

Norsk, særleg nynorsk, dokumentasjonsleksikografi i framtida

Ålov Runde

1 Innleiing

Artikkelen er strukturert slik: I del 1 vil eg seie litt om språkleg dokumentasjonsverksemd generelt. Eg ser nærmare på rammevilkåra for denne typen verksemd, og tek òg opp den situasjonen nynorsken står overfor i dag. I del 2 går eg meir spesifikt inn på dokumentasjons*leksikografisk* verksemd, og nemner nokre av dei viktigaste tilgjengelege ressursane som finst for denne verksemda. I del 3 seier eg litt meir om tilstanden for norsk dokumentasjonsleksikografi i dag. Dette gjer eg med utgangspunkt i prosjektet Norsk Ordbok 2014. Avslutningsvis i del 4 prøver eg å samle trådane, og ser nærmare på nokre av dei utfordringane ein står overfor i framtida.

2 Språkleg dokumentasjonsverksemd

Termen 'dokumentasjonsleksikografi' kan ikkje gjerast greie for isolert, men må setjast i ei større ramme, ei ramme eller ein type verksemd vi kan kalle *dokumentasjonen av norsk språk*. Denne verksemda femner i prinsippet om alt materiale som dokumenterer den språklege tradisjonen, altså eldre materiale, og alt materiale som dokumenterer språket slik det blir brukt i samtida, altså nyare materiale. Språkleg dokumentasjonsverksemd er arbeidet med å samle, systematisere og gjere tilgjengeleg så mykje som mogleg av slikt materiale. Dette materialet er sjølve grunnstommen i den dokumentasjonsleksikografiske verksemda, som mellom anna er arbeidet med å redigere ordbøker og byggje opp elektroniske ordbasar. Språkleg dokumentasjonsverksemd og dokumentasjons*leksikografiverksemd* heng med andre ord tett saman.

Eit sentralt dokument i samband med språkleg dokumentasjonsverksemd er St.meld. nr. 35 (2007-2008): *Mål og meining. Ein heilskapleg norsk språk-politikk*. Kapittel 8 i denne meldinga (heretter kalla Språkmeldinga) seier mykje om denne verksemda. Mellom anna blir det her slege fast at det å dokumentere norsk språk i seg sjølv er ei stor kulturoppgåve og ei viktig nasjonal oppgåve.

Særleg for språk som skal vernast og styrkjast, og som skal stå i mot eit ytre press som i dag kanskje er sterkare enn på lenge, er det spesielt viktig å kunne støtte seg på ein solid dokumentasjon. Her står ordbøker og elektroniske ord-basar heilt sentralt, seier Språkmeldinga.

Denne tankegangen er i og for seg ikkje ny. St.meld. nr. 15 (1968-1969), som byggjer på innstillinga frå Vogt-komiteen og som kom i samband med opprettinga av Norsk Leksikografisk Institutt, målber mykje av den same argumentasjonen når det gjeld synet på verdien av dokumentasjonsverksemd:

”Utforskningen av norsk språk, dets historie, dets skrift-språktradisjoner, talespråkene og de levende målfører i by og bygd, bør stå sentralt i norsk språkforskning og bør støttes planmessig ved alle høyere læresteder hvor det drives vitenskapelig forskning, i første rekke ved våre universiteter og vitenskapelige institutter. Dette er en forutsetning for at den all-menne diskusjon om språklige spørsmål kan bygge på sikker og objektiv informasjon og for at allmennheten kan nyttiggjøre seg forskningens resultater.” (St.meld. nr. 15 *Om språksaka*, s. 6)

Men det politiske bakteppet var litt annleis den gongen:

”på grunn av den langvarige språkstrid i landet, har debatten om språklige spørsmål spilt en stor rolle, men språkstriden har kanskje også vært til hinder for en samlet innsats for en systematisk utforskning av våre skrift-språk, slik de virkelig er, og våre dialekter, de gamle og de nye. Den støtte som myndighetene har gitt til slike tiltak, har ikke stått i noe forhold til omfanget av og intensiteten i språkdiskusjonene. På mange felter av forskning ligger vi etter våre naboland og andre land som det er naturlig å sammenligne oss med.” (op. cit., s. 6)

På 60- og 70-talet hadde språkstriden, som var opprivande og heller lite samlande, lagt mange av premissane for den språkpolitiske diskusjonen. Ser vi på den språkpolitiske situasjonen i dag, er biletet litt annleis. Ideologisk har det vore eit viktig skifte: der språkstriden sette rammene for diskusjonen rundt dokumentasjon av norsk språk på 60- og 70-talet, blir denne verksemda i dag forstått og diskutert i lys av termen ’domenetap’.

