er det ikkje berre syntaksen som varierer i partikkelverba, men au prosodien. I austnorsk og trøndsk vert verb og partikkel ofte realiserte med hovudtrykk på verbet og samansetjingsaksent.⁴ Dette skil seg sterkt frå t.d. svensk, der trykket ligg på partikkelen, og verbet ofte er heilt deaksentuert. I dei austnorske og trøndske dialektane finn me likevel fleire ulike aksentmønster hjå partikkelverb: iblant ber verbet trykk, og i andre fall ber både verb og partikkel eigne trykk. Me ser au dialektale preferansar for ulike mønster. I Tengesdal, Larsson og Lundquist (upublisert manuskript, 2021) ser me au at dialektar som har samansetjingstrykk på partikkelverb relativt sjeldan har element som står mellom verb og partikkel (jf. tabell 1, kategori 5–7). Det verkar med andre ord å finnast korrelasjonar mellom prosodi og leddstillingsmønster.

Me kjenner til fleire syntaktiske, semantiske og informasjonsstrukturelle faktorar som påverkar både partikkelplassering og partikkelprosodi, som me legg fram i avsnitt 3 (sjå au Aa 2015). I mange fall verkar både plassering og trykk å vera dels uføreseielege: Innom same dialektområde finn me eit variabelt trykkmønster i liknande syntaktiske kontekstar. Der finst i minsto to ulike mogelege forklaringsar på dette, anten finst der syntaktiske, semantiske og informasjonsstrukturelle faktorar som påverkar partikkeltrykk som me framleis ikkje kjenner til, eller so er variasjonen dels vilkårleg, og då mogelegvis kopla til sosiolingvistiske faktorar. Western (1921, § 454) nemner at samansetjingstrykk var folkeleg utbreidd, men at dette au vann gradvis meir grunn i den øvre middelklassen på tidleg 1900-talet. Ein kan au tenkja seg at talesnøggleik er kopla til den prosodiske fraseringa (sjå t.d. Tooley, Konopka og Watson 2018).

I denne artikkelen jamfører me trykk i partikkelkonstruksjonar med eit liknande variabelt trykkmønster: trykk på fyrste- eller sistestavinga i sokalla greko-romanske lånord (Fretheim 1969; heretter

⁴ I resten av artikkelen omtalar me «hovudtrykk» som ‘trykk’.

omtala som ‘latinske ord’), som t.d. *inspirera* og *trafikk*. Flestalle norske dialektar har i dag trykk på sistestavinga i stammen i slike ord, men i framfor alt trøndske og til dels austnorske dialektar hamnar trykket på fyrstestavinga, og desse orda får då ei trykkplassering som er identisk med samansetjingsmønsteret hjå partikkelverb. For latinske ord kan me likevel vera visse på at det ikkje er semantiske, syntaktiske eller informasjonsstrukturelle faktorar som styrer variasjonen. Om me finn variasjon innom dialektar eller talarar som har variabelt trykkmønster i latinske ord, kan me vera visse på at i minsto ein del av variasjonen i trykkmønster i partikkelverb ikkje er kopla til innspråklege faktorar. Føremålet med denne artikkelen er fylgjande: (1) å systematisk granska kor stor variasjon der finst i trykkmønster i partikkelverb i tre norske dialektområde; (2) å granska om der finst ein liknande variasjon innom latinske ord, og (3) å vurdere om LIA-korpuset, saman med NDK, kan nyttast som eit verktøy for å granska variasjon og endring kopla til fonetikk og fonologi (her: trykk). I avsnitt 2 legg me fram nokre hovudspor innan norsk prosodisk gransking knytt til spørsmåla våre, etterfylgt av ei framstilling av forskingsspørsmål og metode, resultat og drøfting.

2 Samansetjingsaksent eller fraseaksent?

I dei skandinaviske språka ligg trykket i morfologiske samansetjingar til vanleg på den fyrste trykksterke stavinga til føreleddet, medan trykket i frasar ligg på den siste trykksterke stavinga (Kristoffersen 2003, 125–126). Partikkelverb har ofte typisk frasetrykk i mange norske dialektar, og au i svensk. I austnorske og trøndske dialektar hamnar derimot trykket ofte på den fyrste trykksterke stavinga, nett som i ei samansetjing. Liknande mønster finst au i fleire nordsvenske dialektar (kan henda finst der eit dialektkontinuum med Trondheim som episenter, som spreier seg sørover mot Oslo og nordaust mot austersjøkysten). Sandøy (1976, 13, 1985, 71) oppgjev dessutan at det same mønsteret med samansetjingstrykk gjeld vestover på Nordmøre og Romsdal. Western (1921, § 454) opplyser at dette mønsteret er vanleg i bygdemåla på Austlandet. I desse dialektane finn me ofte

trykk på fyrstestavinga i andre ord og kontekstar der me finn trykk på sistestavinga i t.d. dagsens standardsvensk, men au i vestnorske dialektar. Sandøy (1985, 71) opplyser dessutan at samansetjingstrykk kan nyttast i Indre Sogn, og dels i Stavanger om verb + partikkel ikkje omfattar meir enn to stavingar (jf. 7 i tabell 1). Abrahamsen (2003, 197) nemner at mønsteret finst på Sunnmøre. Skaalbones (2006, 52) fann varierende tonelagsskifte med trykk på verbet i Rana, i mindretalet av falla med verb + partikkel rett attmed, som ho tolkar som ein påverknad frå lågtonedialektar. Sjå Aa (2020, 52) for vidare drøfting.

I tabell 1 legg me fram eit hierarki av aksenteiningar, der me variabelt finn fyrstestavingstrykk i fastlandsskandinaviske dialektar. Hierarkiet skal lesast på fylgjande måte: Om ein dialekt tillèt trykk på fyrstestavinga i dei øvste kategoriane, so tillèt dei au trykk i alle kategoriar under. Til dømes: Alle dialektar som tillèt trykk på fyrstestavinga i monomorfematiske lånord, som *banan* og *tomat*, tillèt au fyrstestavingstrykk i polymorfematiske lånord (*stasjon*, *inspirera*) og partikkelverb. Svensk tillèt berre fyrstestavingstrykk på kategori 9 (jf. 'svart rödstjärt' (Sve), 'svartraud(,)stjert' (Nor)), medan dagsens austnorske bymål gjerne tillèt trykk frå nivå 6–7 og nedover, og vestnorske dialektar frå nivå 7–8 (samansetjingstrykk i partikkelverb ser

<i>Dømeord</i>	
1.	'banan, 'tomat (alle monomorfematiske lånord)
2.	'musikk, 'stasjon (tostava, potensielt komplekse ord)
3.	'inspirera, 'referera (-era-verb)
4.	'svarthest'n, 'nybil'n (modifiserande adjektiv + substantiv, i bundne kontekstar)
5.	'tok jo ut, 'ga itj bort (partikkelverb med lette adverb mellom verb og partikkel)
6.	'tok han ut, 'kjøpte ut (partikkelverb, med trykksvake pronomen mellom verb og partikkel, og fleirstava verb)
7.	'tok ut, 'skreiv opp (einstava verb fylgde av partikkel)
8.	Samansetjing 1: [grå[fluge-snappar]], [svart[raud-stjert]] (artsnamn med adjektivisk modifikasjon av komplekst hovud)
9.	Samansetjing 2: [ørken[steinskvet]], [gul-sporv] (substantivisk modifikasjon av komplekst hovud, adjektivisk modifikasjon av simplekst hovud)

Tabell 1: Hierarki over dialektal toleranse for fyrstestavingstrykk.

ut til å ha spreidd seg utover det austlandske området dei seinare tiåra, sjå Torp 1997).

Me reknar i denne artikkelen med at kontekstane over har noko sams, trass i at dei syntaktisk, morfologisk og leksikalsk skil seg åt kraftig. Utan å føreslå ein noggrann analyse her, reknar me med at der i fleire dialektar finst ein sterk tendens til å leggja trykk på fyrstestavinga i prosodiske domene som inneheld meir enn éi trykksterk staving. I andre dialektar held ein seg unna dette, og trykket hamnar i staden på siste trykksterke staving, som i ein frase. I moderne svensk har dette vorte drive so langt at adjektiv i samansetjingar (her: artsnamn) nærast obligatorisk vert realiserte som syntaktisk frittståande modifikatorar, noko som fører til frasetrykk. Som me påpeikar over, finst der likevel variasjon i fyrstestavingsdialektane: I partikkelverb ser me at partikkelen iblant ber trykk, eller at både partikkel og verb ber trykk; ved adjektivinkorporering ser ein variasjon mellom inkorporering og frasestruktur, og det same gjeld ved samansette artsnamn (t.d. *'gråflugesnappar* – *'grå'flugesopp*). I openbert syntaktisk eller morfologisk samansette kontekstar er det ofte vandt å avgjera om «variasjonen» er regelstyrt, eller om han er meir eller mindre fri: Korrelerer den prosodiske variasjonen med semantisk eller pragmatisk tolking, eller syntaktisk struktur? Når det gjeld partikkelverb, so veit me at fyrstestavingstrykk er mogeleg i fleire ulike syntaktiske og morfologiske kontekstar, som i (5a–d):

- (5) a. Han skal ['skriva ut] to sider. (Verb og Part i VP)
 b. Han ['skreiv ut] to sider. (Verb i V2-posisjon, Part i VP)
 c. I går ['skreiv han ut] to sider. (Verb i V2, Part i VP, trykksvakt subjektspronomen imellom)
 d. I går ['skreiv han dei ut]. (Verb i V2, Part i VP, trykksvake subjekts- og objektspronomen imellom)

Jamvel om samansetjingstrykk er mogeleg når der finst ein viss syntaktisk avstand mellom verb og partikkel (verb i C⁰, partikkel i VP, 5b) og til og med når andre ledd står mellom verb og partikkel (5c–d), so veit me frå korpusstudiar (Tengesdal mfl. 2021) at samansetjingstrykk meir sjeldan vert realisert i kontekstar som (5c–d), enn i (5a). Me finn

likevel ingen skilnad mellom den syntaktiske konteksten i (5b) jamført med (5a) (dvs., ingen effekt av strukturell avstand). Me finn au eit kategorisk fråvære av samansetjingstrykk når ledd med leksikalsk spesifisert trykk står mellom verb og partikkel. Dette gjeld stort sett alle ledd utanom trykksvake pronomen og, i visse dialektar, lette partikkel-liknande adverb (*jo, no*) og enklitisk negasjon (dette gjeld ikkje typisk austnorsk *-kke* (*ha-kke*), men «trøndsk/hedmarks» *itj*), sjå (6).

- (6) a. *I går ['skreiv Mai ut] to sider.
 b. *Mai ['skreiv ikkje ut] to sider.

Samansetjingstrykk forsvinn au når anten verbet eller partikkelen er kontrastiv (*han gjekk 'ut, han kom ikkje 'inn; han 'køyrd inn, han 'flaug ikkje inn*), og au ofte når partikkelen står etter modalverb (*skal 'ut, må 'ut*), men det er noko uklårt om dette har med informasjonsstruktur eller kategori å gjera. Den morfofonologiske forma til verbet synest au å påverka høvet for samansetjingstrykk: Samansetjingstrykk er vanlegare med einstava verb (*gje opp, ta ut*), og underliggjande tostavta verb gjennomgår ofte apokope i samband med samansetjingstrykk (*vaska opp* → *'vask opp*). Me ser au tendensar til at samansetjingstrykk er vanlegare når verb + partikkel har ei høgst leksikalisert tolking (*skjella ut*) jamført med partiklar med ei rein direksjonell tolking (*køyra ut*), og framfor alt då om den direksjonelle partikkelen er fylgd av ein preposisjonsfrase (*køyra ut til hytta*). Det er ingenting i våre data (Tengesdal mfl. 2021) som tyder på at desse gruppene er kategorisk åtskilde; me finn trass alt ei overvekt av samansetjingstrykk uavhengig av om tolkinga er direksjonell eller ikkje-direksjonell.

Kortfatta ser me i minsto ein streng restriksjon på samansetjingstrykk: eit trykksterkt ledd kan ikkje stå mellom verb og partikkel (6a–b). Andre restriksjonar verkar vera meir probabalistiske, men det er vandt å avgjera om einskilde individ har kategoriske system, med tanke på kor mange faktorar som potensielt påverkar trykkmønsteret. Som me skal få sjå nedan, kan ein i talemålskorpus finna partikkelverb utan samansetjingstrykk i dialektar og kontekstar der ein sterkt ventar seg dette. Føremålet med denne granskinga er å jamføra variasjon i trykkmønster i partikkelverb med eit mindre komplekst fe-

nomen, nemleg trykk på fyrstestavinga i latinske ord (1–3 i tabell 1). Me ventar oss ikkje at syntaktiske og semantiske drag påverkar trykk i same utstrekning her, med di me held oss innom orddomenet. Mogeleg variasjon i trykkmønster i desse falla bør difor skuldast ikkje-strukturelle eller sosiolingvistiske faktorar. Om me går tilbake til hierarkiet i tabell 1, so er det nok openbert visse sosiolingvistiske val av trykkmønster her. Til og med lingvistar uttrykkjer gjerne at der finst noko markert, eller kanskje til og med løye, med å leggja trykk på fyrstestavinga i *banan*. Denne oppfatninga verkar vera mindre sterk for latinske ord, og vantar heilt ved partikkelverb. Men lingvistar i vår nærleik har uttrykt at trykk på fyrstestavinga i partikkelverb er noko som høyrer til ein kvardagsleg samtalekontekst. Kristoffersen (2000, 289) nemner au dette, og syner til Alnæs si skildring (1916, 98) om at fyrstestavingstrykk i partikkelverb raskt fekk fotfeste i austnorsk bymål på 1900-talet. Western (1921, § 454) opplyser vidare at fyrstestavingstrykk rådde i austnorske dialektar, men at mønsteret au vann innpass i den danna daglegtaleten i dåtida.

3 Denne granskinga og bakgrunn

3.1 Føremålet med granskinga

Føremålet med denne studien er dels å granska korrelasjon i trykkmønster mellom partikkelverb (5–7 i tabell 1) og latinske ord (1–3 i tabell 1). Det store spørsmålet bak studien er om variasjonen me ser innom ein dialekt med omsyn til partikkeltrykk, kan fangast av eitt sett med strenge reglar (og/eller ‘constraints/vilkår’), eller om reglane til ein viss del er variable. I tidlegare studiar har me funne at der finst sterke preferansar for ulike trykkmønster i ulike dialektområde, men at der framleis finst mykje variasjon. I Tengesdal mfl. (2021) granskar me trykkmønster frå område innom tre fylke (Trøndelag, Buskerud (nord) og Finnmark) i NDK. I Trøndelag finn me ein sterk preferanse for trykk på fyrstestavinga: 223 av 293 (76 %) granska partikkelverb har trykk på fyrstestavinga (‘samansetjingstrykk’). Ein del av falla der samansetjingstrykk vantar, kan forklarast av strenge syntaktiske faktorar: eitt aksentuert ledd, t.d. ein nominalfrase (oftast subjekt)

eller eit setningsadverbial, står mellom verb og partikkel (*då skreiv læraren ut tre sider, læraren skreiv faktisk ut tre sider*, jf. 6a–b), men sjølv når ein ser bort frå desse falla, står framleis ein heil del variasjon att. Ein del av den vidare variasjonen kan forklarast av semantiske eller syntaktiske faktorar. I tabell 2 for Trøndelag gjev me ei førebels oversikt over trykkmønster i tre ulike syntaktiske/semantiske mønster når verb og partikkel står like ved einannan, eller når høgst eitt trykksvakt pronomen står mellom verb og partikkel: (1) metaforisk/ikkje-transparent tolking av verb-partikkel-kombinasjon (*han ga opp*), (2) partikkel med transparent direksjonell tolking (*han gjekk ut*), og (3) partikkel med transparent direksjonell tolking fylgd av ein preposisjonsfrase (*han gjekk ut til hunden*). (Rada «Andre» inneheld fall som av ulike årsaker er vanskelege å klassifisera, t.d. grunna bakgrunnsstøy.)

	<i>Ikkje-trans.</i> N=159	<i>Dir.</i> N=72	<i>Dir. +PP</i> N=20	<i>Andre</i> N=14	<i>Totalt</i> N=265
<i>Samansetjing</i>	91 %	72 %	45 %	57 %	81 %
<i>Partikkel</i>	5 %	11 %	25 %	21 %	9 %
<i>V+partikkel</i>	1 %	7 %	15 %	15 %	4 %
<i>Deaksentuert</i>	2 %	6 %	0 %	0 %	3 %
<i>Andre</i>	1 %	4 %	15 %	7 %	3 %
<i>Totalt</i>	60 %	27 %	8 %	5 %	100 %

Tabell 2: Oversikt over variasjon i trykkmønster i tre ulike syntaktiske/semantiske mønster for Trøndelag, basert på førebelse resultat frå pågåande arbeid (Tengesdal mfl. 2021). (V = verb; Ikkje-trans. = ikkje-transparent; Dir = direksjonell; PP = preposisjonsfrase)

Me ser ein openberr effekt av partikkeltype i resultatata: Samansetjingstrykk er nærast obligatorisk i ikkje-transparente fall (91 %), medan delen samansetjingstrykk søkk i dei fylgjande to kategoriane. Dette er au i tråd med den syntaktiske analysen i Aa (2020). Resultata tyder likevel på at reglane er variable, i minsto på fylkesnivå. I prinsippet kan det syna seg at resultatata er meir kategoriske på individnivå, eller bygdnivå, eller at variasjonen er styrt av finare distinksjonar, men det kan i prinsippet finnast andre faktorar som me ikkje kjenner

til. I denne granskinga jamfører me partikkeltrykk med trykk i latinske ord. Finn me variasjon innom dessa orda, kan me dra slutninga at der ligg variable reglar bakom distribusjonen av dei to formene.

For å granska variasjon i trykkmønster i tid og rom, treng ein ljod-innspelingar frå ulike dialektområde frå ulike tidspunkt, eller frå talarar av ulike aldrar. LIA-korpuset i kombinasjon med NDK (Johannessen mfl. 2009) utgjer altso eit framifrå materiale for ein prosodisk variasjonsstudie, der ein kan fylgja utviklinga i tid og rom. Søkjegrensesnittet legg au godt til rette for å finna høvande døme, og å granska ljodspor akustisk direkte.

3.2 Trykkplasseringsutvikling frå gamalnorsk fram til notida

Kristoffersen og Torp (2016, 107–109, 152–157) legg fram den prosodiske utviklinga frå gamalnorsk fram til notida. For urnordisk kan trykktilordninga forklarast med morfologisk forankring, ved at trykk var fast plassert på rotstavinga. Men frå og med gamalnorsk fangar ein opp generaliseringar ved å nytta ei fonologisk forankring, der trykk var fast plassert på fyrstestavinga, uavhengig av morfologisk status. Dei skriv at sume moderne språkhistoriske framstillingar hevdar at det gamalnorske systemet med fyrstestavingstrykk framleis står ved lag (2016, 152). Men ein slik analyse føreset at ein har eit kategorisk skilje mellom arveord og lånord av t.d. ikkje-germansk opphav, der forklaringa på sistestavingstrykk i ord som *karamell* er at trykkplasseringa er innlånt saman med ordet.

For å unngå dette skiljet, gjev dei ein alternativ analyse som tek utgangspunkt i høgre kant av det prosodiske ordet, og at trykket i norske monomorfematiske ord fell på éi av dei tre siste stavingane i ordstammen. På den måten kan ein au fanga opp trykkplasseringa i germanske arveord som ²*kasta*, som både kan analyserast med trykk på nest siste staving, eller med trykk på fyrste staving. Om ein legg til grunn at sume dialektar kan veksla mellom trykk på sistestaving ([kar.ra.ˈmɛl]) eller på fyrstestaving ([²kar.ra.ˌmɛl]), so føreset analysen deira at det er sistestavingstrykket som er det underliggjande. Eit ord med sistestavingstrykk kan ikkje vekslast til fyrstestavingstrykk dersom det har underliggjande tonelag 2 i den aktuelle dialekten (difor kan trykkplasseringa i *karamell* variera, men ikkje i

metode), og det føreset oftast tyngd i trykkstavinga. Kristoffersen og Torp (2016, 157) meiner at ein kan generalisera ein heilskapleg trykk-analyse til alle norske dialektar, og skriv at vekslinga me observerer i austlandske dialektar, då kan forklarast med ein ekstra, variabel regel som under visse vilkår legg hovudtrykk og tonelag 2 på fyrstestavinga. Rett nok føreset denne analysen eit brot frå det gamalnorske trykksystemet med initialtrykk, men ved hjelp av vilkåra kan ein vidareføra det gamle systemet indirekte, som elles har gått tapt i norsk. Eit interessant spørsmål som spring ut av vår gransking, er om funna våre stør analysen deira, eller om det framleis er aktuelt å rekna med fyrstestavingstrykk som underliggjande, og leksikalsk spesifiserte unntak i austnorsk og trøndske dialektar.

4 Materiale

For å finna ord med potensielt trykksterke avleiingssuffiks i LIA-korpuset og NDK, nytta me dei to CQP-søkjestrengane i (7) i korpusgrensesnittet Glossa (Nøklestad mfl. 2017).

- (7) a. [lemma=".*itet|.*ist|. *ant|. *ent|. *eri|. *isme|. *ikk|. *ment|læ-rarinne|venninne" & lemma!=".*kant|. *kvist|. *blikk"%c & ((pos="subst"))]⁵
 b. [lemma=".*ere" & lemma!="vere|gjere|skjere|bere|avgjere"%c & ((pos="verb"))]

Søkjestrengen i (7a) finn dei vanlegaste trykksterke nominaliserande suffiksa (*kval-itet*, *but-ikk*, *pros-ent* mfl.), og sorterer bort dei vanlegaste uynskte treffa (t.d. *øye-blikk*, som elles stemmer med det regulære uttrykket «lemma=".*ikk"»). Søkjestrengen er basert på suffiks som kjem fram av Norsk referansegrammatikk (Faarlund, Lie og Vannebo 1997, 97–112). Den andre søkjestrengen finn alle verb med det trykksterke verbaliseringssuffikset *-era* (*studera*, *interessera*, *dis-*

⁵ Søkjestrengen for NDK var litt ulik, m.a. måtte «pos="subst"» endrast til «pos="noun"».

kutera osv.). Me får ei høvesvis stor mengd treff frå desse søkjestregane, over 11 000. I tabell 3 gjev me ei oversikt over korleis desse fordelar seg over fylka og dei to korpusa, og innom NDK i dei to ulike aldersgruppene som deltek.

<i>Fylke</i>	<i>NDKu</i> <i>Verb</i>	<i>NDKe</i> <i>Verb</i>	<i>LIA</i> <i>Verb</i>	<i>NDKu</i> <i>Subst.</i>	<i>NDKe</i> <i>Subst.</i>	<i>LIA</i> <i>Subst.</i>	<i>Totalt</i>
<i>Akershus</i>	24	36	30	35	37	38	200
<i>Aust-A.</i>	38	14	128	41	42	205	468
<i>Buskerud</i>	28	30	78	55	70	107	368
<i>Finnmark</i>	48	64	458	96	112	485	1 263
<i>Hedmark</i>	25	34	59	64	87	82	351
<i>Hordal.</i>	36	55	461	73	140	1 106	1 871
<i>M. og R.</i>	65	38	124	128	122	149	626
<i>Nord-T.</i>	31	28	104	53	52	86	354
<i>Nordland</i>	90	82	178	139	133	264	886
<i>Oppland</i>	34	24	102	52	73	82	367
<i>Rogaland</i>	31	38	121	107	81	109	487
<i>S. og F.</i>	69	23	97	53	71	88	401
<i>Sør-T.</i>	49	66	292	100	107	299	913
<i>Telemark</i>	21	61	122	37	27	90	358
<i>Troms</i>	50	75	427	119	105	380	1 156
<i>Vest-A.</i>	24	25	113	45	59	153	419
<i>Vestfold</i>	20	12	5	33	45	2	117
<i>Østfold</i>	17	23	99	46	68	146	399
<i>Totalt</i>	700	728	2 998	1 276	1 431	3 871	11 004

Tabell 3: Alle korpustreff i LIA og NDK frå søkjestregane som elisiterer dei vanlegaste trykksterke nominaliseringssuffiksa og verb med det trykksterke verbaliseringssuffikset *-era*. (NDKu = yngre talarar i NDK; NDKe = eldre talarar i NDK; Subst. = substantiv; Aust-A. = Aust-Agder; Hordal. = Hordaland; M. og R. = Møre og Romsdal; Nord-T. = Nord-Trøndelag; S. og F. = Sogn og Fjordane; Sør-T. = Sør-Trøndelag; Vest-A. = Vest-Agder)

I prinsippet er dette materialet meir enn nok for både ei detaljert synkron og diakron skildring av norske trykkmonster, og materialet er truleg stort nok til å fylla ei heil doktoravhandling. Måla våre i denne artikkelen er mindre, og me kjem berre til å gjeva ei kort skildring av tre fylke: Trøndelag, Hedmark og Finnmark.

Resultata frå dei latinske orda jamfører me med trykkmønster i partikkelverb frå same tre fylke. For å finna partikkelverb, har me nytta same metode som Larsson og Lundquist (2014): Me har søkt på dei fire partiklane *ut*, *inn*, *opp* og *ned*. Me syner m.a. til Aa (2015, 74–75, 2020) for klårgjering av metodiske føremonar. Ved å søkja på desse partiklane, slepp ein unna alle dei irrelevante treffa ein får ved å søkja på transitive preposisjonar (Aa 2020, 44, Larsson og Lundquist 2014, 106), og fordelinga mellom direksjonelle og ikkje-direksjonelle partiklar er ganske jamn (Aa 2020, 48). Før den prosodiske analysen har me annotert den syntaktiske strukturen hjå partikkelverba: Om verb og partikkel står direkte ved sida av einannan, om der finst intervenerande ledd (subjekt, objekt, adverb), og om dei intervenerande ledda er trykksterke (t.d. full NP i staden for trykksvakt pronomen). Me har ekskludert setningar der *ut*, *inn*, *opp* og *ned* ikkje er ein verbpartikkel, men som i staden modifierer eit substantiv eller fungerer som sjølvstendig adverbial, t.d. «*lenger opp gata budde ein mann*», og «*litt lenger ut ligg ei øy*». Me har au ekskludert koordinerande partiklar, av typen «*han sprang opp og ned heile dagen*». I Tengesdal mfl. (2021) gjev me ein meir systematisk analyse av variasjonen i trykkmønster i partikkelverb i Trøndelag, Finnmark og Buskerud. Grunna tids- og plassbrest, gjev me berre ein høgst oversiktleg analyse av partikkelverba i denne artikkelen. Den totale mengda partiklar av typen *ut*, *inn*, *opp* og *ned* i LIA og NDK er fleire tusen berre i dei tre aktuelle fylka. Me fokuserer her i staden på større mønster, men drøftar au variasjon innom individ der me finn det.

Me nytta Glossa i samband med den akustiske analysen. Glossa har innlema ein funksjon som genererer akustiske bylgjeformer og spektrogram som høyrer til korpusegmenta. Grunnfrekvensen (f_0) og andre akustiske parameter er au tilgjengelege, slik at det er råd å gjera ein prosodisk analyse. Me har opna dei relevante korpustreffa i Glossa og annotert prosodisk informasjon i alle falla basert på auditorisk analyse, og i tvilsfall i tillegg ved hjelp av visuell analyse av grunnfrekvensen. Både for dei latinske orda og partikkelverb har me annotert kva for staving som er aksentuert, kva for tonelag ho har; og for partikkelverba, kor langt det prosodiske ordet strekk seg (verb, eventuelle trykksvake element, og partikkel; jf. ‘postleksikalsk tone-

lag 2', sjå t.d. Myrberg og Riad 2015, 127–128, Riad 2018, 352–353). Ut ifrå dette har me utleidd stavinga med hovudtrykket frå den samsvarande aksentuerte stavinga i det prosodiske ordet. Det å avgjera trykk og aksent på einskilde ord i spontantale er ikkje alltid lett, og iblant har me vore uvisse, og notert uvisse med spørjeteikn, eller notert ljodklypp som «uklårt». I Finnmark-dialektane som manglar tonelagsopposisjon, er det meir problematisk å kategorisera, difor har me i tillegg notert om partikkelen er uttala med kvantitetsopposisjon i hald eller ikkje, der dette er mogeleg. Sjølv om grensesnittet legg føringar for kor noggrann og viss prosodisk analyse ein kan utføra, meiner me likevel at både LIA og NDK er godt eigna for å granska prosodisk variasjon og endring for vårt føremål.

5 Resultat: latinske ord

Ved hjelp av søkjestrengane attgjevne i avsnitt 4, har me elisitert til saman 742 korpustreff frå utvalde område i Finnmark, Hedmark og Trøndelag. Me har ekskludert til saman 88 treff av desse; i dei fleste falla handlar det om samansetjingar som inneheld eit latinsk ord (t.d. *sportsjournalist*, *kvalfabrikk*), og, i mindre grad, fall av di tekstsegmenta i korpuset ikkje passa med ljodfilene. Dei resterande 654 treffa kjem frå til saman 123 ulike talarar, frå 32 ulike opptaksstader. I tabell 4 listar me opp dei ulike stadene, fordelt på korpus og fylke.

<i>Fylke</i>	<i>LIA</i>	<i>NDK</i>
<i>Hedmark</i>	Trysil (39), Åsnes (19), Engerdal (21), Åmot (9), Follidal (7), Nord-Odal (7), Ljørdalen (6), Tolga (8), Drevsjø, Rendal, Sør-Odal, Vang, Vinger	Stange (37), Alvdal (34), Kirkenær (34), Rena (24), Dalsbygda (21), Trysil (20), Drevsjø (18)
<i>Trøndelag</i>	Skogn (40), Verdal (28), Høylandet (10), Harran, Leksvik, Nordli, Otterøy, Overhalla	Bjugn (78), Inderøy (32)
<i>Finnmark</i>	Hammerfest (39), Sørøysund (29), Gamvik (17)	Kjøllefjord (37), Hammerfest (28)

Tabell 4: Oversikt over fylka og stadene korpustreffa skriv seg frå (færre enn fem treff er ikkje spesifiserte i parentesar).

I analysen under kjem me fyrst til å slå i hop data frå LIA og NDK, og deretter sjå nærare på einiskilde område/stader/talarar. I analysen av trykk har me koda trykk på fyrstestavinga som «fyrst», og trykk på avleiingssuffikset som «sist» (dette gjeld au fall der trykket fell på den nest siste stavinga, som i *refer'ere* og *ven'ninna*). Me gjev ei oversikt over resultatata i tabell 5.

'σ	Hedmark	Trøndelag	Finnmark	Totalt
<i>Fyrst</i>	202	159	0	361
<i>Fyrst?</i>	7	4	2	13
<i>Sist</i>	89	17	139	245
<i>Sist?</i>	10	10	3	23
<i>Uklårt</i>	2	4	6	12
<i>Totalt</i>	310	194	150	654

Tabell 5: Oversikt over trykkplassering i dei latinske orda, med data både frå LIA og NDK.

I tabell 5 ser me at det frå Finnmark ikkje finst nokre klåre døme på fyrstestavingstrykk. Me finn au få døme på fyrstestavingstrykk i Trøndelag, medan Hedmark er meir blanda, men med ei klår overvekt av fyrstestavingstrykk. For å få ei betre oversikt tek me i tabell 6 bort dei uklåre falla (medrekna dei som er markerte med spørjeteikn), og deler inn resultatata i substantiv og verb (der 'verb' altso berre er ord som er avleidde med *-era*).

	Hedmark	Trøndelag	Finnmark
<i>Substantiv, fyrst</i>	68 % (N=188)	83 % (N=100)	0 % (N=75)
<i>Verb, fyrst</i>	73 % (N=103)	100 % (N=76)	0 % (N=64)

Tabell 6: Oversikt over fyrstestavingstrykk i latinske substantiv og verb, utan uklåre fall.

Finnmark har som sagt kategorisk trykk på suffikset, men i tabell 6 ser me au at trykket i *-era*-verba i Trøndelag kategorisk hamnar på fyrstestavinga. I Hedmark ser me au ein tendens til meir fyrstestavingstrykk på *-era*-verb enn på substantiv. Skilnaden mellom verb og

substantiv kan skuldast ulike faktorar. Det kan tenkjast at nokre ein-skilde ord eller suffiks har bytt status (frå fyrste- til sistestavingstrykk), og det er au mogeleg at nokre talarar har skifta trykkmonster generelt (og at desse av einkvan grunn ikkje produserte nokre *-era*-verb).

Om me fokuserer på Trøndelag, ser me i (8) at ein stor del av suf-fiksa i vår søkjstreng er representerte blant dei 17 falla med sistes-tavingstrykk, og at få ordstammar eller *-former* førekjem meir enn éin gong (med unntak av *prosent*, som førekjem tre gonger med sis-testavingstrykk, og *butikk* er med, både som ubunden eintal og bun-den fleirtal).

(8) **-ist:** *adventistar, artistar, turistar*; **-ikk:** *butikk, butikkane, fab-rikkar, musikk*; **-itet:** *autoritetane, universitetet*; **-ment:** *depar-tementet*; **-ent:** *prosent*; **-ant:** *sersjanten, konfirmantar*; **-inne:** *venninne*

Me ser ikkje nokon sterk effekt av korpus, alder eller talarar. Dei fleste sistestavingstrykka finst hjå dei talarane som snakkar mykje: talaren *bjugn_15* produserer sistestavingstrykk seks av 35 gonger, og *verdal_ntnu_0301* produserer sistestavingstrykk tre av 17 gonger.

I Hedmark spelar opptaksstad ei større rolle (det er ikkje løye, med tanke på at me ser på data frå eit større område her). På nokre stader finn me ei sterk overvekt av sistestavingstrykk: Trysil (35/48), Ljørdalen (4/6), og Folldal (7/7). Om me ser bort frå desse tre sta-dene, ser me at mønsteret ser ut som i Trøndelag. Samanlagt er 38 av 230 fall med sistestavingstrykk. Me ser meir viss systematikk her: Suffikset *-inne* får alltid trykk (samanlagt 12 fall, sjå Kristoffersen og Torp 2016 for diskusjon); stammen *journalist* førekjem berre med sistestavingstrykk (fire gonger). (9) gjev alle ordformer som dukkar opp med sistestavingstrykk (treff frå Trysil, Ljørdalen og Folldal er tekne bort her):

(9) **-ist:** *bassisten, bilister, journalist, journalistane, journalistar, journalister, turister*; **-isme:** *turisme*; **-ikk:** *butikken, musikk, mu-sikken*; **-itet:** *aktivitet, aktiviteter, kvalitetar*; **-ment:** *engasje-ment*; **-ent:** *prosent*; **-ant:** *elefant*; **-inne:** *lærerinne, venninna,*

venninne, venninnene, venninner; -era: benifiserte, levere, risikere, separerte

Der finst tre stader der me berre fann fyrstestavingstrykk: Åmot, Åsnes, og Odal. Fyrstestavingstrykk verkar au å vera stabilt, men til ein viss del variabelt i Alvdal (26/30), Kirkenær (32/34) og Stange (29/34). Større variasjon finst i Dalsbygda, Drevsjø og Rena. Det lyt nemnast at sjølv i Trysil produserer flestalle talarar fyrstestavingstrykk minst éin gong – unnataka er talarar som produserer tre eller færre ord i våre korpustreff.

Me kan also dra slutninga at fyrstestavingstrykk finst hjå alle talarane innom Hedmark og Trøndelag. Det er klårt meir uvanleg i området kring Trysil, og kan henda obligatorisk i områda i Trøndelag. Men der finst inkje i materialet som tyder på at trykk på avleiings-suffikset *ikkje* førekjem i nokre dialektar, i minsto som eit markert alternativ. I dialektområda med trykk på sistestavinga (Nord-Noreg/Finnmark, Vest-Noreg, Sør- og Midt-Sverige), verkar derimot fyrstestavingstrykk i latinske ord å vera heilt fråverande.

6 Jamføring med partikkelverbresultat

For dei latinske orda koda me for to ulike trykkmønster: trykk på fyrstestavinga, eller trykk på sistestavinga, der ‘sistestavinga’ heilt enkelt tyder ‘avleiingssuffikset’ (som i visse fall kan omfatta ei trykksterk staving fylgt av ei trykklett staving). For partikkelverb er trykkmønsteret meir samansett, men i (10) er to mønster som direkte samsvarar trykkmønsteret i latinske ord.

- (10) a. Samansetjingsaksent – hovudtrykk på fyrstestavinga (*'gå-ut*)
 b. Partikkelaksent med deaksentuert verb – hovudtrykk på sistestavinga (*gå 'ut*)

I tillegg til desse mønstera, finn me au fall der både verb og partikkel ber eigne trykk (*'gå 'ut*); dette ser me til vanleg når verb eller partikkel

kel er kontrastive, men der kan vera andre årsaker til dette mønsteret. I tillegg ser det ut til å finnast fall der berre verb ber trykk, og der partikkelen ikkje inngår i same aksenteining som verbet ('gå ut). Nedan går me gjennom utvalde stader frå korpuset, og korrelerer trykk i latinske ord med partikkelverbtrykk.

6.1 Trøndelag

Der er fire stader i Trøndelag som har relativt mange fall av partikkelverb og latinske ord. Desse er Bjugn (NDK), Inderøy (NDK), Skogn (LIA) og Verdal (LIA). I tabell 7 jamfører me trykkmønster i båe konstruksjonane. 17 uklåre fall av latinske ord og åtte uklåre fall av partikkelverb er ikkje tekne med i tabellen eller drøftinga.

Stad	Latinske ord			Partikkelverb			
	N	Fyrst	Sist	N	Verb	Part.	Båe
Bjugn	68	90 %	10 %	58	84 %	12 %	3 %
Inderøy	26	100 %	0 %	51	82 %	6 %	12 %
Skogn	39	95 %	5 %	76	84 %	7 %	9 %
Verdal	27	85 %	15 %	54	89 %	4 %	7 %

Tabell 7: Jamføring av trykkplassering i latinske ord og partikkelverb i Trøndelag.

Me ser at trykk på partikkelen er uvanleg, og at samansetjingsaksent heilt klårt er det vanlegaste mønsteret (altså trykk på verbet). Samanlagt ser det likevel ut som der er meir fyrstestavingstrykk i latinske ord enn verbtrykk i partikkelverb. Det lyt nemnast at syntaktiske og leksikalske faktorar iblant styrer trykkmønsteret her. Til dømes finn me ofte partikkeltrykk ved hjelpeverb/modalverb («*jeg må ned på byen*», 4/5 av partikkeltrykka i Skogn). Hovudtrykk på partikkelen i kontekstar der me ventar oss samansetjingsaksent dukkar likevel opp på alle opptaksstader, sjå (11–13).

- (11) han brukte nå å vere med hjelpe meg å **ta opp** potet om haustane han (verdal_ntnu_0101)
- (12) det var for # liksom bokmålet var # **sett ned** på [latter] (skogn_ntnu_0610)

- (13) og så var det en dag at sauene hadde **brøyta seg ut** # av trøa (ineroey_04gk)

Trykk på både verb og partikkel ser me oftast når der er eitt eller fleire element som står mellom verb og partikkel. I (14–18) ventar me oss i prinsippet samansetjingsaksent i alle falla utanom (15), men trykk hamnar både på verb og partikkel.

- (14) der **flytta vi inn** i august (ineroey_03gm)
 (15) ja det **høres jo litt sånn ut** (bjugn_23)
 (16) så **kom han inn** på Sessilvollen da hos tantene mine (verdal_ntnu_0101)
 (17) da blir dei vel e må **trekke ut** av bygda ja (skogn_ntnu_0703)
 (18) dei må **reise ut** ja # for e det er lite av det (skogn_ntnu_0705)

Som venta finn me samansetjingsaksent når trykksvake pronomen og til og med kortare adverbial står mellom verb og partikkel, som i (19–22).

- (19) **eter nå opp** alle sånne løker (ineroey_04gk)
 (20) så **for vi inn** i Risvika og svømte der (bjugn_15)
 (21) så du veit det # kvaa **kokast nå opp** ho òg da i lag med feittet så ho vart nå ikkje hard da ho vart nå mjuk (verdal_ntnu_0101)
 (22) ho **slepper meg inn** ja (skogn_ntnu_0804)

Oppsummeringsvis finn me altso trykk på fyrstestavinga i latinske ord i om lag like høg grad som samansetjingstrykk på verb i partikkelverb, når me tek omsyn til at visse leksikalske og syntaktiske faktorar hindrar samansetjingstrykk. Men det er viktig å minnast at dagsens «standardmønster», altso trykk på sistestavinga i latinske ord og på partikkel, au finst i Trøndelag, i underkant av 10 % av falla.

6.2 Hedmark

Hedmark er det dialektområdet der me finn mest variasjon når det gjeld trykkmønster i latinske ord, og difor ventar me oss au variasjon i partikkelverb. Variasjonen er i stor grad styrt av stad, og me finn stader der sistestavingstrykk er det dominerande mønsteret (Trysil), og andre stader der fyrstestavingstrykk i varierende styrke dominerer (frå t.d. Kirkenær til Stange). Me fokuserer på opptak frå fem stader i Hedmark: Kirkenær (NDK), Stange (NDK), Engerdal (LIA), Rena (NDK) og Trysil (LIA+NDK). Ti uklåre fall av latinske ord og to uklåre fall av partikkelverb er ikkje tekne med i tabellen eller drøf-tinga.

<i>Stad</i>	<i>Latinske ord</i>			<i>Partikkelverb</i>			
	<i>N</i>	<i>Fyrst</i>	<i>Sist</i>	<i>N</i>	<i>Verb</i>	<i>Part.</i>	<i>Båe</i>
<i>Kirkenær</i>	34	94 %	6 %	47	70 %	11 %	19 %
<i>Stange</i>	34	85 %	15 %	59	73 %	2 %	25 %
<i>Engerdal</i>	21	81 %	19 %	61	57 %	8 %	34 %
<i>Rena</i>	23	65 %	35 %	50	76 %	0 %	24 %
<i>Trysil</i>	53	25 %	75 %	100	33 %	27 %	40 %

Tabell 8: Jamføring av trykkplassering i latinske ord og partikkelverb i Hedmark.

I tabell 8 ser me ein heil del variasjon i trykkmønster. Som venta av resultatata for latinske ord, er det framfor alt Trysil som skyr trykk på verb fylgt av deaksentuert partikkel (samansetjingsaksent). Men me ser at det ikkje absolutt er det typiske partikkeltrykket som tek over etter samansetjingsaksenten. Snarare er det trykk både på verb og på partikkel som vert eit vanlegare mønster. Det er ikkje heilt klårt kva for faktorar som ligg bakom valet av trykkmønster, men me ser ein klår tendens til at «dobbeltrykk» ofte oppstår når verb og partikkel vert åtskilde av eit subjekt eller eit objekt, der desse anten får eigne trykk (NP-ar), eller vert enklitiserte til verb og/eller partikkel (pronomen og lette adverb). Dette skjer til ein viss grad i Trøndelag au, men i monaleg lågare grad. Om me berre fokuserer på fall der grammatikken i prinsippet tillèt samansetjingsaksent med intervenue- rande ledd (lette subjekts- eller objektspronomen), finn me berre 10–

15 % samansetjingsaksent i Hedmark-materialet. Det kan tenkjast at (delar av) Hedmark har utvikla eit system der der berre finst samansetjingsaksent frå nivå 7 (jf. tabell 1). I Trøndelag finn me 50–60 % samansetjingsaksent i same kontekst. Jamvel om me kan føreslå syntaktiske, morfologiske og kan henda leksikalske klårgjeringar til ein del av mønster, er det likevel klårt at der finst ein heil del variasjon innom talarar, både for latinske ord og partikkelverb.

6.3 Finnmark

Som me såg i avsnitt 5, fann me ingen sikre fall med fyrstestavingstrykk i latinske ord frå (Vest-)Finnmark. Me ventar oss då berre partikkeltrykk i partikkelverb, om der no finst ein korrelasjon med trykk i latinske ord. Men resultatane er meir innvikla enn som so, og me har merka under annoteringsarbeidet med partikkelverba i Finnmark at både trykk og aksentmønster er vanskelege å oppfatta, utan at me au utfører meir noggranne prosodiske granskingar. Dette grunnar seg truleg på at dei aktuelle finnmarksdialektane saknar tonelagsdistinksjon, noko som gjer at den typiske «samansetjingsaksenten», som vert realisert med tonelag 2 i austnorske og trøndske partikkelverb, synest å vanta. Til denne artikkelen har me fått tid til å analysere 58 treff frå LIA-korpuset (Hammerfest, Sørøysund og Gamvik), og i tidlegare arbeid (Tengesdal mfl. 2021) har me analysert 116 treff frå NDK (Hammerfest og Kjøllefjord). Tabell 9 syner førebelse funn av trykkmønster frå LIA- og NDK-materialet. Grunna vantande tonologisk analyse er det uklårt om der er grunnlag for eit kategorisk skilje mellom gruppene ‘samansetjingsaksent’ og ‘verbaksent’. Førebels har me difor kategorisert desse som trykk på ‘verb’. Ni uklåre fall er ikkje tekne med i tabellen.

	<i>N</i>	<i>Verb</i>	<i>Part.</i>	<i>Båe</i>	<i>Deaks.</i>
<i>LIA</i>	54	46 %	46 %	7 %	0 %
<i>NDK</i>	111	40 %	40 %	20 %	1 %

Tabell 9: Førebelse funn med fordeling av trykkmønster i partikkelverb i Finnmark.

Me ser at partikkeltrykk er jamfrekvent med trykk på verbet. For det meste verkar ikkje verb og partikkel å måtta danna éi aksenteining, slik det gjer i store delar av Trøndelag og Hedmark. Ein del fall som ofte er vanskelege å avgjera i vårt materiale, har eit fokusert element rett føre verbet, slik at partikkelverbet har postfokal intonasjon. Likevel kan ein ofte høyra trykk på verbet eller partikkelen. På den måten kan au prosodiske faktorar spela inn på resultata våre. Det manglande venta sterke samsvaret mellom sistestavingstrykk i latinske ord og trykk på partikkel i partikkelverb kan ha fleire årsaker. Ein kan tenkja seg at latinske ord ter seg ganske ulikt jamført med meir ålmenne prosodiske preferansar hjå finnmarkstalarane våre, men dette krev ei meir inngåande gransking av trykksystemet o.l. Om det er slik at talarane i Finnmark vantar tonelagsdistinksjon, er det i utgangspunktet heller ikkje venta at den variable regelen til Kristoffersen og Torp (2016) kjem i bruk, i minsto ikkje utan revisjon av tonelagsutfallet. Det verkar ikkje som verb og partikkel utgjer éi prosodisk eining i Finnmark, slik dei gjer i dei andre områda me har granska.

7 Konklusjon

I denne granskinga har me sett eit varierende samsvar mellom trykkplassering i latinske ord og i partikkelverb. Det største samsvaret fann me innom Trøndelag, som hadde om lag like stor mengd trykk på fyrstestavingane på drygt 80 %. Likevel fann me i underkant av 10 % sistestavingstrykk i dei latinske orda og partikkeltrykk. Hedmark hadde meir variasjon, der områda utanom Trysil tedde seg meir som i Trøndelag, med samansetjingstrykk. Trysil hadde aller mest sistestavingstrykk i dei latinske orda og i partikkelverba, men dette var på ingen måte kategorisk. Det er soleis nærliggjande å tenkja at dette kan vera system som er ustabile eller i endring. I Finnmark fann me ikkje samsvaret me hadde venta. Der var berre sistestavingstrykk i dei latinske orda, men partikkelverba hadde mykje meir variabel trykkplassering. Dette kan skuldast mange faktorar, og krev ei meir inngåande gransking. Me mistenkjer at kategorien ‘partikkelverb’

ikkje har ein tilsvarande prosodisk kategori i Finnmark. LIA og NDK utgjer eit godt datamateriale for prosodisk analyse – ein kjem ganske langt med ljdfiler om ein held seg til trykk på stavingsnivået. Søkjegrensensnittet Glossa gjer det au mogeleg å finna relevante treff ganske enkelt, både for latinske ord og partikkelverb.

Referansar

- Abrahamsen, Jardar Eggesbø. 2003. Ein vestnorsk intonasjonsfologi. Doktoravhandling, NTNU.
- Alnæs, Ivar. 1916. *Norsk sætningsmelodi*. Kristiania: H. Aschehoug og Co.
- Fretheim, Thorstein. 1969. Norwegian Stress and Quantity Reconsidered. *Norsk Tidsskrift for Sprogvidenskap XXIII*: 76–96. Attgjeve i *Fonologi/Phonology* (1983), redigert av Ernst Håkon Jahr og Ove Lorentz. Oslo: Novus. 315–334.
- Faarlund, Jan Terje, Svein Lie og Kjell Ivar Vannebo. 1997. *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Jahr, Ernst Håkon. 1990. *Den store dialektboka*. Oslo: Novus Forlag.
- Johannessen, Janne Bondi, Joel Priestley, Kristin Hagen, Tor Anders Åfarli og Øystein Alexander Vangsnæs. 2009. The Nordic Dialect Corpus – an Advanced Research Tool. I *Proceedings of the 17th Nordic Conference of Computational Linguistics NODALIDA 2009*, redigert av Kristiina Jokinen og Eckhard Bick. NEALT Proceedings Series Volume 4.
- Kristoffersen, Gjert. 2000. *The Phonology of Norwegian*. Oxford: Oxford University Press.
- Kristoffersen, Gjert. 2003. *Norsk prosodi*. Bergen: Universitetet i Bergen. <https://bora.uib.no/bora-xmlui/handle/1956/20291>
- Kristoffersen, Gjert og Arne Torp. 2016. Fonologi. I *Norsk språkhistorie I. Mønster*, redigert av Helge Sandøy. Oslo: Novus. 101–211.
- Larsson, Ida og Björn Lundquist. 2014. Objektsplacering vid partikkelverb i norska dialekter och äldre svenska. I *Språk i Norge og*

nabolanda. Ny forskning om talespråk, redigert av Janne Bondi Johannessen og Kristin Hagen. Oslo: Novus. 99–132.

- Myrberg, Sara og Tomas Riad. 2015. The prosodic hierarchy of Swedish. *Nordic Journal of Linguistics* 38-2, 115–147. <https://doi.org/10.1017/S0332586515000177>
- Nøklestad, Anders, Kristin Hagen, Janne Bondi Johannessen, Michał Kosek og Joel Priestley. 2017. A modernised version of the Glossa corpus search system. I *Proceedings of the 21st Nordic Conference on Computational Linguistics (NoDaLiDa)*, redigert av Jörg Tiedemann. 251–254.
- Regjeringen.no. 2019, 19. desember. Nye fylker. <https://www.regjeringen.no/no/tema/kommuner-og-regioner/regionreform/regionreform/nye-fylker/id2548426/>
- Riad, Tomas. 2018. The phonological typology of North Germanic accent. I *Phonological Typology*, redigert av Larry M. Hyman og Frans Plank. 341–388. Berlin, Boston: De Gruyter Mouton. <https://doi.org/10.1515/9783110451931-010>
- Sandøy, Helge. 1976. Laust samansette verb i vestnordisk: Ein samanliknande leddstillingsanalyse for islandsk, færøysk og romsdalsmål. Magisteravhandling, UiO.
- Sandøy, Helge. 1985. *Norsk dialektkunnskap*. Oslo: Novus.
- Skaalbones, Lena Kildal. 2006. Henger de på? Prososyntaktiske trekk ved verbpartikler i en nordnorsk høytone-dialekt. Masteroppgåve, NTNU.
- Tengesdal, Eirik, Ida Larsson og Björn Lundquist. Prosody and word order in Norwegian particle constructions. Upublisert manuskript, sist endra 20. januar 2021. Word-fil.
- Tooley, Kristen M., Agnieszka E. Konopka og Duane G. Watson. 2018. Assessing priming for prosodic representations: Speaking rate, intonational phrase boundaries, and pitch accenting. *Memory & Cognition* 46, 625–641. <https://doi.org/10.3758/s13421-018-0789-5>
- Torp, Arne. 1997. Lågtone – jamvekt – trykkforskyving. I *Nordiska dialektstudier. Föredrag vid Femte nordiska dialektologkonferensen Sigtuna 17–21 augusti 1994*, redigert av Maj Reinhammar. Uppsala: Språk- och folkminneinstitutet. 293–304.

Western, August. 1921. *Norsk riksmåls-grammatikk for studerende og lærere*. Kristiania: H. Aschehoug & Co.

Aa, Leiv Inge. 2015. *The Grammar of Verb-Particle Constructions in Spoken Norwegian*. Doktoravhandling, NTNU.

Aa, Leiv Inge. 2020. *Norwegian Verb Particles*. Amsterdam: John Benjamins.

Korpus

LIA norsk – korpus av eldre dialektopptak:

<https://tekstlab.uio.no/LIA/korpus.html>

Nordisk dialektkorpus/Nordic Dialect Corpus:

<http://www.tekstlab.uio.no/scandiasyn>

English summary

This paper examines the relationship between stress placement in Latin–Greek loanwords (like *butikk* ‘store’) and Norwegian particle verbs (like *gå ut* ‘go out’), where the stress placement variably falls on the initial or ultimate syllable. Because particle verbs vary in part due to language-internal structural factors, we assume that these factors have less impact on stress placement in the loanwords. By examining non-syntactic variation in loanwords, we hope to understand stress placement variation in particle verbs better. Using the LIA and NDC speech corpora, we examined participants from the Norwegian counties Trøndelag, Hedmark, and Finnmark. None of the dialects had categorical stress placement. Trøndelag speakers most frequently produced initial stress both on the loanwords and particle verbs. Hedmark displayed a more variable distribution, still with predominantly initial stress. Finnmark had invariable ultimate stress on the loanwords, but the stress placement on verb and on particle seemed to vary quite equally. Initial stress in particle verbs, i.e. on the verb, was the preferred pattern for all areas except Finnmark. LIA and NDC are suitable for prosodic research.

Eirik Tengedal
Institutt for lingvistiske og nordiske studium
Universitetet i Oslo
eirik.tengedal@iln.uio.no

Björn Lundquist
Institutt for språk og kultur
UiT Norges arktiske universitet
bjorn.lundquist@uit.no



Ka LIA fortæll?

Eit gjensyn med kv-spørsmål i norske dialektar

Øystein A. Vangsnes og Marit Westergaard

Undersøkinga replikerer ein tidlegare korpusbasert studie av *kv*-spørsmål utan V2 i norske dialektar. Den første undersøkinga baserte seg på Nordisk dialektkorpus (NDK), som har opptak frå perioden 2005–2010. Undersøkinga i denne artikkelen har teke utgangspunkt i det norske LIA-korpuset, der materialet er samla inn over ein lang tidsperiode og der balansen i tilfang mellom ulike delar av landet og type opptak er noko annleis enn i NDK. Dataa frå LIA-korpuset stadfestar likevel langt på veg det generelle inntrykket frå NDK, som er (i) at ikkje-V2 først og fremst førekjem med dei korte spørjeorda *kva*, *kven* og *kvar/kor*, og (ii) at ordstillinga er mest utbreidd i nordnorsk og minst utbreidd i austnorsk. Ein mogleg skilnad mellom dei to korpusa gjeld det vestnorske tilfanget med om lag halvparten så mykje ikkje-V2 i LIA (20,5 %) som i NDK (43,3 %). Det kan bera bod om at strukturar med ikkje-V2 er vortne meir vanlege på Vestlandet over tid.

Stikkord: *kv*-spørsmål, V2, ikkje-V2, spørjeord, dialektkorpus, geografisk variasjon

1 Innleiing

I Vangsnes og Westergaard (2014) undersøkte me førekomsten av hovudsetnings *kv*-spørsmål med og utan såkalla V2-ordstilling (dvs. anten fråvær av subjekt/verb-inversjon eller, i subjektsspørsmål, innsetjing av *som* i andreposisjon) i den norske delen av Nordisk dialektkorpus (Johannessen, Priestley, Hagen, Åfarli og Vangsnes 2009) (heretter NDK), altså fordelinga av strukturar som (1a, c) versus (1b, d).

- (1) a. Kor bor du? (V2)
- b. Kor du bor? (ikkje-V2)
- c. Kven kjem? (V2)
- d. Kven som kjem? (ikkje-V2)

Hovudfunna i Vangsnes og Westergaard (2014) kan oppsummerast som følgjer.

- 1) NDK-materialet stadfestar i stor grad den distribusjonen av ordstillinga i (1b) og (1d) som har vore avdekt i forskingslitteraturen frå 1980-talet og framover, som er at ikkje-V2 førekjem hovudsakleg i Nord-Noreg, Trøndelag og på Vestlandet. I tillegg styrkjer data frå korpuset eit bilete av at manglande V2 også førekjem på nordlege delar av Austlandet. Berre på sentrale og sørlege delar av Austlandet og vidare til og med Agder er strukturen uvanleg.
- 2) I spontan tale er det svært få førekomstar av komplekse eller samansette *kv*-uttrykk utan V2 (t.d. *korleis*, *kvifor*, *kva for ein ...*, *kor mange ...*), jamvel om talarar i visse delar av landet godtek slike strukturar i spørjeskjemaundersøkingar.
- 3) I Nord-Noreg er strukturar med ikkje-V2 (ved korte/einstava *kv*-ledd, som i (1b)) langt vanlegare enn strukturar med V2 (som i (1a)). I Trøndelag er fordelinga jamn, medan det er fleire førekomstar av V2 enn ikkje-V2 på Vestlandet sett under eitt.

Materialet i NDK er samla inn på byrjinga av 2000-talet. I dette arbeidet følgjer me opp med å undersøka om det noko eldre materialet i LIA-korpuset gir eit liknande bilete som det me fann i NDK.

Hovudspørsmålet for undersøkinga blir då om resultatata frå LIA-korpuset utfyller eller utfordrar eksisterande kunnskap om fenomenet.

I det neste avsnittet skal me først kort gjera greie for forskingshistorikken på emnet fram til prosjektet Nordisk dialeksyntaks, som etablerte forskingsinfrastrukturane Nordisk syntaksdatabase og Nordisk dialektkorpus rundt 2010 (sjå Vangsnes og Johannessen 2019). Deretter presenterer me hovudfunna frå Vangsnes og Westergaard (2014) og går vidare til å sjå på i kva grad materiale frå LIA-korpuset harmonerer med dei.

2 Litt forskingshistorikk

Nokre av dei mange arbeida som nemner eller har teke føre seg *kv*-spørsmål utan V2 i norske målføre, er Iversen (1918), Elstad (1982), Sandøy (1987), Nordgård (1985, 1988), Åfarli (1986), Taraldsen (1985, 1986), Lie (1992), Nilsen (1996), Fiva (1996), Sollid (2003), Westergaard (2003, 2005, 2009), Vangsnes (2004, 2005), Westergaard og Vangsnes (2005), Reite (2011), Rognes (2011), Vangsnes og Westergaard (2014), Westergaard, Vangsnes og Lohndal (2017), Westendorp (2018), Lohndal, Westergaard og Vangsnes (2020).

Allereie Iversen (1918, 37f) diskuterer syntaksen knytt til dei spørjande pronomena *kæm* og *ka* i Tromsø bymål og seier at dei har:

i visse henseender gjennomgaat en ganske interessant utvikling [...]. Sterkest springer vel den eiendommelige ordstilling i øinene, idet direkte spørjesætninger faar samme ordstilling som indirekte eller avhængige spørjesætninger: *kæm det va'?* *kæm du træfte?* *kæm han har vorre ilam'me?* [...] *ka det e'?* *ka'o e'?* («hvor mange er klokken?»); *ka du hete?* *ka du sa?* *ka du bæer dæ før?*

Iversen (op. cit.) nemner også at når *kæm* og *ka* står som subjekt, må dei vanlegvis følgjast av *som*, og han gir følgjande døme: *ja, kæm som vet?* *kæm som har tadd saksa?* *ka som står på?* I ein merknad peiker han dessutan på at dette gjer at dialekten skil eksplisitt mellom tilfelle der *kv*-ordet er høvesvis subjekt og predikativ som i døma *ka*

petroleum e'? versus *ka som e' petroleum?* Her har «skriftsproget» eit og same uttrykk i begge tilfelle: «Hvad er petroleum?».

Drygt seksti år seinare har Elstad (1982, 32f) ei side om fenomenet i sitt oversyn over kjennemerke ved nordnorske dialektar. Han skriv:

Ei slik ordstilling i spørsmål som: Ka dokker trur? Ka du sei? Kor det går? Kem du fekk det hos?, blir gjerne rekna for særskilt nordnorsk. Og uttrykksmåten er vanleg i alle ekte nordnorske målføre, men han er ikkje lite brukt mange stader lenger sør òg.

Vidare seier Elstad (op. cit.) at «[b]eintfram ordstilling, som vi har her, kan ein berre bruke ved spørjeorda *ka, kem, kor* (–'korleis' eller 'kvar', som somme stader kan heite *karr*).»

Dette siste har seinare, mellom anna av Taraldsen (1985) og Fiva (1996), blitt referert til som *Elstads generalisering*, gjerne presisert som at manglande subjekt/verb-inversjon berre er mogleg om spørjeordet er kort, alternativt einstava.

Me kan elles merka oss at Elstad seier at ordstillinga er «ikkje lite brukt» i dialektar sør for det nordnorske dialektområdet, og Lie (1992) viser at manglande subjekt/verb-inversjon i *kv*-spørsmål førekjem i dei fleste norske dialektar, med unntak av det sentrale austlandsområdet og sørlandskysten.¹ Åfarli (1985, 1986) diskuterer fenomenet for nordmørsk, medan Nordgård (1985, 1988) gjer det også for vestnorsk i tillegg til trøndersk og nordnorsk. Eit viktig funn i deira arbeid er at Elstads generalisering ikkje gjeld i alle dialektområde, og særleg ikkje i nordvestnorske målføre der informantar også tillèt meir komplekse *kv*-uttrykk med manglande V2, sjå følgjande døme frå nordmørsdialekt i Åfarli (1986, 98).

(2) Kåles bil du kjøpte?

¹ Lie (1992: 75) opplyser i ein fotnote at undersøkinga hans vart utført i perioden 1975-1976, altså før Elstad (1982), men ikkje publisert før 16 år seinare på grunn av andre arbeidsoppgåver.

Seinare finn Fiva (1996) at det i tromsødialekten er ein asymmetri mellom subjektsspørsmål og ikkje-subjektsspørsmål ved at Elstads generalisering heilt klart held for ikkje-subjektsspørsmål, medan ein god del informantar tillèt samansette *kv*-uttrykk i subjektsspørsmål av typen *Kor mange eleva som møtte opp i dag?* (Fiva op. cit., 148). Omtrent på same tida avdekker Nilsen (1996) at *kv*-spørsmål med *kv*-uttrykk med meir enn ei staving vert godtekne av særleg eldre talarar av norske målføre i språkkontaktområde i Nord-Troms. Forskinga på fenomenet i dei siste tiåra av 1900-talet kjem altså fram til at det er variasjon mellom norske dialektar om manglande V2 i *kv*-spørsmål er avgrensa til korte *kv*-ledd eller ikkje.

Elstad (1982, 33) seier dessutan eksplisitt om manglande inversjon i *kv*-spørsmål at «[d]enne ordstillinga er ikkje einerådande. Det kan godt heite: Ka sei du? Kor går det? og Kem fekk du det hos? Det synest å vere reint stilistiske omsyn som avgjer om ein bruker den eine eller den andre uttrykksforma». I tråd med dette finn Westergaard (2003, 2005, 2009) i sine undersøkingar av eit materiale med opptak av talarar frå Troms at ikkje-V2 særleg førekjem når subjektet i spørsmålet er gitt informasjon, medan V2 særleg førekjem når subjektet er ny informasjon og når det finitte verbet er kopulaverbet *vera*. Dette funnet har vore viktig ikkje minst for å tilbakevisa ein tanke om at valet mellom V2-struktur og ikkje-V2 kunne handla om eit val mellom ulike grammatikkar (sjå t.d. Taraldsen 1985).

Lie (1992, 65ff) legg elles fram opplysningar som tyder på at ein del av hans informantar berre tillèt manglande V2 i *kv*-spørsmål dersom *kv*-leddet er subjekt, altså til dømes setningar av typen *Kven som har vore her?* men ikkje *Kva du seier* eller *Kor du bur?*. Det verkar som at det er særleg informantar på Sørvestlandet som har gitt tilbakemeldingar av den typen.

3 Data frå ScanDiaSyn om norske *kv*-spørsmål

På byrjinga av 2000-talet visste ein såleis ein del om mikrovariasjonen knytt til hovudsetnings *kv*-spørsmål i norske dialektar. Det var klarlagt at somme dialektar berre tillèt manglande V2 i *kv*-spørsmål

når *kv*-leddet er kort (*kva*, *kven* eller *kor*) medan andre dialektar òg tillèt lengre og meir samansette *kv*-ledd. Vidare var det etablert ei forståing av at valet mellom V2 og ikkje-V2 i *kv*-spørsmål er styrt av informasjonsstrukturelle forhold. Forskinga hadde òg avdekt at dialektar kunne ha ein asymmetri mellom subjektsspørsmål og spørsmål med andre setningsledd som *kv*-ledd, dels ved at avgrensinga til korte *kv*-ledd i somme dialektar ikkje gjeld i subjektsspørsmål, og dels ved at somme dialektar berre tillèt manglande V2 i subjektsspørsmål men ikkje elles.

Dette danna utgangspunkt for dei testsetningane om fenomenet som vart inkluderte i det syntaktiske spørjeskjemaet brukt i prosjektet Nordisk dialektsyntaks (ScanDiaSyn, sjå Vangsnes, 2007a, 2007b). Spørjeskjemaet vart lagt fram for informantar frå 107 forskjellige stader i Noreg, som regel fire personar på kvar stad, og sidan organisert i den opne elektroniske infrastrukturen Nordisk syntaksdatabase (NSD, sjå Lindstad et al. 2009, Vangsnes og Johannessen 2011, 2019). Spørjeskjemaet undersøkte ei rekkje forskjellige forhold, og i utgangspunktet var det berre plass til nokre få testsetningar som gjekk direkte på hovudsetnings *kv*-spørsmål, nærmare bestemt dei følgjande fire:

1. Kven som sel fiskeutstyr her i bygda, då?
2. Kor mange elevar som går på den her skolen?
3. Kva du heiter?
4. Kva tid du gjekk ut av ungdomsskolen, då?

Ein kunne med rette ha ønskt seg fleire setningar, men samla sett kan desse testsetningane gi eit bilete av om talarar frå ein bestemt stad godtek: (i) subjektsspørsmål utan V2 (1, 2), (ii) ikkje-subjektsspørsmål utan V2 (3, 4), (iii) ikkje-V2 med lange *kv*-ledd (2, 4) eller berre korte *kv*-ledd (1, 3).

Som rapportert i Vangsnes og Westergaard (2014, 137ff), og i meir detalj i Westergaard, Vangsnes og Lohndal (2012, 2017), er biletet som NSD-materialet teiknar, langt på veg i tråd med dei indikasjonane den tidlegare forskingslitteraturen hadde gitt. Eit par ting er verdt å nemna. For det første er det slik at sjølv om Elstads genera-

lisering er hovudregelen i området frå Trøndelag og nordover, er det fleire «lommer» som liknar på mønsteret Fiva (1996) hadde rapportert for tromsømålet, nemleg at lange/komplekse *kv*-uttrykk utan V2 vert godtekne i subjektsspørsmål men ikkje elles. Vidare vert manglande V2 ved komplekse *kv*-uttrykk, slik som hjå Åfarli (1985) og Nordgård (1985, 1988), godtekne på Nordvestlandet uavhengig av syntaktisk funksjon på *kv*-leddet, og også ein del plassar i Rogaland slik som også Nordgård (op. cit.) fann hjå dei to informantane han hadde derifrå. Også opplysningane i Lie (1992) om at mange på Sørvestlandet berre godtok ikkje-V2 i subjektsspørsmål, vart attfunne i spørjeskjemadataa frå ScanDiaSyn, rett nok med ein litt nordlegare avgrensing til delar av Hordaland. Til sist synest det for dialektane nord på Austlandet å vera mykje variasjon, og dataa teiknar såleis ikkje noko tydeleg mønster, noko som i viss mon også er i tråd med opplysningane i Lie (1992).

Datainnsamlinga i ScanDiaSyn-prosjektet omfatta også innspelningar med informantane, både samtalar mellom to og to informantar og intervju med dei. Dette materialet med 2-3 timar opptak frå kvar besøkte stad vart transkribert og tilrettelagt i Nordisk dialektkorpus, altså NDK. Vangsnes og Westergaard (2014) undersøkte kva data om *kv*-spørsmål utan V2 ein finn i NDK, og tabell 1 viser dei overordna resultatane frå undersøkinga (med to små justeringar²).

	<i>N</i>	<i>V2</i>	<i>Ikkje-V2</i>
<i>kva</i>	660	284 (43,0 %)	376 (57,0 %)
<i>kven</i>	111	50 (45,0 %)	61 (55,0 %)
<i>kor/kvar</i>	130	68 (52,3 %)	62 (47,7 %)
<i>når/kva tid</i>	79	58 (73,4 %)	21 (26,6 %)
<i>kvifor</i>	47	46 (97,9 %)	1 (2,1 %)
<i>korleis</i>	128	119 (93,0 %)	9 (7,0 %)
<i>kv + XP</i>	177	169 (95,3 %)	8 (4,7 %)
Samla	1332	794 (59,6 %)	538 (40,4 %)

Tabell 1: Fordelinga av V2 og ikkje-V2 i norske *kv*-spørsmål i Nordisk dialektkorpus (Vangsnes og Westergaard 2014).

² Justeringa gjeld at eit døme med *når/kva tid* med ikkje-V2 opphavleg var telt to gonger, og talet er difor 21, ikkje 22, jf. Vangsnes og Johannessen (2019, avsnitt 4.3), og ein reknefeil som gir totaltalet 1332 (snarare enn 1334 som rapportert i Vangsnes og Westergaard 2014).

Tabellen viser eit tydeleg skilje mellom dei korte *kv*-ledda *kva*, *kven* og *kor/kvar* og dei andre ved at det er ei ganske jamn fordeling mellom V2 og ikkje-V2 for dei korte, men ei klar overvekt av V2 for dei andre. Samtidig er det interessant at *kv*-uttrykket for tid, altså *når* i standardspråka, står i ei mellomstilling med 26,6 prosent ikkje-V2, medan ingen av dei andre lange uttrykka har meir enn 7 prosent ikkje-V2. Dei 21 døma omfattar både den einstava forma *når* og ulike former av det samansette uttrykket *kva/kor tid*. Som Vangsnæs og Westergaard (op. cit., 144) diskuterer, har fem av døma det korte *kv*-uttrykket *når*, og dei er alle produserte av fem ulike informantar frå tre ulike stader i det trønderske dialektområdet. Også åtte av dei ni *korleis*-døma viser seg å involvera einstava former, og av dei totalt 39 døma med ikkje-V2 for *når/kva tid*, *kvifor*, *korleis* og *kv* + *XP* står ein såleis att med berre 27 døme som involverer *kv*-uttrykk med meir enn ei staving.³

Det er elles verdt å nemna at det er stor variasjon på tvers av norske talemål når det gjeld sjølve forma på *kv*-orda. Til dømes kan 'kva' ha mellom anna formene *ka*, *ke*, *kæ*, *kå*, *høtt*, *hå*, *å*, *va* og så vidare. Ut over skilnader i tal på stavingar gjer ikkje Vangsnæs og Westergaard (op. cit.) noko forsøk på å undersøka om skilnader i form på segmentnivå spelar noka rolle for ordstillinga.⁴

For dei tre korte *kv*-uttrykka der det er rikeleg med døme og ei jamn fordeling mellom V2 og ikkje-V2 i heile materialet, finn Vangsnæs og Westergaard (op. cit.) forskjellar mellom landslutane. Desse forskjellane går fram av tabell 2, med justeringar frå Vangsnæs og Johannessen (2019), der Agder er slått saman med Vestlandet (i tråd

³ Dei 27 døma med komplekse *kv*-uttrykk stammar frå 11 stader på Vestlandet og i Trøndelag. Vangsnæs og Johannessen (2019, avsnitt 4.3) har undersøkt døma og stadene nærmare opp mot data frå NSD og finn at 24 døme er ytra av informantar frå område der spørjeskjemadataa tilseier at komplekse *kv*-uttrykk vert godtekte. Dei tre attverande døma er frå høvesvis Bergen, Oppdal og Røros, der testsetningane med komplekse *kv*-uttrykk får ein låg skår, også av dei informantane det gjeld, og dette står såleis att som støy i datamaterialet.

⁴ Reite (2011) kom i si undersøking av skedsmokorsmålet fram til at ein føresetnad for at talarane kunne godta ikkje-V2, var at forma på 'kva' var *å* og ikkje *va*. Det er rimeleg å tolka det som at dei to variantane indekserer to ulike register av dialekten, eitt med (i nokon grad) ikkje-V2 og eitt utan.

med firedelinga av det norske dialektlandskapet i Mæhlum og Røyneland 2012).

	<i>N</i>	<i>[kva/kven/kor]</i> + V2	<i>[kva/kven/kor]</i> + ikkje-V2
Nord-Noreg	380	81 (21,3 %)	299 (78,7 %)
Trøndelag	116	50 (43,1 %)	66 (56,9 %)
Vestlandet	219	124 (56,6 %)	95 (43,4 %)
Austlandet	186	147 (79,0 %)	39 (21,0 %)
Heile landet	901	402 (44,6 %)	499 (55,4 %)

Tabell 2: Geografisk fordeling i den norske delen av Nordisk dialektkorpus for ±V2 i *kv*-spørsmål med korte *kv*-uttrykk (basert på Vangsnæs og Westergaard 2014)

Tabellen viser altså at ikkje-V2 dominerer i det nordnorske materialet, det er ganske jamt i det trønderske området, og litt mindre ikkje-V2, men framleis jamt, i vestnorsk. Ikkje uventa er det ein klar overvekt av V2 i austnorske målføre.

Mot denne bakgrunnen skal me no sjå korleis fordelinga av *kv*-spørsmål med og utan V2 er i LIA-korpuset når me gjer dei same søka som vart gjort i NDK.

4 *Kv*-spørsmål i LIA-korpuset

4.1 LIA-korpuset versus NDK og søk etter *kv*-spørsmål

Tabell 3 (neste side) er ei samanlikning av storleik og omfang på den norske delen av Nordisk dialektkorpus og det norske LIA-korpuset.

LIA-korpuset er monaleg større både i talet på ord, talet på personar og talet på stader. I snitt er det likevel om lag dobbelt så mange ord per person i NDK (4 566) som i LIA (2 547). For ord per stad er det meir jamt (18 000 i NDK versus 15 700 i LIA). Me ser elles at det er ein viss skilnad i omfanget frå kvar landslut i dei to korpusa, og særleg gjeld dette Austlandet som har 28,5 prosent av tilfanget i NDK men berre 16,4 prosent av tilfanget i LIA. Både frå Nord-Noreg og særleg Vestlandet er det relative tilfanget større i LIA enn i NDK. Som me skal sjå nedanfor, gir dette nokre utslag for undersøkinga vår.

	Talarar		Ord		Stader	
	NDK	LIA	NDK	LIA	NDK	LIA
Austlandet	132 (30,1 %)	226 (16,4 %)	569 861 (28,5 %)	571 054 (16,4 %)	34	51
Vestlandet	145 (33,1 %)	566 (41,2 %)	613 508 (30,7 %)	133 7241 (38,4 %)	36	100
Trøndelag	52 (11,9 %)	182 (13,2 %)	212 344 (10,6 %)	299 690 (8,6 %)	12	30
Nord-Noreg	109 (24,9 %)	380 (27,7 %)	602 207 (30,1 %)	125 3541 (36 %)	29	41
Anna el. ukategorisert	-	20 (1,5 %)	-	20 021 (0,6 %)	-	0
Til saman	438 (100 %)	1 374 (100 %)	1 997 920 (100 %)	3 481 547 (100 %)	111	222

Tabell 3: Storleiken på den norske delen av Nordisk dialektkorpus versus det norske LIA-korpuset

Ein viktig skilnad mellom korpusa er elles at der materialet i NDK i hovudsak vart samla inn på ein koordinert og systematisk måte i ein relativt kort periode mellom 2005 og 2010, er LIA-korpuset bygd opp av eldre opptak gjort til ulike tider, på ulike måtar og til ulike formål. Såleis er NDK eit korpus med synkront materiale, rett nok med to alderskategoriar, medan LIA-korpuset har ein diakron dimensjon med talarar fødde frå midten av 1800-talet til midten av 1900-talet.

I prinsippet opnar det for at me kan gjera undersøkingar av korleis syntaksen i *kv*-spørsmål har utvikla seg i norske dialektar. I praksis krev det ein stor arbeidsinnsats, for sjølv om det er mogleg å spesifisera både fødselsår og opptaksår i søkegrensesnittet, kan ein per i dag ikkje velja heile periodar, til dømes tiår. Me valde difor å konsentrera oss om å gjenta dei same søka som i Vangsnæs og Westergaard (2014), i denne omgangen utan å fanga eventuelle diakrone aspekt innanfor LIA-materialet.

Formatet for søka i Vangsnæs og Westergaard (2014) kan framstillast som følgjer med spesifikasjonen at det ikkje skal vera nokon ord mellom ledda i søkestrengane:

- (3) a. <segmentinitialt 'KV'> + <'ikkje-verb'>
 b. <#> + <'KV'> + <'ikkje-verb'>

- c. <segmentinitialt 'KV'> + <'verb'>
 d. <#> + <'KV'> + <'verb'>

Bakgrunnen for å bruka taggane <segmentinitialt> og <#> var å unngå flest mogleg tilfelle av underordna *kv*-setningar i trefflista. Talemålsmaterialelet i korpusa har ikkje nokon tagg for kategorien «setning», men «segment» er ei talespråkleg eining som kjem nær dette, og det er større sjanse for at eit segmentinitialt *kv*-ord innleier ei hovudsetning enn ei leddsetning. Taggen # markerer pause, og igjen er det meir sannsynleg at eit *kv*-ord som kjem rett etter ein pause, innleier ei hovudsetning enn ei leddsetning.

Sjølv desse grepa ga ei overgenerering på ca. 40 prosent i NDK-materialet i form av leddsetningar, ellipsar (manglande verb), ekko-spørsmål og så vidare. For undersøkinga av LIA-korpuset gjorde me ei lita endring ved at me la til <?> i søka i (3b) og (3d), og me tillét 0-20 ord mellom <±verb> og <?>. Spørsmålsteiknet vert nytta i transkripsjonen, men naturleg nok berre i hovudsetningsspørsmål slik som i skriftspråket, og ikkje etter underordna spørsmål.⁵

I det som følgjer, presenterer me no først resultatata for dei tre korte *kv*-uttrykka *kva*, *kven* og *kor/kvar* i LIA-korpuset i tabellar, der tala er brotne ned på dei fire hovuddialektområda i Mæhlum og Røyne-

⁵ Det kan sjå ut som at tillegget av spørsmålsteiknet reduserte overgenereringa ein del, samtidig som bruken av taggen «segment» ikkje synest å ha hatt nokon effekt på overgenereringa: Som me ser av bruttoresultata for dei fire søka på *kva* ± V2 i LIA-korpuset, viste i (i), er det massiv overgenerering i søket etter ikkje-V2 i (i-a) med berre 23,3 prosent relevante treff).

- (i) a. <segmentinitialt 'kva'> | 0 | <'ikkje-verb'>
 967 treff; 225 relevante (23,3 %)
 b. <#> | 0 | <'kva'> | 0 | <'ikkje-verb'> | 0-20 | <?>
 115 treff; 83 relevante (72,2 %)
 c. <segmentinitialt 'kva'> | 0 | <'verb'>
 318 treff; 241 relevante (75,8 %)
 d. <#> | 0 | <'kva'> | 0 | <'verb'> | 0-20 | <?>
 98 treff; 92 relevante (93,9 %)

Dette gir oss ein mistanke om at taggen <segment> er nytta annleis i LIA-transkripsjonen enn i NDK.

land (2012), det vil seia Nord-Noreg, Trøndelag, Vestlandet og Austlandet. I del 4.3 ser me så på andre *kv*-uttrykk.

4.2 Korte *kv*-ord

I NDK fann Vangsnes og Westergaard (2014) 660 tilfelle av *kv*-spørsmål innleidde med (pronomenet) *kva*, og fordelinga var 43 prosent V2 versus 57 prosent ikkje-V2. I LIA-korpuset har me funne 638 tilfelle, der 51 prosent har V2 og 49 prosent har ikkje-V2. Eitt av døma utan V2 er dette frå Lavangen i Troms:

- (4) ka e får låv så ikkje di fækk låv te ? (lavangen_uit_0303)
 ‘Kva eg får lov som ikkje dei fekk lov til?’

Kv-spørsmål innleidde med *kven* fann Vangsnes og Westergaard (2014) 111 tilfelle av i NDK, og fordelinga var 45 prosent V2 versus 55 prosent ikkje-V2. I LIA-korpuset har me funne nøyaktig like mange tilfelle, men fordelinga er 56 prosent V2 og 44 prosent ikkje-V2. Eitt av døma utan V2 er dette frå Åsane i Bergen:

- (5) kemm åkka har i gymm ? (aasane_uib_0701)
 ‘Kven de har i gym?’

Undersøkinga til Vangsnes og Westergaard (2014) fann 130 spørsmål etter stad i NDK, altså *kv*-spørsmål innleidde med *kvar/kor*, og fordelinga var 52,3 prosent V2 versus 47,7 prosent ikkje-V2. I LIA-korpuset har me funne 100 tilfelle der 49 prosent har V2 og 51 prosent har ikkje-V2. Eitt av døma utan V2 er dette frå Overhalla i Trøndelag:

- (6) ker a æ frå kjærrinja di ? (overhalla_ntnu_0103)
 ‘Kor ho er frå, kjerringa di?’

Tabell 4 presenterer det samla resultatet for *kv*-spørsmåla med dei tre korte *kv*-orda med vekt på tilfelle med ikkje-V2, både absolutte tal og prosentfordeling.

I NDK fann Vangsnes og Westergaard (2014) totalt 901 tilfelle av *kv*-spørsmål med dei korte spørjeorda *kva*, *kven*, *kor/kvar*, med ei for-

		Nord	Midt	Vest	Aust	Samla
<i>kva</i>	<i>N</i>	270	17	304	47	638
	-V2	235	10	61	4	310
	%	87 %	58,8 %	20,1 %	8,5 %	48,6 %
<i>kven</i>	<i>N</i>	33	0	61	17	111
	-V2	30	0	16	3	49
	%	90,9 %	0 %	26,2 %	17,7 %	44,1 %
<i>kvar/kor</i>	<i>N</i>	54	1	40	0	100
	-V2	43	1	7	0	51
	%	79,6 %	100 %	17,5 %	0 %	51 %
Samla	<i>N</i>	357	18	405	69	849
	-V2	308	11	84	7	410
	%	86,3 %	61,1 %	20,7 %	10,1 %	48,3 %

Tabell 4: *Kv*-spørsmål med korte spørjeord og ikkje-V2 i det norske LIA-korpuset

deling på 44,6 prosent V2 og 55,4 prosent ikkje-V2. I LIA har me altså funne til saman 849 tilfelle der fordelinga er 51,7 prosent V2 versus 48,3 prosent ikkje-V2.

På overordnad nivå er tala for dei to korpusa slåande like, både når det gjeld absolutte tal og prosentfordeling i det samla resultatet. At det er omtrent like mange treff, er interessant med tanke på at LIA-korpuset er nesten dobbelt så stort som NDK, og det fortel oss truleg noko om *kva* type materiale korpusa inneheld: Ein stor del av opptaka til NDK er frie samtalar mellom to og to informantar, og sjølv om det nok også førekjem slike samtalar i LIA-materialet, er det truleg ein større del intervjumateriale der informantane i mindre grad sjølve stiller spørsmål.

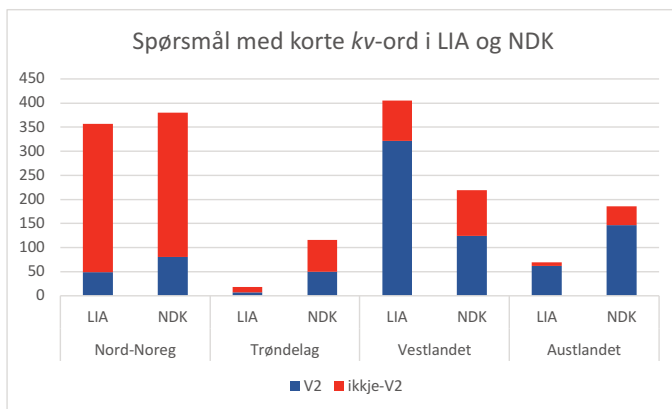
Når me går vidare og samanliknar tala landslut for landslut, kjem det fram fleire skilnader. Tabell 5 på neste side gir ein samla oversikt over den geografiske utbreiinga av treffa i dei to korpusa.

Tala i tabell 5 (neste side) viser for det første at det er ein annan balanse i NDK enn i LIA når det gjeld tilfang frå dei ulike landslutane. Dette kjem tydelegare fram i den grafiske framstillinga i figur 1 (neste side). Som me ser, er det svært få *kva*-spørsmål i LIA frå Trøndelag og ikkje så mange frå Austlandet heller. Den siste kolonnen i tabell 5 viser den prosentvise delen av ord i korpusa henta frå tabell 3 over, og sett mot den har det trønderske materialet berre 2,1 prosent

<i>kv/kven/kor +</i>		<i>V2</i>		<i>ikkje-V2</i>		<i>Samla</i>		% ord ⁶
		<i>n</i>	%	<i>n</i>	%	<i>n</i>	%	
Nord	LIA	49	13,7	308	86,3	357	42,0	36
	NDK	81	21,3	299	78,7	380	42,2	30,1
Midt	LIA	7	38,9	11	61,1	18	2,1	8,6
	NDK	50	43,1	66	56,9	116	12,9	10,6
Vest	LIA	321	79,3	84	20,7	405	47,7	38,4
	NDK	124	56,6	95	43,4	219	24,3	30,7
Aust	LIA	62	89,9	7	10,1	69	8,1	16,4
	NDK	147	79,0	39	21,0	186	20,6	28,5
Samla	LIA	439	51,7	410	48,3	849	100	-
	NDK	402	44,6	499	55,4	901	100	-

Tabell 5: Samanlikning av $\pm V2$ i *kv*-spørsmål med korte *kv*-uttrykk i LIA og Nordisk dialektkorpus

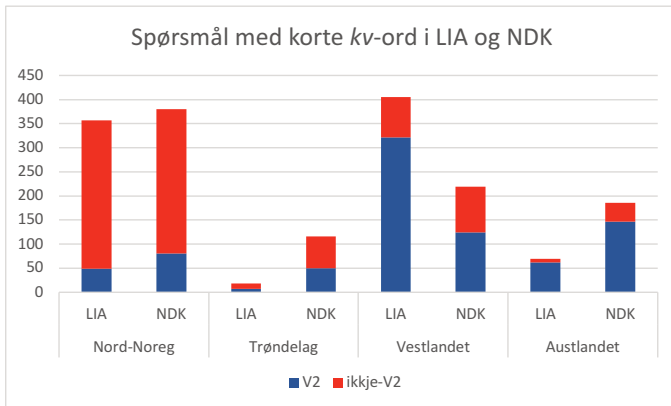
av *kv*-spørsmåla mot 8,6 prosent av det totale ordtilfanget i LIA, medan det er meir balansert med 12,9 prosent *kv*-spørsmål mot 10,6 prosent ordtilfang i NDK. For Austlandet er det lågare del *kv*-spørsmål enn ordtilfang i begge korpusa. For Nord-Noreg er det ein større del *kv*-spørsmål enn ordtilfang i begge korpusa, medan det for Vestlandet er ein større del *kv*-spørsmål enn ordtilfang i LIA-korpuset, men omvendt i NDK.



Figur 1: Den geografiske fordelinga av *kv*-spørsmål med korte *kv*-uttrykk i LIA og Nordisk dialektkorpus

Igjen synest det rimeleg å tru at desse skilnadene i tilfang – denne gongen mellom landslutane – kan skuldast skilnader i kva for type materiale som er transkribert og lagt inn i korpuset. At det er fleire spørsmål i det vestnorske LIA-materialet enn i NDK-materialet, er interessant med tanke på at det ganske omfattande materialet frå Talemålsundersøkelsen i Bergen frå slutten av 1970-talet inngår i korpuset. Mykje av dette materialet er frie samtalar mellom informantar.

Tala i tabell 5 viser elles at sjølv om det er ei ganske lik prosentvis fordeling mellom V2 og ikkje-V2 i dei to korpusa når ein ser heile materialet under eitt, er det nokre skilnader på regionalt nivå. Dette er grafisk framstilt i figur 2. Om me ser bort frå det trønderske og austnorske materialet som er så lite i LIA, kan me særleg merka oss at V2 er langt meir utbreidd i LIA enn i NDK i materialet frå Vestlandet. Omvendt er det noko meir ikkje-V2 i Nord-Noreg i LIA enn i NDK. Det er likevel vanskeleg å utan vidare fastslå om dette speglar ei eventuell diakron utvikling der ikkje-V2 over tid er vorte meir van-



Figur 2: Fordelinga mellom V2 og ikkje-V2 i kv-spørsmål med korte spørjeord i LIA-korpuset og Nordisk dialektkorpus

leg på Vestlandet. Det vil krevja både ein nærmare analyse av kor på Vestlandet tilfanget er frå og frå kva periodar/generasjonar.⁶

4.3 Andre kv-uttrykk

I søka våre i LIA-korpuset fann me 357 kv-spørsmål med komplekse kv-uttrykk, det vil seia kv-adverb som *korleis*, *kvifor*, *når* samt kv-frasar som til dømes *kor mange X*, *kva for ein Y* etc. Dette er kv-uttrykk som vert rekna som «komplekse» i morfologisk og/eller semantisk forstand, sjølv om det òg førekjem einstava variantar som *koss* (for ‘korleis’) og *når* der ulike variantar av *kva tid* og *kor tid* elles er vanlege i målføra.⁷ Berre 23 av treffa med komplekse kv-uttrykk hadde ikkje-V2, noko som tilsvarar 6,4 prosent. Dette er ikkje ulikt funna i Vangsnes og Westergaard (2014), som fann ein andel på 7,2 prosent (39 av 538) komplekse kv-spørsmål utan V2.

Når me ser nærmare på treffa i LIA, har berre fem av dei 58 *når*-spørsmåla ikkje-V2. Fire er frå Nord-Noreg og eitt frå Vestlandet, og alle involverer samansette former: *ke/ka/ko/når* + *ti*. Det vestnorske dømet er dette frå Bergen:

- (7) ko ti de går bussær no ? (bergen_uib_2901)
 ‘Kor tid det går bussar no?’

For *kvifor* finn me til saman berre 11 hovudsetningsspørsmål i LIA, og samtlege har V2.

⁶ Eit raskt og omtrentleg estimat tilseier at om lag 20 prosent av informantane og 24 prosent av tilfanget frå Agder til Møre i korpuset er frå dagens Bergen kommune, og det meste av dette er nok materialet frå Talemålsundersøkelsen i Bergen. No er ikkje kv-spørsmål utan V2 heilt ukjent frå bergensmålet: Lie (1992, 65f) nemner heimlar og informantstadfestingar frå Bergen, men i NSD-materialet blir testsetninga *Kva du heiter?* unisont avvist (skår 1) av alle dei fire informantane frå Bergen, samtidig som testsetninga *Kven som sel fiskeutstyr her i bygda?* vert godteken (skår 5) av alle.

⁷ I mange norske dialektar er det eit skilje mellom kv-uttrykket *når* (*kva tid*, *kor tid*) og subjunksjonen *når* som innleier temporale leddsetningar, jf. også skiljet i dansk mellom *hvornår* (kv-uttrykk) og *når* (subjunksjon).

Sjå elles Vangsnes (2008) for ein analyse av ulike former for ‘korleis’, der einstava former som *koss* vert hevda å ha ein underliggjande kompleks morfologisk struktur.

Måtespørsmål med ulike former av *korleis* er det 122 døme på i LIA-korpuset, og berre to av desse har ikkje-V2, eitt frå nordnorsk og eitt frå vestnorsk. Dei to døma er desse:

- (8) kodann du klæmmstet ann å ? (bergen_uib_2903)
 ‘Korleis du klemstra han då?’
- (9) kolls’n de høvvdde att du fikk bare enn gut ? (lyngen_uit_1201)
 ‘Korleis det høvde at du fekk berre ein gut?’

Kv-spørsmål med utbygde nominalfrasar som *kv*-ledd er det totalt 68 av i våre LIA-søk. Av desse har 11 ikkje-V2, ni frå vestnorsk og to frå nordnorsk, og åtte av dei vestnorske døma er produserte av ein og same informant frå Volda, mellom anna det følgjande.

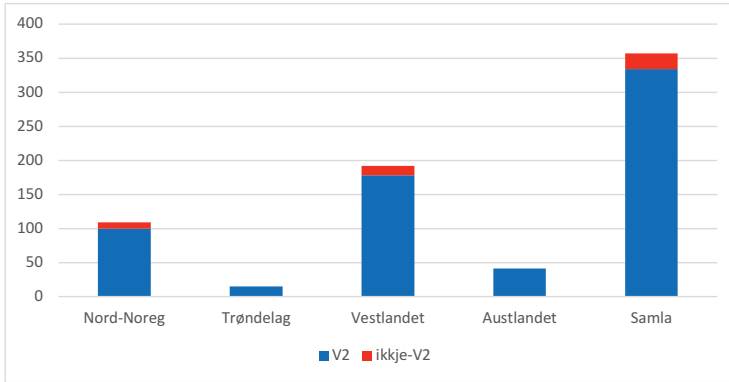
- (10) ka veg du sykkla då ? (volda_uib_0104)
 ‘Kva veg du sykklar, då?’

Gradsspørsmål med *kv*-uttrykk finn me i alt 99 av i LIA. Fem av desse har ikkje-V2, to frå nordnorsk og tre frå vestnorsk. To vestnorske døme er viste her.

- (11) ko mannge så bakkte då ? (fana_uib_0401)
 ‘Kor mange som bakte då?’
- (12) ko mannge leiva me hadde ? (lindaas_uib_0202)
 ‘Kor mange leivar me hadde?’

Me har altså funne til saman 23 døme på komplekse *kv*-uttrykk med ikkje-V2 i LIA-korpuset. Figur 3 (neste side) viser den geografiske utbreiinga av alle dei 357 *kv*-spørsmåla med komplekse *kv*-uttrykk, og som me ser, er delen ikkje-V2 særst liten.⁸

⁸ For alle (dei få) tilfella med ikkje-V2 er det snakk om eit *kv*-uttrykk med to eller fleire stavingar, og når det gjeld einstava variantar av *når* og *korleis*, er dei alt i alt i eit mindretal: 7 av totalt 58 *når*-spørsmål som har den einstava varianten *når*, medan det er 15 av totalt 122 *korleis*-spørsmål som har ein av dei einstava variantane *koss*, *køss*, *kess* eller *voss*. Alle dei einstava variantane står altså i spørsmål med V2-ordstilling.



Figur 3: Kv-spørsmål med komplekse kv-uttrykk i LIA-korpuset fordelt på landslutar

4.4 Oppsummering LIA versus NDK

Når me legg saman tala for korte og komplekse kv-uttrykk i kv-spørsmål med V2 i LIA-korpuset, får me til saman 433 tilfelle. Det samla talet hovudsetningsspørsmål med kv-ledd er 1206.

Me kan no utvida tabell 1 med tal frå LIA-korpuset slik at me får den samla oversikta i tabell 6 (neste side). På overordna nivå er det mykje som er likt. Det er nesten like mange kv-spørsmål i LIA (1206) som i NDK (1332). Det er litt meir V2 i LIA enn i NDK, og dette fordeler seg ganske jamt ut over dei ulike kv-uttrykka. Den største skilnaden mellom korpusa finn me for *når*-spørsmål med 18 prosentpoeng færre ikkje-V2 i LIA, men tala er forholdsvis små. Det er også ein skilnad for *kva*-spørsmål med 8,4 prosentpoeng færre ikkje-V2 i LIA, og her er det rikeleg med døme.

Som me har vore inne på i løpet av framstillinga, er det vanskeleg å seia noko om den diakrone utviklinga av spørsmålet om \pm V2 i kv-spørsmål i norske dialektar utan ein nærmare analyse av tilfanget i LIA versus NDK. Tilfanget i LIA er langt meir heterogent enn i NDK både når det gjeld kronologi og type materiale, og me har også peikt på at det er ein skeivare balanse mellom landslutane enn i NDK. Me har i denne undersøkinga ikkje gått vidare med slike analysar, men me ser heilt klart at det er eit potensial for det, særleg når det gjeld

		<i>N</i>	<i>V2</i>	<i>Ikkje-V2</i>
<i>kva</i>	LIA	638	51,4 %	48,6 %
	NDK	660	43,0 %	57,0 %
<i>kven</i>	LIA	111	55,9 %	44,1 %
	NDK	111	45,0 %	55,0 %
<i>kor/kvar</i>	LIA	100	49,0 %	51,0 %
	NDK	130	52,3 %	47,7 %
<i>når</i>	LIA	58	91,4 %	8,6 %
	NDK	79	73,4 %	26,6 %
<i>kvifor</i>	LIA	11	100,0 %	0,0 %
	NDK	47	97,9 %	2,1 %
<i>korleis</i>	LIA	121	98,3 %	1,7 %
	NDK	128	93,0 %	7,0 %
<i>kv-XP</i>	LIA	167	90,4 %	9,6 %
	NDK	177	95,5 %	4,5 %
Samla	LIA	1206	64,1 %	35,9 %
	NDK	1332	59,6 %	40,4 %

Tabell 6: Samanlikning av *kv*-spørsmål $\pm V2$ i LIA-korpuset og Nordisk dialektsyntaks

dei korte spørjeorda, og særleg når det gjeld Vestlandet og Nord-Noreg, der det er eit rikt tilfang av materiale og der det såleis vil gi meining å sortera materialet nærmare til mindre område innanfor desse regionane. Ikkje minst når det gjeld Vestlandet, kan dette vera av stor interesse i og med at materialet samla i ScanDiaSyn-prosjektet (NDK og NSD-databasen; sjå Westergaard et al. 2012, 2017) tilseier at det er ein god del variasjon innanfor det vestnorske dialektområdet.

5 Konklusjon

I denne undersøkinga har me replikert den korpusundersøkinga av syntaksen i *kv*-spørsmål i norske dialektar som er rapportert i Vangnes og Westergaard (2014). Korpuset den gongen var den norske delen av Nordisk dialektkorpus. I dette arbeidet har me undersøkt det norske LIA-korpuset. Hovudspørsmålet for undersøkinga har vore om resultatata frå LIA-korpuset utfyller eller utfordrar eksisterande kunnskap om fenomenet.