Engelsk har i løpet av dei siste tiåra fått ein særskild posisjon som vår tids internasjonale hjelpespråk. Termen 'domenetap' blir ofte brukt om ei veksande uro for at språket norsk blir trengt til sides og ikkje lenger er i bruk innanfor eit bestemt samfunnsdomene. Om denne tilstanden "smittar" frå eitt domene til eit anna, "kan eit fullverdig norsk språk vere truga", heiter det i Språkmeldinga. Det er med andre ord framleis tale om ein språkstrid; ikkje mellom nynorsk og bokmål, som var der hovudstriden stod på 60- og 70-talet, men mellom norsk og engelsk.

Det er ikkje vanskeleg å tenkje seg eit scenario der eit slikt domenetap kan gjere situasjonen vanskeleg spesielt for nynorsken og dialektane. Nynorsk- og dialektforskinga, forstått som "minoritetsvarietetar", kan i verste fall drukne i det såkalla store biletet, på bekostning av norsk forstått som "majoritetsvarietetet", det vil seie bokmålet. Språkmeldinga nemner då òg særskilt nynorsken si stilling i dette perspektivet, men legg likevel ein viss varsam optimisme for dagen:

"som bokmål er pressa av engelsk, står nynorsk under press frå bokmål. For i den grad bruksområdet for bokmål verkeleg tek til å krympa til fordel for engelsk, vil også nynorsk få stadig mindre olbogerom. Når alt kjem til alt, er det derfor liten tvil om at nynorsk og bokmål står i ein lagnadsfellesskap.[...] Nynorsken har ingen sjanse til å overleva i ein situasjon der bokmålet skulle bli alvorleg truga som samfunnsberande og fullverdig språk. [...] Difor blir verdien av norsk språk sterkare og tydelegare framheva takk vere den nynorske språktradisjonen. Stilt overfor den engelske utfordringa må vi utnytta den samla språkstyrken til nynorsk og bokmål. Til dømes kan det oppstå ein eigen språkleg dynamikk når terminologi skal utviklast både på nynorsk og bokmål. Denne dynamikken kan fungera som ein kulturell konkurranseføremon for norsk i møtet med det store globaliseringsspråket"

(Språkmeldinga kap. 9.1.1.3)

For å oppsummere så langt: det er både politisk forståing for og generell semje om at dokumentasjonen av norsk språk er ei viktig nasjonal, kulturell oppgåve, og semje om at slik dokumentasjonsverksemd er grunnleggjande i språkforskinga. Dette arbeidet er ein

type grunnforskning som har relevans langt ut over det å vere utgangspunkt for redigering av einskildordbøker. I Språkmeldinga blir verdien av slik dokumentasjonsverksemd forstått særleg på bakgrunn av utfordringar knytte til domenetap. Språkmeldinga tek også opp utfordringane for minoritetsnynorsken i forhold til majoritetsbokmålet, og utfordringane for norsk i forhold til globaliseringspråket engelsk.