Om me tek utgangspunkt i hovudfunna frå Vangsnes og Westergaard (op. cit.) vist til i innleiinga, er hovudsvaret at undersøkinga av LIA-materialet langt på veg stadfestar det biletet som kjem fram der. Manglande V2 i *kv*-spørsmål førekjem i alle fire hovuddelar av det norske dialektlandskapet (Mæhlum og Røyneland 2012), og totalfordelinga mellom V2 og ikkje-V2 er omtrent 50/50. Vidare er det som i NDK først og fremst i spørsmål med dei korte *kv*-uttrykka *kva*, *kven* og *kor/kvar* at ein finn ikkje-V2 – i spørsmål med andre, morfologisk og/eller semantisk komplekse *kv*-uttrykk er V2 den langt vanlegaste ordstillinga (alltid > 10 %). LIA-materialet viser elles til liks med NDK-materialet at ikkje-V2 er den vanlegaste ordstillinga i Nord-Noreg og at dette endrar seg etter kvart som ein går sørover til Trøndelag, Vestlandet og vidare til Austlandet. Rett nok har me peikt på at det er ein skeiv balanse i tilfanget i LIA ved at det er langt mindre materiale frå Trøndelag og Austlandet enn tilfellet er i NDK.

Ein potensielt interessant skilnad mellom LIA og NDK gjeld det vestnorske materialet. Her er det samla talet på *kv*-spørsmål i LIA omtrent dobbelt så stort som i NDK, og me ser samtidig at det i prosentpoeng er omtrent halvparten så mange tilfelle med ikkje-V2: 20,5 prosent i LIA mot 43,3 prosent i NDK. Dette kan tyda på at strukturen med ikkje-V2 har vorte vanlegare på Vestlandet over tid. Men for å validera dette må det gjennomførast både geografiske og kronologiske analysar av LIA-materialet, og samanlikningar mellom LIA og NDK må gjerast innanfor mindre avgrensa område på Vestlandet. Det har me ikkje gjort i denne omgangen, og det er òg eit spørsmål om LIA-materialet vil vera stort nok til at ein kan seia noko klart om det.

I det heile er det ei utfordring for denne komparative undersøkinga av Nordisk dialektkorpus og LIA-korpuset at tilfanget i sistnemnde er langt meir heterogent både i tidsdjupn, geografi og materialtype. Der NDK-materialet er samla inn systematisk etter ein felles metode med intervju og samtalar frå fire informantar frå ulike stader jamt fordelt ut over landet, er LIA-materialet sett saman av transkriberte opptak av informantar fødde innanfor eit tidsspenn på over hundre år, og sjølv om det er materiale frå dobbelt så mange stader i landet, varierer det i mengde kor mykje ein har frå dei ulike stadene. Samanlikna med NDK-materialet er det særleg påfallande at det er lite austnorsk materiale i LIA.

Når det vidare er omtrent like mange *kv*-spørsmål i LIA som i NDK trass i at LIA-korpuset i storleik (tal på ord) er nesten dobbelt så stort, fortel det oss truleg noko om ein viktig skilnad når det gjeld type materiale: NDK-korpuset inneheld nok ein større andel frie samtalar der spørsmål mellom deltakarane inngår som ein naturleg del. Me har dessutan sett at det er få døme på *kv*-spørsmål frå Trøndelag, noko me vel å tolka som at dette tilfanget i særleg liten grad omfattar frie samtalar.

Referansar

- Elstad, Kåre. 1982. Nordnorske dialektar. Tove Bull og Kjellaug Jetne (red.): *Nordnorsk. Språkarv og språkforhold i Nord-Noreg*. Oslo: Samlaget, 11–100.
- Fiva, Toril. 1996. Spørresetninger i Tromsø-dialekten. *Nordica Bergensia* 9, 139–155.
- Iversen, Ragnvald. 1918. *Syntaksen i Tromsø bymaal*. Kristiania: Bymaalslagets forlag.
- Johannessen, Janne Bondi, Joel Priestley, Kristin Hagen, Tor Anders Åfarli og Øystein A. Vangsnes. 2009. The Nordic Dialect Corpus – an advanced research tool. Jokinen, Kristiina og Eckhard Bick (red.): *Proceedings of the 17th Nordic Conference of Computational Linguistics NODALIDA 2009. NEALT Proceedings Series 4*.
- Lie, Svein. 1992. Ka du sei? *Maal og Minne* 1992, 62–77.
- Lindstad, Arne M., Anders Nøklestad, Janne Bondi Johannessen og Øystein A. Vangsnes. 2009. The Nordic Dialect Database: mapping microsyntactic variation in the Scandinavian languages. Jokinen, Kristiina og Eckhard Bick (red.): *NEALT Proceedings Series 4*.
- Lohndal, Terje, Marit Westergaard og Øystein A. Vangsnes. 2020. Verb Second in Norwegian: Variation and Acquisition. Rebecca Woods og Sam Wolfe (red.): *Rethinking V2*. Oxford University Press.
- Mæhlum, Brit og Unn Røyneland. 2012. *Det norske dialektlandskapet: Innføring i studiet av dialekter*. Cappelen Damm akademisk.

- Nilsen, Hilde. 1996. *Koffer dæm sir det?* Cand. Philol. avhandling, Universitetet i Tromsø.
- Nordgård, Torbjørn. 1985. *Word Order, Binding and the Empty Category Principle*. Cand. Philol. avhandling, Universitetet i Trondheim.
- Nordgård, Torbjørn. 1988. Omkring ordstilling i hv-spørsmål i norske dialekter. *Skriftserie fra Institutt for fonetikk og lingvistikk, Universitetet i Bergen* 33, 26–37.
- Sandøy, Helge. 1987. *Norsk dialektkunnskap*. Oslo: Novus.
- Sollid, Hilde. 2003. *Dialektsyntaks i Nordreisa: Språkdannelse og stabilisering i møtet mellom kvensk og norsk*. Doktoravhandling, Universitetet i Tromsø.
- Reite, André Midtskogseter. 2011. *Spørjing i skedsmokorsmålet: Undersøking og analyse av leddstilling og spørjeord i interrogative hovudsetningar i talemålet på Skedsmokorset*. Mastergradsoppgåve, NTNU.
- Rognes, Stig. 2011. *V2, V3, V4 (and maybe even more): The syntax of questions in the Rogaland dialects of Norwegian*. Mastergradsoppgåve, Universitetet i Oslo.
- Taraldsen, Knut Tarald. 1985. Ka du trur? Tove Bull og Anton Fjellstad (red.): *Heiderskrift til Kåre Elstad*, Universitetet i Tromsø, Institutt for språk og litteratur, 37–47.
- Taraldsen, Knut Tarald. 1986. On verb second and the functional content of syntactic categories. Hubert Haider og Martin Prinzhorn (red.): *Verb Second Phenomena in Germanic Languages*. Dordrecht: Foris, 7–25.
- Vangsnæs, Øystein A. 2004. On *wh*-questions and V2 across Norwegian dialects: a survey and some speculations. *Working Papers in Scandinavian Syntax* 73, 1–59.
- Vangsnæs, Øystein A. 2005. Microparameters for Norwegian *wh*-grammars. *Linguistic Variation Yearbook* 5. Amsterdam and Philadelphia: John Benjamins, 187–226.
- Vangsnæs, Øystein A. 2007a. ScanDiaSyn: Prosjektparaplyen *Nordisk dialektsyntaks*. Torben Arboe (red.): *Nordisk dialektologi og sociolingvistik*. Peter Skautrup Centeret for Jysk Dialektforskning, Århus Universitet, 54–72.

- Vangsnes, Øystein A. 2007b. Scandinavian Dialect Syntax (before and after) 2005. *Nordlyd* 34:1, 7–24.
- Vangsnes, Øystein A. 2008. Decomposing manner *how* in colloquial Scandinavian. *Studia Linguistica* 62(1), 119–141.
- Vangsnes, Øystein A. og Janne Bondi Johannessen. 2011. Nye nordiske isoglossar: Dialektforskning ved bruk av Nordic Syntactic Judgments Database (NSJD). *Svenska Landsmål och svenskt folkliv*, 69–78.
- Vangsnes, Øystein A. og Janne Bondi Johannessen. 2019. The Nordic research infrastructure for syntactic variation: Possibilities, limitations and achievements. *Glossa: A Journal of General Linguistics*, 4(1), 26. doi: <http://doi.org/10.5334/gjgl.708>
- Vangsnes, Øystein A. og Marit Westergaard. 2014. Ka korpuset fortæll? Om ordstilling i hv-spørsmål i norske dialekter. Janne Bondi Johannessen & Kristin Hagen (red.): *Språk i Norge og nabolanda. Ny forskning om talespråk*. Oslo: Novus, 133–151.
- 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, 1–17.
- Westergaard, Marit R. 2003. Word order in *wh*-questions in a North Norwegian dialect: some evidence from an acquisition study. *Nordic Journal of Linguistics* 26(1), 81–109.
- Westergaard, Marit R. 2005. Optional word order in *wh*-questions in two Norwegian dialects: A diachronic analysis of synchronic variation. *Nordic Journal of Linguistics* 28(2), 269–296.
- Westergaard, Marit. 2009. Microvariation as Diachrony: A view from acquisition. *Journal of Comparative Germanic Linguistics*, 12(1), 49–79.
- Westergaard, Marit R. og Øystein A. Vangsnes. 2005. *Wh*-questions, V2, and the left periphery of three Norwegian dialects. *Journal of Comparative Germanic Linguistics* 8, 117–158.
- Westergaard, Marit, Øystein A. Vangsnes og Terje Lohndal. 2012. Norwegian *som*: The complementizer that climbed to the matrix Left Periphery and caused Verb Second violations. *Festskrift for Luigi Rizzi, CISCL (Centro Interdipartimentale di Studi Cognitivi sul Linguaggio)*.

- Westergaard, Marit, Øystein A. Vangsnes og Terje Lohndal. 2017. Variation and change in Norwegian *wh*-questions: The role of the complementizer *som*. *Linguistic Variation* 17(1), 8–43. doi 10.1075/lv.17.1.02wes.
- Åfarli, Tor. 1985. Absence of V2 effects in a dialect of Norwegian. *Working Papers in Scandinavian Syntax* 22, 1–21. Universitetet i Trondheim.
- Åfarli, Tor. 1986. Some syntactic structures in a dialect of Norwegian. *Working Papers in Linguistics* 3, 93–111. Universitetet Trondheim.

English summary

This study replicates a previous corpus-based investigation of *wh*-questions without V2 in Norwegian dialects. The former investigation was based on the Nordic Dialect Corpus (NDC), which comprises recordings from the period 2005–2010. The investigation in this paper is based on the Norwegian LIA corpus, where the material has been collected over a long period of time and where the balance between different parts of the country as well as type of recordings is somewhat different from the NDC. The data from the LIA corpus nevertheless by and large confirms the general impression from the NDC, which is (i) that non-V2 word order first and foremost is found with the short *wh*-words ‘what’, ‘who’ and ‘where’, and (ii) that the non-V2 order is most common in Northern Norwegian dialects and least common in Eastern Norwegian. A possible difference between the two corpora applies to the Western Norwegian material, which has about half as much non-V2 in LIA (20.5%) as in NDC (43.3%). This may indicate that the non-V2 structure has become more common over the years.

Øystein A. Vangsnes
 Institutt for språk og kultur
 UiT Noregs arktiske universitet
 oystein.vangsnes@uit.no
 og
 Høgskulen på Vestlandet
 (professor II)

Marit Westergaard
 Institutt for språk og kultur
 UiT Noregs arktiske universitet
 marit.westergaard@uit.no
 og
 Noregs teknisk-naturvitenskaplege
 universitet NTNU (professor II)



«Lei niogtredve go byggiiimet»

Om unormerte lån fra norsk i samisk talespråk

Lene Antonsen

Denne artikkelen ser på ord annotert som fremmedspråklig ord eller sitatlån i det nordsamiske talespråkkorpuset LIA Sápmi. De mest frekvente orda blir vurdert mot ordbøker og tekstkorpus. Analysen viser at en del av disse orda i realiteten er etablerte lånord i muntlig språk, og mange av dem finnes også i ordbøker, men orda er fraværende i skriftlig språk. Størstedelen av de annoterte orda er likevel spontanlån hvor det innlånte leksikalske elementet tilpasses samisk morfologi og syntaks.

Stikkord: nordsamisk, talespråk, lånord, normering

1 Innledning

De aller fleste som snakker et samisk språk, er også aktive brukere av majoritetsspråket i landet de bor i. Unntaka finnes blant de aller yngste og eldste. Når samtalepartnere behersker de samme to språka, er det naturlig å blande inn språklige elementer fra det andre språket. Dette bærer opptaka i det nordsamiske talespråskorpuset LIA Sápmi preg av. I denne artikkelen ser jeg nærmere på det materialet som er annotert som fremmedspråklige ord og sitatlån i talespråskorpuset. Jeg vurderer de mest frekvente orda og sammenlikner dem med ordbøker og tekstkorpus, og drøfter hvordan vi skal forstå dem som del av samisk språk.

Det har ikke vært skrevet så mye om unomerte lånord i samisk. Nesheim (1952) og Jernsletten (1985) skrev om norske ord og uttrykk i samisk tale i henholdsvis Lyngen og Spansdalen i Troms, og begge tolker dette som uttrykk for et språkskifte i disse områda. Nesheim skriver at samisk språk stod sterkest i hjemlige omgivelser, men maskiner og nye driftsmåter påvirket bruken av samisk også i hjemmemiljøet. Han nevner også at for tall fra 11 og opp blei det brukt norske tallord (Nesheim 1952). Jernsletten beskriver hvordan bruk av norske ord påvirker språkets syntaks og språkets typologiske karakter ved at avledningssystemet går gradvis ut av bruk (Jernsletten 1985).

I nyere tid har Dannemark (2017) sett på nordsamisk talespråk i majoritetsområdet for samisk språk, og undersøkt hvilke elementer fra andre språk kommunestyrerepresentantene i Kautokeino brukte i perioden 2011–2015. Han intervjuet også åtte representanter om hva de sjøl mener om slik bruk. I et annet forskningsarbeid har han sett på bruken av norske tallord i muntlig språk i Kautokeino når det dominante språket i diskursen er nordsamisk. Han har bedt informanter vurdere i hvor stor grad de mener setninger med norske tallord er akseptable, hvor ofte de sjøl bruker slike setninger, og hvor ofte de har hørt andre bruke dem (Dannemark 2014). Jeg kommer tilbake til Dannemarks konklusjoner i analysen. Det kan også nevnes at det er forsket på hvordan tospråklige norsk- og nordsamiske barn bruker

språka sine i rollelek i barnehagen (Kleemann 2015). Dette arbeidet er interessant, men mindre relevant for denne artikkelen.

Artikkelen er bygd opp slik at jeg i kapittel 2 presenterer LIA Sápmi-korpuset og i kapittel 3 beskriver materialet som er annotert som fremmedspråklige ord og sitatlån. Kapittel 4 drøfter hvordan vi skal forstå dette materialet. Til slutt kommer en konklusjon og noen refleksjoner rundt denne.

2 LIA Sápmi – samisk talespråkkorpus

LIA Sápmi Sámegeiela hállangiellakorpus ‘samisk talespråkkorpus’ inneholder transkriberte opptak hentet fra lydbandsamlinga til Nils Jernsletten, som var professor i samisk språk ved UiT Norges arktiske universitet. Dette korpuset er det eneste samiske talekorpuset som er allment tilgjengelig. Det finnes et allment tilgjengelig korpus av skriftlig tekstmateriale: SIKOR, Samisk Internasjonalt Korpus, eid av UiT og det norske Sametinget.¹ Den nordsamiske delen av tekstkorpuset består av til sammen 32 millioner ord. LIA Sápmi er på langt nær så stort (bare 189 000 ord), men det gir en unik mulighet til å forske på samisk talespråk.

Transkripsjonene i LIA Sápmi er gjort i henhold til gjeldende normert ortografi for nordsamisk, og deretter analysert med UiTs analysator Giella-sme som gir lemmaform og grammatisk informasjon.² Giella-sme analysatoren er bygd for å kunne analysere SIKOR, og har en dekningsgrad for det korpuset på 98 % (Antonsen & Trosterud 2017). I materialet fra LIA Sápmi var det mange ord som ikke blei gjenkjent. Hele 14 % av segmenta inneholdt ord som var ukjente for analysatoren, og for å kunne gi en automatisk grammatisk analyse av hele setninga, var det nødvendig å legge til disse orda til Giella-sme.

¹ <http://gtweb.uit.no/korp>. Jeg viser til versjon 06.11.2018.

² <http://github.com/giellalt/lang-sme>, sme er iso-kode 639-2 for nordsamisk.

Materialet inneholder opptak med 94 informanter fra 19 steder i Sápmi,³ de fleste fra norsk side, se figur 1. Av disse er 33 kvinner. Alle informantene snakker nordsamisk. Opptaka ble gjort i perioden 1960–1987, den yngste informanten er 29 år gammel, men de fleste er eldre enn 60 år. På et par av opptaka er det små samtaler mellom informanter, fordi det tilfeldigvis har vært to informanter til stede da opptaket blei gjort, men i de aller fleste opptaka er det bare en informant og intervjuer, og inngangsspørsmålet til informanten er oftest en oppfordring til å fortelle om seg sjøl og hjemplassen.

Det er stor forskjell på hvilken type og hvor mye materiale som er annotert som fremmedspråklige ord, når man sammenlikner områda med samisk som majoritetsspråk med de andre områda. Til området for samisk som majoritetsspråk regner jeg Kautokeino, Karasjok og Anjøl på finsk side, nært Karasjok. For resten av området vurderer jeg at samisk var mindretallsspråk da opptaka blei gjort: Sør-Varanger, Nesseby, Gamvik, Lebesby, Porsanger, Kvænangen, Kåfjord, Storfjord, Ringvassøya, Levanger, Senja og Skånland.



Figur 1: Informantene i LIA Sápmi materialet markert på kartet ut fra oppgitt bosted.

³ 8.11.2020

Tana kommune er resultatet av sammenslåing av to kommuner i 1964. I den delen av kommunen som opprinnelig utgjorde Polmak kommune, snakket hele 72 % samisk som førstespråk (Rasmussen 2005: 70), og i resten av kommunen 28 %. Dette ville gjort det naturlig å dele informantene fra kommunen etter kommunegrensa før 1964, slik at de som tilhører den gamle Polmak kommune, er med i gruppa med samisk som majoritetsspråk. Dessverre er det ikke nøyaktig stedsinformasjon om alle informantene, og de informantene som har slik informasjon, er plassert utenfor gamle Polmak kommune. Jeg kan derfor ikke lage et skille, og derfor klassifiserer jeg alle informantene fra Tana til å være fra område med norsk som majoritetsspråk.

Korpuset inneholder også litt materiale fra svensk side: Karesuando, Kiruna og Jukkasjärvi. Siden svensk er majoritetsspråket for disse informantene, vil dette materialet i liten grad være synlig i denne analysen.

3 Fremmedspråklige ord og sitatlån i LIA-materialet

Det finnes to transkripsjonsveiledninger for LIA Sápmi,⁴ og det er spesielt to av annoteringstaggene som er interessante for denne artikkelen. Taggen *+h* skal legges til *fremmedspråklige* ord tilpasset samisk, og dette gjelder også for norske stedsnavn med samisk kasusending, hvis navna har en godkjent samisk form. Taggen *+n* skal legges til *nordiske sitatlån*, og også lån fra engelsk og andre fremmedspråk. Dette vil blant annet gjelde stedsnavn som har en godkjent samisk form, men som likevel brukes som sitatlån uten samisk kasusending. Hvis det legges en kasusending til det samme navnet, vil det annoteres med *+h*.

I transkripsjonsveiledningene er det forklart at det bare er lånord som ikke finnes i ordbøker eller er vedtatt som et samisk ord av et samisk normeringsorgan, som skal annoteres med *+h* eller *+n*.⁵ I denne artikkelen skal jeg se nærmere på disse orda, og jeg vil ikke

⁴ http://www.tekstlab.uio.no/LIA/pdf/transkripsjonsrettleiing_lia_samisk.pdf,
http://www.tekstlab.uio.no/LIA/pdf/LIA-ortografijja_transkriberen.pdf.

⁵ Sámi Giellagáldu, fram til 2005 var det Sámi Giellaráddi

skille mellom de to annoteringene, men behandler dem under ett fordi en eventuell tilpassing ofte kommer til syne først når ordet blir bøyd. Det gir et galt bilde å bruke ordet *fremmedspråklig* om materiale fra et språk som informantene har forholdt seg til helt fra barndommen, så jeg velger derfor å kalle orda for *annetspråklig* materiale.

Det skjer flere morfologiske prosesser i samiske ord. Tre typer er viktige for diskusjonen i denne artikkelen: 1) Stadieveksling i konsonantgruppa mellom trykksterk og trykklett stavelse (det er 150 forskjellige mønstre). Et eksempel er ordet *soahki* 'bjørk', som har vekslinga *hk:g* og blir *soagi* i akkusativ-genitiv. 2) Endring i trykklett stavelse til *-ii*, *-e* eller *-o* (fra henholdsvis *-i* og *-u*) gir diftongforenkling i første stavelse, slik at flertall akkusativ-genitiv blir *sogiid* (av *soahki*). 3) I trykklette stavelse kan det skje vokalveksling når man legger til bøyningssuffiks, f.eks. blir *-i*- til *-á*- i *soahkái* 'til bjørka' når suffikset *-i* for illativ legges til *soahki*.

3.1 Navn og metaspråk

Alt annetspråklig materiale er transkribert i henhold til majoritets-språkets ortografi istedenfor samisk uttale. Eksempla i denne artikkelen følger rettskrivingen i LIA Sápmi-materialet. Den norske oversettelsen som jeg har lagt til, er en nokså ordrett gjengivelse av den samiske ytringa.⁶

Av ord og fraser om er merket som annetspråklig materiale, er 11 % navn. De fleste er stedsnavn, men det er også navn på organisasjoner, f.eks. *Den Grønlandske Handeli*, som har fått den samiske illativendelsen *-ii* lagt til. Et annet eksempel er navnet på en studieretning som fantes på det som blei kalt yrkesskole: *Jern og metall*. Ifølge transkripsjonsveiledningen skal slike navn merkes som fremmedspråklig hvis det finnes et samisk parallellnavn. Jeg har valgt å ikke vurdere navna i dette materialet, fordi en del av navna ikke har samisk parallellform, slik som de nevnte eksempla. Det kan også være flere grunner til at informanten bruker den norske formen av et

⁶ Hvis man er ønsker å vite mer om de samiske setningene, kan man bruke morfologisk analysator: <https://giellatekno.uit.no/cgi/d-sme.nob.html>, og ordbok: <https://sanit.oahpa.no>.

stedsnavn istedenfor en samisk form, for eksempel kjenner kanskje ikke informanten til den samiske forma hvis navnet er utafor hjemmeområdet. Metaspråklige sitater er annotert med hermetegn, og dette gjelder også når informanten forklarer et samisk ord med et norsk ord. Også dette ser jeg bort fra.

Materialet inneholder i alt 27 523 segmenter. Når jeg ser bort fra stedsnavn og metaspråklige forklaringer, inneholder 6 % av segmenta fra området med samisk som majoritetsspråk annetspråklig materiale (av 6 820 segmenter), mens 12 % av segmenta fra området utafor (av 20 703 segmenter) inneholder slikt materiale.

3.2 Fraser

I materialet fra området med samisk som majoritetsspråk, er det bare fire fraser. Ett av dem er et sitat med adjektiv og substantiv (1) og fra konteksten forstår man at informanten snakker om et norskspråklig oppslag med teksten *kritisk situasjon*. I neste setning (2) vises det til den norsk-svenske reinbeitekonvensjonen, og jeg tolker at denne viser til den norske betegnelsen som nærmest er som et egennavn å regne, sjøl om det er tilpasset ved at sisteleddet i frasens hode, *reinbeitekommišuvdna*, er samisk.

- (1) [...] čuoččui dan oppslágas **kritisk situasjon** [...]
det sto på oppslaget kritisk situasjon
- (2) **lea svenske norske reinbeitekommišuvdna** mearridan
har den svensk-norske reinbeitekommisjonen bestemt
- (3) **ja de mannen gjennom ett halv ovtta jagi**
og så gikk jeg gjennom ett halvt ett år
- (4) **eai leat nu go dat eará alminnelig farga bohccot**
de ikke er slik som de andre alminnelig fargete reinene

Setning (3) inneholder et mengdebegrep, noe som jeg behandler i kapittel 3.3.4. I denne setninga kan det virke som at informanten starter på norsk og lar tallordet *ett* kongruere med det norske substantivet *år*, men så bestemmer han seg likevel å uttrykke seg på samisk. På norsk starter han med preposisjonen *gjennom*, som i dette tilfellet erstattes med akkusativ-genitivformen på samisk, *jagi* (*jahki* ‘år’ i no-

minativ). Men han bruker ikke det samme tallordet på samisk og norsk. I den siste setninga (4) er frasen et adjektiv med bestemmelse. Samisk inneholder veldig mange ord som betegner fargen på reinsdyras pels, og muligens mangler informanten her et samleord for fargene.

Resten av frasene i materialet er fra minoritetsområdet. Noen få av informantene snakker tydeligvis ikke samisk til daglig, og leter etter de samiske orda. Ofte sier de et ord på norsk, og så kommer det samiske ordet rett etterpå, eller de omformulerer ordet til en frase eller forklaring på samisk. Et par av informantene står dermed for en stor del av segmenta med annetspråklig materiale. Jeg ser bort fra materialet fra disse informantene videre i denne analysen.

En del annetspråklige fraser kan man kalle faste uttrykk. De fleste av dem fungerer som adverbialer. Hver av de følgende frasene brukes av bare én informant: *la meg se*, *ærlig talt*, *javisst ja*, *i det hele*, *ære være*, *å jada*, *til slutt*, *så da*, *nær sagt*, *nei og nei*, *i hvert fall*, *i ett*, *i betraktning*, *hele tiden*, *ellers så*, *av og til*. Følgende fraser blir brukt av minst to informanter: *ja ja*, *å nei*, *så og si*, *for eksempel*, *nei nei*, *å ja*. Frasene *ja ja* og *nei nei* kommer ofte i begynnelsen av eller mellom to segmenter.

Mange av frasene er brukt som et fyllord setninga, som *la meg se* (5). Færre fraser er brukt som nødvendig adverbial, som *til slutt* (6). Det er også eksempler på at fraser blir brukt som adjektiv (7) og at informanten siterer norske ordtak eller faste uttrykk (8), men det er interessant at rekkefølgen av orda i *ulønnsomt direkte* er motsatt av det norske uttrykket. Det etterfølgende ordet er subjunksjonen *ahte* 'at' som her blir brukt i samme posisjon (og kanskje samme betydning) i setninga hvor man på norsk ville hatt infinitivsmarket *å*, og man kan dermed tolke det slik at innlånte fraser fra norsk kan påvirke syntaksen i den samiske setninga.

- (5) *dat leat muhtumat sirdán goit dat la meg se Ákšovutnii ja Ráis- Ráisii ja*
noen har flyttet ihvertfall det la med se til Øksfjord og Nordr- Nordreisa og

- (6) *ja de ja de ohppen **til slutt***
og så og så lærte jeg til slutt
- (7) *dat geaidnorehket gal lea **nokså dårlig** [...]*
dette veiregnskapet riktignok er nokså dårlig [...]
- (8) *naa livčče dat lei šaddat **ulønnsomt direkte ahte viežžat** [...]*
nå det ville blitt ulønnsomt direkte å hente [...]

3.3 Enkeltord

Ifølge transkripsjonsveiledningen skal lånord som står oppført i ordbøker eller i lister over godkjente ord fra et normeringsorgan, ikke annoteres som annetspråklig materiale. I dette kapitlet viser jeg at en del av disse orda likevel finnes oppført i ordbøker. Det er ord som vanligvis oppfattes som innblanding av norsk språk, og det er svært få tilfeller av dem i samiskspråklige tekster.

3.3.1 Verb

Samisk morfologi inkluderer bøyingsformer for verb i ni personer og tall (entall, total og flertall) som alle har presens- og preteritumform (ialt 18 morfosyntaktiske former) i indikativ. Passiv er i samisk en ordavledningsprosess, og også det passive verbet kan bøyes i 18 former. Det er i tillegg tre andre modi: imperativ, kondisjonalis og potensialis, hver av dem med ni morfosyntaktiske former. For nomen skilles det mellom seks kasus, som alle har entall og flertallsformer, bortsett fra for essiv som bare har en form. Dette gir i alt 11 former for nomen.

Sammenlagt 536 verb i materialet er annotert som annetspråklig materiale.⁷ Verbet *bygget* (*t* er infinitivsending) er det verbet som er brukt flest ganger i materialet (55 ganger), fra Nesseby i nord til Skånland i sør, og også i majoritetsområdet Kautokeino. Til sammenlikning er ordet *hukset*, som har samme meningsinnhold, brukt 15 ganger i materialet. Verbet *bygget* finnes i Konrad Nielsens ordbok (KNO) utgitt fra 1932 og utover, og oppført med hans ortografi som *big'gít ~ büg'gít*, med norsk oversettelse *bygge (et hus)*. Nordsamisk

⁷ Annoteringen er ikke helt konsekvent, og tallet innbefatter også noen ord som har fått feil ordklasse, så tallet er ikke nøyaktig.

har ikke hatt noe fonem /y/, og formelt sett er ikke y med i det nord-samiske alfabetet. I samisk-norsk ordbok fra 1995 (SNO), er verbet oppført bare med vokalen *i* som *bigget*, med en pil mot verbet *hukset*, som slik presenteres som et «passeligere ord» (SNO: IX). Uttalen blant noen informanter er /ɥ/, som heller ikke er et nordsamisk fonem, men de aller fleste uttaler fonemet som den norske fonemet /y/. Aikio og Ylikoski (under utg.) skriver at dette fonemet, i tillegg til /ø/ og /ɥ/ er etablerte som marginale fonemer, via lånnord som *fylka* ‘fylke’ og *børsa* ‘børs’, og at samisktalende i Norge og Sverige også ofte beholder norsk /ɥ/ i ord som *busse* ‘buss’, som også kan uttales med fonemet /u/.

I skriftlig språk er *bigget/bygget*-verbet lite brukt. I SIKOR er verbet *hukset* ‘bygge’ brukt 7680 ganger, og *bigget* og *biggen*, som er nominalisering av verbet (‘bygging’), er brukt bare 30 ganger, hvorav 11 tilfeller kommer fra transkribert muntlig språk i en og samme vitenskapelige tekst. Verbet er så lite brukt i skriftlig språk, at det ifølge språkfølelsen hos dem som har transkribert for LIA, er blitt annotert som «fremmedspråklig».

I tabell 1 (neste side) er en oversikt over de ti mest frekvente verba som er merket som fremmedspråklige, men tilpasset samisk, og tabellen viser at fire av dem fins i ordbøker. Verbet *drivet* brukes i forbindelse med arbeid, spesielt om fiske, f.eks. å drive med garn og not (9), men også om jakt. Det er ikke så lett å finne et annet verb på samisk som kan brukes slik om å drive med næringsvirksomhet og med en nøytral betydning. Alternative samiske ord ville også inkludere konnotasjonen om at arbeidet er strevsomt, f.eks. *rassat*, som heller ikke ville kunne brukes i samme type syntaktisk konstruksjon som *drivet*, dvs. transitivt med objekt i akkusativ-genitiv (*sallitbivddu*, akkusativ-genitiv av *sallitbivdu* ‘sildefiske’).

- (9) *ja dat drivejedje seamma áiggi sallitbivddu maid gádde-
nuhtiiguin*
og de drev samtidig sildefiske også med landnøter

Det tredje verbet på lista er transkribert som *prahtet* eller *prátet*. Også dette verbet står i KNO oversatt til *prate*, *skravle*. Verbet er ikke opp-

<i>Ant.</i>	<i>Verb</i>	<i>Informanter fra</i>	<i>Ordbøker</i>	<i>SIKOR</i>
54	<i>bygget</i>	Kau, Kvæn, Leb, Nes, Por, Skå, Sto, Tana	KN: <i>big'git</i> ~ <i>biig'git</i> SNO: <i>bigget</i> → <i>hukset</i>	30
22	<i>drivet</i>	Gam, Kau, Kvæn, Skå, Sto, Tana	–	1 ⁸
18	<i>prahet</i> ~ <i>pråtet</i>	Gam, Kvæn, Tana	KN: <i>pråttit</i> ~ <i>bråttit</i> ~ <i>hråttit</i>	–
13	<i>stemmet</i>	Gam, Kvæn, Nes, SV, Tana	KN: <i>stem'mit</i> SNO: <i>stem'met</i> → <i>jienastit, deaivása</i> <i>doallat</i>	8
12	<i>søket</i>	Kau, Por, Skå, SV, Tana	–	–
12	<i>rømmet</i>	Gam, Kar, Nes, Senja, Tana	–	–
12	<i>klåret</i>	Kar, Kau, Kvæn, Tana	–	5
10	<i>påttet</i> ~ <i>passet</i>	Gam, Kar, Kau, Nes, Por, SV, Tana	KN: <i>pas'sit</i> ~ <i>bas'sit</i>	1
9	<i>trenget</i>	Gam: 2 inf.	–	–
8	<i>fortøvet</i>	Tana: 2 inf.	–	–

Tabell 1: De mest frekvente verba som er merket som *fremmedspråklig materiale som er tilpasset samisk*.

ført i SNO, og det finnes heller ikke eksempler på bruk i samiske setninger i SIKOR. I LIA-materialet er det også brukt andre verb med omtrent samme betydning, som *háleštit* og *hállat*, *hupmat*, *sárdnut*, *hoallat*. Disse frekvente verba brukes i hvert sitt nordsamiske dialektområde.

Verbet *stemmet* brukes i to betydninger, å stemme ved valg og at noe er stemmer overens eller holder stikk. I SNO er verbet *jienastit* oppgitt som mer passende for den første betydninga og flerordsuttrykket *doallat deaivása* 'holde for riktig' for den andre betydninga. Dette stemmer overens med SIKOR, hvor bare verbet *jienastit* er brukt for den første betydninga, mens det for den andre betydninga finnes 8 tilfeller av *stemmet*. Uttrykket *doallat deaivása* finnes 535 ganger.

Verbet *søket* er brukt om å søke om lån, arbeid, pleie og skoleplass. Verbet *ohcat* er brukt 58 ganger, både i tilsvarende sammenhenger som *søket*, i tillegg til om å lete. Verbet *rømmet* 'rømme' er brukt i

⁸ I ei setning i en vitenskapelig tekst, transkribert fra et lydopptak.

samme betydning som på norsk. I materialet er det et verb med samme betydning, *báhtarit*, som er brukt dobbelt så mange ganger (24 ganger). Verbet *pásset ~ passet* er brukt om å passe på noe eller noen. I materialet er også verba *fáktet* og *geahččat* brukt. Sammen dekker de bruksområdet som for *pásset ~ passet*, og de har også samme valens.

To informanter i Nervei i Gamvik kommune bruker det transitive verbet *trenget*. Betydningen dekkes av verbet *dárbbášit*, som er brukt 70 ganger i materialet, blant annet av to andre informanter fra Gamvik. Verbet *fortøyet* er brukt av to informanter fra Tana. En annen informant fra Tana og en fra Gamvik bruker verbet *čatnat* om å fortøye båten. Dette verbet har samme valens som *fortøyet*, men brukes generelt om å binde fast noe, og ikke spesifikt om båt.

Alle de ti verba i tabell 1 brukes av minst to informanter, og alle, bortsett fra to av dem, brukes i to eller flere geografiske områder. Fire verb er i KNO, men bare to av dem i SNO. Bare ett av verba er brukt i særlig grad i tekster. For to verb, *drivet* og *fortøyet*, er det vanskelig å finne et samisk verb med like spesifikk semantikk eller samme valens, men disse to verba er likevel nesten helt fraværende i SIKOR.

Alle annetspråksverb er tilpasset samisk morfologi og bøyes i samme kategorier som andre samiske verb, også de verba som opptrer bare en gang i materialet. Samiske verb har tre stammeklasser: likestavelser-, ulikestavelser- og kontrakte stammer. Verb med endelsen *-et* tilhører den siste stammeklassen og kjennetegnes ved at det ikke forekommer stadieveksling i bøyingsparadigmet. De aller fleste annetspråksverba får *-et* ending og blir dermed kontrakte verb uten stadieveksling, noe som gjør tilpassinga lettere. Kontrakte samiske verb har heller ikke diftonger, og dermed ikke diftongforenkling. Norske diftonger skiller seg fra samiske diftonger, og kan derfor være i stammen uten diftongforenkling, f.eks. *feilet* : *feilii* (infinitiv : 3. person entall preteritum). Noen verb får samisk morfologi via avledningssuffiks, som *synestit*, har fått suffikset *-stit* med diminutiv betydning. Verbet får dermed ulikestavelserstamme, og også denne klassen er uten stadieveksling. Ett annet eksempel er *trivestuvášimmet*, som i grunnform er *trivestuvvat*, med avledningssuffikset *-stuvvat*, som egentlig legger til betydning av å være i ferd med å utføre handlingen (Nickel & Sammallahti 2011: 557). Dette

avledningssuffikset gjør verbet til likestavellesstamme som har stadiesveksling *vv.v* i siste takt, det vil si i avledningssuffikset. Det nevnte eksemplet er bøydd i kondisjonals, 1. person flertall ('vi ville trivdes'), med suffikset *-šeimmet* for kondisjonals, person og tall.

Ett verb skiller seg ut ved at det ikke er tilpasset samisk morfologi. En informant fra Gamvik bruker formen *gåran* både som nektelsesform (10) og presens indikativform (11). Dette betyr at verbet bare har én bøyingsform, men det inngår likevel i den samiske syntaksen. Nektingsformen *gåran* ville for et morfologisk tilpasset verb gitt presensformen **gårana* i setning (11). *Gåran* er en tilpassing av det norske *går an*, som er bøydd for presens, og verbet, som er en frase på norsk, beholder morfologi fra långiverspråket.

- (10) *ii dat ii gåran bivdit dan*
ikke det ikke går an å fiske den
- (11) *ii nu olu gal ahte dat gåran bivdit dan*
ikke så mye riktignok at det går an å fiske den

I alt er 447 verblemmaer merket som annetspråklig materiale. Når man tar bort verba som finnes i ordbøker, utgjør verblemmaene som bare er brukt av én enkelt informant, ca. 50 % av de annetspråklige verba. Dette illustrerer hvordan talerne gjerne tar norske verb og tilpasser dem samisk morfologi og syntaks.

3.3.2 Adverb og konjunksjoner

Ordet *ja* er oftest merket som fremmedspråklig materiale. Det er ikke alltid like lett å avgjøre om ordet er en samisk konjunksjon ('og') eller et lån fra norsk. I noen tilfeller er ordet brukt som bekreftende svar på spørsmålet fra intervjueren. De fleste tilfella ordet er merket som sitatlån, er materialet fra mindretallsområdet.

Det finnes eksempler på bruk av konjunksjonene *men*, *og*, *eller*, og de er typisk brukt i begynnelsen av segmentet, og sjelden for å knytte sammen to nominalfraser.

Bare to av orda i tabell 2 (neste side) er med i ordbøkene. Men alle orda i tabellen er frekvente i talespråkkorpuset og brukt over store deler av det nordsamiske området, også i området med samisk

<i>Ant.</i>	<i>Ord</i>	<i>Informanter fra</i>	<i>Ordbøker</i>	<i>SIKOR</i>
165	<i>ja</i>	Gam, Juk, Kar, Kau, Kvæn, Leb, Por, Senja, Skå, SV, Tana	KNO: <i>ja</i> SNO: <i>ja</i>	? ⁹
130	<i>altså</i>	Gam, Kar, Kåf, Kvæn, Leb, Nes, Skå, SV, Tana	--	-
107	<i>nei</i>	Gam, Kau, Leb, Kvæn, Por, Senja, Skå, SV, Tana	--	-
85	<i>men</i>	Kau, Kvæn, Senja, Skå, SV	--	1
76	<i>áhkkurat ~ akkurat ~ áhkkorat</i>	Gam, Kar, Kvæn, Kåf, Leb, Nes, SV, Tana	SNO: <i>áhkkorat</i>	12 <i>áhkkorat</i>
44	<i>jada</i>	Gam, Leb, Por, Skå, SV, Tana	--	-
32	<i>både</i>	Gam, Kar, Kau, Leb, Nes, Skå, SV, Tana	--	-

Tabell 2: De mest frekvente funksjonsorda som er merket som fremmedspråklige materiale eller tilpasset samisk.

som majoritetsspråk. Orda må derfor kunne karakteriseres som etablerte lånord, men bare i muntlig språk. De finnes ikke i SIKOR, bortsett fra ett tilfelle av konjunksjonen *men* og 12 tilfeller av adverbet *áhkkorat*.

3.3.3 Substantiver og adjektiver

Det mest frekvente adjektivet som er merket som annetspråklig materiale, er *heila ~ heilla ~ heile* (61 ganger). Ordet finnes i ordbøkene: *heilâ (især foran tidsord)* (KNO) og *heila ~ heilla ~ heailla* (SNO). I LIA-materialet modifierer dette adjektivet både ord for tid og andre substantiver, som *heilla máilbmi* 'hele verden'. Deretter følger substantiver som begynner med *tysk*: *tysker, tyskelas, tyskerlas* 'tysker'. De to siste orda har fått avledningssuffikset *-las* som kan legges til et substantiv for å uttrykke tilhørighet (Nickel & Sammallahti 2011: 641). Dessuten finnes *tyskagiella (giella* 'språk') i tillegg til adjektivet *tysk*. Dette ordet finnes ikke i ordbøkene, og heller ikke i LIA-

⁹ Ordet har samme form som den samiske konjunksjonen *ja* og det er vanskelig å vurdere bruken av ordet.

materialet i området hvor samisk er majoritetsspråk. Der brukes de tilpassete lånorda *duiska* og *duiskkalaš*, begge ‘tysker’, som også er vanlige i skriftspråket.

Den siste takta av samiske ord bøyes, og det krever at takta har minst to stavelser. Derfor er det vanlig å tilpasse norske enstavelsesord ved å legge til en ekstra vokal *a* til ordet. Dette gjelder for substantiv som *strømma*, som er en tilpassing av det norske ordet for elektrisk strøm. De morfofonologiske prosessene stadiaveksling og vokalveksling blir ofte ikke realisert i nyere lånord. De norske diftongene skiller seg fra de samiske, og det blir derfor heller ikke diftongforenkling i bøyingsparadigmet. Uten veksling i konsonantsentrum er entallsformene for entall nominativ og akkusativ-genitiv homonyme, og slik blir ikke tilpassinga av substantiv like tydelig som for verb.

Ord som dukker opp ofte, er *skøytá* ‘fiskeskøyte’ og *trålár* ‘tråler’. Det er også mange ord som hører til norsk forvaltning av reindrift, som *områdestivra* og *driftsenheta*. Av disse orda finnes *šö'ita* ‘skøyte’ i KNO, ingen av dem finnes i SNO.

En del av substantiva er internasjonale månedsnavn, lånt fra norsk. Noen ganger opptrer de som sammensetning, som *januarmá-nus* (*mánuš* = *mánnu* ‘måned’ i lokativ kasus), men oftest bare med selve månedsnavnet, ofte med kasusending, som *februarai* (med illativ suffiks *-i*, men uten vokalveksling som ville gitt *-ii*). De gamle samiske benevningene for måneder var basert på 13 måneder og passet dermed ikke til den gregorianske kalenderen. I nyere tid er noen av de gamle månedsnavna tatt i bruk, og andre måneder har fått navn der man ikke vet om de er gamle samiske navn eller om de er resultat av kontakt med finsk (Sergejeva 2001). Det er ingen tilfeller av de samiske månedsnavna i materialet. Alle norske månedsnavn, bortsett fra *februar*, finnes i KNO. De samme navna finnes også i SNO, men da med de samiske månedsnavna oppført som mer passende ord. De norske månedsnavna er også i dag vanlige i muntlig språk, kanskje vanligere enn de samiske navna, men i tekst er de fraværende. I SIKOR finnes de norske månedsnavna nesten bare i automatisk norskspråklig datomerking av avisartikler.

3.3.4 Tallord og tallordsfraser

I resten av det annetspråklige materialet utgjør en fjerdedel tallord eller datoer.

Det er mye tallord i korpuset, fordi informantene forteller mye om hendelser som har skjedd tidligere, og 5 % av tallorda er lånt fra norsk.

- (12) *diet lei niogtredve go byggiimet*
det var niogtredve da vi bygde
- (13) *ja munnos leat søtti nylon ja ođđa ledje rieggát*
og vi to har sytti nylon og nye var ringene

Nesten alle tallord i annetspråksmaterialet er årstall (12) eller brukt om målenheter, og da er ofte hele nominalfrasen lånt, som *søtti nylon* (13) eller *treogseksti år*. Ordenstall, som regnes som adjektiver, finnes med norsk lånord bare i noen fraser med datoer.

Dette stemmer overens med Dannemark sine observasjoner av at for datoer, klokkeslett og målinger er norske tallord en relativt etablert uttrykksmåte i samisk diskurs i Kautokeino i 2014. Det er også en tendens til at substantivet etter et norsk tallord kommer på norsk, noe som tyder på at hele nominalfrasen blir oppfatta som en enhet som blir lånt fra norsk (Dannemark 2014).

4 Hvordan skal vi forstå annetspråklig materiale i LIA Sápmi?

Når både intervjuer og informant behersker både samisk og norsk, kan man forvente mer norsk språkmateriale enn hvis intervjuer hadde kommet fra Finland og ikke behersket norsk tillegg til samisk. Informanten henter leksikalske enheter fra norsk språk sjøl om samtalen går på samisk fordi orda vil bli forstått av samtalepartneren. Orda blir integrert i samisk morfologi og syntaks, noe som spesielt er synlig i verba, og tildels i substantiver. Poplack (2004) skiller mellom etablerte lånord og det hun kaller spontanlån ('nonce borrowing', henta fra Weinrich). Skillet går ved at etablerte lånord er frekvente

og kan brukes av enspråklige talere, mens spontanlån hverken er frekvente eller utbredt blant mange talere, og dermed krever de at samtalepartnerne kjenner begge språka. Dette har de til felles med kodeveksling som Poplack definerer som uttrykk hvor taleren veksler mellom to grammatikker, noe som lettest kommer til syne i flerordsuttrykk (Poplack 2004). En del enkeltord kan passe inn i mottaker-språket uten morfologisk tilpassing, og dermed kan de være vanskelig å bestemme, men dette gjelder ikke for verba i dette materialet. Nesten uten unntak er verba tilpassa samisk morfologi.

Jeg har sjekka de mest frekvente orda mot de samme ordbøkene som Dannemark (2017) bruker i sin artikkel. Han kaller ord som ikke finnes i de nevnte ordbøkene for ‘ad hoc-lån’, som er hans oversettelse av ‘nonce borrowing’. LIA-materialet studert i denne artikkelen viser at mange ord som ikke er i ordbøkene, likevel er frekvente og utbredte. Blant disse er også ord som regnes som fyllord i talen, ord som *altså* og *nei*. De er mye i bruk i hele det nordsamiske området, sjøl om de ikke brukes i tekst, og jeg vil ikke kalle dem for spontanlån, men argumentere for at de er etablerte lånord. Utfordringa er å skille mellom spontanlån og etablerte lånord som ikke står i ordbøker. Poplack og Dion (2012) har ved å se på den diakrone utviklinga av engelsk språkmateriale i tre store korpus med Quebec-fransk, etablert forskjell mellom etablerte lånord, og spontanlån som har den egen-skapaen at de forsvinner igjen. LIA Sápmi-korpuset er for lite til å gjøre en slik undersøkelse.

Noen av orda får bekrefta status som etablerte lånord ved at de er i KNO-ordboka som hadde som formål å beskrive det nordsamiske språket slik det var på 1920–30-tallet. SNO-ordboka fra 1995 bygger på KNO, men har utelatt flere av lånorda, og for andre lånord lagt til en anbefaling om å bruke andre ord som oppfattes som mer genuint samiske. SNO er altså, i motsetning til KNO, en normerende ordbok.

Det blir stadig etablert nye lånord i samisk slik som for eksempel *internationála* og *biologiija*. Disse orda normeres av ordbokforfattere eller av normeringsorganet, eller brukeren tilpasser sjøl ordet etter vedtatte regler for tilpassing av lånord (se normeringsregler for lånord i Sametinget 2016). Mange av de normerte formene inneholder bokstaver og bokstavkombinasjoner som vi ikke finner i samisk ar-

vespråkmateriale, som *y* i *analysa* og *dr* i *drámá* ‘drama’. I motsetning til slike normerte ord, begrenser de omtalte lånorda i LIA-materialet seg til muntlig språk. Orda oppleves som fremmedspråklige i skriftlig form av de som har transkribert materialet, sjøl om de nok så sikkert har hørt orda, og kanskje også bruker dem sjøl i muntlig tale.

Ved at orda ikke kommer med i ordbøkene, blir skrivemåten heller ikke normert. I LIA-transkripsjonene er orda delvis skrevet etter norske rettskrivingsregler, og delvis med flere forskjellige varianter av samisk rettskriving. En del av orda inneholder også bokstaven *ø* som ikke brukes i normert samisk språk. Normering av lånord med lyden *ø* gjøres ved å unngå *ø*-fonen, slik som i for eksempel *direktevra* ‘direktør’, sjøl om mange bruker formen *direktøra* i muntlig språk.

De nevnte normerte lånorda er alle lånord fra gresk og latin, og de regnes som internasjonale lånord. Samisktalende generelt er mer positive til internasjonale ord enn ord lånt inn rett fra norsk, som blir regna som uønska norsk påvirkning (f.eks. Dannemark 2017). Dermed blir ord som *strømna* og *skøytá*, ikke godtatt på samme måte som ord fra gresk og latin, og man leter i større grad etter ord fra samisk språk som kan dekke samme betydning, eller man lager et nytt ord. Det nye ordet kan være oversettingslån, som *el-rávdnji* som direkte oversatt blir ‘el-strøm’ på norsk, eller det kan være en ordavledning. De internasjonale månedsnavnene brukes heller ikke i tekst, sjøl om de brukes mye i dagligtalen. Noen begreper får ikke noe normert samisk ord, slik som det sammensatte ordet *fiskeskøyte*, som jeg ikke finner noe ord for i ordbøkene og derfor ikke vet hvordan jeg skal skrive på samisk.