Mykje av det språklege dokumentasjonsmaterialet finst i dag i samlingar og arkiv ved ulike institusjonar, særleg ved universitet og høgskular rundt om i landet. Ein rapport frå 2007 slår fast at tilhøva rundt desse samlingane gjev grunn til uro (Andersen et al. 2007). Mykje av samlingsmaterialet vart digitalisert i det store Dokumentasjonsprosjektet på 1990-talet, men på langt nær alt. Rapporten nemner fleire utfordringar ein står overfor når det gjeld forvaltninga og vidareutviklinga av dette dokumentasjonsmaterialet. Ëi utfordring er at det etter Dokumentasjonsprosjektet ikkje har skjedd tilsvarande systematisk digitalisering av nyare materiale. Ei anna utfordring er at korkje bokmålstilfanget eller ordtil-fang frå gammal- og mellomnorsk tid har vorte digitalisert tilsvarande dialekt- og nynorsktilfanget. Dessutan er det uklårt kva som vil skje med dialekt- og nynorsktilfanget etter at prosjektet Norsk Ordbok 2014 er avslutta. Rapporten reiser også spørsmål om vidare forvaltning og sikring av det materialet som enno ikkje er digitalisert, og han tek opp generelle spørsmål knytte til ramme-vilkåra for arbeidet med oppbygging og forvaltning av samlingar, arkiv og digitale korpus. Manglande samkøring mellom institusjonane har dessutan ført til delvis fragmentering av dei digitale ressursane, og det er truleg mykje arbeid som står att å gjere for å fullføre digitaliseringa og tilgjengeleggjeringa av alt som finst ved dei ulike institusjonane.

3 Dokumentasjonsleksikografi og dokumentasjonsordbøker

3.1 Allment

Dokumentasjonsleksikografien har, i motsetning til andre typar leksikografi, eit heilt spesielt formål: han skal dekkje det behovet eit kollektiv har for å ta vare på språkleg informasjon for framtida. Dokumentasjonsordbøker er slik sett eit handgripeleg resultat av språkleg dokumentasjonsarbeid. Dei er vitskaplege ordboksverk i fleire band med fleire og fylldigare opplysningar om einskildord enn i andre slags ordbøker. Dermed representerer dei den mest avanserte

katego-rien av ordbøker, og er ei investering i eit produkt som har stor kulturell og vitskapleg verdi. Verdien av denne typen ordbøker minkar ikkje ved bruk, snarare tvert imot – seinare dokumentasjon av den framtidige språkutviklinga kan byggje vidare på det arbeidet som blir gjort i slike ordbøker.

Noreg har ein meir innfløkt språksituasjon å dokumentere enn til dømes svensk og dansk: Vi har to offisielle skriftmål, ein uoffisiell riksmålsvariant, og eit stort spekter av ulike målføre. Men nettopp når det gjeld målføra står Noreg i ei særstilling i dokumentasjons-samanheng, ettersom ein innanfor den nynorske tradisjonen reknar talemålet som ei like god kjelde for nytt ordtilfang som skrift. *Norsk Ordbok* er slik sett eit sjeldan og eineståande verk i europeisk samanheng, fordi det framstiller skriftspråket og talemålet som éi eining. *Norsk Ordbok* er eit nynorskverk, men har likevel eit større nedslagsfelt enn det kanskje kan sjå ut til, understrekar Språkmeldinga:

”Når Norsk Ordbok er fullført i 2014, vil vi ha fått ein etter måten fullnøyande leksikalsk dokumentasjon av den nynorske målforma og det nynorske skriftspråket som vil stå seg i lang tid framover. Det er likevel grunn til å streka under at Norsk Ordbok ikkje kan karakteriserast som eit reint nynorskverk og langt mindre som eit prosjekt berre av interesse for nynorskfolket. Det at Norsk Ordbok i tillegg til det nynorske skriftmålet også dokumenterer norsk talemål i form av dialektord frå heile landet, gjer dette til ein særleg verdifull dokumentasjon av ein felles, nasjonal språk-arv som direkte vedkjem ein mykje større del av norske språk-brukarar enn dei som nyttar nynorsk som skriftmål”

(Språkmeldinga kap. 8.3.5.3)

3.2 Nokre sentrale dokumentasjonsleksikografiske ressursar

3.2.1 Nynorsk

Prosjektet Norsk Ordbok 2014 har eit stort, elektronisk tilgjengeleg kjelde-materiale, som inneheld både skriftlege kjelder og målføre-kjelder (sjå òg Wetås kap. 5.6). Prosjektet er eit samarbeid mellom Kulturdepartementet, Universitetet i Oslo og Det norske Samlaget. *Norsk Ordbok* blir utgjeven av Samlaget.