Det er ihvertfall to grunner til at internasjonale lånord foretrekkes i normeringa: Dels er det et puristisk ønske om å unngå for mye norsk påvirkning, og dels er det også et uttalt politisk ønske om å holde nordsamisk som ett språk på tvers av landegrensene (Sametinget 2012: 27). Dette er motivert ut fra å unngå å normere for mange lånord direkte fra norsk, svensk og finsk, men det er samtidig med på å lage et større gap mellom muntlig og skriftlig samisk. Språket som mange bruker til daglig og som fungerer godt i kommunikasjon med andre samisktalende i nærmiljøet, kan ikke skrives. Når man skriver

samisk, så forventes det at man tar ansvar for en overordna språkplanlegging som mange ikke har forutsetninger for å ta del i, hvis man ikke har god skoleopplæring i samisk språk.

5 Konklusjon

LIA-Sápmi er et viktig materiale fordi korpuset inneholder nord-samisk talespråk i større mengde enn vi har hatt tilgang til tidligere. Det er talespråk fra et stort geografisk område og fra en tidsperiode det ellers ikke er mange samiskspråklige kilder fra. Jeg har spesielt sett på materiale som er annotert som fremmedspråklig, som jeg i denne artikkelen har kalt for *annetspråklig*, og jeg viser at nesten alle verb og substantiver er tilpassa samisk morfologi. En stor del av dette er spontanlån. For eksempel er halvparten av verblemmaene bare brukt av en taler hver. Men analysen viser også at mange ord er så vanlige i talespråket at jeg heller mot at de er etablerte lå nord, men de er fraværende i skriftlige tekster. Det er blant annet en rekke eksempler på at ord som er med i Konrad Nielsens deskriptive ordbok (KNO), ikke er tatt med i den nye ordboka fra 1995 (SNO).

Det er en forholdsvis liten andel av samisktalende som skriver samisk. Også blant dem som har samisk som dagligspråk, er skriftspråket norsk, svensk eller finsk – til og med handlelister skrives på majoritetsspråket. Kanskje oppleves forskjellen mellom det muntlige dagligspråket som man behersker, og det man «har lov» til å skrive, som så stort at mange føler det som et ekstra hinder for å skrive samisk.

Forkortinger

Gam = Gamvik, Juk = Jukkasjärvi, Kar = Karasjok, Kau = Kautokeino, Kvæn = Kvænangen, Kåf = Kåfjord, Leb = Lebesby, Nes = Nesseby, Por = Porsanger, Skå = Skånland, Sto = Storfjord, SV = Sør-Varanger.

Referanser

- Aikio, Ante og Jussi Ylikoski. (under utg.). North Saami. I *The Oxford Guide to the Uralic Languages*, redigert av Marianne Bakró-Nagy, Johanna Laakso og Elena Skribnik. Oxford: Oxford University Press.
- Antonsen, Lene og Trond Trosterud. 2017. Ord sett innafra og utafra – en datalingvistisk analyse av nordsamisk. *Norsk lingvistisk tidsskrift*, 35 (1), 153–185.
- Dannemark, Nils. 2014. Norske talord i samisk tale i Guovdageaidnu. *Maal og Minne* 2014 (2), 131–154.
- Dannemark, Nils. 2017. Om språkbruk i kommunestyremøte i Guovdageaidnu. *Norsk lingvistisk tidsskrift*, 35 (1), 186–219.
- Jernsletten, Nils. 1985. Eksempel på språkskifte. I *Heidersskrift til Kåre Elstad*, redigert av Tove Bull og Anton Fjeldstad, 262–281. Tromsø: Institutt for språk og litteratur, Universitetet i Tromsø.
- Kleemann, Carola Babette. 2015. *Lek på to språk. En studie av kodeveksling og språkalternering i tospråklig rollelek på nordsamisk og norsk i en samisk barnehage*. Doktoravhandling, UiT Norges arktiske universitet.
- KNO = Nielsen, Konrad. 1979 (1932–1962). *Lærebok i lappisk (samisk). Grunnet på dialektene i Polmak, Karasjok og Kautokeino*. Oslo: Universitetsforlaget.
- Nesheim, Asbjørn. 1952. «Samisk og norsk i Lyngen.» – *Sámi cellin – Sameliv. Samisk selskaps årbok 1951–1952*.
- Nickel, Klaus Peter og Pekka Sammallahti. 2011. *Nordsamisk grammatikk*. Karasjok: Davvi Girji.
- Poplack, Shana. 2004. Code-switching. I *Sociolinguistics/Soziolinguistik An international handbook of the science of language* (2nd ed.), redigert av Ulrich Ammon, Norbert Dittmar, Klaus J. Mattheier og Peter Trudgill, 589–596. Berlin: Walter de Gruyter.
- Poplack, Shana og Nathalie Dion. 2012. Myths and facts about loanword development. *Language Variation and Change* 24, 279–315. <https://doi.org/10.1017/S095439451200018X>.
- Rasmussen, Torkel. 2005. *Jávohuvvá ja ealáska. Davvisámegielaigiid demografiija ja buolvaidgaskasaš sirdáseapmi Norggas ja Suo-*

- mas*. Sámegiela váldofágadutkamuš. Romsa: Romssa Universitehta Humanisttalaš fakultehta Sámi ossodat.
- Sametinget. 2012. *Sametingsmelding om samisk språk*.
- Sametinget. 2016. *Riektačállinrávvagat*.
- Sergejeva, Jelena. 2001. Tradisjonell samisk tidsregning: årstider og navn på måneder. I *Sámi jienat. Samiske røyster*, redigert av Vivian Aira og Kristin Jernsletten, 75–96. Tromsø: Nival forlag.
- SNO = Kåven, Berit og Johan Jernsletten, Ingrid Nordal, John Henrik Eira og Aage Solbakk 1995: *Sámi-dáru sátnegirji = Samisk-norsk ordbok*. Kárášjohka: Davvi Girji.

English summary

The LIA-Sápmi corpus of spoken language contains only 189,000 words but is an important material since the corpus contains North Sami spoken language in greater quantities than we have had access to before. The material contains recordings of 94 informants from 19 places in Sápmi, mostly from the Norwegian side of the border. The tapes were recorded during the period 1960–1987.

In the article, I look at words annotated as *foreign language words* or as *quote loans*, both of which I in this article have called *second language words*. Of these almost all verbs and nouns are adapted to Saami morphology and syntax, also when they are spontaneous loans. Many of the words are so common in spoken language that I consider them established loanwords, even though they are absent in written texts. Several words that are included in Konrad Nielsen's descriptive dictionary (KNO), are not accepted in normative written language, and are not included in the new dictionary from 1995 (SNO).

A relatively small proportion of Saami speakers write Saami. Also, among those who have Saami as their everyday language, the written language is Norwegian, Swedish, or Finnish - even shopping lists are written in the majority language. Perhaps the gap between

the spoken language and what one is «allowed» to write is experienced as so great that many feel it as an extra obstacle to writing Saami.

Lene Antonsen
Institutt for språk og kultur
UiT Norges arktiske universitet
lene.antonsen@uit.no



VO – OV-variasjon i nordsamisk

Hva kan LIA Sápmi fortelle oss?

Kristine Bentzen

I denne artikkelen undersøker jeg vekslinga mellom VO- og OV-ordstilling i nordsamisk talemål. Jeg har gjort søk i LIA Sápmi – Sámegeiela hállangiellakorpus. Resultatene mine viser ikke uventet at VO overordnet sett er det mest frekvente mønstret. Men i materialet mitt finner jeg også mange tilfeller av SAuxOV-ordstilling, spesielt i setninger med sammensatte verbformer hvor hovedverbet er i infinitiv (dvs. modalkonstruksjoner), og spesielt i kontekster med et pronominalt objekt. I tillegg finner jeg også noen tilfeller hvor objektet kommer foran både hjelpeverb og hovedverb, men uten å være tematisert til setningsinitial posisjon. Med utgangspunkt i Bentzen & Anderssens (2019) analyse av objektsveksling på norsk, foreslår jeg at OV-ordstilling på nordsamisk kan analyseres som IP-intern tematisering, hvor objekter som er kjent i konteksten kan flytte til en temaposisjon mellom VP og TP. Dette vil forklare OV-mønsteret med sammensatte verbformer. Videre foreslår jeg at det kan finnes en høyere IP-intern temaposisjon, og at objektsflytting til denne er det som gir oss OV med finite verb, og også ordstillingsmønstre hvor objektet kommer foran både hjelpeverb og hovedverb.

Stikkord: VO–OV, informasjonsstruktur, nordsamisk, korpus

1 Innledning

Nordsamisk anses vanligvis for å være et SVO-språk, der den nøytrale ordstillinga typisk er som illustrert i (1):¹

- (1) a. Moai vižže **báhpira.**
 Vi-to henter-1du papir
 ‘Vi to henter papir.’
 b. Dál gal fertet jáddadit **dien.**
 Nå prt må-2sg slukke den
 ‘Nå må du vel slukke den.’

Likevel kan vi både i skrift og tale møte på tilfeller av OV-ordstilling:

- (2) a. Ii kakao gal ábut, go **nugátti** borrá.
 neg-3sg kakao prt trenger når nugatti spiser-3sg
 ‘Man trenger ikke kakao når man spiser nugatti.’
 b. Gallgan go mun de **dien** doallat?
 Skal-1sg prt jeg så den holde
 ‘Skal jeg så holde den?’

I grammatikkbøker om nordsamisk vies som regel ordstilling nokså begrenset oppmerksomhet. Ordstillingsvariasjonen mellom SVO og SOV nevnes i flere verker, men diskuteres ikke omfattende. Både Aikio & Ylikoski (2010, 7) og Valijärvi & Kahn (2017, 238) skriver at SOV-ordstilling forekommer med sammensatte verbformer, hvor objektet da kan stå foran den ikke-finite verbformen, som i (2). Nickel & Sammallahti (2011) og Aikio & Ylikoski (under utgivelse) spesifiserer konteksten noe, og påpeker at OV-ordstillinga spesielt forekommer i forbindelse med infinitivkonstruksjoner (se også Sammallahti 1991). Nickel & Sammallahti (2011, 345) sier for eksempel at objektet står *etter* et finitt hovedverb eller en perfektum partisipp,

¹ De nordsamiske eksemplene i (1) og (2) er voksnes ytringer fra barnespråkkorpuset i DASAGO-prosjektet (<http://site.uit.no/dasago/>). Øvrige eksempler i denne artikkelen er nordsamiske eksempler fra LIA Sápmi når ikke annet er angitt.

som i (3a, b), men at det «står *foran* infinittformen som fungerer som utfyllingsadverbial (konverbadverbial) for hovedverbet...», slik som i (4).

- (3) a. Mun oidnen **Niillasa**.
 Jeg så-1sg Nils
 'Jeg så Nils.'
 b. Mun lean viežžan **čázi**.
 Jeg har-1sg hentet vann
 'Jeg har hentet vann.'
 c. Máhtte dáidá oastit **biilla**.
 Mattis skal-kanskje-3sg kjøpe bil
 'Mattis skal kanskje kjøpe en bil.'
- (4) a. Mun àiggun **čázi** viežžat.
 Jeg skal-1sg vann hente
 'Jeg skal hente vann.'
 b. Mun veigan **doaktára** viežžat.
 Jeg springer-1sg legen hente
 'Jeg springer for å hente legen.'

Videre påpeker Nickel & Sammallahti (2011, 347) at SOV-ordstilling kan brukes for å fremheve eller legge trykk på objektet. Aikio & Ylikoski (under utgivelse) på sin side sier at man gjerne bruker SOV-ordstillinga når objektet er «lett» og referenten til objektet er kjent eller tidligere nevnt i diskursen. Det er altså tilsynelatende både rent syntaktiske og informasjonsstrukturelle faktorer som påvirker variasjonen mellom SVO og SOV i nordsamisk.

Med dette som bakgrunn har jeg gjort søk i LIA Sápmi – Sáme-giela hállangiellakorpus,² som er et korpus av nordsamiske dialekter, for å kartlegge hvordan denne ordstillingsvariasjonen opptrer i faktisk talespråk. I denne artikkelen presenterer jeg noen preliminnære funn fra denne undersøkelsen.

² <http://tekstlab.uio.no/LIA/samisk/index.html>

2 Metode

LIA Sápmi er et talespråkskorpus med nordsamiske dialekter fra 19 ulike steder. Det består av materiale fra 122 informanter og er innspilt i tidsrommet 1960-1987. Korpuset inneholder nesten 190 000 ord.

Ved hjelp av søkegrensesnittet *Glossa*³ har jeg søkt ut kontekster med objekter (dvs. substantiver og pronomener i akkusativ form) foran eller etter finite og infinite verb i dette korpuset, som illustrert i (5):⁴ Merk at negasjon er et hjelpeverb på nordsamisk, og brukes sammen med en egen negasjonsform av hoved verbet (konnegativ). I kontekstene illustrert i (5a,b,e) var subjektet (dvs. substantiv eller pronomen i nominativ form) inkludert i søkestrengen. Nordsamisk er et prodrop-språk, og dersom et pronominelte subjekt er utelatt, er det vanskelig å vurdere om et streng-initialt objekt er et resultat av objektstematisering til setningsinitial posisjon eller en faktisk SOV-kontekst. For å kontrollere for dette er altså et eksplisitt subjekt tatt

³ <https://tekstlab.uio.no/glossa2/saami>

⁴ Jeg har ikke spesifisert maks antall ord mellom søkeordene. De konkrete søkestrengene jeg brukte i *Glossa* var som følger:

For setninger med finite hovedverb:

- 1) Substantiv/pronomen_{NOMINATIV}—V_{PRESENS/PRETERITUM}—**substantiv**_{AKKUSATIV}
- 2) Substantiv/pronomen_{NOMINATIV}—V_{PRESENS/PRETERITUM}—**pronomen**_{AKKUSATIV}
- 3) Substantiv/pronomen_{NOMINATIV}—**substantiv**_{AKKUSATIV}—V_{PRESENS/PRETERITUM}
- 4) Substantiv/pronomen_{NOMINATIV}—**pronomen**_{AKKUSATIV}—V_{PRESENS/PRETERITUM}
- 5) Substantiv/pronomen_{NOMINATIV}—**pronomen**_{AKKUSATIV}+**substantiv**_{AKKUSATIV}—
V_{PRESENS/PRETERITUM}

For setninger med sammensatte verbformer:

- 6) V_{PRESENS/PRETERITUM}—V_{INFINITIV/PERFPART/PRESPART/AKTIOESSIV}—**substantiv**_{AKKUSATIV}
- 7) V_{PRESENS/PRETERITUM}—V_{INFINITIV/PERFPART/PRESPART/AKTIOESSIV}—**pronomen**_{AKKUSATIV}
- 8) V_{PRESENS/PRETERITUM}—**substantiv**_{AKKUSATIV}—V_{INFINITIV/PERFPART/PRESPART/AKTIOESSIV}
- 9) V_{PRESENS/PRETERITUM}—**pronomen**_{AKKUSATIV}—V_{INFINITIV/PERFPART/PRESPART/AKTIOESSIV}
- 10) V_{PRESENS/PRETERITUM}—**pronomen**_{AKKUSATIV}+**substantiv**_{AKKUSATIV}—
V_{INFINITIV/PERFPART/PRESPART/AKTIOESSIV}
- 11) Subst/pron_{NOMINATIV}—**substantiv**_{AKKUSATIV}—V_{PRESENS/PRETERITUM}—
V_{INFINITIV/PERFPART/PRESPART/AKTIOESSIV}
- 12) Subst/pron_{NOMINATIV}—**pronomen**_{AKKUSATIV}—V_{PRESENS/PRETERITUM}—
V_{INFINITIV/PERFPART/PRESPART/AKTIOESSIV}
- 13) Subst/pron_{NOMINATIV}—**pronomen**_{AKKUSATIV}+**substantiv**_{AKKUSATIV}—V_{PRESENS/PRETERITUM}—
V_{INFINITIV/PERFPART/PRESPART/AKTIOESSIV}

Søkestrenger (v), (x) og (xiii) ga treff på demonstrative nomenfraser i OV-kontekster. Søkestrenger (ii) og (vii) ga i tillegg til pronomielle objekter også treff på demonstrative nomenfraser i VO-kontekster.

- (5) *Setninger med finitt hovedverb:*
- a. Substantiv/Pronomen_{NOM}—V_{FIN}—Substantiv/Pronomen_{AKK} (SVO)
 - b. Substantiv/Pronomen_{NOM}—Substantiv/Pronomen_{AKK}—V_{FIN} (SOV)
- Setninger med sammensatte verbformer:*
- c. Au_{XFIN}/Mod_{FIN}/Neg—V_{INFIN}—substantiv/pronomen_{AKK} (SAuxVO)
 - d. Au_{XFIN}/Mod_{FIN}/Neg—Substantiv/Pronomen_{AKK}—V_{INFIN} (SAuxOV)
 - e. Subst/pron_{NOM}—Subst/Pron_{AKK}—Au_{XFIN}/Mod_{FIN}/Neg—V_{INFIN} (SOAuxV)

med i disse kontekstene for å utelukke setninger hvor objektet er tematisert til setningsinitial posisjon. Resultatene ble gjennomgått manuelt og irrelevante treff ble filtrert ut. Deretter ble kontekster med pronominale og nominale objekter telt hver for seg. Videre ble kontekstene med sammensatte verbformer i (5c, d, e) sortert i perfektumskonstruksjoner, modalkonstruksjoner og negative konstruksjoner. I tillegg ble informantene delt i fire grupper etter fødselsår: (i) født før første verdenskrig (49 informanter), (ii) født mellom 1913 og 1939 (38 informanter), (iii) født etter 1940 (11 informanter), og (iv) ukjent fødselsår (24 informanter). Til slutt, ble også opptakssted tatt i betraktning for å avdekke eventuelle dialektforskjeller.

3 Resultater

Som nevnt i innledningen har flere påpekt at OV-mønsteret forekommer oftere i kontekster med sammensatte verbformer enn i kontekster med et finitt hovedverb. Mine undersøkelser i talemålskorpuset støtter opp under dette. Som vi ser i oversikten i (6), er det en sterk dominans for ordstillinga VO i setninger med et finitt hovedverb. Denne dominansen er enda tydeligere dersom objektet er et nomen (kun 7,6 % tilfeller av OV), enn dersom objektet er et pronomen (15,6 % tilfeller av OV).

- (6) *Setninger med finitt hovedverb og nominalt objekt:*
- a. S-V_{FIN}-O_{NOMEN} 232 (92,4 %)
 - b. S-O_{NOMEN}-V_{FIN} 19 (7,6 %)
- Setninger med finitt hovedverb og pronominalt objekt:*
- c. S-V_{FIN}-O_{PRON} 162 (84,4 %)
 - d. S-O_{PRON}-V_{FIN} 30 (15,6 %)

Eksemplene i (7) viser utsagn med de ulike mønstrene fra korpuset.

- (7) a. Mun lohken **dáččagirjjiid**. (SVO_{NOM})
 Jeg leste norske.bøker
 'Jeg leste bøker på norsk.'
- b. Gánddat **muoraid** gudde. (SO_{NOM}V)
 Guttene veden bar
 'Guttene bar veden.'
- c. Ovdalaš áiggi dat sálteje **daid**. (SVO_{PRON})
 Tidligere tid de saltet dem.
 'Før i tida saltet de dem'
- d. Gal mun **dien** dagan. (SO_{PRON}V)
 Prt jeg dette gjør
 'Jeg gjør dette.'

Når vi ser på setninger med sammensatte verbformer under ett, finner vi at OV-mønsteret er langt vanligere enn i setninger med finitte hovedverb (36,6 % OV). Men dersom vi deler opp disse kontekstene i perfektumskonstruksjoner, modalkonstruksjoner og negative konstruksjoner, trer det frem et enda tydeligere mønster. I perfektumskonstruksjoner er det, som ved finitte hovedverb, en klar overvekt av VO. Igjen er tendensen noe mer uttalt ved nominale objekter (16,3 % OV) enn ved pronominal objekter (21,4 % OV), som vist i (8).

(8) *Perfektumskonstruksjoner med nominalt objekt:*

a. S-AUX _{FIN} -V _{PERFPART} -O _{NOMEN}	72	(83,7 %)
b. S-AUX _{FIN} -O _{NOMEN} -V _{PERFPART}	14	(16,3 %)

Perfektumskonstruksjoner med pronominalt objekt:

c. S-AUX _{FIN} -V _{PERFPART} -O _{PRON}	33	(78,6 %)
d. S-AUX _{FIN} -O _{PRON} -V _{PERFPART}	9	(21,4 %)

Eksemplene i (9) illustrerer de ulike mønstrene med data fra korpuset.

Det samme mønsteret finner vi i negative konstruksjoner. Negative setninger med nominalt objekt har VO-ordstilling hele 93,4 % av tida, mens for negative setninger med pronominalt objekt er andelen VO-ordstilling på 76,2 %. Til tross for at OV-ordstillinga er lite brukt i denne setningstypen, finner vi i negative konstruksjoner

- (9) a. Gumpe lei gáskán **gieda**. (SAUXVO_{NOM})
 Ulven hadde bitt hânda
 ‘Ulven hadde bitt hânda.’
- b. Ledje **badjegoadi** ceggen. (SAUXO_{NOM}V)
 Hadde fjellgammen satt.opp
 ‘De hadde satt opp fjellgammen.’
- c. Dat lea čállán **dan** (SAUXVO_{PRON})
 han/hun har skrevet det
 ‘Han/hun har skrevet det.’
- d. Muitalin mun dat gal lean **dáid** báhcán. (SAUXO_{PRON}V)
 Fortalte jeg prt prt har disse skutt
 ‘Jeg fortalte at jeg hadde skutt disse.’

noen få tilfeller hvor objektet forekommer foran både hovedverbet og hjelpeverbet (men altså etter subjektet):

- (10) *Negative konstruksjoner med nominalt objekt:*
- | | | |
|--------------------------------|----|----------|
| a. S-Neg-V-O _{NOMEN} | 57 | (93,4 %) |
| b. S-Neg-O _{NOMEN} -V | 2 | (3,3 %) |
| c. S-O _{NOMEN} -Neg-V | 2 | (3,3 %) |
- Negative konstruksjoner med pronominalt objekt:*
- | | | |
|-------------------------------|-----|----------|
| d. S-Neg-V-O _{PRON} | 115 | (76,2 %) |
| e. S-Neg-O _{PRON} -V | 31 | (20,5 %) |
| f. S-O _{PRON} -Neg-V | 5 | (3,3 %) |

Eksemplene i (11) illustrerer de ulike mønstrene med data fra korpuset:

- (11) a. Ja dat eai dárbbat **dieid** herggjíd. (SNegVO_{NOM})
 Og de ikke trenger de reinene
 ‘Og de trenger ikke de reinene.’
- b. Ii asta **eambo njeallje diimmu** døgnet sove (SNegO_{NOM}V)
 Ikke rekker mer fire timer jándoris nohkat
 ‘Han/hun rekker ikke å sove mer enn fire timer i døgnet.’
- c. Na mun in diede **dan**. (SNegVO_{PRON})
 Prt jeg ikke vet det
 ‘Jeg vet ikke.’
- d. In **dan** ge diede. (SNegO_{PRON}V)
 Ikke-Isg den prt vet
 ‘Jeg vet ikke det heller.’

Når vi så ser på modalkonstruksjonene, finner vi et annet mønster. Som nevnt har både Nickel & Sammallahti (2011) og Aikio & Ylikoski (under utgivelse) påpekt at OV typisk forekommer med infinitivsformer av hoved verbet. Korpusmaterialet viser også dette. Som det fremkommer i (12), har modalkonstruksjoner med nominale objekter ordstillinga S-Mod-O-V hele 41,3 % av tida, og med pronominal objekter er faktisk S-Mod-O-V det dominerende mønstret (50,7 %). I tillegg finner vi også her noen tilfeller hvor objektet kommer foran både hoved verbet og det modale hjelpe verbet:

- (12) *Modalkonstruksjoner med nominalt objekt:*
- | | | |
|---|-----|----------|
| a. S-Mod _{FIN} -V _{INF} -O _{NOMEN} | 119 | (57,7 %) |
| b. S-Mod _{FIN} -O _{NOMEN} -V _{INF} | 85 | (41,3 %) |
| c. S-O _{NOMEN} -Mod _{FIN} -V _{INF} | 2 | (1 %) |
- Modalkonstruksjoner med pronominalt objekt:*
- | | | |
|--|----|----------|
| d. S-Mod _{FIN} -V _{INF} -O _{PRON} | 63 | (45 %) |
| e. S-Mod _{FIN} -O _{PRON} -V _{INF} | 71 | (50,7 %) |
| f. S-O _{PRON} -Mod _{FIN} -V _{INF} | 6 | (4,3 %) |

Eksemplene i (13) illustrerer mønstrene fra korpuset:

- (13) a. ja deinna galgá čiktít **nuohti**. (SModVO_{NOM})
 Og med.det skal bøte nota
 ‘Og med dette skal man bøte nota.’
- b. Dat fertejit **ráidalasa** lágídit. (SModO_{NOM}V)
 De må stigen ordne
 ‘De må ordne stigen.’
- c. De mun ferten oahppat **dan** maiddái. (SModVO_{PRON})
 Så jeg må lære det også
 ‘Så må jeg lære det også.’
- d. Da galggašii **mu** váldit dál. (SModO_{PRON}V)
 Han/hun skulle meg ta nå
 ‘Han/hun skulle ta med nå.’

Det ble kontrollert for informantenes fødselsår, men det så ikke ut til at noen aldersgrupper anvendte OV i større grad enn andre, og jeg konkluderer derfor ut fra disse dataene at denne fordelinga av VO og OV har holdt seg nokså stabil for informanter født fra 1890-tallet og

frem til 1950-tallet. Geografisk tilhørighet ser heller ikke ut til å være en avgjørende faktor.

Resultatene fra denne korpusundersøkelsen bekrefter altså i stor grad det som har blitt beskrevet i grammatikkbøkene. OV i nordsamisk er mulig, men ikke spesielt vanlig, i kontekster med et finitt hovedverb, i perfektumkonstruksjoner og i negativkonstruksjoner. I modalkonstruksjoner derimot, forekommer denne ordstillinga omtrent like ofte som ordstillinga VO.

Disse funnene åpner for flere spørsmål. Hvorfor er OV vanligere foran infinitte enn finitte verb? Hvorfor er det forskjell på perfektumkonstruksjoner og modalkonstruksjoner, hvor OV er mye vanligere i kombinasjon med sistnevnte? Og hvordan kan vi forklare den variasjonen mellom VO og OV vi finner innenfor samme konstruksjonstype, spesielt i modalkonstruksjoner?

4 Diskusjon

Mønsteret for VO-OV-variasjon som vi finner i nordsamisk, har endel klare likhetstrekk med samme fenomen i finsk. Både Vilkuña (1995) og Holmberg (2001, 2017) finner at SOV på finsk typisk brukes når en annen konstituent i setninga er fokuselement, slik som illustrert i (14) (fra Holmberg 2017, 297):

- (14) a. Anne osti **auton**. (finsk)
 Anne kjøpte bil
 'Anne kjøpte en bil.'
- b. ANNE **auton** osti.
 Anne bil kjøpte
 'Det var Anne som kjøpte en bil.'
- c. Milloin Anne **auton** osti?
 Når Anne bil kjøpte
 'Når kjøpte Anne en bil?'
- d. Milloin hän oli **auton** osti?
 Når hun hadde bil kjøpt
 'Når hadde hun kjøpt en bil?'

Holmberg (2001, 2017) argumenterer for at OV-ordstillinga følger av en defokusering av objektet. Ifølge Holmberg (2017) kan OV på finsk genereres på to ulike måter. I konstruksjoner som (14b, c) har finsk en operasjon som tilsvarer objektsveksling på fastlandsskandinavisisk, som Holmberg (1999) analyserer som en defokuseringsoperasjon. Objekter som ikke er fokuselement i setninga må flytte ut av fokusdomenet, dvs ut av VPen. Men som på fastlandsskandinavisisk er slik objektsveksling betinget av Holmbergs Generalisering (HG). I henhold til HG kan objektet kun flytte ut av VPen dersom hovedverbet også gjør det. I fastlandsskandinaviske språk fører objektsveksling til at objektet står foran setningsadverbialer, men ikke foran det finitte hovedverbet, ettersom disse språkene er V2-språk hvor det finitte verbet alltid er andre ledd i setninga. Finsk er derimot ikke et V2-språk. Selv om det finitte hovedverbet har flyttet ut av VPen i (14b, c), så har det flyttet til en lavere posisjon enn det som er landingsposisjonen for det vekslede objektet. Resultatet av objektsvekslinga blir derfor ordstillinga OV. Ordstillinga i (14d) vil derimot ikke kunne forklares på denne måten ettersom vi her vil anta at det infinitte verbet ikke har flyttet ut av VPen. Objektsveksling vil da bryte mot HG. Holmberg (2017) argumenterer for at eksempler som (14d), som kun skal være mulig med initialt fokus på finsk, er et resultat av at verbet er selektert for å følge etter objektet, altså at OV er den underliggende ordstillinga.

Man kunne tenke seg en tilsvarende analyse av dataene jeg har presentert for nordsamisk her. I så tilfelle ville alternativet med OV som underliggende ordstilling være vanligere enn defokuseringsflytting av objektet i nordsamisk. Som vi husker, er det nokså få tilfeller av OV med finitte hovedverb i LIA Sápmi-materialet.

Jeg vil her heller utforske en annen mulig analyse som forsøker å gi ei helhetlig forklaring for alle mønstrene vi finner. I en alternativ analyse av objektsveksling i fastlandsskandinavisisk foreslår Bentzen & Anderssen (2019) at objektsveksling ikke er en form for defokusering, men heller ei slags tematisering. De foreslår at objektsveksling oppstår når objektet har en eller annen form for tematiseringsegenskaper, enten ved at det refererer til noe som er kjent eller gitt i konteksten og del av en felles kontekst for samtale-

partnerene (“familiært” tema) eller at det utgjør et videreført tema i samtalen. Bentzen & Anderssen foreslår at landingsposisjonen til objektet er en IP-intern temaposisjon som må være nokså høyt oppe i IP-omenet ettersom vekslende objekter må stå foran alle slags adverb.

Som vi husker fra avsnitt 2, bemerker Aikio & Ylikoski (under utgivelse) at OV gjerne brukes på nordsamisk når objektet refererer til noe som er kjent eller allerede nevnt i diskursen. En mulighet er dermed å analysere OV-mønsteret i nordsamisk som et resultat av IP-intern tematisering av objektet, på samme måte som objektsveksling i fastlandsskandinaviske. Det vil likevel være en del forskjeller mellom fastlandsskandinaviske objektsveksling og nordsamisk OV. For det første kan alle typer objekter forekomme med OV i nordsamisk, mens objektsveksling i fastlandsskandinaviske kun gjelder for pronominal objekter. Videre er fastlandsskandinaviske objektsveksling som nevnt begrenset av Holmbergs Generalisering (Holmberg 1999). OV i nordsamisk vil derimot ikke være begrenset på denne måten. Dermed vil ikke objektsflytting i nordsamisk nødvendigvis alltid maskeres av påfølgende verbflytting, og kan derfor føre til ordstillinga OV.

Med utgangspunkt i Bentzen & Anderssens (2019) analyse av objektsveksling i norsk vil jeg her foreslå en analyse av VO-OV-veksling i nordsamisk som en form for IP-intern tematisering.⁵ Vi antar at S-(Aux)-V-O er den grunnleggende ordstillinga. Den brukes når objektet er nøytralt med hensyn til tematisk tolkning og dermed forblir i VP, som i (7a, c), (9a, c), (11a, c) og (13a, c). Leddstillinga S-Aux/Mod/Neg-O-V oppstår derimot når objektet har tematiske egenskaper. Objektet flytter da til en IP-intern temaposisjon mellom VP og TP. Under antagelsen om at infinitte verb ikke flytter ut av VP, så vil dette kunne forklare OV i sammensatte verbformer, som i (9b, d), (11b, d) og (13b, d).

⁵ Merk at dette ikke egentlig er inkompatibelt med Holmbergs defokuseringsanalyse. Et spørsmål vil være hva som er motivasjonen for objektsflytting – å komme seg ut av fokusdomenet fordi noe annet i setninga utgjør fokus, eller at objektet så å si har en egen agenda, nemlig at det selv utgjør et tema, og må flytte for å få lisensiert denne tematiske tolkninga.

I konstruksjoner med finitt hovedverb vil slik IP-intern tematisering som regel likevel ikke føre til OV fordi nordsamisk ser ut til å flytte det finitte hoved verbet til T. Det er uavhengige indikasjoner på at finitte verb flytter til en høyere posisjon i IP-domenet i nordsamisk, for eksempel T, i og med at finitte verb som regel står foran adverb, slik som i (15).

- (15) Geasa Ánde luoikkahii **dábálaččat** rudaid? (eget eksempel)
 til.hvem Ánde lånte vanligvis penger
 'Til hvem lånte Ánde vanligvis penger?'

Men som vi så i avsnitt 3, så *kan* OV også brukes med finitte hovedverb. Hvordan kan vi da forklare dette? En mulighet er at flytting av finitte hovedverb til T er valgfritt. Et mulig støtteargument for dette er at talere av nordsamisk også tillater at finitte hovedverb forekommer *etter* adverb i eksempler som det i (15), en indikasjon på manglende verbflytting. Men i et slikt scenario måtte man tenke seg at OV med finitte verb da oppstår når man *både* har et IP-internt tematisert objekt og samtidig benytter muligheten av å la det finitte verbet bli igjen i VP.

Det ville ikke være noen spesielt elegant analyse. Et alternativ er å anta at det fins to IP-interne tematiseringsposisjoner, en mellom VP og TP og en som befinner seg helt øverst i IP-domenet. Den sistnevnte posisjonen vil da være høyere enn posisjonen det finitte verbet flytter til, men lavere enn subjektet (og altså en annen posisjon enn den som benyttes ved setningsinitial tematisering). S-Aux/Mod/Neg-O-V vil dermed involvere tematisering til den lavere IP-interne tematiseringsposisjonen, mens SOV_{FIN} vil involvere tematisering til den høyere tematiseringsposisjonen i IP-domenet. En uavhengig grunn til å anta en slik høyere tematiseringsposisjon i tillegg til den lavere er at det, som nevnt, faktisk finnes tilfeller hvor objektet i en sammensatt verbkonstruksjon plasseres foran både det infinitte og det finitte verbet – men etter subjektet – altså uten å være tematisert til setningsinitial posisjon. (16) viser et eksempel på dette fra LIA Sápmi-korpuset:

- (16) Movt dat **dan** galgá dadjat sámegillii? (SO_{PRON}ModV)
 Hvordan man det skal si samisk.ILL
 'Hvordan skal man si det på samisk?'

Som vi så i avsnitt 3, er det nokså få tilfeller av konstruksjoner som den i (16) i korpuset, og det kan være ei mulig forklaring på hvorfor det også er få tilfeller av OV med finitte verb. Dette kan bety at denne høyere IP-interne temaposisjonen er mindre benyttet enn den lavere. Dermed kan det godt være at det ikke er mer uvanlig med (lav) IP-intern tematisering i konstruksjoner med finitte hovedverb enn i konstruksjoner med sammensatte verbformer. Dersom objektet flytter til den lavere temaposisjonen, vil påfølgende verbflytting til T maskere denne objektsflyttinga og likevel gi VO som ordstilling.

5 Avslutning

Denne undersøkelsen av VO-OV-veksling i nordsamisk basert på talespråksdata fra LIA Sápmi-korpuset har bekreftet det som sies om ordstilling i nordsamiske grammatikkbøker. VO er klart det mest frekvente mønsteret, men OV er også mulig. Det fremkommer også av dataene at OV er mer frekvent i sammensatte verbformer med et verb i infinitiv enn i perfektumskonstruksjoner, negasjonskonstruksjoner og konstruksjoner med et finitt hovedverb. Faktisk er OV mer brukt enn VO i modalkonstruksjoner med pronominale objekter. Jeg har her foreslått en analyse hvor samisk har to IP-interne temaposisjoner, en mellom VP og TP, og en høyt oppe i IP-domenet (men lavere enn subjektsposisjonen). Det kan se ut til at tematisering til den lavere posisjonen er vanligere enn til den høye. Tematisering til den lavere posisjonen vil gi ordstilling Aux-O-V_{PERFPART}, Neg-O-V og Mod-O-V_{INF}, som i (9b, d), (11b, d) og (13b, d), mens tematisering til den høyere posisjonen vil gi S-O-V_{FIN} som i (7b, d) og S-Aux-O-V_{PERFPART}/S-Neg-O-V/S-Mod-O-V_{INF} som illustrert i (16).

Det er fortsatt mange åpne spørsmål som må følges opp i videre forskning. For det første, må det undersøkes nærmere i hvor stor grad det faktisk er samsvar mellom OV og en tematisk tolkning av objek-

tet. Dette vil kreve en grundig analyse av diskursene rundt eksemplene jeg har funnet i korpuset, og det ville også ha vært ønskelig å utføre et eksperiment med morsmålstalere hvor man manipulerer tema i diskursen og undersøker hvilke ordstillingspreferanser dette fører til. Videre vil den foreslåtte analysen kreve at det er ytterligere indikasjoner for de to ulike temaposisjonene, for eksempel ved at de gir ulike tematiske tolkninger. Man kunne muligens tenke seg at den høyere posisjonen har en kontrastiv tolkning, mens den lavere handler om hvorvidt objektets referent er gitt eller et tema i konteksten. Koblet til dette, vil det selvsagt også være viktig å avdekke mer nøyaktig hvor landingsposisjonene til objektet befinner seg i strukturen. Muligens vil man kunne belyse dette aspektet ved å undersøke hvor objekter i nordsamiske OV-konstruksjoner befinner seg i forhold til ulike adverb. Som et siste moment er det også et åpent spørsmål hvorfor OV er så mye mer frekvent i modalkonstruksjoner enn i perfektumskonstruksjoner og negative konstruksjoner. Er det noe som gjør at modalkonstruksjoner oftere enn perfektumskonstruksjoner og negative konstruksjoner har et objekt med tematiske egenskaper? Ytterligere undersøkelser må åpenbart til for å kunne belyse alle disse spørsmålene.

Takk

Tusen takk for innspill fra en anonym fagfelle. Alle feil og mangler står jeg selvsagt for selv.

Referanser

Aikio, Ante & Jussi Ylikoski (under utgivelse). North Sami. I *Oxford Guide to the Uralic Languages*, redigert av Marianne Bakró-Nagy, Johanna Laakso, & Elena Skribnik. Oxford: Oxford University Press.

- Aikio, Ante & Jussi Ylikoski. 2010. The structure of North Saami. Kurs-støtteark, Department of Linguistics, The University of Utah.
- Bentzen, Kristine og 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), 169–88. <https://doi.org/10.1007/s10828-019-09105-w>.
- Holmberg, Anders. 1999. Remarks on Holmberg's Generalization. *Studia Linguistica* 53(1), 1–39.
- Holmberg, Anders. 2000. Deriving OV order in Finnish. I *The Derivation of VO and OV*, redigert av Peter Svenonius, 123–52. Amsterdam: John Benjamins.
- Holmberg, Anders. 2017. The Final-over-Final Condition in a Mixed-Word-Order Language. I *The final-overfinal condition*, redigert av Michelle Sheehan, Theresa Biberauer, Ian Roberts og Anders Holmberg, 297–322. Cambridge, MA: MIT Press.
- Nickel, Klaus P. & Pekka Sammallahti. 2011. *Nordsamisk grammatikk*. Karasjok: Davvi Girji OS.
- Sammallahti, Pekka. 1991. No passing, no xing: Traffic regulations for word order. I *Leikkauspiste: Kirjoituksia kielestä ja ihmisestä*, redigert av Lea Laitinen, Pirkko Nuolijärvi og Mirja Saari, 225–33. Suomalaisen Kirjallisuuden Seura [Finska litteratursällskapet], Helsingfors.
- Valijärvi, Riitta-Liisa & Lily Kahn. 2017. *North Sámi: An Essential Grammar*. London: Routledge.
- Vilkuna, Maria. 1995. Discourse configurationality in Finnish. I *Discourse-configurational languages*, redigert av Katalin É. Kiss, 244–68. Oxford: Oxford University Press.
- Vilkuna, Maria. 1998. Word order in European Uralic. I *Constituent Order in the Languages of Europe*, redigert av Anna Siewierska, 173–233. Berlin: Mouton de Gruyter.

English summary

In this paper I investigate the alternation between VO and OV word order in spoken North Sami, based on data from LIA Sápmi – Sáme-giela hállangiellakorpus. My results show that VO in general is the most frequent word order. However, I also find many instances of SAuxOV order in my material, particularly in sentences with periphrastic verb forms where the main verb is in the infinitive (that is, modal constructions) and where the object is a pronoun. In addition, I also find some cases where the object precedes both the auxiliary and the main verb, but crucially without being topicalized to clause-initial position. Based on the account of Norwegian Object Shift in Bentzen and Anderssen (2019), I suggest that OV word order in North Sami may be analyzed as IP-internal topicalization, where objects that are familiar in the context may move to a thematic position between VP and TP. This will account for the OV pattern with periphrastic tense forms. Furthermore, I suggest that there is an additional higher IP-internal topic position, and that object movement to this position is what results in OV with finite main verbs. This higher IP-internal topic position is also the position involved in patterns where the object precedes both the finite auxiliary and the main verb.

Kristine Bentzen
Institutt for språk og kultur
UiT Norges arktiske universitet
kristine.bentzen@uit.no



Å være eller å bli – det er spørsmålet

Ei sammenligning av verbene VÆRE og BLI i amerikansk og norgensisk

Ragnhild Eik og Anu Laanemets

Artikkelen undersøker bruken av verbene VÆRE og BLI sammen med predikatene *født*, *konfirmert* og *gift* i norsk talt i Amerika sammenlignet med norsk talt i Norge. Datamaterialet er hentet fra korpusene CANS og LIA norsk. Studien tar utgangspunkt i tidligere studier av amerikansk (Kühl 2018, Kühl og Heegård Petersen 2017) som viser at VÆRE i noen grad har tatt over for BLI med disse predikatene, trolig på grunn av kontakten med engelsk. Noe overraskende finner vi at vi ikke kan trekke samme konklusjoner for amerikansk. Grunnen er at fordelingen til VÆRE og BLI i norsk talt i Norge viser seg å være annerledes enn først antatt og også annerledes enn i dansk. Dermed kan vi heller ikke konkludere med at man i amerikansk bruker VÆRE der man i norgensisk ville brukt BLI.

Stikkord: amerikansk; norsk; være; bli; aspekt; overgangspartipp

1 Innledning¹

I skandinavisk talt i Skandinavia uttrykker verbene VÆRE og BLI gjerne et skille mellom en tilstand og en overgang i setninger som i (1a-b).

- (1) a. jeg er gift *tilstand*
 b. jeg ble gift *overgang*

Tidligere studier av arvespråkene amerikadansk (Kühl 2018, Kühl og Heegård Petersen 2017) og amerikavensk (Larsson, Tingsell og Andréasson 2015, Hasselmo 2005) viser derimot at VÆRE i noen grad også blir brukt for å uttrykke overganger, på bekostning av BLI, som i (2a), eller en morfologisk passiv, som i (2b).^{2,3}

- (2) a. Dansk i Amerika
 vi *var gift* i 1906 og vi var gift for halvtreds år.
 (Kühl og Heegård Petersen 2017, 231)
- b. Svensk i Amerika
 De *var byggd* här så vi kunde ha en präst ifrån Sverje.
 (Larsson m.fl. 2015, 378)

Kontekstene i (2a) og (2b) indikerer at det trolig er snakk om overganger, men likevel er verbet VÆRE brukt. Denne utvidede bruken av VÆRE har blitt satt i sammenheng med kontakten med engelsk, etter som engelsk ikke skiller mellom tilstands- og overgangslesninger på samme måte som de skandinaviske språkene.⁴ (2b) kan dermed sam-

¹ Takk til deltakere på LIA- og AcqVA-seminarene for innspill. Vi takker også to anonyme fagfeller for konstruktive forslag.

² «De» i eksempel (2b) er gjengitt som hos Larsson m.fl. (2015), som oversetter det til engelsk med 'it'.

³ Hasselmo (2005, 2133) skildrer utvidet bruk av VÆRE som en generell tendens i amerikaskandinavisk, men ser ut til å basere seg først og fremst på svenske data.

⁴ Engelsk danner generelt passiv med bøyingsformer av *be* både for tilstander og overganger. For overgangslesninger er det imidlertid også mulig å danne passiv

menlignes med den engelske oversettelsen i (3), der ei bøyingsform av *be* ‘være’ er det naturlige valget.

- (3) *It was built* here so that we could have a pastor from Sweden
(oversettelse fra Larsson m.fl. 2015)

Lignende eksempler der VÆRE blir brukt med overgangslesning, finnes også i amerikanorsk som eksemplifisert i (4).

- (4) Norsk i Amerika
Å nei nei jeg *var konfirmert* på engelsk jeg
(CANS)

Derfor undersøker vi i denne artikkelen i hvilken grad en lignende tendens som den man finner i dansk og svensk talt i Amerika, også finnes i norsk talt i Amerika. Vi undersøker altså følgende hypotese:

- (5) *Hypotese*: Talere av amerikanorsk bruker VÆRE i kontekster der man i norgensorsk ville forvente BLI.

Om hypotesen viser seg å stemme, kan man videre undersøke om endringa i amerikanorsk kan skyldes påvirkning fra engelsk.

For å undersøke hypotesen i (5) har vi sammenlignet datamateriale fra *CANS – amerikanordisk talespråkskorpus*, v.2.0 (Johannessen 2015) og *LIA norsk – korpus av eldre dialektopptak*. Vi har brukt det eldre norske dialektmaterialet (dvs. LIA) som sammenligningsgrunnlag (‘*baseline*’), ettersom dette ligger nærmere det talemålet som ble snakket av de norske utvandrerne til Amerika. På den måten er det mulig å belyse om eventuelle særegenheter i det amerikanorske materialet skyldes endringer som har skjedd i den amerikanorske varianten, eller om det er særtrekk som er nedarvet fra eldre norske dia-

med *get* i visse kontekster, e.g. *we got married in June* ‘vi ble gift/giftet oss i juni’, og *he got hit by a car* ‘han ble truffet av en bil’ (se f.eks. Aarts 2014, våre eksempler). Den engelske *get*-konstruksjonen har et mer avgrenset bruksområde enn den norske *bli*-konstruksjonen, jf. *barnet ble født i juni - the child was/*got born in June*.

lekter, se e.g. Lykkes (2020, 100–102) diskusjon om hva som bør være *baseline* i undersøkelser av endringer i arvespråk generelt og i amerikansk spesielt.

Undersøkelsen vår tar utgangspunkt i designet i Kühls (2018) studie av dansk i Danmark og Amerika. På samme måte som Kühl (2018) avgrensner vi undersøkelsen til verbene *VÆRE* og *BLI* sammen med de biografiske overgangspredikatene *født*, *konfirmert* og *gift*. Ved å velge ut de samme predikatene som Kühl (2018) kan vi også sammenligne resultatene våre fra norsk med de danske resultatene. Vi har likevel gjort visse endringer sammenlignet med det danske designet, som vi gjør rede for i del 4.1.

Som det vil komme fram av denne artikkelen, ga studien vår et noe overraskende resultat: Vi kan ikke konkludere med at det er utvidet bruk av *VÆRE* på bekostning av *BLI* i amerikansk sammenlignet med norsk i Norge. Grunnen er at distribusjonen til *VÆRE* og *BLI* i amerikansk (CANS) faktisk viser seg å være nokså lik distribusjonen til *VÆRE* og *BLI* i eldre norsk i Norge (LIA), som igjen er annerledes enn det vi forventet basert på eksisterende skildringer av moderne norsk, og også annerledes enn i dansk.

Artikkelen er strukturert på følgende måte: Vi starter med å presentere Kühls (2018) studie av dansk i del 2. Deretter, i del 3, ser vi på hva grammatikkene skriver om bruken av *VÆRE* og *BLI* i norsk i Norge. Resultatene av vår studie presenteres i del 4 hvor vi legger fram både kvantitative (4.3) og kvalitative (4.4) analyser. I del 5 følger en sammenligning av norsk og dansk fulgt av en generell diskusjon og konklusjon i del 6.

2 Kühls (2018) studie av dansk i Danmark og Amerika

Kühl (2018) undersøker bruken av *var/blev* + *født/konfirmeret/gift* i dansk talt i Danmark sammenlignet med dansk talt i Nord-Amerika. Kühls avgrensning er basert på resultater fra Kühl og Heegård Petersen (2017), som fant at det først og fremst er med disse predikatene *VÆRE* blir brukt på uventet vis i dansk talt i Nord-Amerika. Kühl og Heegård Petersen (2017) fant også at uventet bruk av *VÆRE* kun forekom

i preteritum. Kühl (2018) avgrensner derfor undersøkelsen sin til disse formene, men utvider studien sin til et større datamateriale. Hun sammenligner data fra LANCHART (moderne dansk talespråk i Danmark) og Corpus of North American Danish (dansk talt i Nord-Amerika, heretter CoNAmDa).

Kühl finner at korpuset for amerikadansk har en høyere andel *var* + *født/konfirmeret/gift* enn korpuset for moderne danmarksdansk. Motsatt har korpuset for moderne dansk en høyere andel *blev* + *født/konfirmeret/gift* enn korpuset for amerikadansk. Disse forskyvingene indikerer at *var* kan ha tatt over noen bruksområder for *blev* i amerikadansk sammenlignet med danmarksdansk. Kühl finner også variasjon mellom de ulike predikatene: Mens det er store forskjeller i de to korpuserne for *født* og *konfirmeret*, er forskjellene små for *gift*.