Det viktigaste underlagsmaterialet for *Norsk Ordbok* er:

(a) Skriftspråklege kjelder:

- (i) "Grunnmanuskriptet" (ei samanstøyping av ordbøkene til Aasen, Ross og Schjøtt) og andre ordbøker (særleg fagordbøker)
- (ii) setelmateriale frå ulike litterære kjelder (mellom anna ymse skjønn- og faglitteratur, aviser og tidsskrift, Bibelen, dialekt- og ordsamlingar frå dansketida)
- (iii) det nynorske tekstkorpuset (nær 40 millionar ord; dokumenterer både den eldste nynorsken og nyare nynorsktekster)

(b) Målførekjelder:

- (i) setelmateriale med målføreopplysningar frå heile landet, som aukar kvart år
- (ii) målføresynopsisen (fonologiske og morfologiske opplysningar for visse ord frå heile landet)
- (iii) lokale ordsamlingar, som blir gjort tilgjengelege elektronisk gjennom Ordbokshotellet

3.2.2 Bokmål

Det Norske Akademis store ordbok (NAOB) er ei litterær ordbok over norsk riksmål og bokmål frå 1814 til våre dagar, med hovudtyngda på eldre historiske og litterære kjelder. NAOB skal oppdatere og byggje ut den allereie eksisterande *Norsk Riksmålsordbok*. Ordboka er eigd av *Det Norske Akademi for Sprog og Litteratur*, og blir utgjeven av Akademiet i samarbeid med Kunnskapsforlaget.

Prosjektet *Norsk elektronisk bokmålsordbok* tek sikte på å dokumentere særleg ordtilfang, ordformer og ordbruksmåtar i alle former for moderne bok-mål. Prosjektet er førebels på planleggingsstadiet og har levert ei projektskisse til Kulturdepartementet. Dei rår over eit korpus på om lag 40 millionar ord, *Leksikografisk bokmålskorpus* ved Universitetet i Oslo, ILN.

Tekstlaboratoriet ved Universitetet i Oslo, ILN, rår over ei mengd med ulike talemålskorpus, i tillegg til fleire ein- og fleirspråklege skriftspråkskorpus.

3.2.3 Tilgrensande ressursar

Prosjektet *Uni Digital* ved Universitetet i Bergen driv med forskning og utvikling innanfor mellom anna språk- og informasjonsteknologi. Prosjektet har eit stort aviskorpus, *Norsk aviskorpus*, på nærmare 700 millionar ord, både bokmåls- og nynorsktekster.

Prosjektet *Medieval Nordic Text Archive* (MENOTA-prosjektet) er eit nettverk av arkiv, bibliotek og institutt som arbeider med å ta vare på, formidle og publisere mellomalderstekster og -handskrifter. Det er først og fremst eit digi-taliseringsprosjekt.

Mykje av norsk språk – både bokmål og nynorsk, eldre og nyare, litteratur og målføre – ser ut til å vere relativt godt dokumentert i dei samlingane, arkiva og tekstkorpusa som eg har nemnt her. For ordens skuld vil eg understreke at eg sjølv sagt ikkje har nemnt alle. Men det er éin periode vi saknar: perioden frå ca. 1600-1850, som ofte blir kalla ”det eldste moderne norsk”, er generelt lite utforska og dårleg dokumentert. Prosjektet Norsk Ordbok 2014 har teke initiativ til å få midlar til å gjere noko med denne mangelen.

4 Prosjektet Norsk Ordbok 2014: direkte og indirekte følgjer for tilstanden til norsk dokumentasjonsleksikografi

Skal ein seie noko om norsk dokumentasjonsleksikografi i dag, kjem ein ikkje utanom prosjektet Norsk Ordbok 2014 – det er eit prosjekt som tek plass og merksemd, og det er svært synleg i landskapet. Prosjektet har sidan starten i 2002 fått store statlege løyvingar, det har fått etter måten mykje og god omtale, og det er eit prosjekt med mange tilsette samanlikna med liknande ordboksprosjekt i Norden. Éin negativ konsekvens av at prosjektet er så pass stort og dominerande, er at det kan ta merksemda vekk frå den generelle situasjonen for dei universitetsbaserte leksikografamiljøa. Per i dag er det til dømes berre fire faste leksikografistillingar ved Universitetet i Oslo knytte til instituttet, og det er ikkje utsikter til betring på dette området. Dette er ein uheldig situasjon, som for nokre har ført til ei viss uro for at fagfeltet kan stå i fare for å ”døy” i skuggen av Norsk Ordbok. Korleis vil situasjonen sjå ut etter 2014, når prosjektet er avslutta?

Ein kan snu litt på det siste spørsmålet ovanfor, og spørje seg korleis situasjonen for (ny)norsk leksikografi hadde sett ut dersom prosjektet Norsk Ordbok 2014 *ikkje* hadde vorte realisert i 2002. Det

er ikkje vanskeleg å sjå for seg eit heilt anna scenario: *Norsk Ordbok* hadde neppe vorte ferdigstilt før tidlegast 2060, om i det heile, viss ein skulle ha følgd den utgjevingstakta verket hadde før det vart eit prosjekt i 2002. Eit verk som *Norsk Ordbok* har liten eller ingen verdi, heller ikkje i eit større kulturhistorisk perspektiv, dersom det ikkje blir fullført. Eit anna spørsmål er korleis rekrutteringa til eit slikt verk, med så lang tidshorisont, hadde vore. Dessutan hadde store delar av det arbeidet som har vorte gjort med materialet i Dokumentasjonsprosjektet, vore ein uutnytta ressurs. Alle desse faktorane hadde truleg representert eit mykje større problem – og på sikt fått langt fleire negative konsekvensar – for heile leksikografimiljøet.

Prosjektet Norsk Ordbok 2014 har ført til at eit stort fagmiljø står samla bak eitt stort prosjekt. Ein openberr styrke ved ein slik konsentrasjon er at det banar vegen for framtidige prosjekt bygde over same eller liknande leist. Det at prosjektet er så synleg i leksikografilandskapet, er noko ein kan bruke aktivt på ulike måtar. Rekrutteringa til prosjektet har vore uvanleg god og sterk i heile perioden, noko som har lagt grunnlaget for eit stort og vitalt fagmiljø med høg kompetanse, som ein kan trekkje vekslar på og utnytte også etter 2014. Innanfor rammene av prosjektet har ein dessutan utvikla ei avansert elektronisk plattform og ei rekkje andre dataverkty. Dette har vore eit stort utviklingsløft, der Eining for Digital Dokumentasjon (EDD) ved Universitetet i Oslo har vore tungt inne (sjå òg Wetås s. 58-59). Sjølv ordboka er ein kompleks, søkbar relasjonsdata-base som ein kan trekkje all slags informasjon ut av. Redaktørane skriv ordboksartiklane rett inn i basen ved hjelp av eit redigeringsprogram (ein editor), og EDD har også utvikla ein tilleggsapplikasjon som genererer ferdig sats direkte frå databasen. Dette redigeringsverktøyet har ein klår overføringsverdi til eventuelle framtidige ordboksprosjekt: EDD har utvikla ei maksimumsløysing, spesielt tilpassa redigeringspraksisen i *Norsk Ordbok*, som ein kan bruke som utgangspunkt for nær sagt kva type ordbok ein vil. Den godt skolerte ordboks-staben og det avanserte redigeringsverktøyet er svært viktige ressursar – til saman utgjer dei ein ”infrastruktur” som framtidige ordboksprosjekt kan stø seg på og utnytte. Eg vil òg understreke at sjølv om prosjektet er lausrive frå instituttstrukturen ved Universitetet i Oslo reint formelt, har det likevel mange tangeringspunkt og er godt etablert som eit synleg og sterkt fagmiljø her.