Kühl undersøker også konteksten for de ulike ytringene for å se om *var* blir brukt med overgangslesning i amerikadansk, noe som ville være uventet basert på danmarksdansk. Skillet mellom tilstand og overgang kommer blant annet fram gjennom ulike typer adverbialer, som viser om det dreier seg om et konkret tidspunkt (f.eks. et bryllup) eller en lengre periode (f.eks. et ekteskap). Resultatene fra denne kvalitative delen av undersøkelsen peker i samme retning som de kvantitative resultatene, men nyanserer bildet noe: Med *var/blev født* er mange av eksemplene i CoNAmDa tvetydige mellom lesninger som indikerer tilstand og overgang. Kühl argumenterer her for at det har utviklet seg ulike preferanser, slik at amerikadansk ser ut til å foretrekke *var født*, mens danmarksdansk ser ut til å foretrekke *blev født*, selv om begge valg kan være mulige. Med *var/blev konfirmeret* og *var/blev gift* finner Kühl derimot flere eksempler der amerikadansk bruker *var* selv når *blev* er eneste naturlige valg i danmarksdansk. Kühl konkluderer derfor med at amerikadansk og danmarksdansk i visse tilfeller har utviklet ulike preferanser (dvs. med *født*), mens det i andre tilfeller har skjedd semantiske og grammatiske endringer (dvs. med *gift* og *konfirmeret*).

3 Vekslinga mellom VÆRE og BLI i norsk i Norge

Før vi gjør rede for datamaterialet og designet for vår undersøkelse, ser vi nærmere på det som er skrevet om fordelinga av VÆRE og BLI i norsk i Norge. Vekslinga mellom VÆRE og BLI blir hovedsakelig tatt opp i sammenheng med predikativer og perifrastisk passiv. Vi har konsultert flere grammatikker (Aasen [1864]1965, Beito 1986, Golden m.fl. 1998, Lie 2003, Kulbrandstad og Kinn 2016), men beskrivelsene er sparsomme og overlapper i det store og hele med det som står i Norsk referansegrammatikk (Faarlund m.fl. 1997). Derfor konsentrerer vi oss i det følgende om det som står der.

Om predikativer skriver Faarlund m.fl. (1997, 734) at BLI uttrykker begynnelse eller overgang til den tilstanden som predikativet beskriver, jf. *Bøkene blir dyre*. VÆRE har ikke en slik inkoativ eller påbegynnende komponent, og predikatet uttrykker dermed bare den tilstanden som predikativet beskriver, eksemplifisert ved *Bøkene er dyre*.

I sammenheng med passiv skriver de at passiv med VÆRE skiller seg fra passiv med BLI ved at «være-passiven har presens- og preteritumsformer som bare kan uttrykke en varig tilstand eller et resultat, ikke former som kan uttrykke en avgrenset enkelthandling i nåtid eller fortid.» Dette er illustrert ved *Posten blir/ble utdelt av sekretæren*, som er forskjellig fra *Posten er/var utdelt av sekretæren*. Bare den første kan vise til ei enkelthandling avgrenset i tid (Faarlund m.fl. 1997, 525).

Faarlund m.fl. (1997) påpeker dessuten at valget mellom VÆRE og BLI har ulike semantiske konsekvenser i kombinasjon med ulike typer partisipper. For denne artikkelen er det mest relevant å trekke fram partisipper av overgangsverb, hvor de skriver at VÆRE-passiven kan brukes for å «fokusere på den tilstanden som er et resultat av verbalhandlinga, eller på selve den handlinga som fører fram til tilstanden» (Faarlund m.fl. 1997, 525). Dette blir illustrert med *Veien er åpnet* og *Kjolen er rensset*.

Det er ikke enkelt å hente ut fra disse beskrivelsene nøyaktig hva som skiller VÆRE fra BLI, eller effekten av ulike tempusformer av disse. Det som virker relativt klart, er likevel at VÆRE framhever re-

sultatet eller tilstanden som følger av handlinga, mens BLI framhever selve handlinga.

Formene vi undersøker i denne artikkelen, VÆRE/BLI + *født/konfirmert/gift*, befinner seg i grenselandet mellom passiv og predikativ.⁵ De har til felles at de uttrykker en slags overgang (å komme til live, å bekrefte sin dåpspakt og å inngå et giftemål). Derfor bør vi i utgangspunktet kunne forvente å finne samme bruksmønster med *født*, *konfirmert* og *gift* som med andre overgangspartisipper. På bakgrunn av dette er det også naturlig å anta at *født*, *konfirmert* og *gift*, på samme måte som andre overgangspartisipper, får det vi kan kalle ei resultativ lesning eller tilstandslesning sammen med VÆRE og det vi kan kalle ei punktuell lesning eller overgangslesning sammen med BLI.

Skillet mellom tilstandslesning og overgangslesning kommer også fram gjennom adverbialene som er brukt. Når adverbialet viser til et spesifikt tidspunkt eller sted for handlinga, er det gjerne snakk om en overgang og ikke en tilstand. Slike adverbialer vil derfor typisk brukes sammen med BLI (jf. *Vi ble gift i Sverige* ≠ *Vi var gift i Sverige*). Durative tidsadverbialer understreker derimot handlingas eller tilstandens varighet, og kombineres derfor typisk med VÆRE (*Vi var gift i fire år*/**Vi ble gift i fire år*).

⁵ Ifølge Faarlund m.fl. (1997, 736) kan man skille passivpartisipp fra partisipp som predikativt adjektiv ved at bare førstnevnte kan brukes med formelt subjekt, jf. *Huset er bygt – Det er bygt eit hus (passiv)*. *Huset er nybygd – *Det er nybygd eit hus (predikativ)*. Noen enkle søk på nett viser at *født*, *konfirmert* og *gift* alle kan (i alle fall i noen grad) brukes med formelt subjekt, jf. *Det ble født 59 234 barn i Norge i 2000* (ssb). *Det ble konfirmert 41 632 unge i Den norske kirke i 2004* (aftenposten). *Det ble gift 24 185 par i 1965* (fvn). For *gift* finner vi riktignok bare ett slikt eksempel, og Faarlund m.fl. (1997, 773) ser ut til å regne *gift* som predikativ når de gir *En måned senere var vi gift* som eksempel på «adjektiv avleide av perfektum partisipp».

4 Vår undersøkelse

4.1 Framgangsmåte

Undersøkelsen vår av VÆRE og BLI sammen med *født/konfirmert/gift* i amerikanorsk bygger på korpuset *CANS - amerikanordisk tale-språkskorpus*, v.2.0 (Johannessen 2015). Vi har brukt den norske delen, som består av intervjuer og samtaler med 163 amerikanorske informanter fra 44 steder i USA og Canada. Opptakene er hovedsakelig gjort under relativt strukturerte og ensartede feltarbeid i Midt-vesten i USA og i Canada i perioden 2010 til 2016. De eldste informantene var rundt 100 år og den yngste 37 år på opptakstidspunktet. Korpuset består av i alt 660 707 token.

For å sammenligne de amerikanorske bruksmønstrene med norsk talt i Norge har vi brukt *LIA norsk - korpus av eldre dialektopptak*. Korpuset består av opptak med 1374 informanter fra 226 steder i Norge, i alt 3 481 547 token. Opptakene i LIA-korpuset er gjort til ulike tider og med ulikt formål, og informantenes fødselsår varierer fra midten av 1800-tallet til midten av 1900-tallet.

Som nevnt er det en fordel med LIA at materialet her er nærmere det talemålet som de norske innvandrerne snakket da de dro til Amerika. Dette gir dermed et sammenligningsgrunnlag for å se om det har skjedd endringer i amerikanorsk etter utvandringa til Amerika.

Vi har søkt i hele LIA-korpuset og ikke avgrenset søkene til spesifikke områder. På denne måten representerer LIA-informantene ei større geografisk spredning enn den geografiske herkomsten til informantene i CANS v.2.0, som i stor grad snakker østlandske dalmål (Johannessen og Salmons 2012, 142). Når vi ikke har gjort geografiske avgrensninger, er det fordi preliminaire søk viste at det ville gi altfor få treff. Vi har heller ikke gjort avgrensninger med hensyn til opptakstidspunkt og alder på informantene i LIA-korpuset, noe som innebærer at materialet vårt også inneholder resultater fra noen informanter som er født lenge etter utvandringa til Amerika. Grunnen til at vi ikke har gjort tidsmessige avgrensninger, er at dette er vanskelig og svært tidkrevende å få til med dagens søkegrensesnitt, hvor man foreløpig bare kan velge ut enkeltår og ikke større tidsperioder (se også diskusjon om avgrensningsproblematikk i LIA-korpuset i

Vangsnes og Westergaard, dette volumet). Som det vil komme fram av analysen, gir søkene våre likevel nyttige og interessante resultater.

4.2 Søk

I korpusene søkte vi på formene *føde/konfirmere/gift/gifte*.⁶ Deretter sorterte vi bort irrelevante eksempler og annoterte forekomstene manuelt etter følgende kriterier: type verb (VÆRE eller BLI), tempus, type adverbial (tid, sted, måte, annet). Til forskjell fra Kühl (2018) tok vi med samtlige tempusformer og begrenset oss ikke til preteritum (*var; ble*). Dette gjorde vi for å oppnå et mer helhetlig bilde for norsk, og fordi vi ikke hadde noen grunn til å anta at utfallet for dansk – at uventet bruk av VÆRE kun forekommer i preteritum – nødvendigvis ville gjelde for norsk.

Vi tok stikkprøver av transkripsjonene ved å lytte til en god del eksempler, og sjekket spesielt forekomster av *var* for å forsikre oss om at det faktisk var dette som ble sagt og ikke *vart*.

4.3 Kvantitativ sammenligning

Den overordnede distribusjonen av VÆRE og BLI sammen med de undersøkte predikatene *født, konfirmert* og *gift* viser ikke store forskjeller mellom amerikanorsk og norsk talt i Norge i vårt korpusmateriale (jf. tabell 1).⁷ Andelen av VÆRE er generelt sett høyere enn andelen av BLI i begge korpusene, men her utviser de enkelte predikatene en del variasjon. Slik forekommer *født* i 96 % av tilfellene sammen med VÆRE i CANS-korpuset og tilsvarende 89 % i LIA-korpuset. De tilsvarende tallene for *konfirmert* er 70 % og 58 % i hhv. CANS og LIA, og 74 % og 67 % for *gift*. Vi kan også notere at de relative fre-

⁶ Søkene i CANS ble foretatt i juni 2019 og LIA norsk i november 2019. De eksakte søkestrengene var som følger: [lemma="føde"%c], [lemma="konfirmere"%c], [word="gift"%c] or [word="gifte"%c]. Vi søkte på lemmaene 'føde' og 'konfirmere'. Deretter sorterte vi bort de irrelevante eksemplene. Det viste seg at søk på lemmaet 'gifte' ikke ga oss alle eksempler vi var interessert i, da 'gift' var annotert både som verb-perfektum partisipp og som adjektiv. Derfor søkte vi på ordene 'gift' og 'gifte' og sorterte deretter bort de irrelevante eksemplene.

⁷ Vi har slått sammen BLI og VERTE.

kvensene for VÆRE er litt høyere i amerikanorsk sammenlignet med norsk talt i Norge.

	<i>CANS</i>	<i>LIA</i>
VÆRE <i>født</i>	96% (N=609)	89% (N=512)
BLI <i>født</i>	4% (N=27)	11% (N=63)
VÆRE <i>konfirmert</i>	70% (N=48)	58% (N=74)
BLI <i>konfirmert</i>	30% (N=21)	42% (N=54)
VÆRE <i>gift</i>	74% (N=236)	67% (N=436)
BLI <i>gift</i>	26% (N=81)	33% (N=214)

Tabell 1. Sammenligning av frekvensene av VÆRE/BLI + partisipp i CANS og LIA

På bakgrunn av disse tallene kan vi ikke på et overordnet plan konkludere med at amerikanorsk har en markant utvidet bruk av VÆRE sammenlignet med eldre norsk talt i Norge. Forskjellene i det norske materialet er også mye mindre enn forskjellene i det danske materialet, som vil bli sammenlignet i del 5. En nærmere analyse som også tar hensyn til den semantiske konteksten for ytringene, følger i 4.4.

4.4 Kvalitativ sammenligning

For å få et klarere bilde av hva som ligger bak tallene i 4.3, ser vi nå nærmere på enkeltytringene og den semantiske konteksten for disse. Konteksten, uttrykt blant annet med adverbialer for tid og sted, kan vise om taleren snakker om en tilstand eller en overgang. I det følgende tar vi for oss de tre predikatene *født*, *konfirmert* og *gift* hver for seg og sammenligner måten de blir brukt på i de to korpusene, CANS og LIA. Siden vi er interessert i å undersøke tilfeller der VÆRE kan ha tatt over for BLI, fokuserer vi på ytringer med overgangslesning.

4.4.1 født

I CANS, korpuset for amerikanorsk, finner vi flere eksempler på at *VÆRE født* blir brukt sammen med uttrykk som spesifiserer tid og sted for hendelsen, slik at vi får det vi kan kalle ei punktuell lesning eller overgangsledning. Noen eksempler fra CANS er gitt i (6a-c).

- (6) *VÆRE født* i CANS
- han **var født** i enogåtti # i nærheten av # Kodal (stillwater_MN_01gm)
 - jeg **var** ikke **født** på gården men jeg # jeg væ- vokste opp der (willmar_MN_01gm)
 - når søster mi **var født** så var # første gangen # mor mi var på hospitalet (westby_WI_01gm)

Disse talerne av amerikanorsk bruker altså *VÆRE* i overgangskontekster. (7a-c) viser imidlertid at vi finner tilsvarende eksempler i LIA, korpuset for eldre norsk.

- (7) *VÆRE født* i LIA
- ja # nittenhundre eg **var fødd** i nittenhundreogein eg (maalselv_uio_0401)
 - eg er denne einaste ## som ikkje er døypt Sandvikskyrkja for # eg **var fødd** på Skansen (bergen_uib_0802)
 - men eg var mykje sjuk når eg **var fødd** så dei hadde presten til meg og døypte meg (lavangen_uit_0502)

For (6a-c) og (7a-c) kunne man kanskje forvente at BLI ville være det naturlige valget, ettersom det her dreier seg om hendelser forankret i tid og sted. (8a-c) og (9a-c) illustrerer også at BLI kan være et naturlig valg i parallelle kontekster som dem vi så over.

- (8) *BLI født* i CANS
- ja oldeforeldre **blei født** i # i Norge ja # ja (fargo_ND_04gm)
 - jeg **ble født** i nittenhundreenogtjue (coon_valley_WI_03gm)
 - ikke morfar # han døde før jeg **ble født** (spokane_WA_05gm)

- (9) BLI *født* i LIA
- eg **vart fødd** i Trondheim i Singsakerbakken (trondheim_uib_1702)
 - e han **blei fødd** i # attenfireogåtti (vadsoe_uit_1602)
 - ho **blei fødd** om natta og så såg # såg jo ikkje eg ho før morgonen (vennesla_uib_0201)

Vi tolker dette som at VÆRE og BLI har overlappende bruk både i amerikansk og norgensorsk. Basert på de kvantitative og kvalitative undersøkelserne her kan vi dermed ikke påvise at VÆRE tar over bruksområder for BLI sammen med *født* i amerikansk sammenlignet med norgensorsk.

At man i en talesituasjon kan veksle mellom VÆRE og BLI, kommer også fram i følgende utdrag fra et av intervjuene i CANS, der vi ser hvordan forskeren fra Norge bruker både *er*, *vart* og *var* når hun intervjuer en eldre mann fra Blair om bakgrunnen hans.⁸

- (10)
- INT: # ok # når **er** du **født** e når **vart** du **født**?
- blair_WI_01gm: når er født?
- INT: ja
- blair_WI_01gm: e v- jeg be- i # i French Creek nedmed Ettrick
- INT: du **ble født** i French Creek
- blair_WI_01gm: ja # ut- utpå en farm

Muligheten for å veksle mellom VÆRE og BLI i moderne norsk stemmer med våre egne intuisjoner og med intuisjonen til andre norsktalende lingvister vi har konsultert. Mange oppfatter faktisk VÆRE som det umarkerte valget i norsk, slik at man helst sier *Jeg er født i...* når man beskriver sitt eget oppkomme, eller eventuelt *Hun var født i...* om det er tale om en avdød person. BLI er også mulig i de

⁸ Vekslinga her, og kanskje spesielt bruken av «vart», kan være akkommodasjon fra intervjuerens side. Denne problemstillinga om tilpasning til informantens språk i CANS er blant annet diskutert i Grimstad (2018, 99–100) og av intervjuerne selv i Johannessen og Salmons (2012, 142).

samme kontekstene, men vektlegger trolig selve hendelsesforløpet i større grad. Som utredninga i del 5 vil vise, er *var født* nokså vanlig i norsk og mye vanligere enn i moderne dansk (se tabell 2 i del 5).

Her er det samtidig verdt å peke på at flere av talerne vi har konsultert, oppfatter eksemplene med *var* i preteritum som unaturlige i tilfeller der man snakker om seg selv, som i (6b) og (7a-c).⁹ Vi kommer ikke til bunns i den eksakte distribusjonen til VÆRE og BLI og ulike tempusformer av disse i denne artikkelen, men kan hende har vi her å gjøre med pågående språkendringer og/eller geografisk variasjon i norsk i Norge, noe vi også vil diskutere kort i del 6.

4.4.2 konfirmert

For *konfirmert* finner vi et lignende mønster som for *født*, nemlig at VÆRE *konfirmert* blir brukt sammen med uttrykk for konkret sted og tidspunkt, som indikerer at vi har ei overgangslesning. Dette er illustrert med amerikanorske eksempler i (11a-c) og med norgenske eksempler i (12a-c).

(11) VÆRE *konfirmert* i CANS

- a. ja jeg **var** # **konfirmert** i kirken i Coon Valley her (coon_valley_WI_17gm)
- b. og det var der som jeg møtte mannen min # for vi gikk til s- # t- e ## barneskole ## om søndag # og vi **var konfirmerte** der (chicago_IL_01gk)
- c. kanskje han kanskje det var i åtteogførti han **var** # **konfirmert** (westby_WI_11gm)

⁹ Noen ytterligere eksempler på denne bruken av *var født* kan kommenteres, her fra LIA-korpuset:

bolsoey_uio_0101 ja eg var fødd i attentreognitti
 bolsoey_uio_0102 og eg var fødd i attenseksognitti

Denne samtalen forstås helst som at 1893 er fødselsåret til bolsoey_uio_0101 og 1896 er fødselsåret til bolsoey_uio_0102. Ei alternativ tolkning kunne være at da det ble hhv. 1893 og 1896, var informantene allerede i live, og fødselsåret deres kunne vært f.eks. 1891. Metadata fra korpuset bekrefter imidlertid at den første tolkninga er den riktige, der informantene bruker *var født* selv om de snakker om seg selv.

- (12) *VÆRE konfirmert* i LIA
- så eg for min del **er** både **døpt og konfirmert** i Enebakk kyrkje (fet_uio_0403)
 - ja [latter]nei vi gjekk lenge og las ## i Enge skolehus og så **var** vi **konfirmerte** i Stølekyrkja (etne_ma_01)
 - nei eg var # eg **var konfirmert** i toogtretti og (boemlo_uib_0602)

Igjen finner vi også eksempler der BLI brukes i lignende kontekster i begge korpusene, illustrert i (13a-b) og (14a-b).

- (13) *BLI konfirmert* i CANS
- i det var i det er i S2 hun **ble** e # **konfirmert** og døpt i S2 (harmony_MN_02gk)
 - så da # **ble** han **konfirmert** i ## niogførti trur jeg det var # (westby_WI_11gm)

- (14) *BLI konfirmert* i LIA
- så **vart** eg # **konfirmert** i Lade kyrkje (trondheim_uib_0602)
 - far min d- døydde den sommaren eg **blei konfirmert** og (alta_ntnu_0101)

CANS har i tillegg mange eksempler med *konfirmert* + språk, siden dette er noe talerne blir spurt om. Også i slike kontekster finner vi både *VÆRE* og *BLI*, som illustrert i (15a-b).

- (15) *konfirmert* + språk i CANS
- nei nei nei min e # far **var konfirmert** på norsk (rothsay_MN_01gm)
 - vel de **ble** nå e **konfirmert** på norsk e alle i hop (appleton_MN_01gm)

Disse eksemplene indikerer at *VÆRE konfirmert* brukes med overgangsløsning i begge korpusene. Sett under ett, viser de kvalitative og kvantitative undersøkelsene at vi heller ikke her kan konkludere

med utvidet bruk av *VÆRE* i amerikanorsk sammenlignet med norgesnorsk.

4.4.3 gift

Også med *gift* finner vi eksempler i både CANS og LIA der talerne bruker *VÆRE* med overgangslesning, som når de uttrykker tid og sted for et giftemål. (16) viser eksempler fra CANS, og (17) viser tilsvarende eksempler fra LIA.

(16) *VÆRE gift* i CANS

- a. de **var gift** i Lillehammer kirke i mai i nittenseksti (spokane_WA_05gm)
- b. så det var der jeg møtte mannen min som man sier og vi **var gift** i den kirken også ## vi **var gift** i ## toogførti (chicago_IL_01gk)
- c. e vi **var gift** i # December # nittito og så gikk vi # i # til Norge og (saskatoon_SK_07gk)

(17) *VÆRE gift* i LIA¹⁰

- a. vi **er gifte** på Loppa men vi **var gifte** på kontoret hos presten (loppa_uit_0401)
- b. og eg **var gift** når eg var i # fem og tjuande året og da var eg to år til fjells etter eg **var gift** og da tok eg k- tok eg tenestejenta mi med meg (lindaas_uib_0202)
- c. ja for at vi **var jo gifte** og # i toogfemti og han han var jo lenge etter det # her (torsken_uit_0201)

Eksemplene i (16) og (17) viser at *VÆRE gift* brukes med overgangslesning i begge korpusene. Samtidig finner vi lignende eksempler med *BLI gift* i begge korpusene, som vist i (18) og (19).

¹⁰ I eksempel (17b) er uttalen av den andre forekomsten av «var gift» uklar, og informanten kan ha sagt «vart».

- (18) BLI *gift* i CANS
- a. ja # og han var den første som # **ble gift** i Coon Valley (westby_WI_02gm)
 - b. arbeidde nede der # og de **ble gift** # i nittenhundreogfjorten (gary_MN_01gm)
- (19) BLI *gift* i LIA
- a. var vel antakeleg frå Laksefjorden for der inne **vart** dei **gifte** (gamvik_uit_0103)
 - b. ja sa han # eg **blei gift** i kyrkja og når det kunne føre med ein s- seg slik ein forbannelse så skal eg ikkje komme der meir (lyngdal_uib_0107)

På samme måte som med *født* og *konfirmert* tolker vi dette som at arbeidsdelinga mellom VÆRE og BLI ikke er kategorisk. Her er det relevant å påpeke at vi også finner flere eksempler på individuell variasjon i materialet. Informanten i (16a) sier også (20).

- (20) da # bestemoren min **blei gift** (spokane_WA_05gm)

Vi finner lignende eksempler på individuell variasjon for alle de tre predikatene.

Det er også verdt å trekke fram at VÆRE *gift* + sted kan brukes for å uttrykke at man lever som gift på et sted, altså med tilstandslesning, illustrert i (21).

- (21) eg har ei dotter som **er** i **gift** i Tyskland (...) og så har eg e ei dotter som **er gift** på Kråkerøy i Fre- med Fredrikstad og dei # borna har lært det (finsland_uio_0201)

Eksemplene i (16a-c) og (17a-c) er ikke av denne typen, men ser virkelig ut til å handle om overganger.

Samlet for *født*, *konfirmert* og *gift* kan vi si at eksemplene som er trukket fram i denne delen, indikerer at VÆRE og BLI har nokså overlappende bruk sammen med disse predikatene, både i amerikanorsk

og i norgenorsk. Sett sammen med den kvantitative sammenligninga i 4.3, kan vi dermed ikke konkludere med at de amerikanske talerne i CANS viser uventet bruk av VÆRE sammenlignet med talerne av eldre norske dialekter i LIA. Resultatene våre skiller seg altså fra Kühls (2018) resultater for dansk, og i neste del sammenligner vi de norske og de danske resultatene.

5 Sammenligning av norsk og dansk

Kühl (2018) sammenlignet amerikadanske data med moderne dansk og konkluderte med at det finnes nokså store forskjeller for bruk av formene *var/blev* sammen med *født, konfirmeret* og *gift*. For å sammenligne Kühns resultater med de norske resultatene tar vi i denne delen kun med preteritumsformene *var/ble* sammen med *født, konfirmert* og *gift* fra det norske datamaterialet, ettersom det bare var preteritumsformene som ble undersøkt i den danske studien. Vi supplerer også sammenligninga med tall fra noen preliminære søk i Nordisk dialektkorpus (v. 4.0, heretter NDK, Johannessen m.fl. 2009), som inneholder opptak med moderne norske dialekter. Selv om de absolute forekomstene av søkeordene i NDK riktignok er nokså små, gir dette oss et hint om forholdene i moderne norsk og et bedre sammenligningsgrunnlag med de danske dataene.

Kvantitative data fra de fem korpusene er sammenfattet i tabell 2 på neste side. Data fra CoNAmDa og LANCHART er fra Kühl (2018, 66).

Som det framgår av tabellen, finnes det nokså store (kvantitative) forskjeller mellom amerikadansk og danmarksdansk, mens forskjellene mellom amerikansk og eldre norsk talt i Norge ikke kan sies å være særlig store (jf. også 4.3). Det er interessant å legge merke til at det spesielt er moderne dansk som skiller seg ut med et diametralt annerledes bruksmønster. Her er det gjennomgående verbet *blev* som dominerer over *var* i kombinasjon med de undersøkte predikatene. Tallene for moderne dansk ligner heller ikke mønsteret for moderne norsk. For moderne norsk er det således en nokså ensartet og lik distribusjon mellom *var* og *ble* på tvers av de tre predikatene *født, kon-*

	Amerika- norsk CANS	Eldre norsk LIA	Moderne norsk NDK	Amerika- dansk CoNAmDa	Moderne dansk LANCHART
<i>var født</i>	96% (N=574)	75% (N=189)	52% (N=73)	93% (N=264)	29,4% (N=93)
<i>ble født</i>	4% (N=25)	25% (N=63)	48% (N=67)	7% (N=20)	70,6% (N=223)
<i>var konfirmert</i>	70% (N=44)	57% (N=65)	56% (N=14)	47% (N=28)	1,4% (N=1)
<i>ble konfirmert</i>	30% (N=19)	43% (N=49)	44% (N=11)	53% (N=34)	98,6% (N=73)
<i>var gift</i>	70% (N=173)	54% (N=232)	50% (N=26)	25,4% (N=83)	21,9% (N=89)
<i>ble gift</i>	30% (N=75)	46% (N=195)	50% (N=26)	74,6% (N=244)	78,1% (N=317)

Tabell 2. Sammenligning av frekvensen av *var/ble* + partisipp i norsk og dansk.

firmert og *gift*. På denne måten skiller resultatene for moderne norsk seg fra de to andre norske datasettene (jf. tabell 2).

I den danske studien gjennomførte Kühl også kvalitative sammenligninger og argumenterte for at amerikadansk og danmarksdansk i noen tilfeller har utviklet ulike preferanser (*var/blev født*), mens det i andre tilfeller har skjedd semantisk-grammatiske endringer (*var/blev + konfirmeret/gift*). Også her skiller de norske resultatene seg fra de danske, ettersom både *VÆRE* og *BLI* blir brukt i lignende kontekster med alle tre predikater i amerikansk og norgensk. For predikatet *født* ser *VÆRE* ut til å være det mest naturlige valget i norsk.

6 Generell diskusjon og konklusjon

Vi har tatt for oss hypotesen om at talere av amerikansk bruker *VÆRE* + *født/konfirmert/gift* i kontekster der man i norgensk ville forvente *BLI*. I den kvantitative delen av analysen så vi at det er noe høyere bruk

av VÆRE i CANS, korpuset for amerikanorsk, enn i LIA, korpuset for eldre norsk i Norge, men forskjellene er små. I den kvalitative delen av analysen så vi at talerne i CANS og LIA ser ut til å bruke VÆRE i lignende kontekster. Dermed kan vi ikke påvise uventet bruk av VÆRE i amerikanorsk sammenlignet med eldre norgensorsk på en slik måte at talerne av amerikanorsk bruker VÆRE på bekostning av BLI. Gitt dette resultatet, har vi heller ingen indikasjon på at bruken av VÆRE i amerikanorsk skyldes påvirkning fra engelsk. Amerikanorske former som ved første øyekast kan virke påfallende for moderne talere av norsk, kan like gjerne være nedarvet fra eldre norsk i Norge.

Et forbehold for konklusjonen vår er at vi har sammenlignet de amerikanorske talerne med eldre talere fra hele Norge, mens vi vet at mange av talerne i CANS v.2.0 har talemålet sitt fra dalene på Østlandet (se 4.1). Kanskje finnes det dialektale forskjeller i Norge for bruken av VÆRE og BLI, og kanskje ville slike forskjeller kunne påvirke konklusjonen vår.¹¹ Med avgrensninga vår til predikatene *født*, *konfirmert* og *gift* var datamaterialet for lite til å kunne undersøke dette, men forhåpentligvis vil videre forskning kunne svare på om det finnes dialektale forskjeller i bruken av VÆRE og BLI i norgensorsk mer generelt. Videre ville det være interessant å utforske eventuelle aldersmessige forskjeller og individuell variasjon i valget mellom de to verbene. Slike undersøkelser vil forhåpentligvis også kunne gi oss mer nøyaktige skildringer enn det vi har i dag for arbeidsdelinga mellom VÆRE og BLI og ulike tempusformer av disse i norsk.

Et spesielt interessant resultat i studien vår er at vi ikke gjør de samme funnene som Kühl (2018) og Kühl og Heegård Petersen (2017) gjør i sine studier av dansk i Amerika og Danmark. En viktig forskjell mellom studien av amerikadansk og vår studie av amerikanorsk er imidlertid materialet vi bruker som sammenligningsgrunnlag. Mens det amerikanorske materialet ble sammenlignet med eldre norsk talemål, ble det amerikadanske materialet sammenlignet med

¹¹ Her kan vi igjen trekke inn innspill fra flere norske lingvister som oppfatter at eksemplene vi har sett med VÆRE + *konfirmert* og *gift* klinger spesielt dårlige i urbane østnorske ører. Dette kan passe med at dette talemålet også i mange andre henseender ligger nærmere dansk.

moderne dansk. Det er en stor fordel at vi nå, takket være det nylig ferdigutviklede LIA-korpuset, har mulighet til å sammenligne amerikansk med eldre norske talemål, som ligger tettere opptil det språkstadiet som norsken i Amerika har utviklet seg fra. Det er også interessant at preferansene for VÆRE og BLI ser ut til å være såpass forskjellige i materialet som så langt er undersøkt for moderne norgensorsk og danmarksdansk. Studien vår åpner slik opp mange nye dører for videre forskning.

Referanser

- Aarts, Bas. 2014. Get-passive. I *The Oxford Dictionary of English Grammar*. Oxford: Oxford University Press. DOI: 10.1093/acref/9780199658237.001.0001.
- Beito, Olav T. 1986. *Nynorsk grammatikk. Lyd- og ordlære*. Oslo: Det Norske Samlaget.
- CANS – *The Corpus of American Nordic Speech*. <http://tekstlab.uio.no/norskiamerika/english/corpus.html>.
- Faarlund, Jan Terje, Svein Lie og Kjell Ivar Vannebo. 1997. *Norsk referansegrammatikk*. Oslo: Universitetsforlaget.
- Golden, Anne, Kirsti Mac Donald og Else Ryen. 1998. *Norsk som fremmedspråk: Grammatikk*. (4. utg.). Oslo: Universitetsforlaget.
- Grimstad, Maren Berg. 2018. English-origin Verbs in American Norwegian : A Morphosyntactic Analysis of Mixed Verbs. Doktoravhandling, Norges teknisk-naturvitenskaplige universitet.
- Hasselmo, Nils. 2005. History of the Scandinavian emigrant languages. I *The Nordic languages. An international handbook of the history of the North Germanic languages 2* (HSK 22.2), redigert av Oskar Bandle, Kurt Braunmüller, Ernst Håkon Jahr, Allan Karer, Hans-Peter Naumann og Ulf Teleman, 2127–2141. Berlin: de Gruyter.
- Johannessen, Janne Bondi. 2015. The Corpus of American Norwegian Speech (CANS). I *Proceedings of the 20th Nordic Conference of Computational Linguistics*, NODALIDA 2015, 11.-13.

- mai, 2015, Vilnius, Litauen. NEALT Proceedings Series 23, redigert av Béata Megyesi, 297–300.
- Johannessen, Janne Bondi og Joseph C. Salmons. 2012. Innledning. *Norsk lingvistisk tidsskrift* 30 (2), 139–148.
- Johannessen, Janne Bondi, Joel Priestley, Kristin Hagen, Tor Anders Åfarli og Øystein Alexander Vangsnes. 2009. The Nordic Dialect Corpus - an Advanced Research Tool. I *Proceedings of the 17th Nordic Conference of Computational Linguistics*, NODALIDA 2009, 14.-16. mai, Odense, Danmark. NEALT Proceedings Series 4, redigert av Kristiina Jokinen og Eckhard Bick, 73–80.
- Kulbrandstad, Lars Anders og Torodd Kinn. 2016. *Språkets mønstre*. (4. utg.). Oslo: Universitetsforlaget.
- Kühl, Karoline. 2018. Constructional Change Based on Cross-linguistic Analogy Building in North American Danish. I *Selected Proceedings of the 8th Workshop on Immigrant Languages in the Americas* (WILA 8), redigert av Jan Heegård Petersen og Karoline Kühl, 63–70. Somerville, MA: Cascadilla Proceedings Project.
- Kühl, Karoline og Jan Heegård Petersen. 2017. Var de gift eller blev de gift? Om biografiske overgangsprædikater i amerikadansk. I *16. Møde om Udforskningen af Dansk Sprog*, redigert av Inger Schoonderbeek Hansen, Tina Thode Hougaard og Kathrine Thisted Petersen, 231–245. Aarhus: Aarhus Universitetsforlag.
- Larsson, Ida, Sofia Tingsell og Maia Andréasson. 2015. Variation and Change in American Swedish. I *Germanic heritage languages in North America. Acquisition, attrition and change*, redigert av Janne Bondi Johannessen og Joseph C. Salmons, 359–388. Amsterdam: Benjamins. <https://doi.org/10.1075/silv.18.16lar>.
- LIA norsk – korpus av eldre dialektopptak*. <https://www.hf.uio.no/iln/tjenester/kunnskap/sprak/korpus/talespraksakorpus/lia/>.
- Lie, Svein. 2003. *Innføring i norsk syntaks*. (5. utg.). Oslo: Universitetsforlaget.
- Lykke, Alexander K. 2020. Variation and Change in the Tense Morphology of Heritage Norwegian in North America. Doktoravhandling, Universitetet i Oslo.

NDK – Nordisk dialektkorpus. <http://tekstlab.uio.no/nota/scandiasyn/>.

Vangnes, Øystein Alexander og Marit Westergaard. 2021. *Ka LIA fortæll? Eit gjensyn med kv-spørsmål i norske dialektar*. I denne boka, 155–178.

Aasen, Ivar. [1864]1965. *Norsk Grammatik*. Oslo: Universitetsforlaget.

English summary

The paper investigates the use of the verbs *VÆRE* 'be' and *BLI* 'become' with the predicates *født* 'born', *konfirmert* 'confirmed' and *gift* 'married' in American heritage Norwegian compared to homeland Norwegian. In the Scandinavian languages, *VÆRE* 'be' is typically used with stative readings, whereas *BLI* 'become' is typically used with eventive readings. However, previous studies of American Danish (Kühl 2018, Kühl og Heegård Petersen 2017) have found that *VÆRE* 'be' has to some extent taken over for *BLI* 'become' with these predicates, a change which is argued to be due to contact with English. We investigated whether a similar development has occurred in American Norwegian, comparing data from the Corpus of American Norwegian Speech (CANS) and the LIA corpus of Older Norwegian dialects. Somewhat surprisingly, we find that we cannot establish a similar development in American Norwegian, the reason being that the distribution of *VÆRE* 'be' and *BLI* 'become' in American Norwegian actually turns out to be quite similar to the distribution of *VÆRE* 'be' and *BLI* 'become' in older Norwegian in Norway (LIA). The latter in turn is different from what we expected based on existing descriptions of Norwegian, and also different than in Danish. Thus, we cannot conclude that speakers of American Norwegian use *VÆRE* 'be' in contexts where speakers of homeland Norwegian would have used *BLI* 'become'.

Ragnhild Eik
Institutt for språk og litteratur
Norges teknisk-naturvitenskapelige universitet NTNU
ragnhild.eik@ntnu.no

Anu Laanemets
Institutt for lingvistiske og nordiske studier
Universitetet i Oslo
anu.laanemets@iln.uio.no



Argumentplacering i norskt arvspråk i Amerika

Ida Larsson och Kari Kinn

Den här artikeln undersöker placeringen av subjekt, objekt och partiklar i äldre och yngre amerikanorska (i CANS-korpusen) och i LIA-korpusen, som representerar talspråket i Norge vid tiden för massutvandringen till Amerika. Ett huvudresultat är att ledföljden varierar i samtliga material, och att det delvis är mer variation än i modernt norskt talspråk. Variationen håller sig relativt stabil över flera generationer i Amerika, men vi noterar i likhet med Anderssen och Westergaard (2020) att andelen subjektskifte och objektskifte minskar. Vi föreslår att frekvensförskjutningen delvis beror på ekonomiprinciper som främjar den oskiftade positionen. För subjektskifte är direkt påverkan från engelska en relevant faktor i yngre AmNo, men bara i frågor; för deklarativer föreslår vi att den avgörande förändringen egentligen inte handlar om en preferens för den oskiftade positionen, men en preferens för flyttning av subjektet till C-domänen.

Nyckelord: amerikanorska, arvspråk, objektskifte, verbpartikel, spec-CP, subjektskifte, tvärspråkligt inflytande

1 Inledning¹

I norskan kan placeringen av subjekt och objekt variera. Subjekt har tre möjliga positioner: satsinitialt (1a), efter finit verb men före satsadverbial (1b), eller efter satsadverbial (1c). Pronominella objekt står normalt före satsadverbial (2a), medan icke-pronominella objekt alltid följer ett satsadverbial (2b); alla typer av objekt kan förstås också stå satsinitialt (jfr Andréasson 2010). Vidare står (icke-initiala) objekt antingen efter eller före verbpartiklar; jämför (3a) och (3b).

- | | | |
|-----|--|---------------------|
| (1) | a. Lisa likte ikke boka. | (Initialt subjekt) |
| | b. Derfor likte Lisa ikke boka. | (Skiftat subjekt) |
| | c. Derfor likte ikke Lisa boka. | (Oskiftat subjekt) |
| (2) | a. Lisa likte den ikke . | (Skiftat objekt) |
| | b. Lisa likte ikke boka . | (Oskiftat objekt) |
| (3) | a. Lisa tok med boka . | (Skiftad partikel) |
| | b. Lisa tok boka med . | (Oskiftad partikel) |

I det följande omtalar vi ordföljden som illustreras i (1a) som SV-följd. Ordningen mellan subjekt och satsadverbial brukar omtalas som *subjektskifte* (SS): subjekt som står före adverbial är skiftade, medan subjektet är oskiftat när det följer adverbial. Ordningen mellan objekt och satsadverbial beror på samma sätt på om objektet har skiftat (*objektskifte*, OS) eller inte. När det gäller ordningen mellan partikel och objekt talar vi i det följande om *partikelskifte*.

Den här variationen är omdiskuterad (se Svenonius 2002, Westergaard 2011 om subjekt, Holmberg 1986, Anderssen och Bentzen 2012 om objekt, Larsson och Lundquist 2014, Aa 2020 om verbpartiklar, och många andra). Det har visat sig att argumentplaceringen påverkas av en rad olika samverkande faktorer: argumentets form

¹ Vi är tacksamma för kommentarer från två anonyma granskare. Arbetet finansierades av Högskolan i Østfold, Universitetet i Bergen, Universitetet i Oslo och Norges forskningsråd genom projekten *Variation och förändring i den nordiska verbfrasen* (projektnummer 250755) och *Norwegian across the Americas* (projektnummer 301114).

(pronomen eller inte), informationsstruktur, prosodi, semantik och lexikala faktorer. Dessutom kan man notera både dialektal och (inom)individuell variation.

I den här artikeln undersöker vi hur variationen överförs mellan generationer och i en annan social kontext än den hemlandsnorska, nämligen i amerikanorskt arvspråk (AmNo). Med arvspråk menar vi ett språk som lärs in som modersmål, men i en kontext där ett annat språk dominerar (se t.ex. Rothman 2009, Polinsky 2018); de amerikanorska talarna har norska som modersmål, men engelska som sitt dominerande språk. Det råder numera i stort sett konsensus om att arvspråk som AmNo är *språk* och att arvspråkstalare har internt koherenta grammatiska system (se t.ex. Polinsky 2018). Man kan alltså undersöka AmNo som ett språk i sin egen rätt. Samtidigt är det viktigt att undersöka eventuella skillnader (och likheter) mellan amerikanorskan och den norska som emigranterna förde med sig till USA. Forskningen om arvspråk syftar ofta just till att jämföra arvspråken med en *baseline* (se avsnitt 2), för att på så vis pröva vilken roll faktorer som input, språkanvändning genom livet och social kontext spelar för förändring av ett grammatiskt system. I det här sammanhanget är argumentplacering intressant, inte minst eftersom hemlandsspråket uppvisar komplex variation där många olika faktorer samspekar. Som vi kommer att se har frågan om vad man ska räkna som baseline en viss betydelse för tolkningen av resultaten.

SS och OS i AmNo har tidigare undersökts av Anderssen och Westergaard (2020; A&W i det följande), som visar att det är fler oskiftade subjekt och objekt i AmNo än i den hemlandsnorska som finns representerad i Oslodelen av NoTa-korpusen och i barnriktat tal. Nu ger den nya LIA-korpusen och nya transkriptioner av äldre amerikanorska inspelningar oss möjlighet att mera direkt jämföra argumentplaceringen i dagens AmNo med en baseline som inte nödvändigtvis är helt lik den norska som talas av modersmålstalare i Norge idag. Förutom SS och OS inkluderar vi också partikelskifte; det har inte undersökts tidigare i AmNo. Vi studerar argumentplaceringen i tre material: äldre dialektinspelningar i LIA-korpusen, samt äldre och yngre AmNo i CANS-korpusen (Johannessen 2015).

Nästa avsnitt ger en bakgrund till amerikanorskt arvspråk. I avsnitt 3 presenterar vi vårt material, och avsnitt 4 ger en översikt över resultaten. I avsnitt 5 diskuterar vi resultaten i ljuset av olika typer av förändringsprocesser som har observerats i arvspråk och andra kontexter. Avsnitt 6 avslutar artikeln.

2 Arvspråk

Idag talas AmNo av en liten grupp ättlingar till de hundratusentals emigranter som lämnade Norge för USA och Kanada under 1800-talet och början av 1900-talet (se Haugen 1953 och, för en kort översikt, Johannessen och Salmons 2015, 10). De allra flesta amerikanorska talare är numera 3–4 generationens invandrare, och de är över 70 år gamla. AmNo är som framgått ett så kallat arvspråk, och de som talar AmNo är arvspråkstalare (eller arvstalare). Många av de amerikanorska arvstalarna hade norska som enda förstaspråk, men engelskan introducerades senast vid skolstart, och är nu det starkaste språket för samtliga talare.

Arvspråken kännetecknas bland annat av den specifika inlärningskontexten. Arvstalarna får mindre input av sitt modersmål, och input från färre talare (och inte nödvändigtvis från generationskamrater), än språkbrukarna i hemlandet. De nutida amerikanorska talarna fick dessutom input från vuxna som själva ofta var arvstalare, och vars språkbruk därför högst sannolikt avvek från språkbruket i hemlandet.

Dessutom är arvstalarna tvåspråkiga, och eftersom norskan bara används i begränsad utsträckning kan den delvis gå förlorad senare i livet (så kallad *attrition*). Arvspråket kan dessutom på olika sätt påverkas av det starkare språket, och närma sig engelskan genom tvärspråkligt inflytande (*cross-linguistic influence*), eller bli mindre likt engelska, genom tvärspråklig överkorrigering (*cross-linguistic over-correction*; se Kupisch 2014).

Det är alltså flera olika typer av faktorer inblandade i utvecklingen av ett arvspråk, och det kan vara svårt att skilja dem från varandra. Dessutom är det alltså inte alltid givet vad som ska räknas som en utveckling specifik för arvspråket och vilka som arvstalarna ska jäm-

föras med. Tvärtom är frågan om *baseline*, alltså vilken varietet eller vilka varieteter som det är mest meningsfullt att jämföra arvspråket med, rätt omdiskuterad (se tex Montrul 2016, Polinsky 2018 och många andra). Som Polinsky (2018, 34) noterar kan man förvänta sig att arvstalare förstärker förändringstendenser som redan finns i språket hos första generationens invandrare, och det är viktigt att veta om det språk som emigranterna förde med sig uppvisade variation. I det följande undersöker vi argumentplaceringen i två generationers amerikanorska och i äldre dialektinspelningar från Norge, för att på så sätt etablera en baseline som inkluderar den norska som talades av de första utvandrarerna, samt norskan hos en tidigare generation arvstalare som utgjorde input för dagens amerikanorska arvstalare.

3 Material

I den här studien utnyttjar vi två korpusar: version 3 av CANS-korpusen (Johannessen 2015) samt LIA-korpusen. CANS innehåller inspelningar av amerikanorska och amerikasvenska talare i USA och Kanada. Den norska delen av CANS (746 000 tokens i version 3) innehåller material från 227 talare som de flesta är 3–5 generationens invandrare. Nytt för den tredje versionen är att det ingår äldre inspelningar som gjordes av Einar Haugen och Didrik Arup Seip och Ernst W. Selmer på 1930- och 40-talet. Det handlar om material (totalt 40 900 tokens) från 47 talare, som många är 2–3 generationens invandrare. Inspelningarna har tidigare används av bland annat Hjelde (2011, 2012, 2015) och Riksem (2018), men är först nu tillgängliga i sökbar form.

Dessutom använder vi alltså dialektmaterialen i LIA-korpusen, som bland annat innehåller inspelningar av talare födda så tidigt som 1860-talet. LIA-korpusen är en betydande resurs för studier av variation och förändring i norskan, och den ger oss inte minst tillgång till talspråket under den period då många norrmän emigrerade till Amerika (dvs. sent 1800-tal och tidigt 1900-tal) på ett helt annat sätt än tidigare. Nu kan vi alltså studera språkbruket hos förfäderna till dagens amerikanorska talare på ett sätt som inte var möjligt förut. Ti-

digare studier av AmNo har i hög grad använt dagens språk i Norge som baseline (se Johannessen och Larsson 2015, Lykke 2020 och många andra). För att spåra förändring över flera generationer i ett arvspråk, bör man emellertid ideellt sett börja med de första emigranterna. Med hjälp av CANS och LIA-korpusen kan vi nu göra precis det.

För den här studien har vi gjort ett urval ur korpusarna; se översikten i tabell 1. I LIA-materialet har vi gjort ett urval med inspelningar av 27 talare från det som tidigare var Oppland fylke, eftersom Oppland hade stor utvandring och eftersom dialekterna där (särskilt gudbrandsdalskan) har haft stor betydelse för utvecklingen av dagens amerikanorska (se t.ex. Hjelde 2012, Johannessen och Laake 2012). Talarna är huvudsakligen födda under perioden för massemigrationen, mellan 1889 och 1926. En talare är emellertid yngre, född 1955. Ur det äldre materialet i CANS har vi framför allt inkluderat Haugens inspelningar, men också någon inspelning från Seip och Selmer. Det rör sig om sammanlagt 47 talare, födda mellan 1850 och 1927. Här har vi inte tillgång till riktigt lika detaljerade metadata, men de allra flesta är arvstalare, födda i USA, och de är alltså inte första generationens invandrare. De har huvudsaklig dialektbakgrund från det inre östlandet (dit Oppland hör). Slutligen har vi inkluderat 11 amerikanska arvstalare som spelades in på 2010-talet. De är födda mellan 1918 och 1946, och de har anknytning till Oppland, typiskt genom mor- eller farföräldrar (eller deras föräldrar).

	<i>Dialekt- område</i>	<i>Antal talare</i>	<i>Födelseår</i>	<i>Omfång</i>
<i>LIA</i>	Oppland	27	1889–1926, en talare från 1955	57 900 tokens
<i>Äldre AmNo</i>	Inre Öst- landet	47	1850–1927	40 900 tokens
<i>Yngre AmNo</i>	Anknytning till Oppland	11	1918–1946	73 100 tokens

Tabell 1: Material

I de här materialen har vi sökt på negationen *ikkje/ikke* samt de adverbiala partiklarna *opp*, *ned*, *ut* och *inn* för att på så vis undersöka placeringen av subjekt och objekt visavi negation, och objekt visavi verbpartikel. Irrelevanta träffar har sorterats bort.² I studien av OS har vi endast inkluderat belägg där OS är möjligt – alltså belägg med ett pronominellt (indirekt eller direkt) objekt och huvud verbet utflyttat ur verbfrasen (jfr Holmbergs generalisering, Holmberg 1986); sats-er där subjektet står efter negation och därmed blockerar möjligheten för objektskifte ingår heller inte.³ När det gäller partiklar kan avgränsningen vara svårare. Vi har följt Larsson och Lundquist (2014), och inkluderat belägg med DP-objekt (pronomen och fulla nominalfraser) som skulle kunna placeras före partikeln i norskan. I studien ingår också belägg där partikeln står med en prepositionsfras (t.ex. *henge den opp på veggen*).