Dokumentasjonsverksemd er både kostbart og langvarig arbeid. Universitet og høgskular har i dag tronge økonomiske kår og lite handlingsrom over vanlege budsjett for denne typen verksemd. Ideelt sett skulle sjølvstøtt denne verksemda fullfinansierast av dei institusjonane som faktisk eig samlingane. Men slik røyndomen er i dag, er det først og fremst delvis direktefinansiering frå ulike statlege organ ein kan stø seg på, og då blir ein tvunga til å tenkje prosjekt og prosjektfinansiering. Prosjektet Norsk Ordbok 2014 har opparbeidd ei solid erfaring med slike store og langvarige prosjekt, ei erfaring som framtidige prosjektsøkjjarar òg kan trekkje lærdom frå og bruke i sitt arbeid. Men ein viktig føresetnad her er sjølvstøtt at dei universitetsbaserte leksikografimiljøa ikkje er heilt nedbygde.

Leiinga ved prosjektet har dessutan opparbeidd ein viktig pragmatisk kunnskap i løpet av prosjektperioden: Det er heilt nødvendig å gjere grunnlags-arbeidet skikkeleg, det vil seie dimensjonere arbeidet og omfanget av det. På det grunnlaget må ein så setje seg realistiske mål, som ein i neste omgang må vere tru mot. Og ikkje minst må ein, i den grad det er mogleg, *vise fram* resultata.

5 Utfordringar i framtida

Heilt til slutt vil eg kort nemne nokre av dei utfordringane ein står overfor etter 2014. Når det gjeld dokumentasjonsverksemd generelt er det, som eg alt har vore inne på, mykje som står att. Dei største utfordringane knyter seg til tilgjengeleggjering av materiale, og korleis ein skal samordne og samkøyre det digitale materialet som allereie finst ved ulike universitet og høgskular. Utsiktene til å byggje opp felles, nasjonale arkiv er noko ein òg må arbeide parallelt med i samband med dette. Når ein har framgang med ei betre samordning og samkøyring av allereie eksisterande materiale, vil det i neste omgang opne for gjenbruk av ulike digitale ressursar og meir samarbeid på tvers av institusjonar.

Når det gjeld den dokumentasjons*leksikografiske* verksemda, er det òg mange utfordringar i framtida. Den kanskje mest openberre utfordringa gjeld forholdet mellom papirbasert og databasebasert (eller elektronisk) ordboksverksemd. Ettersom målet for *Norsk Ordbok* først og fremst har vore – og framleis er – å fullføre ordboksverket i papirutgåve, har mykje av tankegangen vore knytt til papirbasert ordboksverksemd. Men grunnlaget for den andre, elektroniske ordboksverksemda er allereie langt på veg på plass. Prosjektet Norsk

Ordbok 2014 har tilgang til og rår over mange verkty som er viktige i dette arbeidet. Som nemnt tidlegare er ordboka ein kompleks database, og redigeringa går føre seg på ei fullelektronisk plattform. I praksis vil det seie at prosjektet Norsk Ordbok 2014 allereie arbeider i eit databasebasert format, som ein i dag og fram til og med 2014 ”konverterer” til eit papirbasert format. Den informasjonen som ligg i basen, kan utnyttast på svært mange ulike måtar. Det neste steget er å gå over til fullelektroniske ordbøker og ordbasar. I arbeidet med dette må òg forlaga kome sterkare på banen.

Mykje er allereie gjort, men det er likevel mykje å ta fatt i. Særleg viktig er det at planlegginga av tida etter 2014 ikkje kan byrje i 2013, det må pågå parallelt med det arbeidet som blir gjort no. Framtida har i grunnen alt byrja.

Litteratur

- Andersen, Per Thomas et al. ”Nordisk i nåtid. Innstilling fra oppfølgingsutvalget etter evalueringen av nordisk forskning i språk og litteratur.” Noregs forskingsråd, 2007.
- St.meld. nr. 15 (1968-1969). *Om språksaka* (kapittel II: Norsk leksikografisk institutt. Skiping og arbeidsoppgåver, s. 6-13)
- St.meld. nr. 35 (2007-2008). *Mål og meining. Ein heilskapleg norsk språkpolitikk*. Nettversjonen av dokumentet finst på:
<http://www.regjeringen.no/nn/dep/kud/Dokument/proposisjon-ar-og-meldingar/stortingsmeldingar/2007-2008/stmeld-nr-35-2007-2008-.html?id=519923>