Alla belägg har annoterats manuellt. Vi skiljer mellan pronomen och andra typer av argument; som pronomen räknar vi här enbart personliga pronomen som inte är modifierade. När det gäller subjektspacering har vi vidare delat in beläggen efter satstyp. Till bisatserna räknas enbart af-satser (dvs. satser med satsadverbial före finit verb). Entydiga exempel på inbäddad V2 förs till de deklarativa huvudsatserna, eftersom de har tre möjliga positioner för subjektet. I AmNo har en del inbäddade fa-satser (dvs. satser med finit verb före satsadverbial) sorterats bort, eftersom AmNo ibland har V-till-I-flyttning av verbet (se Larsson och Johannessen 2015).⁴ Som frågor har vi räknat V1-satser (inklusive frågeformade konditionalsatser)

² För SS och OS gäller det t.ex. belägg med konstituentnegation eller satsinitial negation. När det gäller subjektspacering är förstas subjektlösa satser (subjektrelativer, imperativer, andrakonjunkter och fragment) också irrelevanta, liksom satser med *wh*-flyttat subjekt.

³ Ett oskiftat indirekt objekt blockerar också skifte av ett direkt objekt; jfr (i) och (ii). Några sådana belägg har vi emellertid inte noterat.

(i) Han ga **ikke** Lisa **den**.

(ii) *Han ga **den ikke** Lisa.

⁴ I den äldre amerikanorskan har vi noterat 6 sådana belägg, samtliga med följden subjekt-verb-negation; i den yngre amerikanorskan finner vi 13 belägg. Ett exempel ges i (i).

(i) nå begynner det å bli roligere nå når de e server ikkje noe

(Äldre AmNO, blair_WI_25gm)

och självständiga hv-satser (inklusive exklamativer). När det gäller objektskifte har vi enbart undersökt pronominella objekt (personliga pronomen och reflexiver). Vi har uteslutit *det* i satsanteciperande konstruktioner, men inkluderat andra expletiver i objektsposition (t.ex. *hun har det ikke godt*).⁵ För partikelkonstruktionerna har vi skilt pronominella objekt (omodifierade personliga pronomen) från andra typer av objekt; reflexiver är inte inkluderade. Dessutom har vi annoterat partikelkonstruktionerna för betydelse; vi har då framför allt skilt klart direktionella partiklar från andra (med metaforisk eller idiomatisk betydelse). Vi har också särskilt noterat om en partikel konstrueras med en direktionell PP. Några oklara exempel är bortsorterade.

4 Resultat

I det här avsnittet presenterar vi resultaten för undersökningen av subjektspacering (4.1), OS (4.2) och partikelskifte (4.3), och relaterar dem till tidigare studier av AmNo och hemlandsnorska. Eftersom de flesta av beläggen i korpusen har pronominella argument > subjekt, och OS endast rör pronominella objekt, fokuserar vi särskilt på pronomenplacering, men när det gäller partikelskifte diskuterar vi även icke-pronominella objekt.

4.1 Subjektplacering

Norska har som bekant tre möjliga positioner för subjekt (jfr (1) ovan): satsinitialt (i spec-CP/fundamentet), och två positioner som vi antar är i T-domänen (eller mittfältet med Didrichsens terminologi), nämligen före och efter satsadverbial. Vi har inkluderat alla tre positionerna i vår undersökning (men begränsar oss som framgått till satser med negation). Frekvensen av initiala subjekt kan vara relevant för SS, eftersom den påverkar antalet potentiella kontexter för skifte

⁵ En satsanteciperande konstruktion är t.ex. *Hun vet ikke det at vi skal ha party* (coon_valley_WI_06gm); *det* kan här analyseras som en determinativ som tar bisatsen som komplement. När det gäller de expletiver som ingår i undersökningen räknar vi dem till den typ som tar nominal antecedent (se tabell 5) även om de saknar innehåll, baserat på deras syntaktiska distribution.

– SS är ju bara möjligt när subjektet står i T-domänen. Vi fokuserar som nämnt på pronomiella subjekt, eftersom det är få icke-pronomiella subjekt i materialet.⁶ Som vi kommer att se i det följande kan vi notera två utvecklingar i AmNo: andelen initiala subjekt ökar, och bland subjekten i T-domänen minskar andelen skiftade pronomen. För enkelhetens skull omtalar vi subjekten i T-domänen (före eller efter negation) som *TP-subjekt* i det följande.

Tabell 2 ger en översikt över placeringen av pronomiella subjekt i deklarativa huvudsatser i de tre materialen. I LIA-materialet inleds 75 % av alla deklarativer med subjekt.⁷ I äldre AmNo kan vi notera en något högre andel (79 %), men skillnaden är inte signifikant.⁸ Däremot är det en signifikant högre andel SV-följd i yngre AmNo än i äldre AmNo (jfr. Westergaard och Lohndal 2019), 93 %.⁹

	<i>Satsinitialt</i>	<i>S-Neg</i>	<i>Neg-S</i>	<i>N</i>
<i>LIA</i>	76 %	22 %	2 %	530
<i>Äldre AmNo</i>	79 %	15 %	6 %	279
<i>Yngre AmNo</i>	92 %	5 %	3 %	730

Tabell 2. Placering av pronomiella subjekt i deklarativa huvudsatser i de tre materialen.

Kontexterna för SS (och oskiftade subjekt) är alltså mycket färre i yngre AmNo än i de äldre materialen, eftersom fler subjekt är sats-

⁶ I LIA-materialet är 25 av subjekten icke-pronomiella. Av dessa är 15 (60 %) initiala, 1 (4 %) skiftat och 9 (36 %) oskiftade. I det äldre amerikanska materialet finner vi 14 icke-pronomiella subjekt. 7 av dem står initialt, 2 är skiftade och 5 är oskiftade. I det yngre amerikanska materialet finns 15 icke-pronomiella subjekt. 12 av dessa är satsinitiala, och 3 är oskiftade.

⁷ Siffran liknar vad som noteras för pronomiella subjekt i deklarativa huvudsatser av Keihl Olsen (2019:49) i en studie av subjektspacering i norskt skriftspråk, nämligen 66,78 % i ett material från perioden 1910–1925 och 78,65 % i perioden 1985–2013.

⁸ $P = 0.3207$; för samtliga tester har vi använt funktionen *prop.test()* i *ctest*-paketet i R (R Core Team 2020).

⁹ $P << 0.005$.

initiala. Det kan delvis förklara minskningen i andelen skiftade subjekt i T-domänen från 22 % i LIA till 5 % i yngre AmNo. Vi kan anta att ökningen av andelen subjekt i spec-CP särskilt sker på bekostnad av skiftade subjekt i T-domänen. Den skiftade positionen kan precis som den satsinitiala sägas vara omarkerad, och båda positionerna innehåller i stor utsträckning topikala subjekt (se t.ex. Keihl Olsen 2019).

Däremot kan inte ökningen av andelen satsinitiala subjekt helt förklara att det blir vanligare med oskiftade subjekt i äldre AmNo, med en ökning från 2 % i LIA till 6 % i äldre AmNo; skillnaden är signifikant.¹⁰ Det ser alltså ut som om preferensen för SS är något svagare i den äldre AmNo än i LIA-materialet. Samtidigt är det knappast fler oskiftade subjekt i yngre AmNo (3 %) än i LIA (2 %) – här gör sig ökningen i SV-följd gällande. När vi inkluderar de satsinitiala subjekten ser vi alltså inte den påtagliga ökning av andelen oskiftade subjekt som noterats i nutida AmNo av A&W (se avsnitt 1), som endast inkluderar TP-subjekt i sin undersökning.

Om vi nu enbart tittar på satser som har subjekt i T-domänen, blir resultaten som väntat mer lika vad som rapporteras av A&W; se tabell 3, som visar en gradvis nedgång i andelen skiftade subjekt (S-Neg), och en motsvarande gradvis ökning av andelen oskiftade subjekt (Neg-S).

	<i>S-Neg</i>	<i>Neg-S</i>	<i>N</i>
<i>LIA</i>	91 %	9 %	190
<i>Äldre AmNo</i>	75 %	25 %	79
<i>Yngre AmNo</i>	52 %	48 %	141
<i>Totalt</i>	75%	25 %	410

Tabell 3. Pronominella subjekt i T-domänen i de tre materialen (alla typer av satser).

Utifrån enbart tabell 3 kan det vara närliggande att anta att ökningen av Neg-S beror på att talarna alltmer föredrar oskiftade subjekt. Vår tolkning är emellertid att nedgången i SS inte först och främst är ett

¹⁰ $P = 0.02869$ för skillnaden mellan andelen Neg-S i LIA och äldre AmNo.

resultat av att talarna undviker att flytta subjektet. Tvärtemot tror vi att nedgången framför allt beror på att preferensen blir allt starkare för att flytta subjektet längre – till spec-CP – i alla fall i deklarativer i den yngsta perioden. Förändringen kan illustreras med exemplen i (4). LIA-exemplet i (4a) inleds med ett topikaliserat objekt; subjektet står i T-domänen, och är skiftat. Yngre AmNo har istället SV-följd.

- (4) a. det trur **eg ikkje** (LIA, sel_uio_0301)
 b. no **jeg** trur **ikke** det (yngre AmNo, blair_WI_01gm)

I yngre AmNo är det alltså spec-CP som “tömmer” subjektspositionen före negation i T-domänen. Ökningen av andelen Neg-S i tabell 3 är då huvudsakligen ett epifenomen som följer av att färre subjekt står i T-domänen, och att särskilt en viss typ av subjekt flyttar vidare därifrån. Till det kommer en mycket mindre ökning av andelen oskiftade subjekt, som vi kan observera redan i äldre AmNo.

A&W föreslår att ökningen av andelen oskiftade subjekt beror på påverkan från engelska. De noterar att många oskiftade subjekt förekommer i frågor, där engelskan gärna har följden Neg-S (*isn't she clever?*). Också i våra amerikanorska material varierar andelen SS något med satstyp; se tabell 4 som inkluderar alla satser med subjektet i T-domänen (men inga satser med SV-följd), fördelat efter satstyp.

	<i>Skifte i deklarativer</i>	<i>Skifte i bisatser</i>	<i>Skifte i frågor</i>
<i>LLA</i>	90 % (n=130)	93 % (n=42)	94 % (n=18)
<i>Äldre AmNo</i>	73 % (n=59)	71 % (n=14)	100 % (n=6)
<i>Yngre AmNo</i>	62 % (n=53)	71 % (n=42)	24 % (n=46)

Tabell 4. Pronominella subjekt i T-domänen indelade efter satstyp. Andel skiftade subjekt av samtliga TP-subjekt.

I yngre AmNo finns en hel del exempel på påhågsfrågor som i (5). Dessa följer, som A&W påpekar, ett mönster som liknar engelska (*she's clever, isn't she?*): subjektet är oskiftat.

- (5) a. du kan klare det, kan **ikke du**?
 (yngre AmNo, coon_valley_WI_07gk)
 b. var fjorten tusen, var **ikke det**?
 (yngre AmNo, coon_valley_WI_03gm)

Det finns emellertid inga exempel på oskiftade subjekt i frågor i vårt äldre amerikanorska material. Däremot kan vi som sagt se en minskning av andelen skiftade subjekt i deklarativer och bisatser på det här stadiet; här verkar påverkan från engelska mindre relevant, eftersom engelska inte har Neg-S-följd i de kontexterna (*that she isn't clever*).

Ytterligare ett argument för att andelen subjekt i spec-CP påverkar fördelningen av TP-subjekten får vi när vi jämför deklarativa huvudsatser med bisatser i tabell 4. I yngre AmNo har andelen skifte i deklarativer minskat ytterligare (från 73 till 62 %), medan mängden skifte i bisatser framstår som stabil (71 % i bägge perioderna). Vi antar att deklarativa huvudsatser och bisatser har samma subjekspositioner i T-domänen (m.a.o. en position före och en efter negation).¹¹ Skillnaden i andelen SS mellan satstyperna låter sig förklaras med att topikala subjekt i allt högre grad står i spec-CP i deklarativer, vilket minskar andelen skiftade subjekt i T-domänen (jfr ovan). Bisatser påverkas inte av detta, eftersom spec-CP inte är tillgänglig. Också A&W noterar en skillnad mellan olika satstyper (60 % skifte av pronomina i huvudsatser mot 86,1 % i bisatser).¹²

Om vi bortser från påhågsfrågor i yngre AmNo, och flyttning av allt fler subjekt till spec-CP, kan vi notera att det i hög grad är samma faktorer som bestämmer placeringen av subjekspronomen i de olika materialen. De exempel vi finner på följden negation–pronomen kan

¹¹ Det kan emellertid också vara skillnad mellan olika typer av bisatser, och t.ex. villkorsbisatser har gärna fler oskiftade subjekt än andra typer av satser; det gäller även för danskan, där subjektskifte annars är obligatoriskt (se Bentzen 2014 och referenser där).

¹² Vi förstår det som att huvudsatserna hos A&W inkluderar frågor.

ofta förklaras utifrån faktorer som är kända från modern hemlandsnorska. Det gäller alla tre materialen. Ett fåtal oskiftade pronomen är fokuserade; se (6):

- (6) a. da sette dei nokre store rauk men det har **ikkje eg** vore med på (LIA, brandbu_ma_01)
 b. det er E1n som e #kjører nedi der [...] og da har **ikke du** borti der å gjøre du (äldre AmNo, spring_grove_MN_18um)
 c. ja, det skjønner **ikke jeg** heller (yngre AmNo, coon_valley_WI_03gm)

I samtliga material återstår ändå en del variation, och SS framstår inte som helt obligatoriskt. I exemplet i (7) är samtliga subjektspronomen obetonade och ofokuserade.

- (7) a. ho var ikkje fastgjorda ennå ser du # så da gjekk **ikkje ho** fram og ikkje tilbake (LIA, brandbu_ma_01)
 b. jeg gjorde det # mye før jeg var gift men nå # har **ikke jeg** kommet ut for mye (äldre AmNo, blair_WI_23um)
 c. ja da er **ikke det** stille (CANS 2010, westby_WI_03gk)

Sammanfattningsvis kan vi alltså notera tre förändringar när det gäller subjektsplacering: en ökning av andelen SV-följd i yngre AmNo, en ökning av andelen oskiftade subjekt i deklarativa huvudsatser framför allt i äldre AmNo, samt en ökning av mängden oskiftade subjekt i yngre amerikanorska påhångsfrågor. Vi återkommer till de här förändringarna i avsnitt 5. Först tittar vi närmare på placeringen av objekt visavi negation (4.2) och verbpartikel (4.3).

4.2 Objektskifte

Också objektskifte beror på syntaktiska och informationsstrukturella faktorer, och prosodi. I norskan är OS bara möjligt med lätta pronomen, och huvud verbet måste ha flyttat ut ur verbfrasen (Holmberg 1986); jfr (8a) och (8b). Fokuserade objektspronomen skiftar inte. Vidare skiftar som regel inte pronomen med icke-nominal antecedent

(9a) eller typ-referens (9b) (Anderssen och Bentzen 2012, Andréas-son 2013).

- (8) a. Han likte **den ikke**.
 b. Han vil {***den**} **ikke like** {**den**}.
- (9) a. Synes du [jeg skal kjøpe denne]? – Nei, jeg synes **ikke det**.
 b. Alle har visst [blå sykkel] i år, men jeg har **ikke det**.

I praktiken förekommer skiftade objekt därför bara i V2-satser utan hjälpverb, med negation eller andra satsadverbial, och med pronomi-nellt objekt med nominal antecedent; potentiella kontexter för ob-jektsskifte är således sällsynta.

I modernt hemlandsnorskt talspråk skiftar ca 87 % av objektspronomen med nominal antecedent, jämfört med omkring 5 % för pronomen med icke-nominal antecedent (Bentzen et al. 2013, A&W 2020). A&W finner en lägre andel skiftade objekt med nominal antecedent i modern AmNo. I deras material, som är hämtat från CANS version 2, skiftar 61 % (25/41) av dessa objektspronomen (se tabell 8 hos A&W).¹³

Vårt material från yngre AmNo visar också en tendens till en minskning av andelen skiftade objekt; se tabell 5. Andelen OS med nominal antecedent är faktiskt till och med något lägre (52 %) än den som noteras hos A&W, men notera att vårt urval från yngre AmNo också ingår som en del i A&Ws material.¹⁴

Vi noterar att det emellertid inte bara är yngre AmNo som verkar skilja sig från modern hemlandsnorskt talspråk: äldre AmNo och LIA har också lägre andel objektskifte (60 respektive 67 %) än modern hemlandsnorska generellt (runt 87 % skifte, som nämnt). Vi måste

¹³ När det gäller objektspronomen med icke-nominal antecedent är det å andra sidan ingen nedgång i andelen skifte i amerikanorskan, utan tvärtom en ökning till 11 % (11/100) skiftar i A&Ws material. För oss framstår det som troligt att förändringen åtminstone delvis beror på att färre objekt står satsinitialt i yngre amerikanorska än i hemlandspråket (eftersom preferensen för SV-följd är så mycket starkare; se avsnitt 4.1 ovan och jämför nedan).

¹⁴ Det är möjligt att våra principer för excerpering och annotering kan ha påverkat resultatet något. Det framgår till exempel inte helt klart hur A&W har analyserat formella objekt (se avsnitt 3 och fotnot 3 om hur vi hanterar dem).

	<i>OS med nominal antecedent</i>	<i>OS med icke- nominal antecedent</i>	<i>Totalt</i>
<i>LIA</i>	67 % (n=6)	23 % (n=13)	37 % (N=19)
<i>Äldre AmNo</i>	60 % (n=5)	0 % (n=10)	20 % (N=15)
<i>Yngre AmNo</i>	52 % (n=25)	4 % (n=23)	29 % (N=48)

Tabell 5. Skifte av pronomen med nominal och icke-nominal antecedent.

förstås reservera oss för att de absoluta talen är väldigt låga, eftersom kontexterna för OS är så sällsynta.

När vi undersöker beläggen från ett mer kvalitativt perspektiv, får vi intryck av att distributionen av OS i yngre AmNo präglas av kontinuitet från tidigare generationer snarare än förändring. I alla tre materialen finner vi att objekt med icke-nominal antecedent eller typpreferens bara skiftar i undantagsfall. Lätta pronomen med nominal antecedent skiftar å andra sidan som väntat; se (10).

- (10) a. men eg minnest **henne ikkje** veit du (LIA, lesja_ma_01)
 b. ... hun rokket **han ikke**
 (äldre AmNo, spring_grove_MN_18um)
 c. men han fikk **henne ikke** (yngre AmNo, westby_WI_03gk)

Vi kan emellertid notera exempel med oskiftade objekt i samtliga material, i kontexter som till synes uppfyller kriterierna för skifte, och där vi utifrån modernt norskt talspråk generellt skulle vänta oss ett skiftat pronomen. Några exempel ges i (11).

- (11) a. og ## eg kom der heilt uforløyves så han såg **ikkje meg** da
 (LIA, biri_uio_0202)
 b. han # kom på meg uvitende og # så **ikkje meg**
 (äldre AmNo, blair_WI_23um)
 c. vi måtte gå da veit du for de # e kjørte **ikke oss** # på med bil
 (yngre AmNo, coon_valley_WI_06gm)

Objekten i (11) har alla nominal antecedent, och de är trycklätta och icke-fokuserade. Det är förväntat att vi finner oskiftade pronomen av den här typen, givet resultaten hos A&W, men våra data från LIA och äldre AmNo tyder på att OS inte var fullt obligatoriskt i baseline (med exempel som (11a)). Tvärtom verkar det tidigare ha varit något mer tillåtet med oskiftade objektspronomen i Oppland än vad det är generellt i norskt talspråk idag.¹⁵ Variationen har upprätthållits under flera generationer i Amerika. Tendensen till oskiftade pronomen har där förstärkts med tiden (jämför A&W, se också Polinsky 2018, 34). Däremot ser inte utvecklingen ut att bero på någon grundläggande amerikanorsk innovation i det grammatiska systemet.

När det gällde subjektplacering kunde vi notera att den kraftiga ökningen av andelen subjektsinitiala sats påverkar fördelningen av subjekten i T-domänen. Det är möjligt att den starka preferensen för SV-följd i yngre AmNo också påverkar fördelningen mellan skiftade och oskiftade objekt. Jämför exemplen i (4) igen, upprepade i (12) nedan:

- (12) a. **det** trur eg **ikkje** (LIA, sel_uio_0301)
 b. no jeg trur **ikke det** (yngre AmNo, blair_WI_01gm)

I LIA-exemplet i (12a) står objektet initialt, men i den amerikanorska SV-satsen i (12b) är det istället oskiftat. Det är således möjligt att ökningen av andelen oskiftade objekt i yngre AmNo åtminstone delvis beror på att det är färre satsinitiala objekt. Det skulle i så fall vara ännu ett exempel på att ökningen av andelen SV-följd skapar en kedjereaktion med konsekvenser för andra syntaktiska fenomen. För OS är konsekvensen motsatt den vi såg för SS: det blir fler potentiella kontexter för objektskifte, inte färre. Det verkar också som om detta gynnar den oskiftade positionen: samtidigt som färre objekt topikaliserar, ökar andelen oskiftade objekt. En närliggande hypotes är att

¹⁵ En anonym granskare noterar är det inte är helt oväntat att LIA-materialet från Oppland uppvisar mer variation än modernt norskt talspråk från andra delar av landet. Resultaten från ScanDiaSyn-undersökningen tyder också på att äldre talare i Oppland ibland godtar oskiftade pronomen i högre utsträckning än yngre (se den nordiska syntaxdatabasen, Lindstad et al. 2009).

objekt som flyttar fram till spec-CP (vilket alltså sker i högre grad i hemlandsnorska än i modern AmNo) särskilt tömmer den oskiftade positionen. En sådan analys verkar rimlig med tanke på exempel som i (12), där objektet har icke-nominal antecedent. Det är möjligt att det också gäller mer generellt, men vi lämnar den frågan till framtida forskning.

4.3 Partikelplacering

I de flesta moderna norska dialekter står nominalfrasobjekt vanligen efter partikel, men framförställda objekt förekommer också, om än mindre ofta. Pronomen står endast i undantagsfall efter partikel (Larsson och Lundquist 2014). Ordföljden kan delvis bero på vad det är för typ av partikel: Icke-direktionella (metaforiska) partiklar står något oftare före objekt än direktionella partiklar. Om partikeln följs av en PP är följden objekt–partikel vanlig, även med icke-pro-nominella objekt.

Också i våra material placeras pronomen oftast före partiklar; se tabell 6 som inkluderar samtliga excerperade partikelkonstruktioner med objekt (med direktionella och icke-direktionella partiklar, och konstruktion med partikel + PP; jämför avsnitt 3 ovan). Variationen är emellertid återigen större än vad vi skulle vänta oss mot bakgrund av modern hemlandsnorska, i samtliga material. I LIA står 18 % av alla pronomen som ingår i partikelkonstruktioner efter partikeln; motsvarande siffra för följden partikel–pronomen i Oslo-delen av den nordiska ordföljdsdatabasen (Lundquist et al. 2019) är endast 1 % (4/443); även om beläggen är jämförelsevis få i LIA-materialet är det alltså en påtaglig skillnad. Variationen tycks bibehållas i AmNo, men eventuellt minska marginellt; vi noterar 12 % pronomen efter partikel i äldre AmNo och 16 % i det yngre materialet.

	<i>Partikel-pro</i>	<i>Partikel-NP</i>
<i>LIA</i>	18 % (n=38)	97 % (n=73)
<i>Äldre AmNo</i>	12 % (n=25)	80 % (n=46)
<i>Yngre AmNo</i>	16 % (n=45)	94 % (n=66)

Tabell 6. Följden partikel–objekt i konstruktion med verbpartikel.

Exempel på variationen ges i (13)–(14). Exempelen i (13) illustrerar den normala ordföljden med pronomen före partikel i de tre materialen. I (14) ser vi exempel på den motsatta ordföljden, som alltså är mindre vanlig.

- (13) a. ja det var som regel i mars det for ein måtte ha køyrt **det ut** med det det var så mykje snø at det var føre (LIA, brandbu_ma_01)
- b. og dette kjøttet ble tørt så hengte vi **det opp** # om sommeren så brukte vi å koke suppe på det (äldre AmNo, coon_valley_WI_51gk)
- c. varm han opp # bare varm **han opp** litt? (yngre AmNo, westby_WI_06gm)
- (14) a. og så gav han # han sukkerbetane dama og han åt **opp dei** (LIA, brandbu_ma_01)
- b. blir kornet right da må vi harske korn og få **inn det** og (äldre AmNo, blair_WI_24um)
- c. det blåste så mye tre # blåste ned # så vi # hadde [uninterpretable] # pynta **opp det** [laughter] nedi skogen (yngre AmNo, coon_valley_WI_04gm)

I modern hemlandsnorska är det normalt större variationsmöjligheter i placeringen av icke-pronominella objekt än pronomina objekt (se t.ex. Larsson och Lundquist 2014). I Oslo-delen av den nordiska ordföljdsdatabasen föregår icke-pronominella objekt en partikel i 34 % av fallen (138/401); om man bortser från direktionella partiklar med PP, står 20 % (54/275) av nominalfraserna före partikel. Också i våra material kan vi belägga ordföljdsvariation; jämför (15) som har partikeln före objektet och (16) som har följden objekt–partikel. Notera att (16c) har partikel med en direktionell PP; då är följden objekt–partikel den vanliga också i hemlandsnorskan.

- (15) a. og så drakk dei **opp brennevinet hans** (LIA, biri_uio_0104)
- b. da plukka **opp mye høy** (äldre AmNo, westby_WI_22gm)
- c. så # e de e # satte **opp e # tenter** (yngre AmNo, sacred_heart_MN_01gk)

- (16) a. kom til å køyre # sånn **bommen ned** (LIA, sel_uio_0301)
 b. ha oss to guttungene til å pile **den veden opp**
 (äldre AmNo, iola_WI_09gm)
 c. en handle # som # stikker **tobakken inn** i inn i # inn i den e
 ## rullen (yngre AmNo, westby_WI_11gm)

Variationen är emellertid betydligt mer begränsad än vad vi kunde vänta oss från den moderna hemlandsnorskan, och det gäller samtliga material. Om man bortser från belägg med en direktionell PP (som (16c)), står 90 % (i äldre AmNo) eller 97 % (LIA, yngre AmNo) av alla icke-pronominella objekt efter partiklar.

Sammanfattningsvis kan vi knappast notera någon förändring i AmNo när det gäller partikelplacering. Det är variation redan i baseline, och det gäller faktiskt framför allt pronominella objekt. Variationen tycks bevaras i flera generationer i arvspråket.¹⁶

4.4 Sammanfattning

I det föregående har vi sett att subjektspacering, objektskifte och partikelplacering uppvisar olika utvecklingslinjer i de tre materialen. När det gäller subjektspacering kan vi notera en förändring som tycks specifik för arvspråket: medan andelen SS i LIA-materialet ligger mycket nära det moderna hemlandsspråket, har AmNo en lägre andel. Vi föreslår att nedgången av andelen SS i deklarativer i yngre AmNo kan knytas till en ökning av andelen subjekt i spec-CP. När det gäller *ja/nej*-frågor antar vi med A&W att påverkan från engelskan kan vara relevant, men också det gäller bara i yngre AmNo. Generellt är det annars samma faktorer som påverkar SS i hemlandet och AmNo, och SS framstår inte som helt obligatoriskt i någon av varieteterna. Det är likväl något fler oförklarade oskiftade subjekt i (äldre) AmNo; vi uppfattar detta som en frekvensförskjutning som har med arvspråkets speciella förutsättningar att göra.

¹⁶ Polinsky (2018, 53) observerar att arvspråkstalare av engelska ofta avviker från baseline när det gäller bruket av partiklar, oavsett vad de har som dominerande språk. I det avseende är det överraskande att de norska partikelkonstruktionerna framstår som så stabila.

När det gäller objektskifte kan vi med A&W notera att modern AmNo har fler oskiftade pronomen än modernt norskt talspråk. Det är möjligt att den högre andelen subjekt i spec-CP (på bekostnad av topikaliserade objekt) spelar en viss roll också här. Samtidigt noterar vi oskiftade objektspronomen redan i LIA-materialet. AmNo uppvisar alltså åtminstone i viss mån stabilitet i förhållande till baseline – men baseline framstår inte som helt identisk med den moderna hemlandsnorskan. När det gäller partikelplacering har vi funnit en hög grad av stabilitet. Det gäller inte minst placeringen av pronomen visavi partikel, och också här kan vi notera mer variation i baseline än vad som väntas utifrån den moderna hemlandsnorskan. När det gäller icke-pronominella objekt, är det å andra sidan mindre variation än väntat.

5 Diskussion

Vår undersökning av argumentplacering hos talare med bakgrund i Oppland visar att det är variation redan i baseline: varken (obetonade) pronomina subjekt eller objekt står helt obligatoriskt efter negation i vårt material från LIA-korpusen; pronomina objekt står också ibland efter en verbpartikel. Variationen i LIA tycks åtminstone delvis större än i de moderna norska dialekterna, och den bevaras också i hög grad i AmNo, över flera generationer. Ett huvudresultat i vår undersökning är stabilitet – vi kan inte observera någonting i den amerikanska grammatiken som inte också kan beläggas i LIA. Samtidigt kan vi notera ett par förskjutningar i variationen. I det följande diskuterar vi olika faktorer som kan ha bidragit till utvecklingen i AmNo.

5.1 Grammatisk komplexitet och syntaktisk ekonomi

Det är välbekant att grammatisk komplexitet och syntaktisk ekonomi kan spela roll för språkinlärning och i utvecklingen av arvspråk (se t.ex. Anderssen et al. 2018). Det är också känt att arvstalare ofta undviker komplexa strukturer (t.ex. bisatser) i sin språkproduktion (se t.ex. Larsson och Johannessen 2015, och Polinsky 2018, 58), och det

påverkar inte minst den input som arvstalare får i tredje eller fjärde generationen. Syntaktisk ekonomi är något delvis annat – man brukar tänka sig att ekonomiprinciper vägleder barnets språktillägnande (se t.ex. van Gelderen 2011, 21ff.). Anderssen et al. (2010) och Anderssen et al. (2012) noterar t.ex. att norska barn tenderar att ha fler oskiftade subjekt och objekt än vuxna och de föreslår att det beror på att det är mindre ekonomiskt att flytta argument än att låta det stå kvar efter adverbiallet. Om ekonomiprinciper är inblandade i utvecklingen av amerikanorskan, förväntar man sig därmed mindre skifte av subjekt, objekt och (beroende av analysen av ordföljden i partikelkonstruktioner) partiklar. Till viss grad är det också vad vi ser: äldre AmNo har något lägre andel SS än LIA, utan att det kan förklaras med tvärspråkligt inflytande eller preferens för satsinitiala subjekt (jfr nedan och se vidare 5.3). Både i äldre och yngre AmNo har vi också kunnat observera att något färre objekt skiftar, i linje med vad man kan förvänta utifrån principer om syntaktisk ekonomi. I partikelkonstruktioner ser vi emellertid ingen förändring.

När det gäller subjektspacering i yngre AmNo har vi visat att förändringen i första hand handlar om en preferens för att flytta subjektet hela vägen till den satsinitiala positionen, alltså längre än vad som sker vid SS. Likväl kan man argumentera för att detta har med ekonomi att göra: V2-följden är relativt stabil i AmNo (t.ex. Eide och Hjelde 2018, Lohndal och Westergaard 2019), och i den mån det kan noteras avvikelser handlar det typiskt om att det står två led före det finita verbet snarare än att platsen står tom. Ett led måste alltså flytta fram till den satsinitiala positionen – och det mest ekonomiska är att flytta subjektet, som i utgångspunkten står högst i strukturen.

Ekonomiprinciper kan alltså antas styra barns språktillägnande och därmed påverka hur det grammatiska systemet kommer att utvecklas över generationer. Särskilt i en arvspråkskontext där input är mer begränsat kan de leda till förändring i en åtminstone delvis förutsägbar riktning.¹⁷ Det kan emellertid också tänkas att komplexitet

¹⁷ En anonym granskare frågar varför ändringarna sker när de gör (jfr t.ex. Walkden 2017 och referenser där om "the actuation problem"), om ekonomi och komplexitet är principer som alltid gäller. Vi föreställer oss som framgått att

i en mer ytlig bemärkelse (jfr ovan) spelar in. Det är som noterat känt att arvstalare i sin språkliga produktion undviker konstruktioner som är krävande ur processeringssynpunkt, och det är möjligt att det är så vi ska förstå ökningen av SV-följd: man kan tänka sig att subjektsinitiala satser föredras eftersom de är enklast att processa. Om det är så vi ska förstå den kraftiga ökningen av andelen SV-följd rör det sig alltså inte om någon underliggande förändring av grammatiken.

Det ska tilläggas att varken ekonomi eller komplexitet är helt okontroversiella begrepp, och även om man skiljer på dem i teorin kan de vara svåra att särskilja i praktiken. Man kan som framgått tänka sig att det är mindre ekonomiskt ur snäv syntaktisk synvinkel att topikalisera ett objekt (från verbfrasen till spec-CP) än att spetsställa ett subjekt (från T-domänen till spec-CP), men objektsinitiala strukturer är också mer krävande ur processeringssynpunkt (jfr Polinsky 2018, 241). Argumentplacering ger oss därför inte utan vidare någon god möjlighet att pröva vilken faktor som är relevant. Istället skulle t.ex. studier av resumptiva pronomen eller nollargument kanske ge information: det kan tänkas vara mindre ekonomiskt med realiserade (resumptiva) pronomen, men det gör satserna lättare att processa. Som Polinsky (2018, 69) noterar är det vanligt att arvstalare använder resumptiva pronomen just för att undvika strukturer som är svåra att processa. Det skulle kunna tala för att ökningen av SV-följd snarare har med komplexitet och processering att göra än syntaktisk ekonomi.

5.2 Variation och frekvens i input

Det är högst sannolikt att också ett fenomenens frekvens spelar viss roll både vid inläring och för hur en struktur bibehålls under talarens liv. Som A&W påpekar skulle man därför kunna förvänta sig att SS

barn styrs av ekonomiprinciper vid språktillägnet, men med tillräcklig input tillägnas mindre ekonomiska strukturer. Arvspråksinläring utmärks av begränsad input, och vi kan därför förvänta oss förändring. När det gäller komplexitet är frågan en delvis annan, och det är inte alls givet att komplexitet alltid spelar samma roll för alla talare. Tvärtom väntar vi oss att arvspråkstalare på ett annat sätt än hemlandstalare undviker komplexa strukturer i sin språkliga produktion (se t.ex. Polinsky 2018:58).

och OS beter sig olika i arvspråket – SS har hög textfrekvens (och bör vara vanligt förekommande i talarnas input), OS låg. Vi behöver då komma ihåg att hur vi räknar frekvens beror på vilka kategorier vi antar att barnet tar hänsyn till i språkinläringen. Även om man som A&W utgår från att barn är känsliga för ganska fina grammatiska distinktioner (t.ex. pronomen – icke-pronomen), finns det olika tänkbara alternativ (svaga–starka pronomen, pronomen med nominal – icke-nominal antecedent) som måste utvärderas empiriskt.¹⁸

När det gäller subjektspacering har vi sett att andelen SS minskar något i äldre AmNo, trots relativt hög textfrekvens. I det här fallet kan det se alltså se ut som om syntaktisk ekonomi har större betydelse än textfrekvens; vi ska då också komma ihåg att SS inte är helt obligatoriskt ens i baseline. När det gäller OS, å andra sidan, drar både (låg) frekvens och ekonomi i samma riktning: OS har låg textfrekvens och är “oekonomiskt” eftersom det kräver flyttning av objektet (se 5.1).

I yngre AmNo ser vi en kraftig ökning av andelen satsinitiala subjekt. Här kan man på samma sätt som vid OS argumentera för att frekvens och ekonomi påverkar utvecklingen på samma sätt: SV-satser är som nämnt syntaktiskt ekonomiska, och samtidigt frekventa även i baseline (se tabell 2). (VS-satser är å andra sidan inte heller särskilt ovanliga i baseline, men de är alltså mindre ekonomiska.)

5.3 Direkt och indirekt tvärspråkligt inflytande: CLI och CLO

Också tvärspråklig påverkan (CLI) från engelskan eller överkorrigering (CLO) kan vara inblandad i utvecklingen av ett arvspråk (se avsnitt 2), delvis beroende på hur lika eller olika de inblandade språken är med avseende på det aktuella fenomenet. Här är det viss skillnad mellan de tre olika fenomenen som undersöks i den här studien. När det gäller partikelplacering är engelska och norska relativt lika: pro-

¹⁸ Andraspråksinlärare är inte nödvändigtvis känsliga för samma kategorier som modersmålsinlärare. I studier av SS hos andraspråkstalare av norska, verkar det istället som om pronomen och fulla nominalfraser hanteras på samma sätt – båda typerna av subjekt placeras före negation (Anderssen et al. 2018, Westendorp och Lundquist kommande). För L2-inlärare tycks också förstaspråket kunna påverka preferenserna.

nominella objekt står normalt före partikeln (*throw it out* inte **throw out it*), medan icke-pronominella objekt antingen står före eller efter (*throw the garbage out/throw out the garbage*). Ordningen mellan pronomen och partikel er emellertid mer kategorisk i engelskan än i norskan, medan det i modern hemlandsnorska är ovanligare med icke-pronominella objekt före partikel än vad det är i engelska (jämför Larsson och Lundquist 2014 med Haider 2020, fn. 7). Den stabilitet som vi kan notera med avseende på partikelplacering i AmNo kan därmed delvis bero på konvergens med engelska, men samtidigt ser vi att variationen – som delvis avviker från engelska – också bevaras över generationerna.

När det gäller subjektspacering har engelskan huvudsakligen SV-följd, om man bortser från frågor där subjektet ofta, men inte alltid, följer negation (jämför A&W): *Didn't you/Lisa go?* men också *Did Lisa/you not go?* Vi antar i likhet med A&W att CLI spelar en roll för den lägre andelen SS i frågor i yngre AmNo. CLI har samtidigt ganska begränsad betydelse, som bara kan observeras i det yngsta materialet, och vi har inte anledning att tro att påverkan sprider sig till andra kontexter. I deklarativer har vi visat att ändringen i subjektspacering först och främst handlar om en ökad andel SV-följd, snarare än en preferens för oskiftade subjekt. Lohndal och Westergaard (2019) tolkar i och för sig ökningen av SV-följd som CLI, eftersom engelsk huvudsakligen har SV-följd. Vi vill emellertid inte i nuläget dra den slutsatsen; i 5.1 ovan nämnde vi att ekonomi och/eller undvikande av komplexitet också skulle kunna vara förklaringar. För att reda ut orsakssambanden krävs andra typer av data än från produktion av amerikanorska, t.ex. från undersökningar av norska (eller andra nordiska språk) som arvspråk i kontakt med andra språk än engelska, och språk som i mindre utsträckning sätter subjektet i den initiala positionen.

När det gäller OS saknar engelskan helt motsvarighet – objekt är alltid verbfrasinterne (om de inte har topikaliserats). Vi kan notera att huvud verbet följer satsadverbialiet i engelska (*he always read the newspaper* inte **he read always the newspaper*), och att inte heller norskan tillåter OS när verbet står kvar i verbfrasen. Det är alltså inte uppenbart vad påverkan från engelskan skulle innebära för bruket av

OS i AmNo (jfr Anderssen och Westergaard 2020).¹⁹ Utvecklingen mot mindre OS är därför inte något entydigt exempel på vare sig CLI eller CLO. Däremot kan den ses som en förstärkning av ett mönster som syns som variation redan i baseline (Polinsky 2018, 34).

5.4 Sammanfattning

Sammanfattningsvis betraktar vi alltså utvecklingen i AmNo som frekvensförskjutningar vilka i princip kan förstås inom ramen för ett stabilt system som tillåter variation. Vi noterar inga säkra tecken på CLI bortsett från subjeksplacering i frågor, och ingenting tyder direkt på CLO, även om vi ser att tendenser till variation i baseline kan förstärkas. Komplexitet och ekonomi spelar sannolikt roll för riktningen på förändringen. Det vi framför allt kan se är ändå att variationen kan spåras tillbaka till baseline, och vi kan notera en hög grad av stabilitet.

6 Avslutning

Tillsammans med de äldre inspelningarna i CANS, gör LIA-korpusen det möjligt att undersöka norska som arvspråk i Nordamerika över flera generationer, från tiden för massutvandringen på 1800- och 1900-talet, fram till idag. Det här gör korpusarna till en ovärderlig resurs. I den här artikeln har vi använt dem för att undersöka placering av subjekt och objekt visavi negation och partiklar. Vi har jämfört argumentplaceringen i tre material: LIA (dvs. äldre hemlandsnorska), äldre AmNo och yngre AmNo.

Ett huvudresultat är att det är variation i alla materialen, och delvis mer variation än väntat utifrån vad vi vet om modernt hemlands-

¹⁹ Variation i argumentplacering beror sannolikt åtminstone delvis på prosodiska preferenser (se t.ex. Ertshik-Shir et al 2020), och man kunde möjligen tänka sig prosodiskt motiverad påverkan eller överkorrigering. Vi förväntar oss dock snarast att prosodi bidrar till ett stabilt system, eftersom prosodi generellt inte tycks påverkas i så stor grad i arvspråk, åtminstone inte jämfört med L2-språk (Polinsky 2018:187). Vi vet emellertid ännu ganska lite om amerikanorsk prosodi.

norskt talspråk. Det innebär att något av det som ser ut som särdrag i AmNo (särskilt när det gäller objektskifte och partikelskifte), från ett modernt norskt perspektiv, inte är uttryck för någon grundläggande grammatisk innovation i AmNo; alla fenomenen kan också beläggas i det äldre norska dialektmaterialet. Variationen är relativt stabil över generationerna (och det gäller särskilt partiklarna), och det är samma faktorer som bestämmer ledföljden i arvspråket som i hemlandet. Resultaten visar att arvspråkstalarna är känsliga för fina grammatiska distinktioner. I det avseendet är de som andra modersmålstalare, och inte alls som L2-inlärare.

Vi har likväl observerat några frekvensförskjutningar som är specifika för AmNo. För det första är det lite mindre SS och OS redan i äldre AmNo, vilket vi kan tolka som ett uttryck för generella ekonomiprinciper som motverkar flyttning. I yngre AmNo finner vi ännu färre exempel på SS, men den utvecklingen är begränsad till frågor och deklarativer. I frågor handlar det troligtvis om CLI från engelska som främjar den oskiftade positionen i T-domänen (jfr A&W). När det gäller deklarativer, å andra sidan, har vi argumenterat för att den avgörande förändringen inte handlar om preferensen för oskiftade subjekt, utan en preferens för flyttning av subjektet till C-domänen, vilket visar sig i en kraftig ökning av andelen SV-följd. Den här ökningen har tidigare analyserats i termer av CLI från engelska (Lohndal och Westergaard 2019). Vi föreslår i stället att ekonomi kanske kan vara en bidragande orsak även här, men vi kan heller inte utesluta att också faktorer som har med processering att göra påverkar arvspråket. Studier av nordiska språk som arvspråk i kontakt med andra majoritetspråk än engelska bör kunna belysa den frågan vidare.

Referenser

- Aa, Leiv Inge. 2020. *Norwegian Verb Particles*. Amsterdam: John Benjamins.
- Anderssen, Merete och Kristine Bentzen. 2012. Norwegian object shift as IP-internal topicalization. *Nordlyd* 39 (1), 1–23.

- Anderssen, Merete, Kristine Bentzen och Yulia Rodina. 2012. Topicality and complexity in the acquisition of Norwegian object shift. *Language Acquisition* 19 (1), 39–72.
- Anderssen, Merete, Kristine Bentzen, Yulia Rodina och Marit Westergaard. 2010. The acquisition of apparent optionality: The word order of subject and object shift constructions in Norwegian. I *Variation in the Input: Studies in the Acquisition of Word Order*, redigerad av Merete Anderssen, Kristine Bentzen och Marit Westergaard, 241–270. Dordrecht: Springer.
- Anderssen, Merete, Björn Lundquist och Marit Westergaard. 2018. Cross-linguistic similarities and differences in bilingual acquisition and attrition: Possessives and double definiteness in Norwegian heritage language. *Bilingualism: Language and Cognition*, 21 (4), 748–764.
- Anderssen, Merete och Westergaard, Marit. 2020. Word order variation in heritage languages: Subject shift and object shift in Norwegian. I *Lost in Transmission: The role of attrition and input in heritage language development*, redigerad av Bernhard Brehmer och Jeanine Treffers-Daller, 100–124. Amsterdam: John Benjamins.
- Andréasson, Maia. 2010. Object shift or Object Placement in General. I *Proceedings of LFG10*, redigerad av Miriam Butt och Tracy Holloway King, 26–42. Stanford: CSLI Publications.
- Andréasson, Maia. 2013. Object shift in Scandinavian languages: The impact of contrasted elements. *Nordic Journal of Linguistics*, 36 (2), 187–217.
- van Baal, Yvonne. 2020. *Compositional definiteness in American heritage Norwegian*. Doktoravhandling, Universitetet i Oslo.
- Bentzen, Kristine, Merete Anderssen och 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.org/10.1017/S0332586513000218>
- Bentzen, Kristine. 2014. Subject Shift. *Nordic Atlas of Language Structures (NALS)* 1, 332–343. <https://dx.doi.org/10.5617/nals.5402>

- Eide, Kristin Melum och Arnstein Hjelde. 2018. Om verbplassering og verbmorfologi i amerikanorsk. *Maal og Minne* 2018 (1), 25–69.
- Erteschik-Shir, Nomi, Gunlög Josefsson och Björn Köhnlén. 2020. Variation in Mainland Scandinavian Object Shift: A Prosodic Analysis. *Linguistic Inquiry* 2020. https://doi.org/10.1162/ling_a_00393
- van Gelderen, Elly. 2011. *The Linguistic Cycle. Language Change and the Language Faculty*. Oxford: Oxford University Press.
- Haider, Hubert. 2020. VO-/OV-Base Ordering. I *The Cambridge Handbook of Germanic Linguistics*, redigerad av Michael T. Putnam och B. Richard Page, 339–364. Cambridge: Cambridge University Press.
- Haugen, Einar. 1953. *The Norwegian Language in America: A Study in Bilingual Behavior*. Philadelphia: University of Philadelphia Press.
- Hjelde, Arnstein. 2011. Some phonological changes in a Norwegian dialect in America. I *Language contact across the North Atlantic*, redigerad av Per Sture Ureland och Iain Clarkson, 283–296. Berlin: De Gruyter.
- Hjelde, Arnstein. 2012. ”Folkan mine, dæm bære snakka norsk” - Norsk i Wisconsin frå 1940-talet og fram til i dag. *Norsk lingvistisk tidsskrift* 30 (2), 183–203.
- Hjelde, Arnstein. 2015. Changes in a Norwegian dialect in America. I *Germanic Heritage Languages in North America. Acquisition, Attrition and Change*, redigerad av Janne Bondi Johannessen och Joseph C. Salmons, 283–298. Amsterdam: John Benjamins.
- Holmberg, Anders. 1986. *Word order and syntactic features in the Scandinavian languages and English*. Doktorsavhandling, Stockholms universitet.
- Johannessen, Janne Bondi. 2015. The Corpus of American Norwegian speech. I *Proceedings of the 20th Nordic conference of computational linguistics, NODALIDA 2015, May 11–13, 2015, Vilnius, Lithuania*, redigerad av Béata Megyesi, 297–300. NEALT.

- Johannessen, Janne Bondi och Signe Laake. 2012. Østnorsk som norsk fellesdialekt i Midtvesten. *Norsk lingvistisk tidsskrift* 30 (2), 365–380. <http://urn.nb.no/URN:NBN:no-46109>
- Johannessen, Janne Bondi och Ida Larsson. 2015. Complexity matters: On gender agreement in Heritage Scandinavian. *Frontiers in Psychology* 6, 1–20.
- Johannessen, Janne Bondi, Joel Priestley, Kristin Hagen, Tor A. Åfarli och Øystein A. Vangsnes. 2009. The Nordic Dialect Corpus – an advanced research tool. I *Proceedings of the 17th Nordic Conference of Computational Linguistics NODALIDA 2009*, redigerad av Kristiina Jokinen och Eckhart Bick, 73–80. Odense: NEALT.
- Johannessen, Janne Bondi och Joseph C. Salmons 2015. The study of Germanic heritage languages in the Americas. I *Germanic Heritage Languages in North America. Acquisition, Attrition and Change*, redigerad av Janne Bondi Johannessen och Joseph C. Salmons, 1–17. Amsterdam: John Benjamins.
- Keihl Olsen, Ida. 2019. Subjektplassering i moderne norsk. En korpusbasert analyse av leddstillingsvariasjon i bokmål det siste hundreåret. Masteroppgave, Universitetet i Oslo.
- Kupisch, Tanja. 2014. Adjective placement in simultaneous bilinguals (German–Italian) and the concept of crosslinguistic overcorrection. *Bilingualism: Language and Cognition*, 17 (1), 222–233.
- Larsson, Ida and Björn Lundquist. 2014. Objektsplacering vid partikelverb i norske dialekter och äldre svenska. I *Språk i Norge og nabolanda*, redigerad av Janne Bondi Johannessen and Kristin Hagen, 99–131. Oslo: Novus.
- Larsson, Ida och Janne Bondi Johannessen. 2015. Incomplete acquisition and verb placement in heritage Scandinavian. I *Moribund Germanic Heritage Languages in North America: Theoretical Perspectives and Empirical Findings*, redigerad av Richard S. Page och Michael T. Putnam, 153–189. Leiden: Brill Academic Publishers.
- Lindstad, Arne Martinus, Anders Nøklestad, Janne Bondi Johannessen och Øystein A. Vangsnes. 2009. The Nordic Dialect Database: Mapping Microsyntactic Variation in the Scandinavian Language

- ges. I *Proceedings of the 17th Nordic Conference of Computational Linguistics NODALIDA 2009. NEALT Proceedings Series*, volym 4, redigerad av Kristiina Jokinen och Eckhard Bick, 283–286.
- Lundquist, Björn, Ida Larsson, Maud Westendorp, Eirik Tengesdal, och Anders Nøklestad. 2019. Nordic Word Order Database: Motivations, Methods, Material and Infrastructure. *Nordic Atlas of Language Structures (NALS)* 4(1), 1–33. <https://doi.org/10.5617/nals.7529>
- Lykke, Alexander K. 2020. *Variation and change in the tense morphology of heritage Norwegian in North America*. Doktoravhandling, Universitetet i Oslo.
- Montrul, Silvina. 2016. *The Acquisition of Heritage Languages*. Cambridge: Cambridge University Press.
- Polinsky, Maria. 2018. *Heritage Languages and their Speakers*. Cambridge: Cambridge University Press.
- R Core Team. 2020. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing. Vienna, Østerrike.
- Riksem, Brita Ramsevik. 2018. *Language mixing in American Norwegian noun phrases. An exoskeletal analysis of synchronic and diachronic patterns*. Doktoravhandling, NTNU.
- Rothman, Jason. 2009. Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism* 13 (2), 155–163.
- Svenonius, Peter. 2002. Subject positions and the placement of adverbials. I *Subjects, Expletives and the EPP*, redigerad av Peter Svenonius, 201–242. Oxford: Oxford University Press.
- Walkden, George. 2017. The actuation problem. I *The Cambridge Handbook of Historical Syntax*, redigerad av Adam Ledgeway och Ian Roberts, 403–424. Cambridge: Cambridge University Press.
- Westendorp, Maud och Björn Lundquist. Kommande. Tverrspråklig innflytelse fra L1 i tilegnelsen av argumentplassering i L2 norsk og svensk. *Norsk Lingvistisk Tidsskrift*.

- Westergaard, Marit. 2011. Subject positions and information structure: The effect of frequency on acquisition and change. *Studia Linguistica* 3, 299–332.
- Westergaard, Marit och Terje Lohndal. 2019. Verb second word order in Norwegian heritage language: Syntax and pragmatics. In *Variable properties in language: Their nature and acquisition*, redigerad av David. W. Lightfoot och Jonathan Havenhill, 91–102. Washington, D.C.: Georgetown University Press.

English summary

This paper investigates the placement of subjects, objects and verbal particles in the CANS corpus, representing earlier and present-day American Norwegian, as compared to the LIA corpus, representing Norwegian spoken in Norway around the time of mass emigration (late 19th and early 20th century). There is considerable variation in all of the three subcorpora studied; more so than in present-day, spoken Norwegian. This variation is relatively stable across generations in America, in particular with regard to the placement of verbal particles; however, we do note (similar to Anderssen och Westergaard 2020) a decline in subject shift and object shift. For subject shift in present-day AmNo, cross-linguistic influence from English is relevant, but only in questions. For declaratives in present-day AmNo, there is a significant rise in subject-initial clauses. On our account, this bleeds the shifted subject position, but in principle leaves the unshifted position unaffected.

Ida Larsson
 Institutt for språk, litteratur og kultur
 Høgskolen i Østfold
 ida.larsson@hiof.no

Kari Kinn
 Institutt for lingvistiske,
 litterære og estetiske studier
 Universitetet i Bergen
 kari.kinn@uib.no



Adjektivkongruens i amerikanorsk

Brita Ramsevik Riksem, Terje Lohndal og Tor A. Åfarli

I arvespråket amerikanorsk finst det mykje språkblanding. Denne artikkelen ser nærare på attributive engelske adjektiv i nominalfrasar som har norsk struktur og norsk substantiv, og vi undersøker om desse engelske adjektiva har norsk bøyning i eit nyare korpusmateriale (CANS). Resultata syner at dei fleste engelske adjektiva står saman med eit engelsk substantiv, og det er altså ein kontekst der ein verken ventar eller finn norsk bøyning. Resultata viser også at i dei få tilfella der eit engelsk attributiv adjektiv står saman med eit norsk substantiv, er det heller ikkje norsk bøyning, noko som er uventa. Vi godtgjer at det ikkje har skjedd ei tydeleg endring frå tidlegare generasjonar av amerikanorsktalarar, dokumentert i Haugen (1953) og Hjelde (1992). Artikkelen diskuterer kvifor adjektiva oppfører seg annleis enn substantiv og verb og foreslår ein formell analyse av fråveret av norsk bøyning på engelske attributive adjektiv i dei relevante kontekstane i det nyare CANS-materialet.

Stikkord: adjektiv; amerikanorsk; bøyning; språkblanding

1 Innleiing¹

Amerikanorsk er språket til nordmenn som utvandra til Amerika i løpet av 1800-talet og tidleg på 1900-talet og etterkomarane deira, og i dag er det framleis nokre som har amerikanorsk som sitt førstespråk. Eit slikt språk blir ofte kategorisert som eit arvespråk (eller nedarvingsspråk), nemleg eit språk talaren tileignar seg som barn, og som ofte blir brukt i heimen, men som ikkje er det dominante språket i samfunnet elles (sjå t.d. Rothman 2009 eller Johannessen og Salmons 2012, som diskuterer omgrepet på norsk). I tilfellet med amerikanorsk er det engelsk som er det dominerande språket i samfunnet, og for talarane av amerikanorsk i dag er engelsk også deira dominante språk.

Amerikanorsk er eit *norsk* språk, men den massive kontakten med engelsk har sett sitt preg på språket mellom anna ved at talarane ofte nyttar engelske ord medan dei snakkar norsk. Dette er det vi refererer til som språkblanding, altså bruken av leksikalske og grammatiske element frå ulike språk i same ytring (sjå t.d. Muysken 2000). Språkblanding av denne sorten er ikkje eit nytt fenomen i amerikanorsk,² og det mest utbreidde mønsteret for språkblanding er at dei engelske orda blir integrerte i den norske strukturen og opptrer med norske suffiks, funksjonsord og i ei norsk ordstilling. Eit døme på dette er gitt i (1) kor eit engelsk verb og substantiv opptrer med norske bøyingsuffixs i ei elles norsk setning.

(1) Jeg **teach**-a første **grad[e]**-en (coon_valley_WI_07gk)³

¹ Delar av denne artikkelen blei skrivne som ein del av prosjektet *MultiGender* ved Senter for grunnforskning (CAS) i 2019/2020. Takk til to fagfeller og redaktørane for gode og konstruktive innspel.

² Faktisk vart språkblanding i amerikanorsk kommentert allereie på 1800-talet: «They do not bother about keeping the two languages separate [...]. Instead, they eliminate one word after the other from their Norwegian and substitute English words in such a way that the Norwegian will soon be completely forgotten» (Ole Munch Ræder i 1847, sitert i Haugen 1953: 13).

³ Alle døma, der ikkje anna er nemnt, er henta frå CANS-korpuset. Informantkoden syner syner stad, stat, informantnummer og om det er kvinne ('k') eller mann ('m').

Dette mønsteret er registrert i fleire undersøkingar av amerikanorsk heilt frå rundt 1900 og fram til i dag (Flaten 1900–1904, Haugen 1953, Hjelde 1992, Grimstad, Lohndal og Åfarli 2014, Åfarli 2015, Grimstad 2018, Grimstad, Riksem, Lohndal og Åfarli 2018, Riksem 2017, 2018a, Riksem, Grimstad, Lohndal og Åfarli 2019). Haugen (1953: 440) skildrar blandingsmønsteret slik: «A single form is usually imported and is then given whatever endings the language requires to make it feel like a proper word and to express the categories which this particular language requires its words to express». Desse tidlegare studiane av språkblanding i amerikanorsk har hovudsakleg tatt føre seg verb og substantiv, men gitt skildringa til Haugen og nyare forskning skulle ein vente å finne det same mønsteret i tilfelle med engelske adjektiv i amerikanorsk. Sidan adjektiv kongruerer med andre element i nominalfrasen i norsk, kunne vi vente engelske adjektiv med norsk bøyning. Denne forventninga undersøker vi i denne artikkelen. Eit døme er gitt i (2).

(2) Det er ikke noen **smalle** farmer (westby_WI_06gm)

I framhaldet fokuserer vi på attributive adjektiv, mellom anna fordi det er i den konteksten vi skulle vente slik bøyning. I predikativ kontekst er avstanden større, og sidan arvespråksbrukarar ofte har større problem med slike kontekstar (jf. mellom anna Polinsky 2018), ser vi ikkje primært på det her. Vi undersøker adjektivbøyning på engelske adjektiv i den norske delen av det amerikanordiske korpuset *CANS – Corpus of American Nordic Speech* (Johannessen 2015). I del 2 presenterer vi tidlegare undersøkingar av engelske adjektiv i amerikanorsk og dessutan av nominalkongruens i amerikanorsk i dag, og i del 3 presenterer vi forskningsspørsmåla våre. I del 4 går vi gjennom metoden for undersøkingane før vi legg fram resultatata i del 5 og analyserer desse i del 6. Del 7 er konklusjonen.

2 Bakgrunn

I korte trekk blir attributive adjektiv i norsk bøygde i samsvar med substantivet det står til, og vi deler formene i svak og sterk bøyning avhengig av om substantivet står i høvesvis bunden eller ubunden form. Den svake bøyinga har berre éi form (*-e* i dei fleste varietetar), medan den sterke bøyinga har ulike former avhengig av genus og numerus. Det typiske mønsteret for sterk bøyning er inga ending ved maskulint eller feminint genus, *-t* ved nøytrum og *-e* ved fleirtal, men det finst også bøyingsklasser utan genusskilnadar eller med komplett synkretisme mellom formene, t.d. *stille* og *rosa* (Faarlund, Lie og Vannebo 1997).

Når det gjeld det amerikanorske samfunnet og språket, gjorde Einar Haugen ei omfattande undersøking av dette på 1930- og 40-talet (Haugen 1953). Han registrerer mellom anna bruken av engelske ord i amerikanorsk og kommenterer at få av dei engelske adjektiva får norsk bøyning. Samanlikna med substantiv og verb ser adjektiv ut til å vere mindre påverka av norske reglar, og mange av adjektiva får ikkje bøyning i det heile. Enkelte adjektiv kunne likevel få *e*-ending: *next*, *busy*, *easy*, *smart*, medan andre vart registrert med nøytrum-sendinga *-t*: *common*, *ripe* (Haugen 1953: 454). Vidare kunne ei handfull adjektiv få både *-e* og *-t* (Haugen 1953: 454):

- | | | |
|-----|----------|---|
| (3) | a. nice | > <i>nais</i> , <i>naist</i> , <i>nai'se</i> ⁴ |
| | b. rough | > <i>råff</i> , <i>råft</i> , <i>råf'fe</i> |
| | c. tough | > <i>tåff</i> , <i>tåft</i> , <i>tåf'fe</i> |
| | d. cheap | > <i>kjipp</i> , <i>kjipt</i> , <i>kjip'pe</i> |

I Haugen sitt materiale finst det altså døme, som i (3), der engelske adjektiv får norsk bøyning til liks med det ein typisk finn for engelske substantiv og verb, men det overordna mønsteret er likevel at engelske adjektiv ikkje får ei slik bøyning.

⁴ Merk at Haugen brukar ein lyd nær transkripsjon av sine døme. Vi har ikkje endra på den her.

På 1980-talet gjorde Arnstein Hjelde nye undersøkingar av amerikansk basert på 28 informantar, og då særleg talarar som hadde bakgrunn frå Inn-Trøndelag (Hjelde 1992). Ifølgje Hjelde har inntrøndsk generelt lite bøyning på adjektiv; berre nøytrumsforma *-t* blir brukt systematisk i sterk bøyning. Funna til Hjelde viser at det ikkje er uvanleg at nøytrumsendinga blir utelaten i amerikansk – også når det er snakk om norske adjektiv: *kvar år, rau hår* (Hjelde 1992: 83). På same måte som hos Haugen har også få av dei engelske adjektiva i hans materiale bøyning, men det finst nokre døme på engelske adjektiv som får bøyning: *rough, specialt*. Med andre ord får Hjelde same resultat som Haugen: Generelt er det lite bøyning på engelske adjektiv når dei blir tekne inn i amerikansk, men der er likevel nokre få døme kor den norske bøyninga blir brukt.⁵

Eit nærliggande spørsmål er vidare om det kan vere adjektivkongruensen i seg sjølv som er i ferd med å forsvinne i amerikansk. I ein nyare studie av nominalkongruens i amerikansk finn Johannessen og Larsson (2015) tvert imot at denne i hovudsak er stabil. Johannessen og Larsson (2015) ser rett nok ikkje på engelske adjektiv, men viser at kongruensen generelt er stabil når det gjeld attributive adjektiv i amerikansk. I undersøkingar av 34 talarar finn dei at 88% av dei relevante tilfella har den venta bøyninga. Blant dei resterande tilfella med avvik er dette typisk registrert på determinativet og ikkje på adjektivet. På individnivå finn dei også ein klar individuell variasjon mellom talarane. Majoriteten, dvs. 20 av 34 talarar, produserer utelukkande dei venta formene for nominalkongruens i norsk, og dei talarane som produserer feil former, har også ein del rette former. Ingen talarar ser dermed ut til å ha eit kongruenssystem som har kollapsa (Johannessen og Larsson 2015). Ut frå dette må vi konkludere med at kongruensen er stabil når det gjeld norske adjektiv i amerikansk. Det kan med andre ord ikkje vere sjølve bøyningssystemet som

⁵ Informantane i Haugen (1953) har dialektbakgrunnar frå heile Noreg, men dei fleste har bakgrunn frå Aust-Noreg. Likskapane mellom resultatata i Haugen (1953) og Hjelde (1992) kan såleis ikkje forklarast med samsvar i dialektbakgrunn frå Trøndelag. På same måte som Haugen (1953), har informantane i CANS, som vi undersøker vidare i artikkelen, dialektbakgrunnar frå heile landet, men hovudsakleg frå Aust-Noreg.

smuldrar opp ettersom ein då skulle vente å finne avvik også ved bruken av norske adjektiv.

3 Forskingsspørsmål

Forskinga på amerikanorsk viser altså at den generelle regelen for bruk av engelske verb og substantiv er at dei får norsk bøyning, men som både Haugen (1953) og Hjelde (1992) viser, ser ikkje dette ut til å gjelde dei engelske adjektiva ettersom berre eit fåtal av desse får bøyning. Samtidig viser Johannessen og Larsson (2015) at adjektivkongruensen i seg sjølv ser i hovudsak ut til å vere stabil. Vi stiller dermed følgjande forskingsspørsmål:

- Kva kongruens syner attributive engelske adjektiv (numerus, genus) i CANS-materialet?
- Kva for skilnader er det mellom data i Haugen (1953), Hjelde (1992) og CANS?
- Korleis kan vi forklare eventuelle endringar over tid?

4 Metode

I undersøkinga vår har vi nytta versjon 2 av CANS. Korpuset inneheld transkriberte opptak av 163 talarar med norsk som arvespråk som samla har produsert ca. 660 000 ord. Korpuset består av samtalar mellom arvespråksbrukarar av norsk, og desse samtalanene er ortografisk og semi-fonetisk transkriberte.

Vi har søkt i CANS for å identifisere engelske adjektiv i det nye norske materialet.⁶ Der har vi nytta søkefunksjonane som er lista opp i (4), der ‘X’ tyder eit element som ikkje finst i norske ordbøker.

⁶ Vi har også sett på dei transkriberte opptaka frå før krigen, men der finst det omtrent ikkje engelske adjektiv føre norske substantiv. Eit av særst få døme er gitt i (i).

(i) så var de der i # **twenty** år tror jeg ... (coon_valley_WI_49uk)
 Dette stadfestar generaliseringa i Haugen (1953).

- (4) a. adjektiv + X
 b. det + adj_X + N
 c. engelsk segment
 d. X med bruk av statistikkfunksjonen for å identifisere aktuelle adjektiv

(4a) identifiserer alle element som både er kategorisert som adjektiv og som element som ikkje finst i norske ordbøker, altså typisk engelske adjektiv. I (4b) får ein fram former med artikkel følgd av eit engelsk adjektiv og eit norsk substantiv. Kategorien i (4c) er ein eigen kategori i korpuset som får fram segment som berre er på engelsk. Til sist er (4d) ein funksjon som får opp ei liste over dei mest frekvente engelske orda i korpuset, der ein må identifisere ulike ordklasser manuelt.

Det var naudsynt å sortere funna manuelt. Vi lytta også på alle døma for å dobbeltsjekke transkripsjonen og høyre på uttalen. Ei utfordring i arbeidet er sjølve kategoriseringa av adjektiv i CANS. Til dømes fekk vi 2 014 treff på ‘adjektiv+X’, men berre 12 av desse synt seg å vere relevante, altså ha engelsk adjektiv med norsk substantiv etter. Dette skuldast den taggaren som er brukt på korpusmaterialet. Resultata og generaliseringane våre i del 5 er baserte på dei resultata vi fekk fram gjennom søkekriteria i (4). Vi utelukkar ikkje at andre søkemetodar vil kunne finne ytterlegare døme.

5 Resultat

Blant dei 2 014 treffa er det sær mange førekomstar av engelske adjektiv saman med engelske substantiv. To døme er gitt i (5).

- (5) a. jeg trur en som er smart kunne farme på en **small farm**
 (coon_valley_WI_14gm)
 b. ja det var jeg fikk en **golden parachute**
 (albert_lea_MN_01gk)

Meir generelt syner det seg at dei fleste engelske adjektiva står saman med eit engelsk substantiv. I slike høve vil vi ikkje vente norsk bøy-ing. Vi ser difor ikkje meir på desse, men tar føre oss dei døma der vi har engelske adjektiv med norske substantiv. (6) syner alle døma vi fann i korpuset (eit par førekjem fleire gonger).

- (6) a. en **young** gutt (portland_ND_01gm)
 b. en **small** by (sunburg_MN_11gk)
 c. det var ## **cheap** land her (glasgow_MT_01gm)
 d. jeg var e dette gam- dette # age # the **next** # **next** e neste år (sacred_heart_MN_01gk)
 e. my **eldest** søster (vancouver_WA_01gm)
 f. my **eldest** bror og søster (vancouver_WA_01gm)
 g. and # so # the **eldest** # datter # og **youngest** datter (vancouver_WA_01gm)
 h. så da s- # kjøpte jeg # **electric** pumpe som jeg satte inn i basementet da veit du (coon_valley_WI_20gm)
 i. han hadde **hard** time det trur jeg (coon_valley_WI_02gm)

For fleire av døma i (6) ville vi ikkje vente synleg bøyingsuffixs, men i (c) kunne vi etter norske reglar ha venta suffikset *-t* (nøytrum), og i (e)-(g) ville vi etter norske reglar ha venta svak bøyning på adjektiva. Ein kan også stille spørsmål ved kor engelske nokre av adjektiva er, som *youngest* og kan hende *electric*. Begge formene er uttala på ein tydeleg norsk måte. Her kan det anten vere tale om norske former som har fått sett inn ein ekstra vokal (*youngst* > *youngest*), eller blandingsformer mellom norsk og engelsk (*youngst* vs. *youngest*). Eit anna poeng er at det engelske adjektivet *eldest* allereie er bøygd på engelsk vis, og det er uklart om ein kan vente ei norsk bøyingsending utanpå denne. Vi har tatt med (6i) også: Her er adjektivet uttala på tydeleg engelsk vis medan *time* er uttala på norsk måte. Likevel syner konteksten at *time* ikkje tyder det same som det norske *time*. I alle høve ville vi uansett ikkje vente synleg bøyning på adjektivet i denne konteksten.

Nærare gransking syner at det finst svært få døme på bøyning av attributive adjektiv. Vi har berre vore i stand til å finne to i heile kor-

puset, nemleg dei i (7).⁷ Her er substantivet til og med engelsk, men med norsk uttale og bøyning i (7a, b).

- (7) a. Det er ikke noen **smalle** farmer (westby_WI_06gm)
 b. det er nå **smalle** farmer bønn # digre farmene går og renter de (westby_WI_06gm)

Vi tar likevel eit atterhald om at vi ikkje har greidd å identifisere alle førekomstane i korpuset på grunn av tagginga i korpuset, jf. diskusjonen i del 4.

Oppsummert kan vi seie at dei fleste engelske attributive adjektiva står saman med eit engelsk substantiv. Det er for få døme på kongruens med eit norsk substantiv til at vi kan seie noko om korleis ein slik kongruens eventuelt fungerer. Likevel ser tendensen ut til å vere ganske lik det Haugen (1953) og Hjelde (1992) fann: i) Engelske adjektiv opptre sjeldan saman med norske substantiv, og ii) når engelske adjektiv opptre saman med norske substantiv, har dei nesten aldri bøyning. Dette peikar tydeleg på ein skilnad mellom adjektiv på den eine sida og verb og substantiv på den andre. Spørsmålet blir korleis vi skal forklare det, noko vi ser på i neste del.

6 Analyse

I denne delen vil vi gi ein analyse av funna i del 5. Vi ser først på grammatiske endringar generelt i amerikanorsk og korleis desse kan analyserast. Deretter ser vi på kvifor adjektiv ser ut til å vere spesielle, før vi gir syntaktiske analysar av relevante data.

⁷ Ein potensiell hypotese kunne ha vore at informantane generaliserer alle adjektiv til bøyingsklassa med komplett synkretisme i norsk og dermed ikkje har synlege bøyingsendringar. Det er likevel usikkert om desse talarane har hatt tilstrekkeleg med eksponering for dei ulike bøyingsklassene til å gjere ei slik generalisering. Om dei gjorde ei slik generalisering, ville ein vente at dei heldt seg til det mest typiske mønsteret (som ein default). Slik er det for genus, der fleirtalet av dei engelske orda får maskulint genus. Parallelt skulle ein vente at generaliseringa ved adjektiv skulle gå mot den typiske sterke bøyninga.

6.1 Litt om grammatiske endringar i amerikanorsk

Vi er i ein relativt god situasjon når det gjeld diakrone data for amerikanorsk. Kjeldegrunnlaget går attende til ordlistene i Flaten (1900–04) og Flom (1900–04, 1903, 1926). Seip og Selmer (1931) gjorde også feltarbeid i 1931, men det var først med Einar Haugen sitt feltarbeid på 1930- og 1940-talet og dei to omfattande voluma Haugen (1953) at amerikanorsk verkeleg blei grundig dokumentert. Haugen hadde 260 informantar som grunnlag for arbeidet sitt. Etter Haugen blei det ikkje gjort mykje før Arnstein Hjelde sitt feltarbeid på 1980-talet, publisert som Hjelde (1992). Og frå 2010 har vi altså CANS.

Tidlegare forskning har synt at det er endringar i språkbruken frå Haugen til CANS. Riksem (2017, 2018a) finn tre hovudendringar i nominalfrasane (basert på 50 talarar i CANS). Desse er gitt i (8).

- (8) a. Auka bruk av engelsk fleirtals *-s*
 b. Utelating av funksjonelle suffiks, både i fleirtal og i bunden form
 c. Bruk av det engelske determinativet *the*

Endringane i (8) er deskriptive skildringar. Eit viktig analytisk spørsmål er kva som faktisk endrar seg (Riksem 2017, 2018a): Er det berre morfologien (eksponentane), eller skjer det endringar i dei underliggjande syntaktiske strukturane. Ei tilleggsutfordring handlar om dei dataa som finst tilgjengelege. Haugen (1953) har data frå 1.-4.-generasjon i ein situasjon der norsk blei brukt i samfunnet. CANS har derimot data frå 2.-5.-generasjon der norsk blir brukt mykje mindre enn tidlegare. Vi manglar påliteleg informasjon om input til informantane til både Haugen og CANS, og også om barndommen meir generelt og korleis den tidlege bruken av norsk var. Dette gjer at vi vanskeleg kan sjå på endringar i eit einskild individ. Dei endringane vi ser, er altså endringar på gruppenivå.

Vi ser no nærare på det analytiske spørsmålet som vi nemnde ovanfor, sidan det i prinsippet er aktuelt for analysen vår av korleis engelske adjektiv blir integrerte i elles norske nominalfrasar. Spørsmålet er altså kva som endrar seg: overflatemorfologien (eksponentane) eller dei underliggjande strukturane.

Forskning innanfor andrespråksfeltet har foreslått det som blir kalla *Missing Surface Inflection Hypothesis* (MSIH, Lardiere 2000, Prévost og White 2000, sjå også Slabakova 2016 og Putnam, Perez-Cortez og Sánchez 2019). Denne hypotesen går ut på at bøyingsmorfologien er spesielt utsett medan den syntaktiske strukturen ikkje endrar seg. Riksem (2017, 2018a) appliserer dette på amerikanorsk og lanserer to moglege strategiar for utelating av bøyingsmorfologi: i) utelate funksjonelle suffiks generelt, eller ii) setje inn ein engelsk eksponent, t.d. suffikset *-s* for fleirtal, i staden for norsk bøying. Eit problem med MSIH er likevel at han ikkje gir klare prediksjonar for kva tid ein kan vente at morfologien vil forsvinne.

Det andre alternativet går ut på at det er sjølve den syntaktiske strukturen som endrar seg. Dette alternativet er knytt til anna forskning som syner at det kan vere strukturelle skilnader i arvespråk samanlikna med ‘baseline’ (f.eks. Scontras, Fuchs og Polinsky 2015, Yager et al. 2015, Polinsky 2016, 2018). Meir formelt tyder dette at syntaktiske trekk eller trekkmatriser blir endra, svekka, eller rett og slett fjerna frå strukturen. Slike endringar kan gjerne vere gradvise og blir stundom skildra som ‘reanalyse’ av grammatikarar (f.eks. Polinsky 2011, Putnam og Sánchez 2013). For amerikanorsk er til dømes nominalfrasar med engelsk determinativ og norsk substantiv (som *the by*) noko ein kunne vente dersom syntaktiske trekk blir endra eller forsvinn heilt.

Som Riksem (2017) syner, er det ikkje enkelt å avgjere kva for ein av desse analysane som er best. I analysen vår nedanfor av den manglande adjektivbøyinga i amerikanorsk skal vi argumentere for at forklaringa er å finne i den syntaktiske kompleksiteten som er involvert ved adjektivkongruens, og at dette gjer at nettopp adjektivbøyinga er særleg sårbar i eit arvespråk som amerikanorsk. Det er altså ikkje primært den underliggande syntaktiske strukturen som har endra seg, men realiseringa av overflatemorfologien er blitt svekka som resultat av underliggande syntaktisk kompleksitet.

6.2 Syntaktiske analysar

Isolert sett er det nærliggande å meine at det ikkje burde vere så merkeleg at adjektiva manglar kongruensbøying i det nyare CANS-ma-

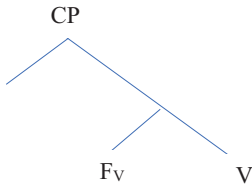
terialet. Ein kunne tenkje seg at det kjem av transfer eller interferens frå engelsk, sidan engelsk heilt manglar bøyning av adjektiv (når vi ser bort frå gradbøyning). I ein større samanheng er det likevel uventa at engelske adjektiv som blir blanda inn i amerikanorsk og står saman med eit norsk substantiv, nesten heilt manglar kongruensbøyning, sidan adjektiv i norsk som hovudregel blir bøygd, og hovudregelen for andre ordklassar med bøyning er at dei engelske orda som blir blanda inn, får norsk bøyning. Her er det altså ein ulikskap når ein samanliknar med substantiv og verb som blir blanda inn frå engelsk. Desse får som regel norsk bøyning. Kvifor skulle nettopp adjektiva vere meir 'sårbare' på den måten at dei ikkje får norsk bøyning?

Vi skal no prøve å forklare kvifor adjektiva er spesielle. Det første vi vil peike på, er at adjektiv får kongruensbøyinga si 'indirekte' ved at dei rettar seg etter bøyinga til andre innhaldsord. Kongruensbøyinga til adjektivet har såleis ikkje opphav i adjektivet eller adjektivfrasen sjølv. Nærare bestemt blir adjektiva bøygd via kongruens med eit anna nominalt element i den aktuelle strukturelle konfigurasjonen, der det nominale elementet anten er eit substantiv eller ein nominalfrase. Dette står i motsetnad til verb og substantiv, som får bøyninga si 'direkte', dvs. frå grammatiske eller funksjonelle trekk som er spesialtileigna for nettopp verbet eller substantivet. Innanfor generativ grammatikk blir dette vanlegvis analysert ved at bøyinga til verb og substantiv kjem frå eit grammatisk funksjonselement i den 'utvida projeksjonen' (*extended projection*, jf. Grimshaw 1991, 2005) til høvesvis verbstammen og substantivstammen.⁸

Når det gjeld verb, inneheld dette grammatiske funksjonselementet eit tempustrekk, sjå (9) der $F_V=[\text{tempus}]$.

⁸ Legg likevel merke til at substantiv som blir brukte predikativt, kan få 'indirekte' bøyning frå predikasjonssubjektet.

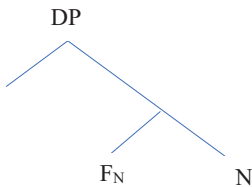
(9)



V får tempus via flytting til F_V (eller via sonde – mål-mekanismen; *probe – goal*, sjå t.d. Chomsky 2000: 122–124, Grimstad et al. 2018), dvs. ‘direkte’ frå det spesialtileigna bøyingstrekket i den utvida projeksjonen sin.

Når det gjeld substantiv, inneheld det grammatiske funksjonselementet F_N (for norsk) trekk for genus, numerus og definit. Dette er skissert i (10).

(10)



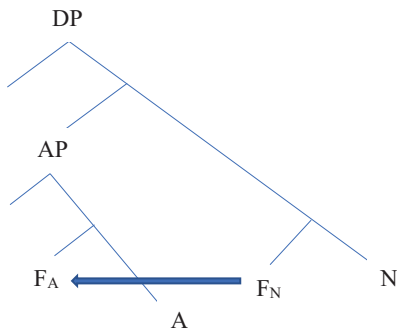
N får genus-, numerus- og definitbøying via flytting til F_N (eller via sonde – mål-mekanismen), dvs. ‘direkte’ frå det spesialtileigna bøyingstrekket i den utvida projeksjonen sin.⁹

Dersom vi no ser på attributive adjektiv, er dei inkluderte i nominalfrasen, dvs. DP, men dei er ikkje del av den utvida projeksjonen til N i DPen. Attributive adjektiv er derimot fakultative modifikatorar internt i DP (sjå t.d. Julien 2005), og A (= adjektivet) har sin eigen utvida projeksjon av A internt i AP (= adjektivfrasen). Og som vi sa

⁹ Det kan stillast spørsmål ved om genus skal sidestillast med numerus og definit når det gjeld tilordning av bøyingstrekk. Argumentasjon for at dette er ein rimeleg analyse, finst i Nygård og Åfarli (2015) og i Riksem (2018b).

ovanfor, får attributive adjektiv bøyingsstrek via ‘indirekte’ kongruens mellom dei relevante bøyingsstrekka til N og (dei uspesifiserte) bøyingsstrekka i den utvida projeksjonen til A. Denne analysen er skissert i (11).

(11)

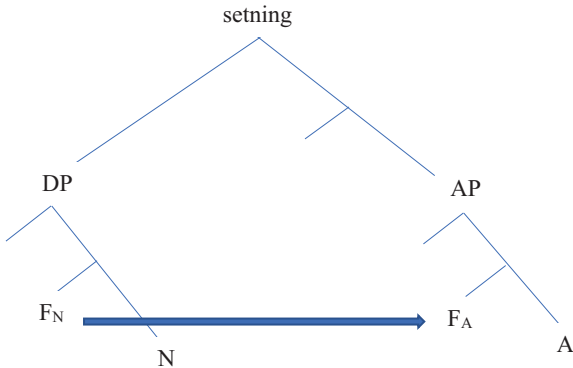


Sjå no på den ‘direkte’ kongruensen med omsyn til høvesvis F_V/V og F_N/N i (9) og (10) samanlikna med den ‘indirekte’ kongruensen i (11). Det er rimeleg å rekne med at ord som verb og substantiv som får bøying via ‘direkte’ kongruens, dvs. via ein direkte relasjon med spesialtileigna funksjonelle trekk, har meir robust bøying enn ord som attributive adjektiv som får bøyingsstrek via ‘indirekte’ kongruens, dvs. via ein indirekte relasjon med andre innhaldsord. Vi meiner at dette er ei plausibel forklaring på kvifor engelske verb og substantiv som blir blanda inn i amerikanorsk, har ein sterk tendens til å få (norsk) bøying, medan engelske adjektiv som blir blanda inn i ein attributiv posisjon i ein (norsk) DP, i langt mindre grad får (norsk) bøying. Den indirekte relasjonen er meir kompleks og dermed meir sårbar, og dette er altså årsaka til at (den norske) kongruensbøyinga på attributive adjektiv blanda inn frå engelsk berre førekjem sporadisk i det eldre amerikanorske materialet og er nesten heilt borte i det nyare CANS-materialet.

Også når adjektivet står i predikativ posisjon, får det bøying via indirekte kongruens, men når det står predikativt, er det heile DPen

adjektivet står til, som bestemmer bøyingstrekka (sjå t.d. Åfarli og Vangsnes 2020). Dette er skissert i (12).

(12)



Sjølv om det er heile DPen som bestemmer bøyingstrekka, har likevel desse trekka sitt opphav i F_N internt i DPen. Både i (11) og (12) finn vi altså ein indirekte kongruensrelasjon mellom F_N i DP og F_A i AP. Vi ventar såleis på basis av analysen vår at bøyinga til predikative adjektiv som er blanda inn frå engelsk, er akkurat like sårbar og fråverande som bøyinga til attributive adjektiv som er blanda inn. Dette har vi ikkje sett på i denne omgangen, men det er ein opplagt prediksjon som bør undersøkast i framtidig arbeid.

7 Konklusjon

CANS gjer det mogleg å utføre store undersøkingar av grammatiske mønster hos arvespråksbrukarar. I denne undersøkinga har vi brukt CANS-materialet til å utforske om amerikanorske arvespråksbrukarar bøyer adjektiv henta frå engelsk på norsk måte slik dei stort sett utstyrer verb og substantiv som er blanda inn frå engelsk, med norsk bøying. I del 3 stilte vi tre forskningsspørsmål:

- Kva kongruens syner attributive engelske adjektiv (numerus, genus) i CANS-materialet?
- Kva for skilnader er det mellom data i Haugen (1953), Hjelde (1992) og CANS?
- Korleis kan vi forklare eventuelle endringar over tid?

For det første fann vi at dei aller fleste engelske attributive adjektiva i materialet står saman med eit engelsk substantiv og derfor ikkje får (norsk) bøyning. Dette er som venta. Det er ganske få døme i materialet der eit engelsk adjektiv står saman med eit norsk substantiv, men i dei døma som finst, kan vi seie at kongruens mellom engelsk adjektiv og norsk substantiv praktisk tala har forsvunne i CANS-materialet. Dette passar godt med mønstra i Haugen (1953) og Hjelde (1992), men utviklinga har gått vidare i CANS: engelske adjektiv blir brukte sær sjeldan med norske substantiv, og i dei få høva det skjer, har adjektiva altså så godt som aldri bøyning. I det eldre materialet finn ein litt bøyning i slike tilfelle.

Forklaringa vår på kvifor engelske adjektiv i praksis manglar norsk bøyning medan verb og substantiv ikkje gjer det, ligg i ein generell/prinsipiell skilnad på korleis kategoriane får bøyning. Verb og substantiv får bøyning via direkte kongruens med eit spesialtileigna funksjonelt element i den utvida projeksjonen til høvesvis verbstammen og substantivstammen. Dette funksjonelle elementet inneheld dei relevante bøyingsegenskapane til verbstammen og substantivstammen. Adjektiv får derimot bøyning via indirekte kongruens med anten eit substantiv (attributive adjektiv) eller ein nominalfrase/DP (predikative adjektiv). Teknisk sett skjer dette ved at dei uvaluerte bøyingstrekka i adjektivets utvida projeksjon blir tilordna bøyingstrekka i det substantivet/nominalfrasen som utløyser kongruensen. Det er altså ein større avstand mellom det som gir og det som får bøyning ved adjektiv enn ved verb og substantiv. Det er denne meir komplekse indirekte kongruensen ved adjektiva som forklarar kvifor bøyinga av engelske adjektiv i amerikanorsk er spesielt sårbar og faktisk praktisk tala er forsvunnen i det nyare CANS-materialet.

Referansar

- Chomsky, Noam. 2000. Minimalist inquiries: The framework. I Martin, Roger, David Michaels & Juan Uriagereka (red.): *Step by step: Essays on Minimalist syntax in honor of Howard Lasnik*, 89–155. Cambridge, Mass.: The MIT Press.
- Faarlund, Jan Terje, Svein Lie & Kjell Ivar Vannebo. 1997. *Norsk referasegrammatikk*. Oslo: Universitetsforlaget.
- Flaten, Nils. 1900–1904. Notes on the American-Norwegian with vocabulary. *Dialect Notes Vol. 2*, 1900–1904, 115–126.
- Flom, George T. 1900–1904. English elements in Norse dialects of Utica, Wisconsin. *Dialect Notes Vol. 2*, 1900–1904, 257–268.
- Flom, George T. 1903. The Gender of English Loan-Nouns in Norse Dialects in America; A Contribution to the Study of the Development of Grammatical Gender. *The Journal of English and Germanic Philology* 5-1, 1–31.
- Flom, George T. 1926. English Loanwords in American Norwegian as Spoken in the Koshkonong Settlement, Wisconsin. *American Speech* 1-10, 541–558
- Grimshaw, Jane. 1991. Extended Projections. Ms., Brandeis University.
- Grimshaw, Jane. 2005. *Words and Structure*. Stanford, CA: CSLI.
- Grimstad, Maren Berg. 2018. English-origin verbs in American Norwegian. A morphosyntactic analysis of mixed verbs. Doktoravhandling, NTNU.
- Grimstad, Maren Berg, Terje Lohndal & Tor A. Åfarli. 2014. Language mixing and exoskeletal theory: A case study of word-internal mixing in American Norwegian. *Nordlyd* 41-2, 213–237.
- Grimstad, Maren Berg, Brita Ramsevick Riksem, Terje Lohndal & Tor A. Åfarli. 2018. Lexicalist vs. exoskeletal approaches to language mixing. *The Linguistic Review* 35-2, 187–218.
- Haugen, Einar. 1953. *The Norwegian Language in America: A Study in Bilingual Behavior* (Vol. I og II). Philadelphia: University of Philadelphia Press.
- Hjelde, Arnstein. 1992. *Trøndsk talemål i Amerika*. Trondheim: Tapir.

- Johannessen, Janne Bondi. 2015. The Corpus of American Norwegian Speech (CANS). I Megyesi, Beáta (red.) *Proceedings of the 20th Nordic Conference of Computational Linguistics (NODALIDA 2015)*, 279–300. NEALT Proceedings Series 23.
- Johannessen, Janne Bondi & Ida Larsson. 2015. Complexity Matters: On Gender Agreement in Heritage Scandinavian. *Frontiers in Psychology* 6. <https://doi.org/10.3389/fpsyg.2015.01842>.
- Johannessen, Janne Bondi & Joseph Salmons (Red.). 2012. *Norsk Lingvistisk Tidsskrift* 30-2, *Norsk i Amerika* (temanummer).
- Julien, Marit. 2005. *Nominal Phrases from a Scandinavian Perspective*. Amsterdam: John Benjamins.
- Lardiere, Donna. 2000. Mapping Features to Forms in Second Language Acquisition. I Archibald, John (red.) *Second Language Acquisition and Linguistic Theory*, 102–129. Malden, MA: Blackwell.
- Muysken, Pieter. 2000. *Bilingual Speech*. Cambridge: Cambridge University Press.
- Nygård, Mari & Tor A. Åfarli. 2015. On the structure of gender assignment. *Indian Linguistics* 76-1–2, 67–76.
- Polinsky, Maria. 2011. Reanalysis in adult heritage language: A case for attrition. *Studies in Second Language Acquisition* 33, 305–328.
- Polinsky, Maria. 2016. Structure vs. use in heritage language. *Linguistics Vanguard* 2-1, 1–14. <https://doi.org/10.1515/lingvan-2015-0036>.
- Polinsky, Maria. 2018. *Heritage Languages and Their Speakers*. Cambridge: Cambridge University Press.
- Prévost, Philippe & Lydia White. 2000. Missing surface inflection or impairment in second language acquisition? Evidence from tense and agreement. *Second language research* 16-2, 103–133.
- Putnam, Michael T. & Liliana Sánchez. 2013. What's so incomplete about incomplete acquisition? A prolegomenon to modeling heritage language grammars. *Linguistic Approaches to Bilingualism* 3-4, 478–508.
- Putnam, Michael T., Silvia Perez-Cortez & Liliana Sánchez. 2019. Differential Access: Asymmetries in Accessing Features and Bu-

- ilding Representations in Heritage Language Grammars. *Languages* 4-4. <https://doi.org/10.3390/languages4040081>.
- Riksem, Brita Ramsevik. 2017. Language Mixing and Diachronic Change: American Norwegian Noun Phrases Then and Now. *Languages* 2-2. <https://doi.org/10.3390/languages2020003>.
- Riksem, Brita Ramsevik. 2018a. Language mixing in American Norwegian noun phrases. An exoskeletal analysis of synchronic and diachronic patterns. Doktoravhandling, NTNU.
- Riksem, Brita Ramsevik. 2018b. Language Mixing in American Norwegian Noun Phrases. *Journal of Language Contact* 11-3, 481–524.
- Riksem, Brita Ramsevik, Maren Berg Grimstad, Terje Lohndal & Tor A. Åfarli. 2019. Language mixing within verbs and nouns in American Norwegian. *The Journal of Comparative Germanic Linguistics* 22-2, 189–209.
- Rothman, Jason. 2009. Understanding the nature and outcomes of early bilingualism: Romance languages as heritage languages. *International Journal of Bilingualism* 13-2, 155–163.
- Scontras Gregory, Zuzanna Fuchs & Maria Polinsky. 2015. Heritage language and linguistic theory. *Frontiers in Psychology* 6. <https://doi.org/10.3389/fpsyg.2015.01545>.
- Seip, Didrik Arup & Ernst W. Selmer. 1931. Voksrull-opptak av norsk-amerikanske informantar. Tilgjengeleg frå <http://www.tekstlab.uio.no/norskiamerika/opptak/seip-selmer.html>.
- Slabakova, Roumyana. 2016. *Second Language Acquisition*. Cambridge: Cambridge University Press.
- Yager, Lisa, Nora Hellmold, Hyoun-A Joo, Michael T. Putnam, Eleonora Rossi, Catherine Stafford & Joseph Salmons. 2015. New Structural Patterns in Moribund Grammar: Case Marking in Heritage German. *Frontiers in Psychology* 6. <https://doi.org/10.3389/fpsyg.2015.01716>.
- Åfarli, Tor A. 2015. Hybrid Verb Forms in American Norwegian and the Analysis of the Syntactic Relation between the Verb and its Tense. I Johannessen, Janne Bondi & Joseph Salmons (red.) *Germanic Heritage Languages in North America: Acquisition, Attrition and Change*, 161–177. Amsterdam: John Benjamins.

Åfarli, Tor A. og Øystein Vangsnes. 2020. Formell og semantisk adjektivkongruens i norsk. I Hagen, Kristin, Arnstein Hjelde, Karine Stjernholm & Øystein Aleksander Vangsnes (red.) *Bauta: Janne Bondi Johannessen in memorian. Oslo Studies in Language* 11-2, 527–540.

English summary

Language mixing is common in the heritage language American Norwegian. In this article, we take a closer look at English attributive adjectives mixed into noun phrases with a Norwegian structure and which contains a Norwegian noun. Our main research question is whether or not these English adjectives show Norwegian inflection in the recent Corpus of American Nordic Speech (CANS). Our results indicate that most of the English adjectives co-occur with an English noun, that is, in a context where Norwegian inflection is not expected and not found. The results also indicate that, contrary to expectation, in the few cases where an English adjective co-occur with a Norwegian noun, there is still no Norwegian inflection on the adjective. We demonstrate that there is no significant change compared to previous generations of American Norwegian, as documented in Haugen (1953) and Hjelde (1992). The article discusses why adjectives behave differently from nouns and verbs and suggests a formal analysis of the absence of Norwegian inflection on English attributive adjectives in the relevant contexts in the newer CANS material.

Brita Ramsevik Riksem
Institutt for lærarutdanning
Noregs teknisk-naturvitskaplege universitet NTNU
brita.riksem@ntnu.no

Terje Lohndal
Institutt for språk og litteratur
Noregs teknisk-naturvitskaplege universitet NTNU
terje.lohndal@ntnu.no
og
UiT Noregs arktiske universitet (professor II)

Tor A. Åfarli
Institutt for språk og litteratur
Noregs teknisk-naturvitskaplege universitet NTNU
tor.aafarli@ntnu.no



«Ein må tenke seg om, ser du veit du, kva ord ein skal bruke»

Pragmatiske partiklar i amerikansk

Åshild Søfteland og Arnstein Hjelde

Hovudfokus i denne artikkelen er på bruken av dei pragmatiske partiklane *ser du*, *veit du* og *trur eg* i CANS – Corpus of American Nordic Speech. Desse blir samanlikna med bruk av engelsk *you see*, *you know*, *I think* og norsk *du veit*, *du ser*, *eg trur* i same datamateriale. Artikkelen gir ei oversikt over frekvens, syntaktisk posisjon og overordna pragmatisk analyse, saman med drøfting av nokon metodologiske problemstillingar. Hovudfunnet er at amerikanskktalarane nyttar moglegheitene som desse pragmatiske partiklane gir som støtte for informasjonsflyten i samtalesituasjonen, og kombinerer både norske og engelske bruksmønster saman til ein større ressurs.

Stikkord: pragmatisk markør; pragmatisk partikkel; amerikansk; korpusdata; kodeveksling; informasjonsflyt; flagging

1 Innleiing

Fleire har peikt på at i arvespråk blir pragmatiske markørar lett henta inn frå majoritetsspråket (Matras 2009), og med det etter kvart nokså omfattande korpuset *CANS – Corpus of American Nordic Speech* (Johannessen 2015), opnar det seg nå også svært gode moglegheiter til å forske på pragmatiske markørar i norsk arvespråk, slik som:

- (1) å såmme tid just *you know* bærre få morro *væit du* (Rushford-01gm)
- (2) næi *du ser* minn fars fammli kåmm ifrå Jevvnaker Hadelann dær (Chicago-01gk)
- (3) hann binnt å farrmen i nittenhundreogatten *I think* å farrme heile live (Portland-01gm)

I denne artikkelen er fokuset på bruk av frasane *ser du*, *veit du* og *trur eg* som pragmatisk partikkel samanlikna med informantanes bruk av engelsk *you see*, *you know*, *I think* og dei syntaktisk tilsvarende norskspråklege variantane *du ser*, *du veit*, *eg trur*. Dei amerikanske dataa vil til ein viss grad bli samanlikna med (moderne) europansk og (amerikansk-)engelsk, men skal primært få stå på eigne bein. Det store fleirtalet av bruksmåtane i CANS er målspråksnære (i norsk eller engelsk), og det mest påfallande med materialet er heller i kor stor grad amerikansk-talarane tar alle moglege ressursar i bruk i kommunikasjonen. Hovudmålet med artikkelen er å gi eit innblikk i datamaterialet spesielt med tanke på overordna bruksmønster, syntaktisk plassering og kva som skjer rundt partiklane, i tillegg til nokon sentrale metodiske problemstillingar.

2 Pragmatisk markør og pragmatisk partikkel

Gjennom prosjektet *NOT – Norwegian Tag Particles* har det dei siste åra blitt forska mykje på pragmatiske partiklar i norsk, primært etterstilte ‘tags’. NOT-prosjektet inneheldt blant anna ei stor dialekt-geografisk kartlegging av 46 slike partiklar, tilgjengeleg i *Småord-databasen* (Borthen et al. 2019). Denne databasen vil vere

ein nyttig referanse i utforskinga av amerikanorsk. Informantgrunnlaget i Småord-databasen er i stor grad overlappende med *Nordisk dialektkorpus* (NDK) (Johannessen et al. 2009), som vi også har som referansepunkt.

I tråd med Fraser (1999) definerer Borthen (2018, 226) *pragmatisk markør* som «et lingvistisk tegn (leksem, bøyingsendelse, ordstillingsmønster, intonasjonsmønster, ...) som brukes til å kommunisere noe som går utover ytringens grunnleggende proposisjonelle innhold». Pragmatiske markører seier altså noko om ein proposisjon og kan bestå av ulike typar lingvistiske teikn, inkludert intonasjon og bøyning, og også både føre- og etterstilte ord og frasar. *Pragmatisk partikkel* (jf. Borthen 2018, 227) er ein undertype av pragmatisk markør. Desse er *leksikalske* pragmatiske markører, små og trykklette, men sjølvstendige. Pragmatiske partiklar har gjerne avbleika semantisk tyding om ein samanliknar med det etymologiske opphavet, slik som *sjø* samanlikna med opphavet *ser du* i bokstaveleg tyding (jf. Søfteland og Borthen 2018). Ein viktig underkategori her er ‘tag particle’, etterstilte partiklar som vedheng eller hale (som i *Æ e trønder æ, sjø*). Ein vanleg bruksmåte for mange av frasane vi skal sjå på i denne artikkelen, er etterhengt (som i *Berre for moro, veit du*), men dei er alle i utgangspunktet fleirordsuttrykk og kanskje for tunge til å kallast partikkel. Borthen (2018, 228) argumenterer for at fleirordsuttrykk likevel kan reknast som pragmatiske partiklar, så sant dei er trykklette og har avbleika tyding, og dette synet vil vi følgje her. Eit viktig aspekt er også at språkleg einingar som dei vi skal sjå på her, er ikkje-obligatoriske – noko som også kjem fram gjennom intonasjonsmønster.¹

¹ Blant anna skriv Maschler & Schiffrin (2015, 197) at prototypiske diskursmarkører står i «intonation-unit [...] initial position», enten ved byte av talar eller, ved same talar: «immediately following any intonation contour other than continuing intonation [unless it follows another marker in a cluster]».

3 Pragmatiske partiklar i arvespråk og kodeveksling

Pragmatiske markørar blir ofte drøfta i samband med kodeveksling og arvespråk. Matras (2009, 137) seier at «[t]here is widespread agreement in the literature on bilingual speech concerning the high frequency of L2-insertion of discourse markers» og Polinsky (2018, 293) legg til grunn at «heritage speakers follow standard bilingual practices in code-switching between their two languages». For Poplack (t.d. 1980), som prøvde å finne fram til syntaktiske reglar og restriksjonar for kodeveksling, var ‘tags’ ein viktig kategori i teoribygginga. Desse omfatta, men avgrensa seg ikkje til, pragmatiske markørar. Sentrale funn frå denne tidlege forskinga er at det å kodeveksle tags er spesielt vanleg hos meir usikre tospråklege, mens balanserte tospråklege i tillegg vekslar meir setningsinternt. Teorien til Poplack er at kodeveksling primært skjer der det er grammatisk overlapp mellom språka, slik at plasseringa er grammatisk i begge språk, noko som òg krev høg grammatisk kompetanse i begge språka. Ved tag-veksling krevst det mindre grammatisk kompetanse i begge språk, sidan «(t)ags are freely moveable constituents which may be inserted almost anywhere in the sentence without fear of violating any grammatical rule» (Poplack 1980, 589). Den tidlege forskinga til Poplack analyserte også *flagging* av kodeveksling – ulike typar metalingvistiske kommentarar, omsetting, pausemarkering og liknande rundt vekslinga. Dette kan vere for å ramme inn eller rette merksemda mot språkbyte – eller for å forklare, presisere, få fram «the apt expression», altså *det treffande uttrykket* (jf. Rosignoli 2011, 52). Som vi skal sjå i del 5, viser også CANS-informantane tendensar til å bruke dei aktuelle pragmatiske partiklane til dette.

Dei siste tiåra har det komme ei aukande mengde forskning på ulike grammatiske sider ved arvespråk, med spesielt stor auke dei siste åra for amerikanorsk, men til nå har det i liten grad inkludert pragmatiske partiklar. Eitt unntak frå dette er Moquin og Salmons (2020). Dei analyserer data frå CANS, og samanliknar desse med funn frå Pennsylvania Dutch, data frå ulike norske talespråkskorpus og akseptabilitetstestar frå (europa-)norskspråklege. Det er særleg to einingar som blir studert: *well/vel* og *you know/veit du*. Dei norske

variantane er frekvente, men dei har tatt opp i seg noko av den engelske bruksmåten ('pattern borrowing', Moquin & Salmons 2020, 307): I amerikansk-engelsk er *well* gjerne setningsinitial, mens *vel* i Noreg særleg er medial. Funna frå CANS viser at amerikansk *vel* er frekvent i initial posisjon (44 %), mens den i liten grad er initial i dei europanske korpusa (under 1 %). Tilsvarende endringar for *veit du* finn dei ikkje; her ser plasseringa ut til å vere relativt intakt. Derimot meiner dei å sjå ei pragmatisk-semantisk endring etter engelsk mønster ved at denne partikkelen i større grad markerer «a plea for understanding» samanlikna med den europanske hovudbetydninga «to insist on the truth of a statement» (Moquin & Salmons 2020, 315). Språkkontakten har såleis ført til «new and additional richness in the discourse marking system of American Norwegian» (Moquin & Salmons 2020, 316).

Klintborgs (1999) studie av amerikansk inneheld ein eigen del om pragmatiske markørar i vid forstand, inkludert t.d. *and* og *yes*. Mange av problemstillingane er generelle, men nokre detaljar om bl.a. *well*, *you know* og *you see* er med i analysen. Han viser til tidlegare forskning som vurderer om *you know* pressar ut *vet du*, men fleire av informantane hans brukar begge (Klintborg 1999, 140). Det kan også nemnast at (*you*) *see* er mykje meir frekvent enn *you know* i hans materiale. Klintborg trekker vidare inn at pragmatiske markørar gjerne dukkar opp i klynger (Klintborg 1999, 145). Han peikar òg på at grensa mellom proposisjonell og ikkje-proposisjonell tyding kan vere uklar, t.d. ved ytringsinitial *I mean*, slik at det blir vanskeleg å avgjere om eit uttrykk er pragmatisk markør eller ikkje. (Dette blir drøfta for våre data i del 5.)

Når det gjeld pragmatiske markørar, kan vi òg sjå til forskning på amerikansk: Salmons (1990) analyserer blant anna bruken av engelsk *well* og *you know* i samtaler på tysk mellom tysk-engelske tospråklege. Han argumenterer for at det som kan sjå ut som kodeveksling, heller er lån, og at diskursmarkeringssystema i tysk og engelsk har falle saman i arvespråket til desse språkbrukarane, mellom anna fordi mange vanlege tyske diskursmarkørar er fråverande i datamaterialet. I Salmons' data er *you know* spesielt frekvent

mellom setningar, og han nemner også nøling og sjølvreparasjon som bruksmåtar. Begge delar ser vi i (4) (Salmons 1990, 460):

- (4) ... Ohh *you know*, ich wollte jetzt really dich mal frage ... fragen, *you know*, ich dachte, Sie kämen von Deutschland ... ('... ohh, you know, I really wanted to ask you, you know, I thought you were from Germany ...')

4 Tidlegare studiar av dei aktuelle partiklane/frasane

Vi kjenner ikkje til nokon større arbeid om *ser du*, *veit du* eller *trur eg* i europansk (men sjå Borthen et al. 2021 for eit oversiktsbilete). Likevel er *sjø* som analysert i Søfteland og Borthen (2018) truleg parallell nok med *ser du* til at det kan reknast som høgst relevant. Den overordna analysen der er at ytringar med *sjø* oppfordrar lyttaren til å tolke innhaldet i ytringa som ei støtte til informasjon som er tilgjengeleg i konteksten, og til å tolke ytringa i lys av at han/ho har blitt kjent med dette innhaldet ved eit tidspunkt *t* (Søfteland og Borthen 2018, 278). Dette gjer at ytringar med *sjø* lett kan bli tolka som ei forklaring eller stadfesting. Det gjer også at ytringar med *sjø* kan fungere som eit hint om at informasjonen som blir gitt, er ny for tilhøyraren, eller som eit teikn på at talaren reknar informasjonen som kjent og/eller at samtalepartnarane er einige. Eit utgangspunkt for vidare analyse her kan vere at det same gjeld for toordsuttrykket *ser du* (jf. Søfteland og Borthen 2018, 253).

Sidan vi analyserer pragmatiske partiklar i ein norsk-amerikansk setting, er det også klart relevant å sjå på beskrivingar av engelske bruksmåtar av dei tilsvarande *you see*, *you know* og *I think*. Dette er eit stort forskingsfelt internasjonalt, spesielt når det gjeld *you know*. Eitt av dei første større bidraga her er Östman (1981), og ein av dei overordna analysane hans er at *you know* er ein leksikalsk 'hedge' som gjerne kommenterer på ei enkelt leksikalsk eining, og at talaren med *you know* ønsker å få lyttaren til å akseptere innhaldet i ytringa som presupponert. Ei følge av dette er at talaren brukar *you know* nettopp når han tenker at lyttaren kanskje *ikkje* 'veit'. Ved å legge til

you know ber talaren eksplisitt om merksemd, med mål om samarbeid for å få tolka ein proposisjon som bakgrunnskunnskap (Östman 1981, 17).

Samarbeid om kva som skal reknast som kjent informasjon i samtala er også sentralt i Schiffrin (1987). Ho analyserer den overordna funksjonen til *you know* som å sikre at lyttaren heng med i relevant informasjonsflyt: «*y'know* marks the speaker as an information provider, but one whose successful fulfillment of that role is contingent upon hearer attention» (Schiffrin 1987, 290). Det kan vere mange ulike årsaker til at talaren treng slik merksemd frå lyttaren, og dermed opptrer *you know* også i mange ulike språklege omgavnader (Schiffrin 1987, 295). Ho framhevar at *you know* kan peike mot både kommande og føregåande informasjon, og forklarar dette med at frasen er ei redusert form av *do you know*-spørsmål, som kan ha rekkevidde både framover og bakover i diskursen (Schiffrin 1987, 287). Maschler og Schiffrin (2015) gir ei drøfting av forskingslitteraturen med vekt på korleis og kvifor *you know* er notorisk vanskeleg å analysere. Dette er blant anna fordi intonasjonen ikkje nødvendigvis følgjer prototypiske mønster for pragmatiske markørar, og fordi *you know* gjerne koplar saman meir abstrakte einingar i diskursen (Maschler og Schiffrin 2015, 204).

Også Erman (1987) er relevant å trekke inn i vår samanheng, sidan ho med utgangspunkt i korpusdata ser på både *you know*, *you see* og *I mean* i detalj, komparativt mot kvarandre og med hovudvekt på plasseringa dei har i ytringar.² Det klart vanlegaste er *medial* posisjon (ca. 80 %) (Erman 1987, 50). Det er *you see* som blir mest bruka finalt av dei tre (15 %) og *I mean* som blir mest bruka initialt (9 %). *You know* blir spesielt ofte bruka inni setningsledd/frasar som ei avventing mens talaren leitar etter ord (Erman 1987, 52). Alle tre uttrykka blir ofte bruka mellom setningar ('connective'). Erman konkluderer om den spesielt mangesidige bruken av *you know* med at uttrykket «could occur between complete propositions as well as wit-

² Merk at Erman analyserer plassering i *ytringar*, gjerne *turar* i ei samtale, og ikkje i *setningar* som syntaktisk eining slik vi gjer i vår analyse. Sjå vidare drøfting av dette i del 5 – og i Maschler og Schiffrin (2015, 203ff.).

hin smaller units of discourse, fulfilling, in fact, completely different functions» (Erman 1987, 75).

Til slutt kan vi nemne Aijmer (1997) som analyserer pragmatikaliseringa av *I think* frå opphavet som [subjekt+hovudverb] i oversetning: *I think that Bill is at home* > *I think Bill is at home* > *Bill is at home I think* (Aijmer 1997, 7). Studien inneheld òg detaljerte analysar av bruken av *I think* med utgangspunkt i komparative undersøkingar av engelsk og svensk. Sidan funksjonen som oversetning og pragmatikalisert bruk sameksisterer i dag, drøftar ho òg korleis ein kan skilje mellom dei, spesielt i språk som svensk og norsk med obligatorisk V2. Vi har to moglegheiter for ordrekkefølge ved førestilte frasar som her: *Ärlig talat, han är inte vidare trevlig* og *Ärlig talat är han inte vidare trevlig*. I førstnemnde variant er *ärlig talat* eit «speech-act adverbial», og vi får ikkje inversjon (Aijmer 1997, 4). I den andre varianten med predikatsadverbial får vi det. Merk at prososodien vil vere ulik i dei to tilfella, noko som er spesielt relevant for analysane våre av toordsuttrykk med initial plassering i del 5. I tillegg framhevar Aijmer posisjon som ein viktig analysefaktor for dei mange ulike bruksmåtene av *I think*. Generelt sett er det her to grunnleggande hovudfunksjonar som er i spel: «indicating facts» og «expressing the state of the speaker» (Aijmer 1997, 10); *I think* kan på den eine sida signalisere objektivitet og sikkerheit (*eg meiner klart*), og på den andre sida estimat og usikkerheit (*eg trur kanskje*) (Aijmer 1997, 21). Det er den sistnemnde funksjonen som primært gjeld ved bruk som pragmatisk partikkel i vår definisjon.

5 Korpusdata: Metodiske og strukturelle problemstillingar

Arbeidet med denne artikkelen starta opphavleg med søk i CANS etter alle dei 46 pragmatiske partiklane i Småord-databasen (Borthen et al. 2019). Her er det mange interessante funn å ta tak i. Til dømes såg vi tilsvarande bruksmønster for *vel* i CANS som Moquin og Salmons (2020) fann i same materiale (jf. del 3): *Vel* er om lag like frekvent i amerikanorsk som i europansk, men i sistnemnde har det primært medial (og final) plassering, mens det i mykje større grad

står initialt i amerikanorsk, tilsvarende bruksmønsteret for *well* i engelsk. For partiklane *sjø* (kjent partikkel frå trøndersk målområde) og *damma* (kjent partikkel frå Mjøs-området) fann vi derimot store kvantitative forskjellar samanlikna med (moderne) europansk: *Sjø* er høgfrekvent i NDK (ca. 400 døme, 2,00/10.000 tokens) og nesten fråverande i CANS (3 døme, 0,043/10.000 tokens), mens *damma* er lågfrekvent i NDK (ca. 20 døme, 0,10/10.000 tokens) og dukkar opp mykje oftare i CANS (ca. 130 døme, 1,85/10.000 tokens). Vidare forskning må vise om dette primært har å gjere med talet på informantar med relevant dialektbakgrunn i dei to korpusa, eller om det kan skuldast meir metodiske forhold: Også varianten *se* kjent frå Østfold er lågfrekvent i NDK (jf. Søfteland og Borthen 2018), så kanskje kan det vere negative haldningar til dialektar i «semi-periferien» til Oslo som spelar inn i europansk (jf. Mæhlum 2009, 16), slik at pragmatisk partiklar som indeksikalisering tilhøyrer til desse områda (*damma*, *se*) i mindre grad blir ytra i opptakssituasjonen enn dei som indeksikalisering tilhøyrer til dialektområde med høgare status (*sjø*). Arvespråkstarar som informantane i CANS vil ein derimot gjerne rekne med har avgrensa eller manglande register (jf. Chevalier 2004), slik at indeksikalisering av enkeltord også kan vere fråverande, og ikkje noko som blir påverka av den elles litt kunstige opptakssituasjonen.

Til hovudstudien vår her valde vi altså ut tre toordsuttrykk (verb+subjekt) som har ein tilsvarende variant i engelsk (subjekt+verb) og også ein norsk variant med same ordrekkefølge som i engelsk (jf. tabell 1):

Norsk VS	Engelsk SV	Norsk SV
ser du	you see	du ser
veit du	you know	du veit
trur eg	I think	(eg trur) ³

Tabell 1. Oversikt over dei undersøkte to-ordsuttrykka

³ Uttrykket *eg trur* (SV) er heilt marginal som pragmatisk partikkel i norsk, men det er nokre døme i både NDK og CANS.

Dette utvalet gjer at vi kan analysere bruken av pragmatiske partiklar i amerikanorsk i både ein- og tospråklege ytringar, på norsk og engelsk. Vidare gjer utvalet det mogleg å samanlikne bruk av same partikkel i ulike posisjonar. Forskjell i ordstilling mellom dei to språka spelar også inn, sidan etterstilt VS, som er vanleg i norsk, er ugrammatisk i engelsk (**Bill is at home, think I*).

Som nemnt vil vi i utgangspunktet analysere det amerikanorske datamaterialet mest mogleg frittstående/sjølvtendig. Derfor gir vi også døme frå CANS i beskrivinga av det vi reknar som protypiske bruksmåtar i norsk og engelsk. Slik vi les tidlegare forskning på desse og tilsvarande pragmatiske partiklar, er VS i final posisjon det mest frekvente i europansk:

- (5) e banne ikjett så mye menn minn bror bottpå dæ enn M1 de e gutt somm kann banne *se u* (Flom-02gm)
- (6) å så mæsste tå rømm jiffa se itte hell førr de va itte dømm kunn tte finne jennte ra *væit du* (CoonValley-06gm)
- (7) gammLeste gutten minn hann æter litt ludefissk hann bærre ti pLise mei *tru jæ* (Portland-01gm)⁴

VS i initial posisjon er sjeldan, men vi skal sjå nokre døme på det frå CANS i del 6 som også må kunne reknast som grammatiske i europansk.

Setningsinitial SV er klart mindre frekvent, men det finst enkelte døme på dette i NDK om ein leitar, og døme (8)–(10) frå CANS må også seiast å vere grammatiske i europansk. (Prosodien viser her nokså klart at toordsuttrykka er brukt som pragmatisk partikkel.)

- (8) *du ser # em # væiane jikk runnt dissa sekksjona* (Stillwater-01gm)
- (9) *du vett mor hadde de fint ho va sjukkepleiåske ho hadde pennger menn allt va dårlekt i Nærrge* (Saskatoon-01gk)

⁴ L representerer retrofleks flapp (r) i CANS-notasjonen, mens R blir bruka om retrofleks kontinuuant (ɹ). # markerer ei kort pause i talestraumen. I presentasjonen av døma her og i det følgjande er ofte nøling, pauseteikn og gjentakning utelate.

- (10) *je tru* far minn hann ville gå baksjatt te Nàrrge (Portland-01gm)

Etterstilt SV kan sjå ut til å vere fråverande i NDK, kanskje fordi det prinsipielt ville medføre V2-brot. (Sjå vidare diskusjon av prosodi og V2-brot i CANS lenger nede.)

I engelsk er det berre rekkefølga SV som er aktuell, og slik vi les den relevante forskingslitteraturen, er både initial og medial plassering av pragmatiske partiklar relativt meir frekvent i engelsk enn i norsk generelt sett (og final tilsvarande mindre frekvent). Døme (11)–(13) viser nokon reint engelskspråklege døme frå CANS med dei aktuelle toordsuttrykka, med høvesvis final, medial og initial plassering (sjå drøfting av *medial* plassering lenger ned).

- (11) he was the one who was supposed to come home and farm *you see* (Hatton-01gm)
 (12) we're second generation here *you know* our folks were all born here (CoonValley-06gm)
 (13) *I think* Kvande that is a family name (NorthBattleford-01gm)

Datainnsamlinga i CANS blei gjort som enkle toordssøk utan annan spesifikasjon, men med både *veit/vet* og *trur/tror* sidan begge former er bruka i den ortografiske transkripsjonen. (Berre *see* aleine er ikkje inkludert). Søka er avgrensa til norsk arvespråk, utan den svensk-språklege delen av korpuset. Dette ga 205 informantar totalt, og 700.837 tokens i CANS v.3.0. Tabell 2 (neste side) viser talet på treff brukt som pragmatisk partikkel og talet på treff totalt (i parentes). For *veit/vet du* og *jeg trur/tror* har vi analysert berre eit utval av treffa (høvesvis 501 og 201).

Her ser vi blant anna at *veit/vet du* er klart mest frekvent (3528 treff totalt), og at svært mange av desse er pragmatisk partikkel (487 av dei 501 som er analysert, 97 %). Den engelske varianten *you know* er også høgfrekvent, og med tilsvarande prosentandel (598 av 619, 97 %). Som venta er det få døme på *jeg trur/tror* som pragmatisk partikkel, men det er nokså mange døme på dette blant *trur/tror jeg*

	<i>Norsk VS</i>	<i>Engelsk SV</i>	<i>Norsk SV</i>
SER DU	59 (66)	5 (8)	13 (37)
VEIT DU	487 av 501 (3528)	598 (619)	133 (196)
TRUR EG	283 (463)	37 (94)	22 av 201 (1010)

Tabell 2: Talet på treff i CANS før og etter sortering etter bruk som pragmatisk partikkel.

(283, 61%). SV-varianten *du veit/vet* er også relativt frekvent som pragmatisk partikkel (133, 69%).

For å kunne analysere pragmatiske markører i detalj må ein sjå på både syntaktisk posisjon, uttale og betydning (jf. t.d. Maschler og Schiffrin 2015, 202). I vår samanheng gjeld dette for det første nett-opp avgjerda om eit treff i korpuset faktisk kan reknast som pragmatisk partikkel – semantisk avbleika, sjølvstendig, som eit tillegg til det proposisjonelle innhaldet som det står til (jf. del 2). I ein del tilfelle kan semantikken i konteksten klargjere at det er den konkrete, og ikkje pragmatiske/avbleika, betydninga som er i bruk (her: *å sjå*):

- (14) a. *du ser* denna brua uti der # kenn se å høu ho e
(CoonValley-06gm)
b. *du ser* domm tåLa på phone å så tjøre nå hæile tia
(Westby-03gk)

Vi såg i del 4 at Aijmer (1997) drøftar V2-regelen i skandinavisk som eit mogleg verktøy for å disambiguere i analysen. Samanlikna med tradisjonelle eller kategoriske skildringar av norsk syntaks, skulle bl.a. *du vet / du ser* i (15a–b) vere klare tilfelle av pragmatisk partikkel (ikkje oversetning), mens *trur jeg* i (15c–d) ikkje skulle kunne vere pragmatisk partikkel:

- (15) a. *du vet* de ær ikke alle såmm kann fikkse tåssjk
(Chicago-01gk)
b. *du ser* je kRager ikke omm nåssjke språke mitt
(Stillwater-01gm)

- c. menn i tutus'nåni *tru jæi* vi tokk en sånn kjøretur
(Kalispell-02uk)
- d. i går *tru e* re va sækks daLe (Westby-01gm)

Men tidlegare syntaktisk forskning på CANS har vist at det er ein god del V2-brot i materialet (Eide og Hjelde 2015; Westergaard og Lohndal 2019). Dette kan t.d. illustrerast med (16a–b):

- (16) a. *du vet* nå alle nåssjkene kåmm i sammen demm bestanndi
snakkte nåssjk (Chicago-01gk)
- b. *du væit* nå du kåmm te denni kanntRi du må lære enngg'lsk
(Decorah-01gm)

I praksis er det som oftast prosodien som må avgjere i tvilstilfelle, noko som er vanskeleg i analysar av spontantale generelt (utan tekniske verktøy), og enda meir i andrespråks- eller arvespråksdata der informantane ikkje nødvendigvis følgjer same prosodiske mønster som i målspråket (jf. Polinsky 2018, 148). Om frekvenstala i tabell 2 kan vi legge til at i tvilstilfelle er det analyse som pragmatisk partikkel som er valt.

Vi har sett at *posisjonen* til pragmatiske markørar er viktig i analysar i faglitteraturen, og tabell 3 (neste side) gir ei oversikt over plasseringa til dei partiklane vi ser på her, *final*, *initial* eller *medial* (bruken av desse omgrepa blir forklart i dei kommande avsnitta):

Her ser vi blant anna at det, som forventa, er mykje variasjon i bruken av *you know*. Det er variasjon i dei andre kategoriane òg, men både *ser du*, *veit du* og *trur jeg* har klart mest final plassering, mens både *du ser*, *du veit* og *jeg trur* har mest initial plassering. Bruken av engelsk *I think* som pragmatisk partikkel har litt variasjon, men er mest final (jf. Aijmer 1997).

Som det kom fram i del 4, er det mange ulike måtar å analysere *posisjonen* til pragmatiske markørar på i litteraturen. Erman (1987) analyserer plassering i ytringa/turen. Schiffrin (1987) inkluderer også større, tilsvarande einingar i analysen (t.d. *you know what I mean*). Vi analyserer derimot berre toordsuttrykka i funksjon som pragmatisk partikkel, og når det gjeld plassering, ser vi på posisjon i forhold til

	<i>Norsk VS</i>	<i>Engelsk SV</i>	<i>Norsk SV</i>
SER DU	<i>ser du</i> (N=59)	<i>you see</i> (N=5)	<i>du ser</i> (N=13)
	Final: 95 %	Final: 60 %	Final: 7,7 %
	Initial: 5,1 %	Initial: 40 %	Initial: 92 %
	Medial: –	Medial: –	Medial: –
VEIT DU	<i>veit/vet du</i> (N=487)	<i>you know</i> (N=598)	<i>du veit/vet</i> (N=133)
	Final: 96 %	Final: 62 %	Final: 5,3 %
	Initial: 1,4 %	Initial: 20 %	Initial: 89 %
	Medial: 2,7 %	Medial: 18 %	Medial: 5,3 %
TRUR EG	<i>trur/tror jeg</i> (N=283)	<i>I think</i> (N=37)	<i>jeg trur/tror</i> (N=22)
	Final: 97 %	Final: 68 %	Final: 18 %
	Initial: –	Initial: 19 %	Initial: 77 %
	Medial: 3,2 %	Medial: 14 %	Medial: 4,5 %

Tabell 3: Partiklane analysert etter syntaktisk plassering.

den proposisjonen det fungerer som ein kommentar til. I dei fleste CANS-døma vi har vist til nå, står toordsuttrykka til ei full *setning*, og enten rett føre denne eller avsluttande rett etterpå, slik som i døme (8)–(10) og (5)–(7).

I mange tilfelle er det derimot klart meir komplisert enn dette. Mellom anna er det slik at setninga ofte held fram etter det som elles kan seiast å vere ein pragmatisk partikkel i final posisjon:

- (17) a. menn de itte så mannge att *trur e* te snakke me
(Outlook-06gm)
- b. va itte fært lanngt du reist ifrå heme *veit du* denn tia
(Westby-01gm)
- c. ja såmmin *tru i* så reise biene hass te California
(Sunburg-13gk)

Når det ikkje er fulle setningar, kan vi seie at partiklane står til ein proposisjon som er uttrykt som eit setningsfragment (jf. Søfteland og Borthen 2018, 257), men i (17c) og i dei klaraste døma på *flagging*, må ein kanskje likevel seie at det ikkje er proposisjonar partikelen står til, men enkeltord. Vi kjem tilbake til dette i del 6.

Vidare er det slik at det i ein god del tilfelle er umogleg å avgjere om partikkelen er *final* (peikar mot det som er sagt rett føre) eller *initial* (peikar mot det som kjem), og det er her kategorien *medial* kjem inn, som i (18):

- (18) a. hann kamm te høre de væit du viss omm vi skulle ha danns
you know enn kâmm ti finn ut de likkevæR (CoonValley-06gm)
- b. vi hadde stor grønn saker *you know* vi vekk sa mat'n (Albert-Lea-01gk)
- c. kusiner menn ikke nær kusiner *du vet* fømilie (Chicago-01gk)

Dette gjeld i spesielt stor grad for bruken av *you know* som i (18a), noko som har bidratt til dei mange diskusjonane om statusen til dette uttrykket også internasjonalt (jf. del 4). Om frekvenstala i tabell 3 kan vi legge til at det ofte er vanskeleg å dra grensa mellom setningar med framhald etter partikkelen (jf. 17a-c) og medial bruk (jf. 18a-c): I tvilstilfelle er det medial som er valt.

6 Korpusdata: Bruksmåtar for SER DU, VEIT DU og TRUR EG

Etter oversikta som er gitt i del 5, skal vi her sjå litt meir detaljert på ulike bruksmåtar. Først to aspekt som i større eller mindre grad er relevant for alle variantane, *klynger* og *flagging*, og deretter nokon preliminære tankar om vidare analysar inndelt etter dei tre hovudtypane.⁵

6.1 Klynger

Det er mange døme på *klynger* av pragmatiske partiklar for alle tre hovudkategoriane (jf. t.d. Klintborg 1999). Det aller mest påfallande er *veit du* i kombinasjon med *da* (832 treff totalt):

⁵ I det følgjande brukar vi KAPITEL når vi omtalar alle variantar av eit toordsuttrykk samla, t.d. SER DU om *ser du* + *du ser* + *you see*.

- (19) a. sku ta båt'n te Amerika *da veit du* så varrt de # kåmm dømm på å så hellste på rømm såmm ræiste *da veit du* (CoonValley-02gm)
 b. å hænn brukkte på å hanndle me fissk ifrå Nårkje *da vet du* # femm million punn i åre (Westby-01gm)

Det er ein del døme på norsk *veit du* i klynge med engelsk *you know*, altså ei dobling av partikkelen på dei to språka. I følgje Maschler og Schiffrin (2015, 200) er det vanleg at slike kjem rett etter kvarandre som i (20a–c). Enda oftare står dei i ulike posisjonar i setninga (20d):

- (20) a. ja vi hadd itte kLee *veit du you know* (Billings-01gm)
 b. ho pLannter mannge tå di bLåmmer såmm bare læsta æin år *veit du you know* (Fargo-02gm)
 c. hann ha så mye græier såmm far hadde *you know veit du* (Stillwater-01gm)
 d. menn *you know* når ru høre i dag *veit du* # goodness me (Hatton-04gk)

Klyngene er primært med VS-variantane, og mest i final posisjon, spesielt der dei står inntil kvarandre. Men det er mange ulike klynge-variantar i materialet, bl.a. som vist i (21a–c):

- (21) a. så a må tænsje se omm *ser du veit du* hå oR n ska beruke (Westby-02gm)
 b. å så da bynnte je å vara skolelærer her i FaRgåo *you know* før sekksten år *tru jei veit du you know* (Fargo-02gm)
 c. *ser du* denn nye banngken såmm kjem opp i Kon Valli # dømm ræiv ne denna byggingen *væit du* # fårr å bygge denn nye banngken (CoonValley-02gm)

6.2 Flagging

I del 3 nemnde vi *flagging* som eit mogleg relevant aspekt for analysen, spesielt med tanke på kodeveksling. Flagging er når ulike språklege trekk rammar inn eller retter merksemda mot eit språkbyte (jf. Rosignoli 2011). I materialet er dette spesielt framståande for *you*

know, som ofte dukkar opp i ytringar med veksling til engelsk ved enkeltord/frasar som talaren ser ut til å ikkje komme på på norsk:

- (24) a. ja de bi lidd'le kallt kjue kjue *you know* # twenty degrees below (Wanamingo-04gk)
 b. menn ee # de va Re- ee # *you know* respectable kara all i hop (Appleton-01gm)
 c. fækk # fækk se på alle # ee alle ee # the memorials *you know* (Billings-01gm)

Vi finn også ein god del døme på dette med både *du veit*, *veit du*, *trur jeg* og *I think*:

- (25) a. så då fekk hann gå her *du vett* homestead (Saskatoon-01gk)
 b. å så radiologist *du veit* å de ja ja denn æ både dåkkter å de (Coon Valley-32gm)
 c. å vi playde ee eller ball *veit du* played ball (Rushford-01gm)
 d. (Fem av dem. Fem hva?) e veit tj hå ru # grandchildren *I think s-* # de e noko anna ru kalle # [latter] sekkembarn or something like that # *væit du* (Sunburg-01gm)
 e. jiffte fært young *tru je* unng *tru je* ee (Hatton-06gm)

Som det også kjem fram i (25d), skjer det same rundt *norske* ord som talarane kanskje ikkje er heilt sikre på (26a–c). Ein variant av dette er ved namn på personar som talaren er usikker på om lyttaren kjenner til (26d):

- (26) a. menn nå ru menn nå ru ser i we say bLa # *you know* vi-avis'n (Sunburg-03gm)
 b. hennge n åpp å pLukke tå løva å så ha de ti en bonnt bonnt # en bonnt *da veit du* (CoonValley-02gm)
 c. cousin minn har har ei ee nåssjk bok u *veit du* på ee dictio-nary på ee Nårje *veit du* (Coon Valley-32gm)
 d. æi hukksa sje hann M19 *du væit* bare eit år enn jikk på skurn æ tru e (Hatton-03gm)

Om ein ser på døma i (24)–(26) som ulike sider av same sak, kan det kanskje ikkje kallast flagging på den måten omgrepet er mest bruka i faglitteraturen, og det framstår òg som at den primære funksjonen er å sjekke forståinga mellom talar og lyttar. Samtidig er det *you know* ved usikkerheit rundt *engelske* ord som er det mest frekvente, gjerne i ein elles for det meste norskspråkleg kontekst. Vidare forskning trengst for å finne ut korleis dette best kan analyserast.

6.3 Preliminær analyse av *SER DU*

I tabell 2 såg vi at *VS ser du* er klart mest frekvent i denne gruppa, og primært i final posisjon. Initial er også mogleg (27a). Og *you see* kan stå før og etter norskspråklege ytringar (27b–c):

- (27) a. ja åsså *ser du* vi # hadde brækkfest hema ra vett du
(Blair-17gm)
b. *you see* ee me leve på en faRm å sysstro mi så ønngRe
(Outlook-02gk)
c. menn ho britt sjuk denn gånngen *you see*
(NorthBattleford-01gm)

Bruksmåtene for *SV du ser* kan sjå ut til å spegle bruken av *you see* i engelsk, men dei fleste eller alle døma med *du ser* må seiast å vere velfungerande konstruksjonar i europansorsk òg, som i (28a–d). (Også her peikar uttala mot at uttrykka er brukt som pragmatisk partikkel.)

- (28) a. em ## *du ser* nårr # nå jænnerasjon minn æ borrtte så
(Westby-05gm)
b. vi hadde kjøtt sømm vi ee *du ser* ee vi sjlakhta åmm
hæusst'n (CoonValley-51gk)
c. næi # *du ser* eg behøvvde tje de (Saskatoon-01gk)
d. de va ba jennter fårr *du se* # gutt'ne kunne gå når såmm
hellst (Chicago-01gk)

I del 4 såg vi at ytringar med *sjø* i trøndersk gjerne inneheld ei forklaring, men at *sjø* også kan fungere som eit hint frå talar til lyttar om at ein er einige (Søfteland og Borthen 2018). Mange av dei 59

døma på *ser du* står etter ei forklaring (jf. 29a–b), men det er òg tilfelle der vi må rekne med at talaren veit at lyttaren kjenner innhaldet frå før (jf. 29c – samtale mellom jamnaldringar):

- (29) a. (I: Så far din snakka engelsk heime?) å ja hænn va hæLLv enngels hann *se ru* (CoonValley-17gm)
 b. (I: Og for å få den til å vekse, kva brukte de da?) yeast # vi ha mæ yeast i re *ser ru* # so it rises yeah (Sunburg-13gk)
 c. te gammLare ru bLir de forrtare går de *se ru* (CoonValley-06gm)

Ein preliminær analyse er at *ser du* i CANS kan sjå ut til å fungere mykje tilsvarande *sjø* i trøndersk, men at det er spesielt mykje forklaring, truleg fordi mange av dei aktuelle døma er svar på spørsmål frå ein intervjuar som i (29a–b). *Du ser* kan sjå ut til å delvis likne bruken av engelsk *you see*, men det er for få døme på begge delar i dette materialet til at vi kan seie noko sikkert.

6.4 Preliminær analyse av VEIT DU

Alle dei tre variantane innanfor kategorien VEIT DU er høgfrekvente i materialet. Vi har sett at *veit du* primært har final plassering og *du veit* primært initial, men det er også ein viss variasjon, t.d. *veit du* initialt (30a) og *du veit* finalt (30b):

- (30) a. å *veit du* # i desse da- lanngge gonngen em de var ein sinn omm en jinnte jikk me en gutt ifrå anndre tjørrka (Blair-04gk)
 b. ja domm brukkt å taRa litt nåssjk dama menn de taRa de itte nå no meir ama *du væi* (Blair-01gm)

You know er høgfrekvent som pragmatisk partikkel i materialet. Ein del av dei er i einspråkleg engelske ytringar, men dei står blant anna ofte mellom to setningar på norsk, ein bruksmåte som er kjent frå andre analysar av arvespråkssituasjonar (jf. Poplack 1980):

- (31) a. menn dæ rarrrt me re nå remm vekks opp slikk å fløtt ifrå heimen *you know* de kan bli slekk att ru ser emm ikkje så åffte vett du (Sunburg-07gm)
- b. å mor skreik *you know* å denn stakkasj lit'n gutt'n hann va bare fire år gammal å dær ær n sekksti fot opp (Gary-01gm)

I faglitteraturen (jf. del 4) kjem det fram at *you know* ofte er vanskeleg å analysere, og det gjeld i stor grad også her. I døme som (31a) er *you know* analysert med *medial* plassering i tabell 3; det er umogleg å bestemme om *you know* er knytt intonasjonsmessig mest til det som er sagt føre eller til det som kjem etter, og semantisk kan det vere begge delar (jf. Schiffrin 1987). *You know* står også ofte inni setningar og inni setningsledd (jf. bl.a. diskusjonen rundt døma i (17)):

- (32) a. så otherwise em du vill ikke gå *you know* ti de fåRR de skulle ta fåRR lennge menn (Portland-02gk)
- b. somme tier att moro å hatt # *you know* e bror ee sisster elli hårrt (Westby-12gm)
- c. va ei stunn de va sju hundre fåRk ell de va sju hundre fammli de kjømm itt e i høu ## *you know* somm livvde på einn gaR ein farrm (Westby-01gm)

Du veit er også frekvent, i dei fleste tilfella i initial posisjon, ofte med ny informasjon etter:

- (33) a. *du veit* vess du ee whispera åt nåån i skoLe såmm hadde nåssjk de (Appleton-01gm)
- b. well *du veit* nå re va heitt da va re haRt (Sunburg-12gk)
- c. å *du veit* vi vi vå- hann far våkks åpp i i Vessfåll i næren av Vessfåll så jæi har denn dialekt'n (Stillwater-01gm)

Schiffrin (1987) analyserer bruk av *you know* aller mest som ein måte å sikre kommunikasjonen i samtala på. Moquin og Salmons (2020) seier at det er forskjell på engelsk og norsk – at *you know* er ei bønne om forståing, mens *veit du* insisterer på sanninga i ei utsegn – men

at *veit du* hos amerikanorsktalarane har tatt opp i seg meir av den engelske bruksmåten. Dette stemmer godt med våre analysar. Hovudinntrykket er at det store fleirtalet av døme på alle variantar i VEIT DU-kategorien primært fungerer som ein sjekk for informasjonsflyten; dei hintar om at talaren gir informasjon, men at forståinga av denne er avhengig av at lyttaren følgjer med (jf. Schiffrin 1987, 290). Døma i (34) kan illustrere ulike variantar av dette:

- (34) a. mais å korn # å høy *veit du* nå vi mjøRRke mye kue må ru ha mye høy å my pasture grass (Decorah-01gm)
 b. ja # e tru kanne- e kudd hæ- e kurra # off da # si # de bRi enngelst e nåRsk i samens *veit du* (Harmony-01gk)
 c. ja m- messta # ee # ee # strull å # Råosetts å # menn rumm-græut # å ru vet # å læffse å fLattbrø # alt detta der # lutfisk (Westby-03gk)
 d. vi har ikkje råkkå nåRske aviseR sjøR menn *u veit* ee # e e æin såmm brinnga # brinnga remm hit te culture-hus næ vi # omm frædag (Sunburg-03gm)

Vi kan seie at både *veit du* og *du veit* i dette materialet i stor grad tilsvavar bruken av *you know*.

Samanliknar vi treff på *veit du* i CANS mot NDK, er det over dobbelt så frekvent (5,03 mot 1,89/1000 tokens). Ei mogleg forklaring på det høge talet på VEIT DU-variantar generelt kan vere at informantane er usikre på situasjonen språkleg og kommunikativt. Dei blir under opptaka oppfordra til å snakke (berre) norsk med kjende og ukjende sjølv om det naturlege for dei i settinga ville vere engelsk, og mange av dei har ikkje snakka norsk på veldig mange år.

6.5 Preliminær analyse av TRUR EG

I del 5 såg vi at VS *trur/tror jeg* er nokså høgfrekvent i datamaterialet, og nesten berre i final posisjon. Eit interessant poeng her er at ein god del elles tvitydige døme med tanke på plassering kan bli eintydiggjort om ein ser på *kva som er usikkert*. I (35a) er det mest sannsynleg at usikkerheita ligg i tidsomfanget *tri dager* og ikkje i om den omtalte personen faktisk dreiv med omsetting. Døme (35b) er deri-

mot mogleg å analysere både som at det er *langrenn* det er noko usikkert med, eller det at det ikkje er populært. Denne er dermed kategorisert som *medial* i tabell 3.⁶

- (35) a. ja # hann hell på tri tri dager *tru e* [latter] me å s- translate (CoonValley-04gm)
 b. (I: Er det både liksom slalåm og langrenn?) lanngrenn *tru jæi* ær ikke så velldi populært da (KalisPELL-02uk)

Vi har sett at Aijmer (1997) trekker fram to hovudbetydningar for *I think*, (sikker) påstand og usikker (informasjons)status. For bruk som pragmatisk partikkel er det primært sistnemnde som er relevant, spesielt når dei står i final posisjon. Dette stemmer godt for døma på VS *trur/tror jeg* i CANS, og det er mange ulike typar usikkerheit som kjem fram: Usikker på tidspunkt i eige liv (36a), på ei oppleving i framtida (36b) og på allmenne situasjonar eller hendingar i fortida (36c–d):

- (36) a. je var Sunny Side de va treåføRRti *trur je* førr da reiste je te høgskuurn (Blair-04gk)
 b. nå ska vi ræse te å besikk dæmm e NåRRge så de ska være moLo *tru e* (CoonValley-12gm)
 c. å på sjlutt'n på trædvetaLe så va re allt enngelst *tru e* (Fargo-01 gm)
 d. da vi jiffa åss så va skuurn borrtte allt *tru e* # consolidated with Westby and then they tore it down and sold it (CoonValley-06gm)

Også det at *trur/tror jeg* er så frekvent, kan vere eit resultat av at informantane er usikre i kommunikasjonssituasjonen språkleg, og desutan av at dei får spørsmål frå intervjuar om hendingar langt bakover i tid.

⁶ Eit spørsmål her blir om *trur jeg* kan stå i initial posisjon i det heile i norsk, men det er mogleg med *veit du* (jf. 30a) – og det er mogleg med *I think/you know/you see*, noko som truleg kan påverke bruken i amerikanorsk.

Vidare er det òg ein del døme på *I think*, nokon i einspråkleg engelske ytringar, men også mange i primært norskspråklege ytringar:

- (37) a. (I: Og du husker ikke at mor og far snakka norsk heller?)
 næi tRu ikke de baRe enn- # baRe enngelsk # *I think*
 (Spokane-07gm)
- b. hann ha binnt å farrmen i 1918 *I think* å farrme heile live
 (Portland-01gm)

Også dette stemmer godt med analysen til Aijmer (1997) og kommentarane om *trur jeg* ovanfor.

Som nemnt i del 5 er det eit par døme på *jeg trur* som pragmatisk partikkel i materialet (t.d. 10). I tabell 2 er dette 22 døme totalt av dei 201 som er analysert, men dei fleste av desse er tvilstilfelle. Den høgfrekvente bruken av *jeg trur* som vanleg setningsinnleiar bidrar truleg til blokkering av utviklinga av denne ordkombinasjonen til frittstående pragmatisk partikkel.

7 Konklusjon

Bruken av dei omtalte pragmatiske partiklane i CANS er på mange måtar tilsvarande bruken i europansk, men amerikanorsktalarane må seiast å ha eit større repertoar å spele på. I tillegg til dei typisk etterstilte (VS-)bruksmåtene frå europansk, tar dei også i bruk dei liknande engelskspråklege uttrykka og deira varierte bruksmåtar, samtidig som dei tilpassar og utvidar meir marginale norskspråklege variantar i retning engelsk bruk, i betydning og/eller frekvens. Spezielt er ulike variantar av VEIT DU svært høgfrekvent. Ei mogleg forklaring kan vere at desse partiklane gjerne blir bruka som ein sikkerheitsfunksjon for informasjonsflyten, noko informantane kan trenge om dei kommunikativt sett er utrygge på situasjonen, og som truleg kan bli forsterka av at dei snakkar med ein europansk-talar. Tilsvarande kan det vere nyttig for dei å kunne bruke pragmatiske partiklar som flagging av språkbyte til engelsk for ord dei ikkje hugsar på norsk, eller markering av norske ord dei er usikre på. Det har

komme fram fleire metodiske problemstillingar ved analysane i artikkelen, blant anna knytt til intonasjon, V2-spørsmål og tvitydig semantikk. I den vidare forskinga bør det gjerast meir detaljerte pragmatisk-semantiske analysar, også komparativt mot engelsk og eldre norske dialektar, inkludert diakrone perspektiv på CANS-data og analysar i lys av kodevekslingsteori.

Referansar

- Aijmer, Karin. 1997. I think – an English modal particle. I *Modality in Germanic languages: Historical and comparative perspectives*, redigert av Toril Swan og Olaf J. Westvik, 1–47. Berlin/Boston: De Gruyter.
- Borthen, Kaja. 2018. Store spørsmål om små ord. Innledning til temanummer om pragmatiske partikler i norsk. *Norsk Lingvistisk Tidsskrift* 36 (2), 225–247.
- Borthen, Kaja, Anders Nøklestad, Åshild Søfteland, Perlaug M. Kveen og Elena Karagjosova. (2018). Database for etterstilte småord i norsk (Småord-databasen). Tekstlaboratoriet, UiO. <http://tekstlab.uio.no/not>
- Borthen, Kaja, Åshild Søfteland, Perlaug M. Kveen, Elena Karagjosova og Anders Nøklestad. Finalpartiklar i norske talemål. Ei undersøking av variasjon knytt til geografi, alder og kjønn. *Maal og Minne* 1-2021, 1–63.
- Chevalier, Joan F. 2004. Heritage language literacy: Theory and practice. *Heritage Language Journal* 2 (1), 1–19.
- Eide, Kristin og Arnstein Hjelde. 2015. Verb Second and Finiteness Morphology in Norwegian Heritage Language of the American Midwest. I *Moribund Germanic Heritage Languages in North America*, redigert av B. Richard Page og Michael T. Putnam, 64–101. Leiden/Boston: Brill.
- Erman, Brit. 1987. *Pragmatic Expressions in English. A study of You know, You see and I mean in Face-to-face Conversation*. Stockholm: Almqvist & Wiksell International.

- Fraser, Bruce. 1999. What are discourse markers? *Journal of Pragmatics* 31, 931–952.
- Johannessen, Janne B., Joel Priestley, Kristin Hagen, Tor A. Åfarli og Øystein A. Vangsnes. 2009. The Nordic dialect corpus: An advanced research tool. I *Proceedings of the 17th Nordic Conference of Computational Linguistics*, redigert av Kristiina Jokinen og Eckhard Bick, 73–80. Odense: Northern European Association for Language Technology (NEALT).
- Johannessen, Janne B. 2015. The Corpus of American Norwegian Speech (CANS). I *Proceedings of the 20th Nordic Conference of Computational Linguistics*, redigert av Beáta Megyesi, 297–300. Vilnius: Northern European Association for Language Technology (NEALT).
- Klintborg, Staffan. 1999. *The transcience of American Swedish*. Lund: Lund University Press.
- Maschler, Yael og Deborah Schiffrin. 2015. Discourse markers: Language, meaning, and context. I *The handbook of discourse analysis*, redigert av Deborah Tannen, Heidi E. Hamilton og Deborah Schiffrin, 189–221. Chichester, UK: Wiley Blackwell.
- Matras, Yann. 2009. *Language contact*. Cambridge: Cambridge University Press.
- Moquin, Laura og Joseph Salmons. 2020. American Norwegian discourse marking: Convergence, detachability, pragmatic change. I *Bauta – Janne Bondi Johannessen in memoriam*, redigert av Kristin Hagen, Arnstein Hjelde, Karine Stjernholm og Øystein A. Vangsnes, 303–318. *Oslo Studies in Language* 11(2).
- Mæhlum, Brit. 2009. Standardtalemål? Naturligvis! En argumentasjon for eksistensen av et norsk standardtalemål. *Norsk lingvistisk tidsskrift* 27 (1), 7–26.
- Polinsky, Maria. 2018. *Heritage languages and their speakers*. Cambridge: Cambridge University Press.
- Poplack, Shana. 1980. Sometimes I'll start a sentence in spanish y termino en espanol: toward a typology of code-switching. *Linguistics* 18 (7–8), 581–618.
- Rosignoli, Alberto. 2011. *Flagging in English-Italian code-switching*. Doktoravhandling, Bangor University.

- Salmons, Joseph. 1990. Bilingual discourse marking: code switching, borrowing, and convergence in some German-American dialects. *Linguistics* 28 (3), 453–480.
- Schiffrin, Deborah. 1987. *Discourse markers*. Cambridge: Cambridge University Press.
- Søfteland, Åshild og Kaja Borthen. 2018. Æ e trønder æ, sjø. Den pragmatiske partikkelen ‘sjø’ i midtnorske dialektar. *Norsk lingvistisk tidsskrift* 36 (2), 249–280.
- Östman, Jan O. 1981. *‘You Know’: A discourse-functional study*. Amsterdam: John Benjamins.
- Westergaard, Marit og Terje Lohndal. 2019. Verb second word order in Norwegian heritage language: Syntax and pragmatics. I *Variable properties in language: Their nature and acquisition*, redigert av David W. Lightfoot og Jonathan Havenhill, 91–102. Washington, DC: Georgetown University Press.

English summary

The main focus in this article is on the use of the pragmatic particles *ser du*, *veit du* og *trur eg* in CANS – Corpus of American Nordic Speech. These are compared with the use of English *you see*, *you know*, *I think* and Norwegian *du veit*, *du ser*, *eg trur* in the same material. The article provides an overview of frequency, syntactic position and overall pragmatic analysis, together with a discussion of some methodological issues. The main finding is that the American Norwegian speakers use the possibilities that these pragmatic particles provide as support for the information flow in the conversational situation, and combine both Norwegian and English usage patterns together into a larger resource.

Åshild Søfteland
 Institutt for språk, litteratur og
 kultur
 Høgskolen i Østfold
 ashild.softeland@hiof.no

Arnstein Hjelde
 Institutt for språk, litteratur og
 kultur
 Høgskolen i Østfold
 arnstein.hjelde@hiof.no